POWER GENERATION, OPERATION, AND CONTROL

Third Edition

Allen J. Wood • Bruce F. Wollenberg
Gerald B. Sheblé

WWW. LINK AVAILABLE

WILEY

www.mepcafe.com

www.mepcafe.com

POWER GENERATION, OPERATION, AND CONTROL

www.mepcafe.com

www.mepcafe.com

POWER GENERATION, OPERATION, AND CONTROL

THIRD EDITION

Allen J. Wood

Bruce F. Wollenberg

Gerald B. Sheblé

IEEE

WILEY

www.mepcafe.com

Cover illustration: Xcel Energy
Copyright © 2014 by John Wiley & Sons, Inc. All rights reserved

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at http://www.wiley.com/go/permission.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data

Wood, Allen J., author.
Power generation, operation, and control. – Third edition / Allen J. Wood,
Bruce F. Wollenberg, Gerald B. Sheblé.
pages cm
Includes bibliographical references and index.
ISBN 978-0-471-79055-6 (hardback)
1. Electric power systems. I. Wollenberg, Bruce F., author. II. Sheblé, Gerald B., author. III. Title.
TK1001.W64 2013
621.31–dc23

2013013050

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

www.mepcafe.com

Allen Wood passed away on September 10, 2011, during the preparation of this edition. Al was my professor when I was a student in the Electric Power Engineering Program at Rensselaer Polytechnic Institute (RPI) in 1966. Allen Wood and other engineers founded Power Technologies Inc. (PTI) in Schenectady, NY, in 1969. I joined PTI in 1974, and Al recruited me to help teach the course at RPI in 1979. The original text was the outcome of student notes assembled over a 5 year period from 1979 to 1984 and then turned over to John Wiley & Sons. Allen Wood was my professor, my mentor, and my friend, and I dedicate this third edition to him.

BRUCE F. WOLLENBERG

I dedicate this work to my family, my wife Yvette Sheblé, my son Jason Sheblé, my daughter Laura Sheblé, and grandson Kiyan, as they helped me so much to complete this work.

GERALD B. SHEBLÉ

www.mepcafe.com

www.mepcafe.com

CONTENTS

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

PREFACE TO THE THIRD EDITION

It has now been 17 years from the second edition (and a total of 28 years from the publishing of the first edition of this text). To say that much has changed is an understatement. As noted in the dedication, Allen Wood passed away during the preparation of this edition and a new coauthor, Gerald Sheblé, has joined Bruce Wollenberg in writing the text. Dr. Sheblé brings an expertise that is both similar and different from that of Dr. Wollenberg to this effort, and the text clearly shows a new breadth in topics covered.

The second edition was published in 1996, which was in the midst of the period of "deregulation" or more accurately "reregulation" of the electric industry both in the United States and worldwide. New concepts such as electric power spot markets, Independent System Operators (ISOs) in the United States, and independent generation, transmission, and distribution companies are now common. Power system control centers have become much larger and cover a much larger geographic area as markets have expanded. The U.S. government has partnered with the North American Electric Reliability Corporation (formerly the North American Electric Reliability Council) and has begun a much tighter governance of electric company practices as they affect the system's reliability and security since the events of 9/11.

We have added several new chapters to the text to both reflect the increased importance of the topics covered and broaden the educational and engineering value of the book. Both Sheblé and Wollenberg are professors at major universities and have developed new examples, problems, and software for the text. Both Wollenberg and Sheblé are consultants and expert witnesses to the electric energy industry. We hope this effort is of value to the readers.

Today, students and working engineers have access to much more information directly through the Internet, and if they are IEEE members can access the very extensive IEEE Explore holdings directly from their home or office computers. Thus, we felt it best not to attempt to provide lists of references as was done in earlier editions.

We would like to extend our thanks to those students who provided excellent programming and development skills to difficult problems as they performed research tasks under our direction. Among them are Mohammad Alsaffar and Anthony Giacomoni at the University of Minnesota; George Fahd, Dan Richards,

www.mepcafe.com

Thomas Smed, and David Walters at Auburn University; and Darwin Anwar, Somgiat Dekrajangpetch, Kah-Hoe Ng, Jayant Kumar, James Nicolaisen, Chuck Richter, Douglas Welch, Hao Wu, and Weiguo Yang at Iowa State University; Chin-Chuen Teoh, Mei P. Cheong, and Gregory Bingham at Portland State University; Zhenyu Wan at University of South Wales.

Last of all, we announce that we are planning to write a sequel to the third edition in which many of the business aspects of the electric power industry will be presented, along with major chapters on topics such as extended auction mechanisms and reliability.

BRUCE F. WOLLENBERG
GERALD B. SHEBLÉ

www.mepcafe.com

PREFACE TO THE SECOND EDITION

It has been 11 years since the first edition was published. Many developments have taken place in the area covered by this text and new techniques have been developed that have been applied to solve old problems. Computing power has increased dramatically, permitting the solution of problems that were previously left as being too expensive to tackle. Perhaps the most important development is the changes that are taking place in the electric power industry with new, nonutility participants playing a larger role in the operating decisions.

It is still the intent of the authors to provide an introduction to this field for senior or first-year graduate engineering students. The authors have used the text material in a one-semester (or two-quarter) program for many years. The same difficulties and required compromises keep occurring. Engineering students are very comfortable with computers but still do not usually have an appreciation of the interaction of human and economic factors in the decisions to be made to develop "optimal" schedules, whatever that may mean. In 1995, most of these students are concurrently being exposed to courses in advanced calculus and courses that explore methods for solving power flow equations. This requires some coordination. We have also found that very few of our students have been exposed to the techniques and concepts of operations research, necessitating a continuing effort to make them comfortable with the application of optimization methods. The subject area of this book is an excellent example of optimization applied in an important industrial system.

The topic areas and depth of coverage in this second edition are about the same as in the first, with one major change. Loss formulae are given less space and supplemented by a more complete treatment of the power-flow-based techniques in a new chapter that treats the optimal power flow (OPF). This chapter has been put at the end of the text. Various instructors may find it useful to introduce parts of this material earlier in the sequence; it is a matter of taste, plus the requirement to coordinate with other course coverage. (It is difficult to discuss the OPF when the students do not know the standard treatment for solving the power flow equations.)

The treatment of unit commitment has been expanded to include the Lagrange relaxation technique. The chapter on production costing has been revised to change the emphasis and introduce new methods. The market structures for bulk power transactions have undergone important changes throughout the world. The chapter

www.mepcafe.com

on interchange transactions is a "progress report" intended to give the students an appreciation of the complications that may accompany a competitive market for the generation of electric energy. The sections on security analysis have been updated to incorporate an introduction to the use of bounding techniques and other contingency selection methods. Chapter 13 on the OPF includes a brief coverage of the security-constrained OPF and its use in security control.

The authors appreciate the suggestions and help offered by professors who have used the first edition, and our students. (Many of these suggestions have been incorporated; some have not, because of a lack of time, space, or knowledge.) Many of our students at Rensselaer Polytechnic Institute (RPI) and the University of Minnesota have contributed to the correction of the first edition and undertaken hours of calculations for homework solutions, checked old examples, and developed data for new examples for the second edition. The 1994 class at RPI deserves special and honorable mention. They were subjected to an early draft of the revision of Chapter 8 and required to proofread it as part of a tedious assignment. They did an outstanding job and found errors of 10 to 15 years standing. (A note of caution to any of you professors that think of trying this; it requires more work than you might believe. How would you like 20 critical editors for your lastest, glorious tome?)

Our thanks to Kuo Chang, of Power Technologies, Inc., who ran the computations for the bus marginal wheeling cost examples in Chapter 10. We would also like to thank Brian Stott, of Power Computer Applications, Corp., for running the OPF examples in Chapter 13.

ALLEN J. WOOD
BRUCE F. WOLLENBERG

www.mepcafe.com

PREFACE TO THE FIRST EDITION

The fundamental purpose of this text is to introduce and explore a number of engineering and economic matters involved in planning, operating, and controlling power generation and transmission systems in electric utilities. It is intended for first-year graduate students in electric power engineering. We believe that it will also serve as a suitable self-study text for anyone with an undergraduate electrical engineering education and an understanding of steady-state power circuit analysis.

This text brings together material that has evolved since 1966 in teaching a graduate-level course in the electric power engineering department at Rensselaer Polytechnic Institute (RPI). The topics included serve as an effective means to introduce graduate students to advanced mathematical and operations research methods applied to practical electric power engineering problems. Some areas of the text cover methods that are currently being applied in the control and operation of electric power generation systems. The overall selection of topics, undoubtedly, reflects the interests of the authors.

In a one-semester course it is, of course, impossible to consider all the problems and "current practices" in this field. We can only introduce the types of problems that arise, illustrate theoretical and practical computational approaches, and point the student in the direction of seeking more information and developing advanced skills as they are required.

The material has regularly been taught in the second semester of a first-year graduate course. Some acquaintance with both advanced calculus methods (e.g., Lagrange multipliers) and basic undergraduate control theory is needed. Optimization methods are introduced as they are needed to solve practical problems and used without recourse to extensive mathematical proofs. This material is intended for an engineering course: mathematical rigor is important but is more properly the province of an applied or theoretical mathematics course. With the exception of Chapter 12, the text is self-contained in the sense that the various applied mathematical techniques are presented and developed as they are utilized. Chapter 12, dealing with state estimation, may require more understanding of statistical and probabilistic methods than is provided in the text.

The first seven chapters of the text follow a natural sequence, with each succeeding chapter introducing further complications to the generation scheduling problem and new solution techniques. Chapter 8 treats methods used in generation system

www.mepcafe.com

planning and introduces probabilistic techniques in the computation of fuel consumption and energy production costs. Chapter 8 stands alone and might be used in any position after the first seven chapters. Chapter 9 introduces generation control and discusses practices in modern U.S. utilities and pools. We have attempted to provide the "big picture" in this chapter to illustrate how the various pieces fit together in an electric power control system.

The topics of energy and power interchange between utilities and the economic and scheduling problems that may arise in coordinating the economic operation of interconnected utilities are discussed in Chapter 10. Chapters 11 and 12 are a unit. Chapter 11 is concerned with power system security and develops the analytical framework used to control bulk power systems in such a fashion that security is enhanced. Everything, including power systems, seems to have a propensity to fail. Power system security practices try to control and operate power systems in a defensive posture so that the effects of these inevitable failures are minimized. Finally, Chapter 12 is an introduction to the use of state estimation in electric power systems. We have chosen to use a maximum likelihood formulation since the quantitative measurement–weighting functions arise in a natural sense in the course of the development.

Each chapter is provided with a set of problems and an annotated reference list for further reading. Many (if not most) of these problems should be solved using a digital computer. At RPI, we are able to provide the students with some fundamental programs (e.g., a load flow, a routine for scheduling of thermal units). The engineering students of today are well prepared to utilize the computer effectively when access to one is provided. Real bulk power systems have problems that usually call forth Dr. Bellman's curse of dimensionality—computers help and are essential to solve practical-sized problems.

The authors wish to express their appreciation to K. A. Clements, H. H. Happ, H. M. Merrill, C. K. Pang, M. A. Sager, and J. C. Westcott, who each reviewed portions of this text in draft form and offered suggestions. In addition, Dr. Clements used earlier versions of this text in graduate courses taught at Worcester Polytechnic Institute and in a course for utility engineers taught in Boston, Massachusetts.

Much of the material in this text originated from work done by our past and current associates at Power Technologies, Inc., the General Electric Company, and Leeds and Northrup Company. A number of IEEE papers have been used as primary sources and are cited where appropriate. It is not possible to avoid omitting, references and sources that are considered to be significant by one group or another. We make no apology for omissions and only ask for indulgence from those readers whose favorites have been left out. Those interested may easily trace the references back to original sources.

We would like to express our appreciation for the fine typing job done on the original manuscript by Liane Brown and Bonnalyne MacLean.

This book is dedicated in general to all of our teachers, both professors and associates, and in particular to Dr. E. T. B. Gross.

ALLEN J. WOOD
BRUCE F. WOLLENBERG

www.mepcafe.com

ACKNOWLEDGMENT

I am indebted to a number of mentors who have encouraged me and shown the path toward development: Homer Brown, Gerry Heydt, Pete Sauer, Ahmed El-Abiad, K Neal Stanton, Robin Podmore, Ralph Masiello, Anjan Bose, Jerry Russel, Leo Grigsby, Arun Phadke, Saifur Rahman, Aziz Fouad, Vijay Vittal, and Mani Venkata. They have often advised at just the right time with the right perspective on development. My coauthor, Bruce, has often provided mentorship and friendship over the last several decades. I have had the luxury of working with many collaborators and the good fortune of learning and of experiencing other viewpoints. I especially thank: Arnaud Renaud, Mark O'Malley, Walter Hobbs, João Abel Peças Lopes, Manuel Matos, Vladimiro Miranda, João Tomé Saraiva, and Vassilios G. Agelidis.

Gerald B. Sheblé

www.mepcafe.com

About the companion website

The University of Minnesota offers a set of online courses in power systems and related topics. One of the courses is based on this book. For further information, visit

http://www.cusp.umn.edu

and click on the link for the course.

A companion site containing additional resources for students, and an Instructor's site with solutions to problems found in the text, can be found at

http://www.wiley.com/go/powergenoperation.

www.mepcafe.com

1

INTRODUCTION

1.1 PURPOSE OF THE COURSE

The objectives of a first-year, one-semester graduate course in electric power generation, operation, and control include the desire to:

1. Acquaint electric power engineering students with power generation systems, their operation in an economic mode, and their control.
2. Introduce students to the important "terminal" characteristics for thermal and hydroelectric power generation systems.
3. Introduce mathematical optimization methods and apply them to practical operating problems.
4. Introduce methods for solving complicated problems involving both economic analysis and network analysis and illustrate these techniques with relatively simple problems.
5. Introduce methods that are used in modern control systems for power generation systems.
6. Introduce "current topics": power system operation areas that are undergoing significant, evolutionary changes. This includes the discussion of new techniques for attacking old problems and new problem areas that are arising from changes in the system development patterns, regulatory structures, and economics.

Power Generation, Operation, and Control, Third Edition. Allen J. Wood, Bruce F. Wollenberg, and Gerald B. Sheblé.
© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

www.mepcafe.com

1.2 COURSE SCOPE

Topics to be addressed include

1. Power generation characteristics
2. Electric power industry as a business
3. Economic dispatch and the general economic dispatch problem
4. Thermal unit economic dispatch and methods of solution
5. Optimization with constraints
6. Optimization methods such as linear programming, dynamic programming, nonlinear optimization, integer programming, and interior point optimization
7. Transmission system effects
   a. Power flow equations and solutions
   b. Transmission losses
   c. Effects on scheduling
8. The unit commitment problem and solution methods
   a. Dynamic programming
   b. Lagrange relaxation
   c. Integer programming
9. Generation scheduling in systems with limited energy supplies including fossil fuels and hydroelectric plants, need to transport energy supplies over networks such as pipelines, rail networks, and river/reservoir systems, and power system security techniques
10. Optimal power flow techniques
11. Power system state estimation
12. Automatic generation control
13. Interchange of power and energy, power pools and auction mechanisms, and modern power markets
14. Load forecasting techniques

In many cases, we can only provide an introduction to the topic area. Many additional problems and topics that represent important, practical problems would require more time and space than is available. Still others, such as light-water moderated reactors and cogeneration plants, could each require several chapters to lay a firm foundation. We can offer only a brief overview and introduce just enough information to discuss system problems.

1.3 ECONOMIC IMPORTANCE

The efficient and optimum economic operation and planning of electric power generation systems have always occupied an important position in the electric power industry. Prior to 1973 and the oil embargo that signaled the rapid escalation in fuel

www.mepcafe.com

prices, electric utilities in the United States spent about 20% of their total revenues on fuel for the production of electrical energy. By 1980, that figure had risen to more than 40% of the total revenues. In the 5 years after 1973, U.S. electric utility fuel costs escalated at a rate that averaged 25% compounded on an annual basis. The efficient use of the available fuel is growing in importance, both monetarily and because most of the fuel used represents irreplaceable natural resources.

An idea of the magnitude of the amounts of money under consideration can be obtained by considering the annual operating expenses of a large utility for purchasing fuel. Assume the following parameters for a moderately large system:

Annual peak load: 10,000 MW

Annual load factor: 60%

Average annual heat rate for converting fuel to electric energy: 10,500 Btu/kWh

Average fuel cost: \$3.00 per million Btu (MBtu), corresponding to oil priced at 18\$/bbl

With these assumptions, the total annual fuel cost for this system is as follows:

Annual energy produced: 10^7 kW × 8760 h/year × 0.60 = 5.256 × 10^{10} kWh

Annual fuel consumption: 10,500 Btu/kWh × 5.256 × 10^{10} kWh = 55.188 × 10^{13} Btu

Annual fuel cost: 55.188 × 10^{13} Btu × 3 × 10^{-6} \$/Btu = \$1.66 billion

To put this cost in perspective, it represents a direct requirement for revenues from the average customer of this system of 3.15 cents/kWh just to recover the expense for fuel.

A savings in the operation of this system of a small percent represents a significant reduction in operating cost as well as in the quantities of fuel consumed. It is no wonder that this area has warranted a great deal of attention from engineers through the years.

Periodic changes in basic fuel price levels serve to accentuate the problem and increase its economic significance. Inflation also causes problems in developing and presenting methods, techniques, and examples of the economic operation of electric power generating systems.

1.4 DEREGULATION: VERTICAL TO HORIZONTAL

In the 1990s, many electric utilities including government-owned electric utilities, private investor–owned electric utilities were “deregulated.” This has had profound effects on the operation of electric systems where implemented. This topic is dealt with in an entire chapter of its own in this text as Chapter 2.

1.5 PROBLEMS: NEW AND OLD

This text represents a progress report in an engineering area that has been and is still undergoing rapid change. It concerns established engineering problem areas (i.e., economic dispatch and control of interconnected systems) that have taken on new

www.mepcafe.com

importance in recent years. The original problem of economic dispatch for thermal systems was solved by numerous methods years ago. Recently there has been a rapid growth in applied mathematical methods and the availability of computational capability for solving problems of this nature so that more involved problems have been successfully solved.

The classic problem is the economic dispatch of fossil-fired generation systems to achieve minimum operating cost. This problem area has taken on a subtle twist as the public has become increasingly concerned with environmental matters, so "economic dispatch" now includes the dispatch of systems to minimize pollutants and conserve various forms of fuel, as well as to achieve minimum costs. In addition, there is a need to expand the limited economic optimization problem to incorporate constraints on system operation to ensure the "security" of the system, thereby preventing the collapse of the system due to unforeseen conditions. The hydrothermal coordination problem is another optimum operating problem area that has received a great deal of attention. Even so, there are difficult problems involving hydrothermal coordination that cannot be solved in a theoretically satisfying fashion in a rapid and efficient computational manner.

The post–World War II period saw the increasing installation of pumped-storage hydroelectric plants in the United States and a great deal of interest in energy storage systems. These storage systems involve another difficult aspect of the optimum economic operating problem. Methods are available for solving coordination of hydroelectric, thermal, and pumped-storage electric systems. However, closely associated with this economic dispatch problem is the problem of the proper commitment of an array of units out of a total array of units to serve the expected load demands in an "optimal" manner.

A great deal of progress and change has occurred in the 1985–1995 decade. Both the unit commitment and optimal economic maintenance scheduling problems have seen new methodologies and computer programs developed. Transmission losses and constraints are integrated with scheduling using methods based on the incorporation of power flow equations in the economic dispatch process. This permits the development of optimal economic dispatch conditions that do not result in overloading system elements or voltage magnitudes that are intolerable. These "optimal power flow" techniques are applied to scheduling both real and reactive power sources as well as establishing tap positions for transformers and phase shifters.

In recent years, the political climate in many countries has changed, resulting in the introduction of more privately owned electric power facilities and a reduction or elimination of governmentally sponsored generation and transmission organizations. In some countries, previously nationwide systems have been privatized. In both these countries and in countries such as the United States, where electric utilities have been owned by a variety of bodies (e.g., consumers, shareholders, as well as government agencies), there has been a movement to introduce both privately owned generation companies and larger cogeneration plants that may provide energy to utility customers. These two groups are referred to as independent power producers (IPPs). This trend is coupled with a movement to provide access to the transmission

www.mepcafe.com

system for these nonutility power generators as well as to other interconnected utilities. The growth of an IPP industry brings with it a number of interesting operational problems. One example is the large cogeneration plant that provides steam to an industrial plant and electric energy to the power system. The industrial-plant steam demand schedule sets the operating pattern for the generating plant, and it may be necessary for a utility to modify its economic schedule to facilitate the industrial generation pattern.

Transmission access for nonutility entities (consumers as well as generators) sets the stage for the creation of new market structures and patterns for the interchange of electric energy. Previously, the major participants in the interchange markets in North America were electric utilities. Where nonutility, generation entities or large consumers of power were involved, local electric utilities acted as their agents in the marketplace. This pattern is changing. With the growth of nonutility participants and the increasing requirement for access to transmission has come a desire to introduce a degree of economic competition into the market for electric energy. Surely this is not a universally shared desire; many parties would prefer the status quo. On the other hand, some electric utility managements have actively supported the construction, financing, and operation of new generation plants by nonutility organizations and the introduction of less-restrictive market practices.

The introduction of nonutility generation can complicate the scheduling–dispatch problem. With only a single, integrated electric utility operating both the generation and transmission systems, the local utility could establish schedules that minimized its own operating costs while observing all of the necessary physical, reliability, security, and economic constraints. With multiple parties in the bulk power system (i.e., the generation and transmission system), new arrangements are required. The economic objectives of all of the parties are not identical, and, in fact, may even be in direct (economic) opposition. As this situation evolves, different patterns of operation may result in different regions. Some areas may see a continuation of past patterns where the local utility is the dominant participant and continues to make arrangements and schedules on the basis of minimization of the operating cost that is paid by its own customers. Centrally dispatched power pools could evolve that include nonutility generators, some of whom may be engaged in direct sales to large consumers. Other areas may have open market structures that permit and facilitate competition with local utilities. Both local and remote nonutility entities, as well as remote utilities, may compete with the local electric utility to supply large industrial electric energy consumers or distribution utilities. The transmission system may be combined with a regional control center in a separate entity. Transmission networks could have the legal status of "common carriers," where any qualified party would be allowed access to the transmission system to deliver energy to its own customers, wherever they might be located. This very nearly describes the current situation in Great Britain.

What does this have to do with the problems discussed in this text? A *great deal*. In the extreme cases mentioned earlier, many of the dispatch and scheduling methods we are going to discuss will need to be rethought and perhaps drastically revised. Current practices in automatic generation control are based on tacit

www.mepcafe.com

assumptions that the electric energy market is slow moving with only a few, more-or-less fixed, interchange contracts that are arranged *between interconnected utilities*. Current techniques for establishing optimal economic generation schedules are really based on the assumption of a single utility serving the electric energy needs of its own customers at minimum cost. Interconnected operations and energy interchange agreements are presently the result of interutility arrangements: all of the parties share common interests. In a world with a transmission-operation entity required to provide access to many parties, both utility and nonutility organizations, this entity has the task of developing operating schedules to accomplish the deliveries scheduled in some (as yet to be defined) "optimal" fashion within the physical constraints of the system, while maintaining system reliability and security. If all (or any) of this develops, it should be a fascinating time to be active in this field.

1.6 CHARACTERISTICS OF STEAM UNITS

In analyzing the problems associated with the controlled operation of power systems, there are many possible parameters of interest. Fundamental to the economic operating problem is the set of input–output characteristics of a thermal power generation unit. A typical boiler–turbine–generator unit is sketched in Figure 1.1. This unit consists of a single boiler that generates steam to drive a single turbine–generator set. The electrical output of this set is connected not only to the electric power system, but also to the auxiliary power system in the power plant. A typical steam turbine unit may require 2–6% of the gross output of the unit for the auxiliary power requirements necessary to drive boiler feed pumps, fans, condenser circulating water pumps, and so on. In defining the unit characteristics, we will talk about *gross* input versus *net* output. That is, gross input to the plant represents the total input, whether measured in terms of dollars per hour or tons of coal per hour or millions of cubic feet of gas per hour, or any other units. The net output of the plant is the electrical power output available to the electric utility system. Occasionally, engineers will develop gross input–gross output characteristics. In such situations, the data should be converted to net output to be more useful in scheduling the generation.

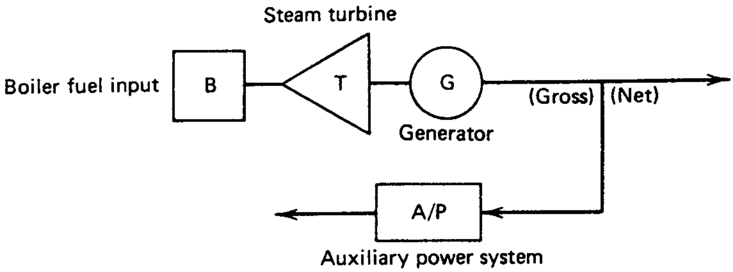


FIGURE 1.1 Boiler–turbine–generator unit.

www.mepcafe.com

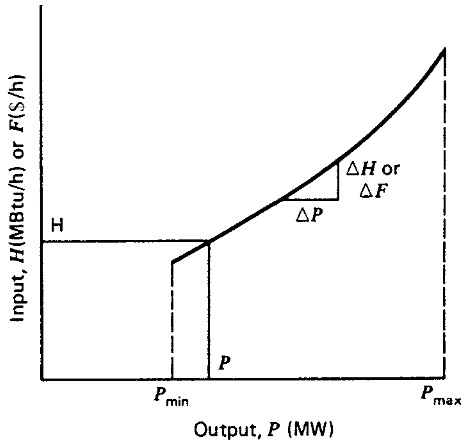


FIGURE 1.2 Input–output curve of a steam turbine generator.

In defining the characteristics of steam turbine units, the following terms will be used:

$$H = \text{Btu per hour heat input to the unit } (\text{or MBtu/h})$$
$$F = \text{Fuel cost times } H \text{ is the \$ per hour } (\$/\text{h}) \text{ input to the unit for fuel}$$

Occasionally, the $/h operating cost rate of a unit will include prorated operation and maintenance costs. That is, the labor cost for the operating crew will be included as part of the operating cost if this cost can be expressed directly as a function of the output of the unit. The output of the generation unit will be designated by P, the megawatt net output of the unit. Figure 1.2 shows the input–output characteristic of a steam unit in idealized form. The input to the unit shown on the ordinate may be either in terms of heat energy requirements [millions of Btu per hour (MBtu/h)] or in terms of total cost per hour ($/h). The output is normally the net electrical output of the unit. The characteristic shown is idealized in that it is presented as a smooth, convex curve.

These data may be obtained from design calculations or from heat rate tests. When heat rate test data are used, it will usually be found that the data points do not fall on a smooth curve. Steam turbine generating units have several critical operating constraints. Generally, the minimum load at which a unit can operate is influenced more by the steam generator and the regenerative cycle than by the turbine. The only critical parameters for the turbine are shell and rotor metal differential temperatures, exhaust hood temperature, and rotor and shell expansion. Minimum load limitations are generally caused by fuel combustion stability and inherent steam generator design constraints. For example, most supercritical units cannot operate below 30% of design capability. A minimum flow of 30% is required to cool the tubes in the furnace of the steam generator adequately. Turbines do not have any inherent overload

www.mepcafe.com

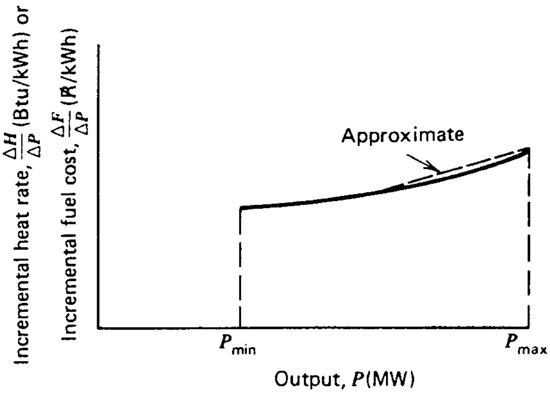


FIGURE 1.3 Incremental heat (cost) rate characteristic.

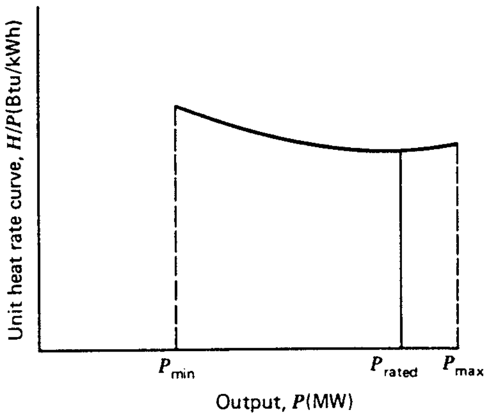


FIGURE 1.4 Net heat rate characteristic of a steam turbine generator unit.

capability, so the data shown on these curves normally do not extend much beyond 5% of the manufacturer's stated valve-wide-open capability.

The incremental heat rate characteristic for a unit of this type is shown in Figure 1.3 This incremental heat rate characteristic is the slope (the derivative) of the input–output characteristic ($\Delta H/\Delta P$ or $\Delta F/\Delta P$). The data shown on this curve are in terms of Btu/kWh (or \$/kWh) versus the net power output of the unit in megawatts. This characteristic is widely used in economic dispatching of the unit. It is converted to an incremental fuel cost characteristic by multiplying the incremental heat rate in Btu per kilowatt hour by the equivalent fuel cost in terms of \$/Btu. Frequently, this characteristic is approximated by a sequence of straight-line segments.

The last important characteristic of a steam unit is the unit (net) heat rate characteristic shown in Figure 1.4. This characteristic is H/P versus P. It is proportional to the reciprocal of the usual efficiency characteristic developed for

www.mepcafe.com

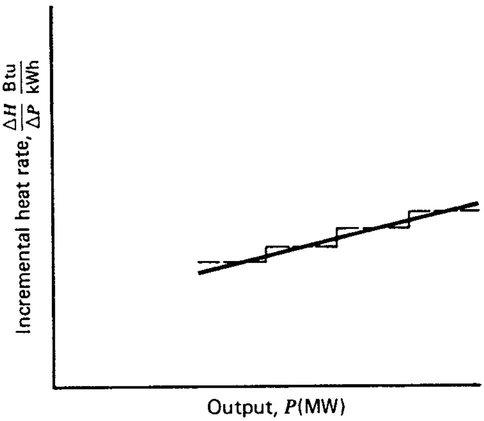


FIGURE 1.5 Approximate representations of the incremental heat rate curve.

machinery. The unit heat rate characteristic shows the heat input per kilowatt hour of output versus the megawatt output of the unit. Typical conventional steam turbine units are between 30 and 35% efficient, so their unit heat rates range between approximately 11,400 Btu/kWh and 9,800 Btu/kWh. (A kilowatt hour has a thermal equivalent of approximately 3412 Btu.) Unit heat rate characteristics are a function of unit design parameters such as initial steam conditions, stages of reheat and the reheat temperatures, condenser pressure, and the complexity of the regenerative feed-water cycle. These are important considerations in the establishment of the unit's efficiency. For purposes of estimation, a typical heat rate of 10,500 Btu/kWh may be used occasionally to approximate actual unit heat rate characteristics.

Many different formats are used to represent the input–output characteristic shown in Figure 1.2. The data obtained from heat rate tests or from the plant design engineers may be fitted by a polynomial curve. In many cases, quadratic characteristics have been fit to these data. A series of straight-line segments may also be used to represent the input–output characteristics. The different representations will, of course, result in different incremental heat rate characteristics. Figure 1.5 shows two such variations. The solid line shows the incremental heat rate characteristic that results when the input versus output characteristic is a quadratic curve or some other continuous, smooth, convex function. This incremental heat rate characteristic is monotonically increasing as a function of the power output of the unit. The dashed lines in Figure 1.5 show a stepped incremental characteristic that results when a series of straight-line segments are used to represent the input–output characteristics of the unit. The use of these different representations may require that different scheduling methods be used for establishing the optimum economic operation of a power system. Both formats are useful, and both may be represented by tables of data. Only the first, the solid line, may be represented by a continuous analytic function, and only the first has a derivative that is nonzero. (That is, d^2F/d^2P equals 0 if dF/dP is constant.)

www.mepcafe.com

At this point, it is necessary to take a brief detour to discuss the heating value of the fossil fuels used in power generation plants. Fuel heating values for coal, oil, and gas are expressed in terms of Btu/lb or joules per kilogram of fuel. The determination is made under standard, specified conditions using a *bomb calorimeter*.

This is all to the good except that there are *two* standard determinations specified:

- The higher heating value of the fuel (HHV) assumes that the water vapor in the combustion process products condenses and therefore includes the latent heat of vaporization in the products.
- The lower heating value of the fuel (LHV) does not include this latent heat of vaporization.

The difference between the HHV and LHV for a fuel depends on the hydrogen content of the fuel. Coal fuels have a low hydrogen content with the result that the difference between the HHV and LHV for a fuel is fairly small. (A typical value of the difference for a bituminous coal would be of the order of 3%. The HHV might be 14,800 Btu/lb and the LHV 14,400 Btu/lb.) Gas and oil fuels have a much higher hydrogen content, with the result that the relative difference between the HHV and LHV is higher; typically on the order of 10 and 6%, respectively. This gives rise to the possibility of some confusion when considering unit efficiencies and cycle energy balances. (A more detailed discussion is contained in the book by El-Wakil, [reference 1].)

A uniform standard must be adopted so that everyone uses the same heating value standard. In the United States, the standard is to use the HHV *except that engineers and manufacturers that are dealing with combustion turbines* (*i.e., gas turbines*) *normally use* LHVs *when quoting heat rates or efficiencies*. In European practice, LHVs are used for all specifications of fuel consumption and unit efficiency. In this text, HHVs are used throughout the book to develop unit characteristics. Where combustion turbine data have been converted by the authors from LHVs to HHVs, a difference of 10% was normally used. When in doubt about which standard for the fuel heating value has been used to develop unit characteristics—*ask*!

1.6.1 Variations in Steam Unit Characteristics

A number of different steam unit characteristics exist. For large steam turbine generators the input–output characteristics shown in Figure 1.2 are not always as smooth as indicated there. Large steam turbine generators will have a number of steam admission valves that are opened in sequence to obtain ever-increasing output of the unit. Figure 1.6 shows both an input–output and an incremental heat rate characteristic for a unit with four valves. As the unit loading increases, the input to the unit increases and the incremental heat rate decreases between the opening points for any two valves. However, when a valve is first opened, the throttling losses increase rapidly and the incremental heat rate rises suddenly. This gives rise to the discontinuous type of incremental heat rate characteristic shown in Figure 1.6. It is possible to use this

www.mepcafe.com

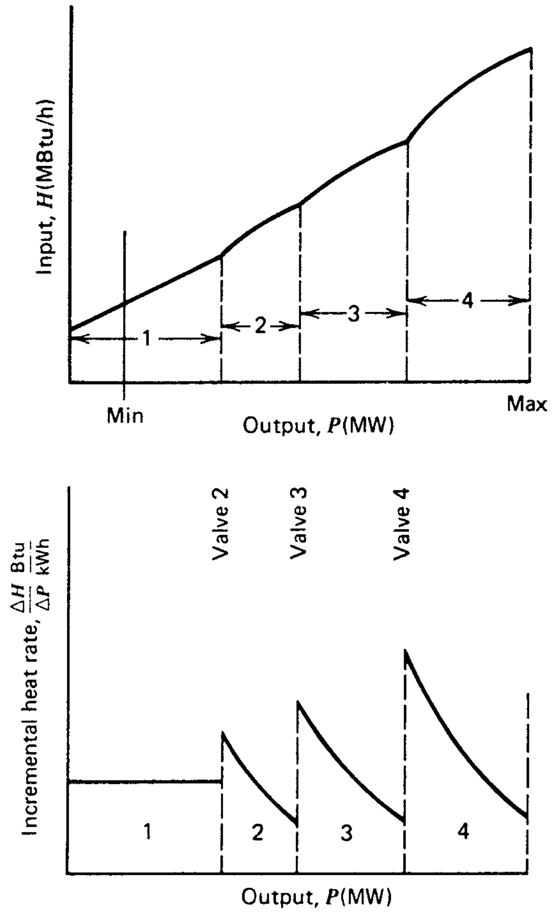


FIGURE 1.6 Characteristics of a steam turbine generator with four steam admission valves.

type of characteristic in order to schedule steam units, although it is usually not done. This type of input–output characteristic is nonconvex; hence, optimization techniques that require convex characteristics may not be used with impunity.

Another type of steam unit that may be encountered is the *common-header plant*, which contains a number of different boilers connected to a common steam line (called a common header). Figure 1.7 is a sketch of a rather complex common-header plant. In this plant, there are not only a number of boilers and turbines, each connected to the common header, but also a "topping turbine" connected to the common header. A *topping turbine* is one in which steam is exhausted from the turbine and fed not to a condenser but to the common steam header.

A common-header plant will have a number of different input–output characteristics that result from different combinations of boilers and turbines connected to the header. Steinberg and Smith (reference 2) treat this type of plant quite extensively. Common-header plants were constructed originally not only to provide a large

www.mepcafe.com

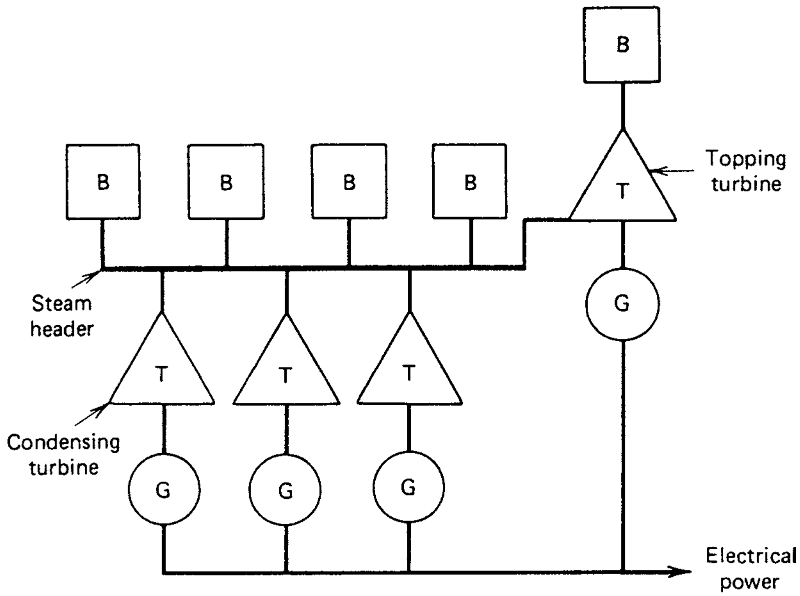


FIGURE 1.7 A common-header steam plant.

electrical output from a single plant but also to provide steam sendout for the heating and cooling of buildings in dense urban areas. After World War II, a number of these plants were modernized by the installation of the type of topping turbine shown in Figure 1.7. For a period of time during the 1960s, these common-header plants were being dismantled and replaced by modern, efficient plants. However, as urban areas began to reconstruct, a number of metropolitan utilities found that their steam loads were growing and that the common-header plants could not be dismantled but had to be expected to provide steam supplies to new buildings.

Combustion turbines (gas turbines) are also used to drive electric generating units. Some types of power generation units have been derived from aircraft gas turbine units and others from industrial gas turbines that have been developed for applications like driving pipeline pumps. In their original applications, these two types of combustion turbines had dramatically different duty cycles. Aircraft engines see relatively short duty cycles where power requirements vary considerably over a flight profile. Gas turbines in pumping duty on pipelines would be expected to operate almost continuously throughout the year. Service in power generation may require both types of duty cycle.

Gas turbines are applied in both a simple cycle and in combined cycles. In the simple cycle, inlet air is compressed in a rotating compressor (typically by a factor of 10–12 or more) and then mixed and burned with fuel oil or gas in a combustion chamber. The expansion of the high-temperature gaseous products in the turbine drives the compressor, turbine, and generator. Some designs use a single shaft for the turbine and compressor, with the generator being driven through a suitable set of

www.mepcafe.com

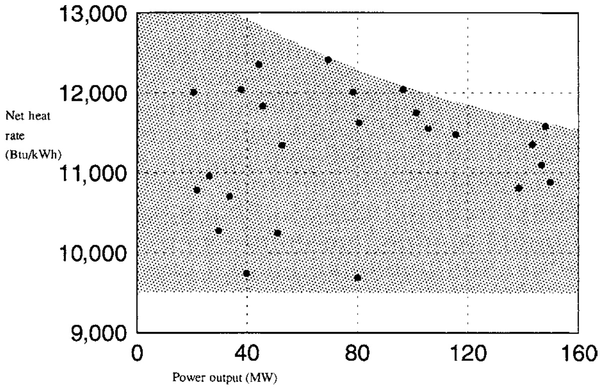


FIGURE 1.8 Approximate net heat rates for a range of simple cycle gas turbine units. Units are fired by natural gas and represent performance at standard conditions of an ambient temperature of 15°C at sea level. (Heat rate data from reference 1 were adjusted by 13% to represent HHVs and auxiliary power needs).

gears. In larger units the generators are driven directly, without any gears. Exhaust gases are discharged to the atmosphere in the simple cycle units. In combined cycles, the exhaust gases are used to make steam in a heat-recovery steam generator (HRSG) before being discharged.

The early utility applications of simple cycle gas turbines for power generation after World War II through about the 1970s were generally to supply power for peak load periods. They were fairly low-efficiency units that were intended to be available for emergency needs and to insure adequate generation reserves in case of unexpected load peaks or generation outages. Full-load net heat rates were typically 13,600 Btu/kWh (HHV). In the 1980s and 1990s, new, large, and simple cycle units with much improved heat rates were used for power generation. Figure 1.8 shows the approximate, reported range of heat rates for simple cycle units. These data were taken from a 1990 publication (reference 3) and were adjusted to allow for the difference between lower and higher heating values for natural gas and the power required by plant auxiliaries. The data illustrate the remarkable improvement in gas turbine efficiencies achieved by the modern designs.

1.6.2 Combined Cycle Units

Combined cycle plants use the high-temperature exhaust gases from one or more gas turbines to generate steam in HRSGs that are then used to drive a steam turbine generator. There are many different arrangements of combined cycle plants; some may use supplementary boilers that may be fired to provide additional steam. The advantage of a combined cycle is its higher efficiency. Plant efficiencies have been reported in the range between 6600 and 9000 Btu/kWh for the most efficient plants. Both figures are for HHVs of the fuel (see reference 4). A 50% efficiency would correspond to a net heat rate of 6825 Btu/kWh. Performance data vary with specific

www.mepcafe.com

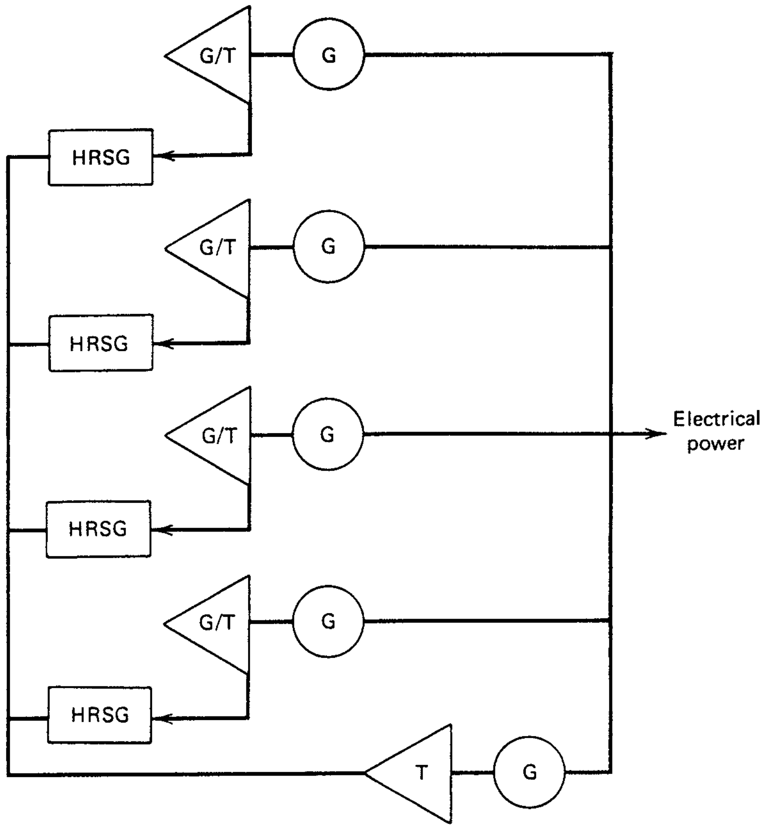


FIGURE 1.9 A combined cycle plant with four gas turbines and a steam turbine generator.

cycle and plant designs. Reference 2 gives an indication of the many configurations that have been proposed.

Part-load heat rate data for combined cycle plants are difficult to ascertain from available information. Figure 1.9 shows the configuration of a combined cycle plant with four gas turbines and HRSGs and a steam turbine generator. The plant efficiency characteristics depend on the number of gas turbines in operation. The shape of the net heat rate curve shown in Figure 1.10 illustrates this. Incremental heat rate characteristics tend to be flatter than those normally seen for steam turbine units.

1.6.3 Cogeneration Plants

Cogeneration plants are similar to the common-header steam plants discussed previously in that they are designed to produce both steam and electricity. The term "cogeneration" has usually referred to a plant that produces steam for an industrial process like an oil refining process. It is also used to refer to district heating plants. In the United States, "district heating" implies the supply of steam to heat buildings

www.mepcafe.com

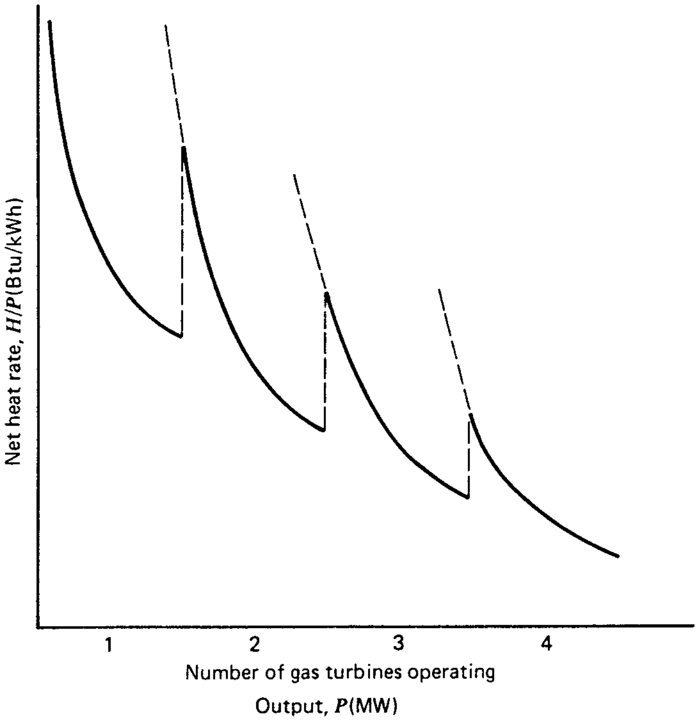


FIGURE 1.10 Combined cycle plant heat rate characteristic.

in downtown (usually business) areas. In Europe, the term also includes the supply of heat in the form of hot water or steam for residential complexes, usually large apartments.

For a variety of economic and political reasons, cogeneration is assuming a larger role in the power systems in the United States. The economic incentive is due to the high-efficiency electric power generation "topping cycles" that can generate power at heat rates as low as 4000 Btu/kWh. Depending on specific plant requirements for heat and power, an industrial firm may have large amounts of excess power available for sale at very competitive efficiencies. The recent and current political, regulatory, and economic climate encourages the supply of electric power to the interconnected systems by nonutility entities such as large industrial firms. The need for process heat and steam exists in many industries. Refineries and chemical plants may have a need for process steam on a continuous basis. Food processing may require a steady supply of heat. Many industrial plants use cogeneration units that extract steam from a simple or complex (i.e., combined) cycle and simultaneously produce electrical energy.

Prior to World War II, cogeneration units were usually small sized and used extraction steam turbines to drive a generator. The unit was typically sized to supply sufficient steam for the process and electric power for the load internal to the plant.

www.mepcafe.com

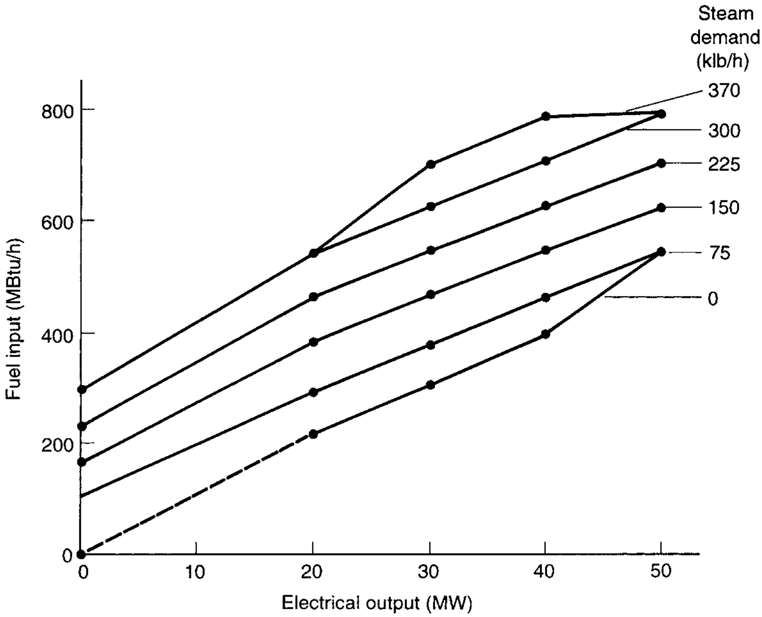


FIGURE 1.11 Fuel input required for steam demand and electrical output for a single extraction steam turbine generator.

Backup steam may have been supplied by a boiler, and an interconnection to the local utility provided an emergency source of electricity. The largest industrial plants would usually make arrangements to supply excess electric energy to the utility. Figure 1.11 shows the input–output characteristics for a 50-MW single extraction unit. The data show the heat input required for given combinations of process steam demand and electric output. This particular example is for a unit that can supply up to 370,000 lb/h of steam.

Modern cogeneration plants are designed around combined cycles that may incorporate separately fired steam boilers. Cycle designs can be complex and are tailored to the industrial plant's requirements for heat energy (see reference 2). In areas where there is a market for electric energy generated by an IPP, that is a non-utility-owned generating plant, there may be strong economic incentives for the industrial firm to develop a plant that can deliver energy to the power system. This has occurred in the United States after various regulatory bodies began efforts to encourage competition in the production of electric energy. This can, and has, raised interesting and important problems in the scheduling of generation and transmission system use. The industrial firm may have a steam demand cycle that is level, resulting in a more-or-less constant level of electrical output that must be absorbed. On the other hand, the local utility's load may be very cyclical. With a small component of nonutility generation, this may not represent a problem. However, if the IPP total generation supplies an appreciable portion of the utility load demand, the utility may have a complex scheduling situation.

www.mepcafe.com

1.6.4 Light-Water Moderated Nuclear Reactor Units

U.S. utilities have adopted the light-water moderated reactor as the “standard” type of nuclear steam supply system. These reactors are either pressurized water reactors (PWRs) or boiling water reactors (BWRs) and use slightly enriched uranium as the basic energy supply source. The uranium that occurs in nature contains approximately seven-tenths of 1% by weight of ^{235}U. This natural uranium must be enriched so that the content of ^{235}U is in the range of 2–4% for use in either a PWR or a BWR.

The enriched uranium must be fabricated into fuel assemblies by various manufacturing processes. At the time the fuel assemblies are loaded into the nuclear reactor core, there has been a considerable investment made in this fuel. During the period of time in which fuel is in the reactor and is generating heat and steam, and electrical power is being obtained from the generator, the amount of usable fissionable material in the core is decreasing. At some point, the reactor core is no longer able to maintain a critical state at a proper power level, so the core must be removed and new fuel reloaded into the reactor. Commercial power reactors are normally designed to replace one-third to one-fifth of the fuel in the core during reloading.

At this point, the nuclear fuel assemblies that have been removed are highly radioactive and must be treated in some fashion. Originally, it was intended that these assemblies would be reprocessed in commercial plants and that valuable materials would be obtained from the reprocessed core assemblies. It is questionable if the U.S. reactor industry will develop an economically viable reprocessing system that is acceptable to the public in general. If this is not done, either these radioactive cores will need to be stored for some indeterminate period of time or the U.S. government will have to take over these fuel assemblies for storage and eventual reprocessing. In any case, an additional amount of money will need to be invested, either in reprocessing the fuel or in storing it for some period of time.

The calculation of “fuel cost” in a situation such as this involves economic and accounting considerations and is really an investment analysis. Simply speaking, there will be a total dollar investment in a given core assembly. This dollar investment includes the cost of mining the uranium, milling the uranium core, converting it into a gaseous product that may be enriched, fabricating fuel assemblies, and delivering them to the reactor, plus the cost of removing the fuel assemblies after they have been irradiated and either reprocessing them or storing them. Each of these fuel assemblies will have generated a given amount of electrical energy. A pseudo-fuel cost may be obtained by dividing the total net investment in dollars by the total amount of electrical energy generated by the assembly. Of course, there are refinements that may be made in this simple computation. For example, it is possible by using nuclear physics calculations to compute more precisely the amount of energy generated by a specific fuel assembly in the core in a given stage of operation of a reactor.

Nuclear units will be treated as if they are ordinary thermal-generating units fueled by a fossil fuel. The considerations and computations of exact fuel reloading schedules and enrichment levels in the various fuel assemblies are beyond the scope of a one-semester graduate course because they require a background in

www.mepcafe.com

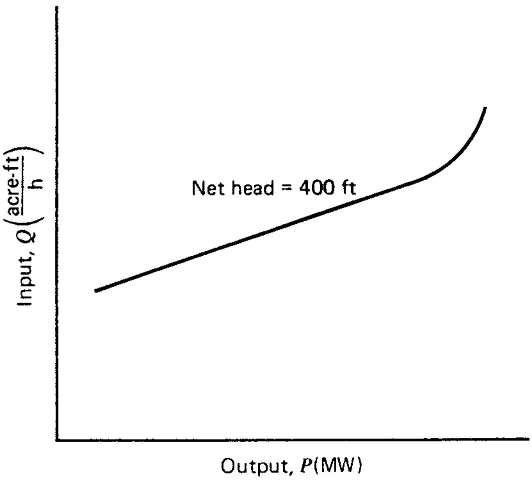


FIGURE 1.12 Hydroelectric unit input–output curve.

nuclear engineering as well as detailed understanding of the fuel cycle and its economic aspects.

1.6.5 Hydroelectric Units

Hydroelectric units have input–output characteristics similar to steam turbine units. The input is in terms of volume of water per unit time; the output is in terms of electrical power. Figure 1.12 shows a typical input–output curve for hydroelectric plant where the net hydraulic head is constant. This characteristic shows an almost linear curve of input water volume requirements per unit time as a function of power output as the power output increases from minimum to rated load. Above this point, the volume requirements increase as the efficiency of the unit falls off. The incremental water rate characteristics are shown in Figure 1.13. The units shown on both these curves are English units. That is, volume is shown as acre-feet (an acre of water a foot deep). If necessary, net hydraulic heads are shown in feet. Metric units are also used, as are thousands of cubic feet per second (kft^3/s) for the water rate.

Figure 1.14 shows the input–output characteristics of a hydroelectric plant with variable head. This type of characteristic occurs whenever the variation in the storage pond (i.e., forebay) and/or afterbay elevations is a fairly large percentage of the overall net hydraulic head. Scheduling hydroelectric plants with variable head characteristics is more difficult than scheduling hydroelectric plants with fixed heads. This is true not only because of the multiplicity of input–output curves that must be considered, but also because the maximum capability of the plant will also tend to vary with the hydraulic head. In Figure 1.14, the volume of water required for a given power output decreases as the head increases. (That is, dQ/d head or dQ/d volume is negative for a fixed power.) In a later section, methods are discussed that have been

www.mepcafe.com

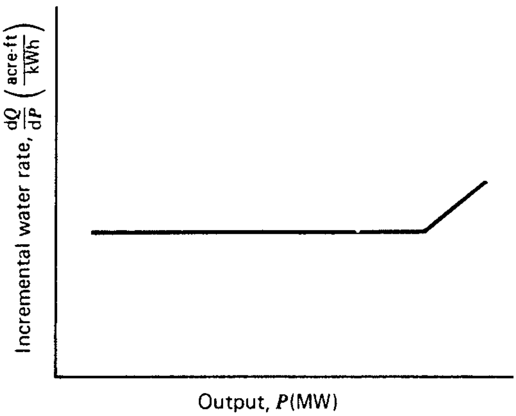


FIGURE 1.13 Incremental water rate curve for hydroelectric plant.

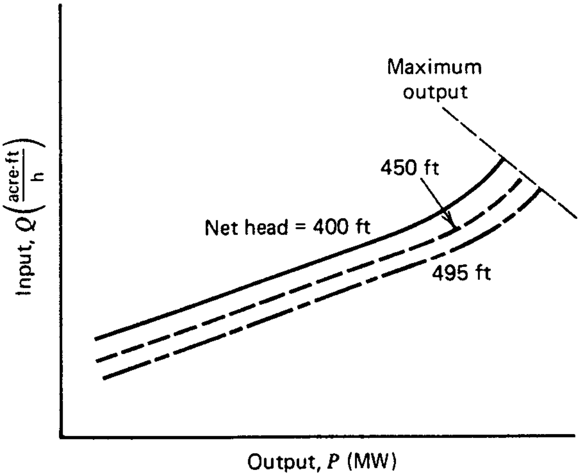


FIGURE 1.14 Input–output curves for hydroelectric plant with a variable head.

proposed for the optimum scheduling of hydrothermal power systems where the hydroelectric systems exhibit variable head characteristics.

Figure 1.15 shows the type of characteristics exhibited by pumped-storage hydroelectric plants. These plants are designed so that water may be stored by pumping it against a net hydraulic head for discharge at a more propitious time. This type of plant was originally installed with separate hydraulic turbines and electric-motor-driven pumps. In recent years, reversible, hydraulic pump turbines have been utilized. These reversible pump turbines exhibit normal input–output characteristics when utilized as turbines. In the pumping mode, however, the efficiency of operation tends to fall off when the pump is operated away from the rating of the unit. For this reason, most plant operators will only operate these units in the pumping mode at a fixed

www.mepcafe.com

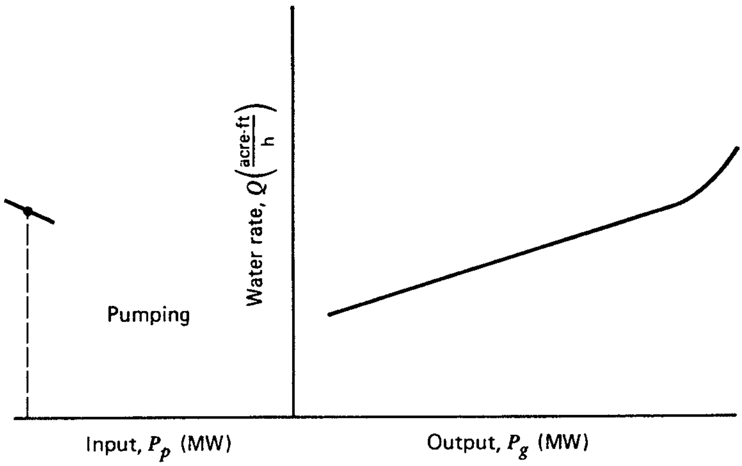


FIGURE 1.15 Input–output characteristics for a pumped-storage hydroplant with a fixed, net hydraulic head.

pumping load. The incremental water characteristics when operating as a turbine are, of course, similar to the conventional units illustrated previously.

The scheduling of pumped-storage hydroelectric plants may also be complicated by the necessity of recognizing the variable-head effects. These effects may be most pronounced in the variation of the maximum capability of the plant rather than in the presence of multiple input–output curves. This variable maximum capability may have a significant effect on the requirements for selecting capacity to run on the system, since these pumped-storage hydroplants may usually be considered as spinning-reserve capability. That is, they will be used only during periods of highest cost generation on the thermal units; at other times, they may be considered as readily available ("spinning reserve"). That is, during periods when they would normally be pumping, they may be shut off to reduce the demand. When idle, they may be started rapidly. In this case, the maximum capacity available will have a significant impact on the requirements for having other units available to meet the system's total spinning-reserve requirements.

These hydroelectric plants and their characteristics (both the characteristics for the pumped-storage and the conventional-storage hydroelectric plants) are affected greatly by the hydraulic configuration that exists where the plant is installed and by the requirements for water flows that may have nothing to do with power production. The characteristics just illustrated are for single, isolated plants. In many river systems, plants are connected in both series and in parallel (hydraulically speaking). In this case, the release of an upstream plant contributes to the inflow of downstream plants. There may be tributaries between plants that contribute to the water stored behind a downstream dam. The situation becomes even more complex when pumped-storage plants are constructed in conjunction with conventional hydroelectric plants. The problem of the optimum utilization of these resources involves the complicated problems associated

www.mepcafe.com

with the scheduling of water as well as the optimum operation of the electric power system to minimize production cost. We can only touch on these matters in this text and introduce the subject. Because of the importance of the hydraulic coupling between plants, it is safe to assert that no two hydroelectric systems are exactly the same.

1.6.6 Energy Storage

Electric energy storage at the transmission system level where large amounts of electric energy can be stored over long time periods is very useful. When the prices of electric energy are low (for example at night), then it is useful to buy electric energy and then sell it back into the system during high-priced periods. Similarly, if you are operating renewable generation sources such as wind generators that cannot be scheduled, then it would be useful to store electric energy when the wind is blowing and then release it to the power system when most advantageous. Last of all, if there are seasonal variations such as in hydro systems, we would like to store energy during high runoff periods and then use it later when runoff is lower.

Parameters of electric energy storage (reference 5)

- *Available energy capacity*, W_{op}: The quantity of stored energy that is retrievable as electric power.
- *Rated power*, P_{rated}: The nameplate value for the rate at which electric energy can be continually stored or extracted from the storage system, usually given in kilowatts (kW) or megawatts (MW). Also referred to as the discharge capacity.
- *Discharge time*, $t_{storage}$: The duration of time that the energy storage system can supply rated power, given as $t_{storage} = (W_{op}/P_{rated})$.
- *Energy density*: Available energy capacity per unit mass, given in Wh/kg.
- *Power density*: Rated power per unit mass, given in W/kg.
- *Round-trip efficiency*, $\eta_{round\text{-}trip}$: The overall efficiency of consuming and later releasing energy at the point of common coupling with power grid. Also known as AC–AC efficiency, round-trip efficiency accounts for all conversion and storage losses and can be broken into charging and discharging efficiencies: $\eta_{round\text{-}trip} = \eta_{charge}\eta_{discharge} \cong \eta^2_{one\text{-}way}$.
- *Cycle life*: The maximum number of cycles for which the system is rated. The actual operating lifespan of the battery is either the cycle life or the rated lifespan, whichever is reached first.

List of technologies used in electric power energy storage:

- Pumped hydro
- CAES (compressed air energy storage)
- Flywheel
- SMES (superconducting magnetic energy storage)
- Lead-acid battery

www.mepcafe.com

- NaS battery
- Li-ion battery
- Metal-air battery
- PSB flow battery
- VRB flow battery
- ZnBr flow battery
- Fuel cells
- Ultra capacitors

Applications grouped by storage capacity and response time

Very short	0–20 s (1–4 MW or >20 MW)	End user protection
Short	10 min to 2 h (up to 2 MW)	End use reserves
Long	1–8 h (greater than 10 MW)	Generation, load leveling, ramp following
Very long	1–7 days (greater than 1 MW)	Seasonal and emergency backup, renewable backup

For this text, we are mainly interested in the last two for large transmission system applications. The types of storage technologies that make up the long and very long storage time categories are pumped storage, compressed air storage, as well as some of the battery types. However, for this text we shall deal mainly with pumped storage and compressed air since they are proven technologies that have scaled to large installations that can be used on the transmission system itself.

1.7 RENEWABLE ENERGY

Renewable energy is energy that comes from natural resources such as sunlight, wind, rain, tides, and geothermal heat, which are renewable (naturally replenished).[1] A renewable resource is a natural resource with the ability of being replaced through biological or natural processes and replenished with the passage of time.[2]

Renewable fuels are those fuel sources that can be burned in conventional generation systems such as boiler–turbine–generators, gas turbine generators, and diesel generators.

Organic plant matter, known as biomass, can be burned, gasified, fermented, or otherwise processed to produce electricity.[3]

Geothermal energy extracts steam directly from the earth and uses it to power turbine–generator units.

[1] http://en.wikipedia.org/wiki/Renewable_energy
[2] http://en.wikipedia.org/wiki/Renewable_resource
[3] http://www.acore.org/what-is-renewable-energy/

www.mepcafe.com

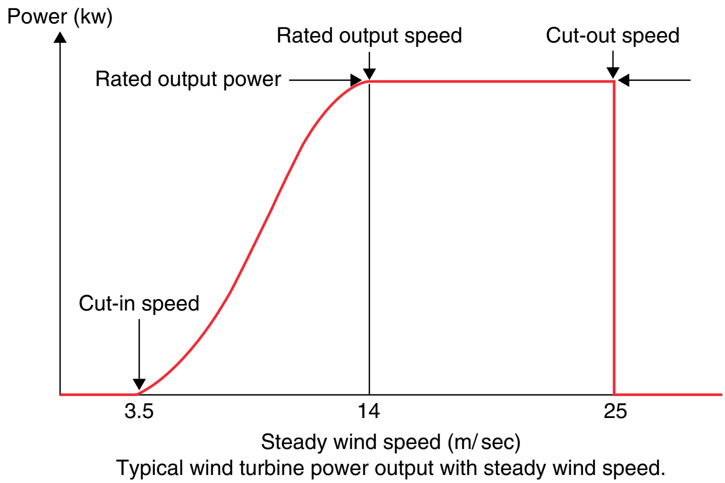


FIGURE 1.16 Typical wind turbine power output with steady wind speed.

Ocean energy can also be used to produce electricity. In addition to tidal energy, energy can be produced by the action of ocean waves, which are driven by both the tides and the winds. Because of their link to winds and surface heating processes, ocean currents are considered as indirect sources of solar energy.[4] In this case, ocean energy is converted through direct action of water on a turbine in the same manner as a hydroelectric plant turbine, although the shape and characteristics of the turbine for extracting energy from the oceans is different.

1.7.1 Wind Power

By far the most common renewable electric generation system is the wind generator. In the past 10 years, wind generation has advanced to the point that it is now quite economical to build and operate large sets of wing generators often called wind farms. In addition, wind generators are now being developed specifically to be placed in the ocean near the shore where strong and almost constant winds blow.

Figure 1.16 shows a sketch a how the power output from a wind turbine varies with steady wind speed. (This figure and the following paragraphs up to the equation for available power are taken from http://www.wind-power-program.com/turbine_characteristics.htm)

1.7.2 Cut-In Speed

At very low wind speeds, there is insufficient torque exerted by the wind on the turbine blades to make them rotate. However, as the speed increases, the wind turbine will begin to rotate and generate electrical power. The speed at which the turbine first

[4] ibid

www.mepcafe.com

starts to rotate and generate power is called the cut-in speed and is typically between 3 and 4 m/s.

1.7.3 Rated Output Power and Rated Output Wind Speed

As the wind speed rises above the cut-in speed, the level of electrical output power rises rapidly as shown. However, typically somewhere between 12 and 17 m/s, the power output reaches the limit that the electrical generator is capable of. This limit to the generator output is called the rated power output and the wind speed at which it is reached is called the rated output wind speed. At higher wind speeds, the design of the turbine is arranged to limit the power to this maximum level and there is no further rise in the output power. How this is done varies from design to design but typically with large turbines, it is done by adjusting the blade angles so as to keep the power at the constant level.

1.7.4 Cut-Out Speed

As the speed increases above the rate output wind speed, the forces on the turbine structure continue to rise and, at some point, there is a risk of damage to the rotor. As a result, a braking system is employed to bring the rotor to a standstill. This is called the cut-out speed and is usually around 25 m/s.

1.7.5 Wind Turbine Efficiency or Power Coefficient

The available power in a stream of wind of the same cross-sectional area as the wind turbine can easily be shown to be

$$\text{Available power in watts} = \frac{1}{2}\rho U^3 \frac{\pi d^2}{4}$$

where

U is the wind speed in m/s
ρ is the density of air in kg/m^3
d is the rotor diameter in m

We will talk in later chapters on the problems that wind generation presents due to its “nondispatchable” nature—simply meaning that we cannot order wind generation to be “on” during certain hours or “off” during others since it depends on the wind, which we do not have control over. The result is a strong interest in programs that use metrological data to predict wind speed, direction, location, and time of day.

In addition, it is apparent from recent data that large numbers of wind generators do have the ability to produce a smoother wind generation output than a single wind generator. Figure 1.17 shows this quite clearly.

www.mepcafe.com

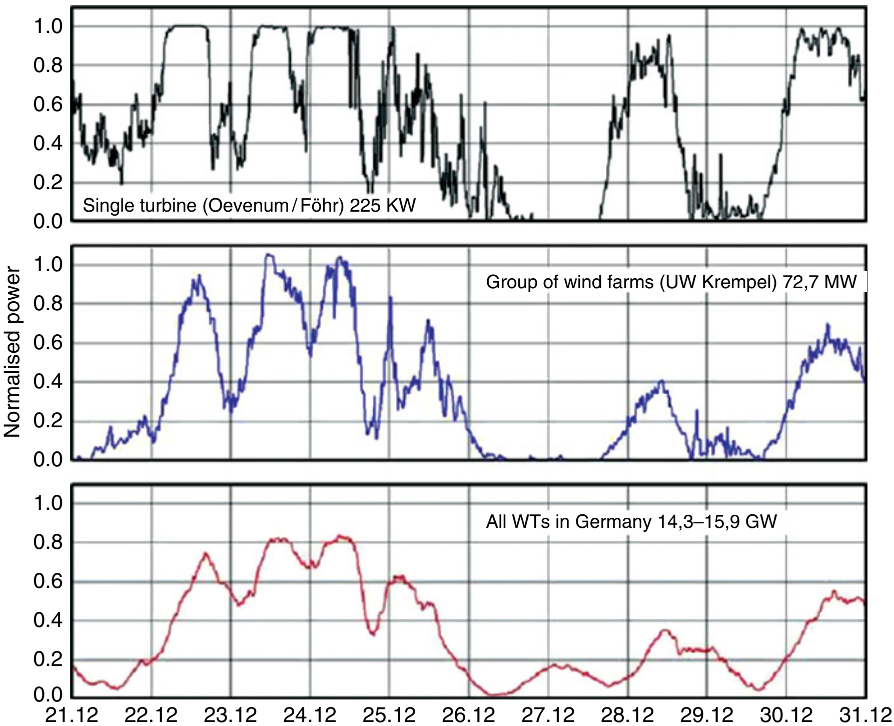


FIGURE 1.17 Example of time series of normalized power output from a single wind generator, a group of wind generators and all wind generators in Germany (21–31.12.2004) (reference 6).

This figure is taken from reference 6 and shows the output of a single wind generator (top), a group of wind farms (center), and the entire fleet of wind generators in Germany (bottom). Obviously, the wind generators taken as a large group overcome the very unpredictable and noisy output of a single wind generator.

1.7.6 Solar Power

Solar power comes in two varieties with respect to generation of electricity: photovoltaic and concentrated solar power.

Photovoltaic sources use cells that depend on the "photovoltaic" effect to convert incident sunlight into direct current (DC) electric power. The DC power is then converted to AC electric power at the system frequency where this is connected by power electronics converters. Small arrays of photo cells can be placed on the roof of a single home and supply electric power to that home or large numbers of arrays can be arranged in fields and wired to supply power directly to the electric system.

Concentrated solar power (also called *concentrating solar power*, *concentrated solar thermal*, and *CSP*) systems use mirrors or lenses to concentrate a large area of

www.mepcafe.com

sunlight, or solar thermal energy, onto a small area. Electrical power is produced when the concentrated light is converted to heat, which drives a heat engine (usually a steam turbine) connected to an electrical power generator.[5]

Obviously, both of these sources depend on the availability of sunlight and like wind generators cannot be dispatched. However, the CSP units can produce some electric energy after the sun has gone down due to the storage of heat in its steam generators.

APPENDIX 1A
Typical Generation Data

Up until the early 1950s, most U.S. utilities installed units of less than 100 MW. These units were relatively inefficient (about 950 psi steam and no reheat cycles). During the early 1950s, the economics of reheat cycles and advances in materials technology encouraged the installation of reheat units having steam temperatures of 1000°F and pressures in the range of 1450–2150 psi. Unit sizes for the new design reheat units ranged up to 225 MW. In the late 1950s and early 1960s, U.S. utilities began installing larger units ranging up to 300 MW in size. In the late 1960s, U.S. utilities began installing even larger, more efficient units (about 2400 psi with single reheat) ranging in size up to 700 MW. In addition, in the late 1960s, some U.S. utilities began installing more efficient supercritical units (about 3500 psi, some with double reheat) ranging in size up to 1300 MW. The bulk of these supercritical units ranged in size from 500 to 900 MW. However, many of the newest supercritical units range in size from 1150 to 1300 MW. Maximum unit sizes have remained in this range because of economic, financial, and system reliability considerations.

Typical heat rate data for these classes of fossil generation are shown in Table 1.1. These data are based on U.S. federal government reports and other design data for U.S. utilities (see *Heat Rates for General Electric Steam Turbine-Generators* 100,000 *kW and Larger*, Large Steam Turbine Generator Department, G.E.).

The shape of the heat rate curve is based on the locus of design "valve-best-points" for the various sizes of turbines. The magnitude of the turbine heat rate curve has been increased to obtain the unit heat rate, adjusting for the mean of the valve loops, boiler efficiency, and auxiliary power requirements. The resulting approximate increase from design turbine heat rate to obtain the generation heat rate in Table 1.1 is summarized in Table 1.2 for the various types and sizes of fossil units.

Typical heat rate data for light-water moderated nuclear units are as follows:

Output (%)	Net Heat Rate (Btu/kWh)
100	10,400
75	10,442
50	10,951

[5] http://en.wikipedia.org/wiki/Concentrated_solar_power

www.mepcafe.com

TABLE 1.1 Typical Fossil Generation Unit Heat Rates

Fossil Unit—Description	Unit Rating (MW)	100% Output (Btu/kWh)	80% Output (Btu/kWh)	60% Output (Btu/kWh)	40% Output (Btu/kWh)	25% Output (Btu/kWh)
Steam—coal	50	11,000	11,088	11,429	12,166	13,409[a]
Steam—oil	50	11,500	11,592	11,949	12,719	14,019[a]
Steam—gas	50	11,700	11,794	12,156	12,940	14,262[a]
Steam—coal	200	9,500	9,576	9,871	10,507	11,581[a]
Steam—oil	200	9,900	9,979	10,286	10,949	12,068[a]
Steam—gas	200	10,050	10,130	10,442	11,115	12,251[a]
Steam—coal	400	9,000	9,045	9,252	9,783	10,674[a]
Steam—oil	400	9,400	9,447	9,663	10,218	11,148[a]
Steam—gas	400	9,500	9,548	9,766	10,327	11,267[a]
Steam—coal	600	8,900	8,989	9,265	9,843	10,814[a]
Steam—oil	600	9,300	9,393	9,681	10,286	11,300[a]
Steam—gas	600	9,400	9,494	9,785	10,396	11,421[a]
Steam—coal	800–1,200	8,750	8,803	9,048	9,625[a]	
Steam—oil	800–1,200	9,100	9,155	9,409	10,010[a]	
Steam—gas	800–1,200	9,200	9,255	9,513	10,120[a]	

[a]For study purposes, units should not be loaded below the points shown.

www.mepcafe.com

TABLE 1.2 Approximate Unit Heat Rate Increase over Valve-Best-Point Turbine Heat Rate

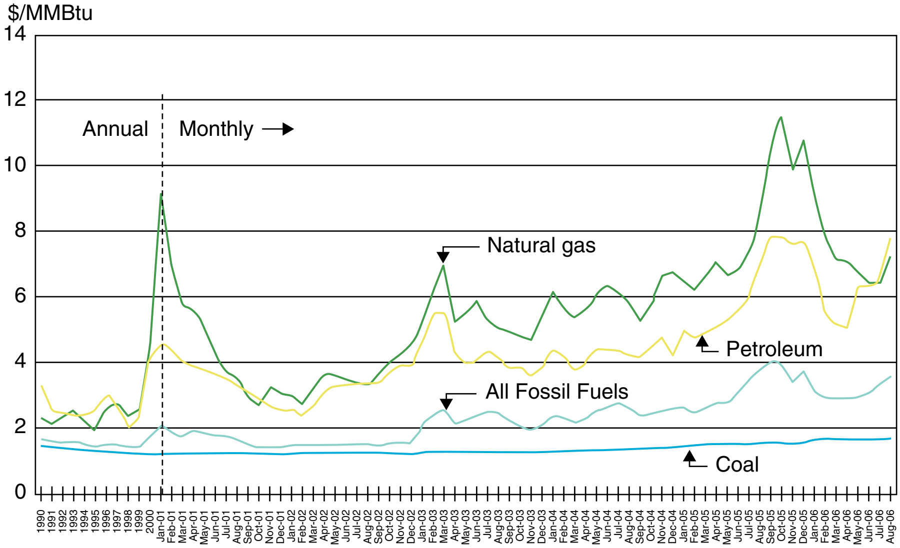
Unit Size (MW)	Coal (%)	Oil (%)	Gas (%)
50	22	28	30
200	20	25	27
400	16	21	22
600	16	21	22
800–1200	16	21	22

These typical values for both PWR and BWR units were estimated using design valve-best-point data that were increased by 8% to obtain the net heat rates. The 8% accounts for auxiliary power requirements and heat losses in the auxiliaries.

Typical heat rate data for newer and larger gas turbines are discussed earlier. Older units based on industrial gas turbine designs had heat rates of about 13,600 Btu/kWh. Older units based on aircraft jet engines were less efficient, with typical values of full-load net heat rates being about 16,000 Btu/kWh.

APPENDIX 1B
Fossil Fuel Prices

As can be seen in Figure 1.18, the prices for petroleum and natural gas have varied over time, sometimes peaking for short periods of times (several months). The price of coal is relatively constant over the past two decades.



Data Source: DOE/EIA

FIGURE 1.18 Fossil fuel prices 1990–August 2006 [reference 7].

www.mepcafe.com

APPENDIX 1C
Unit Statistics

In North America, the utilities participate in an organization known as the North American Electric Reliability Council (NERC) with its headquarters in Princeton, New Jersey. NERC undertakes the task of supporting the interutility operating organization that publishes an operating guide and collects, processes, and publishes statistics on generating units. NERC maintains the *Generating Availability Data System* (GADS) that contains over 25 years of data on the historical performance of generating units and related equipment. This information is made available to the industry through special reports done by the NERC staff for specific organizations and is also issued in an annual report, the *Generating Availability Report*. These data are extremely useful in tracking unit performance, detecting trends in maintenance needs, and in planning capacity additions to maintain adequate system generation reserves. The GADS structure provides standard definitions that are used by the industry in recording unit performance. This is of vital importance if collected statistics are to be used in reliability and adequacy analyses. Any useful reliability analysis and prediction structure requires three essential elements:

- Analytical (statistical and probability) methods and models,
- Performance measures and acceptable standards,
- Statistical data in a form that is useful in the analysis and prediction of performance measures.

In the generation field, GADS performs the last two in an excellent fashion. Its reputation is such that similar schemes have been established in other countries based on GADS.

Table 1.3 contains typical generating unit data on scheduled maintenance requirements, the "equivalent forced outage rate" and the "availability factor" that were taken from an NERC summary of generating unit statistics for the period 1988–1992. For any given, specified interval (say a year), the NERC definitions of the data are:

$$\begin{aligned}\text{Equivalent forced outage rate} &= (\text{forced outage hours} + \text{equivalent forced}\\ &\quad \text{derated hours}) \div (\text{forced outage hours} + \text{hours}\\ &\quad \text{in service} + \text{equivalent forced derated hours}\\ &\quad \text{during reserve shutdown})\\ \text{Availability factor (AF)} &= \text{available hours} \div \text{period hours}\end{aligned}$$

Scheduled maintenance requirements were estimated from the NERC data using the reported "scheduled outage factor," the portion of the period representing scheduled outages.

www.mepcafe.com

TABLE 1.3 Typical Maintenance and Forced Outage Data

Unit Type	Size Range (MW)	Scheduled Maintenance Requirement (Days/Year)	Equivalent Forced Rate (%)	Availability Factor (%)
Nuclear	All	67	18.3	72
Gas turbines	All	22	—	91
Fossil-fueled steam	1–99	31	7.2	88
	100–199	42	8.0	85
	200–299	43	7.2	85
	300–399	52	9.5	82
	400–599	47	8.8	82
	600–799	45	7.6	84
	800–999	40	5.8	88
	≥1000	44	9.0	82

From *Generating Unit Statistics* 1988–1992 issued by NERC, Princeton, NJ.

	Effective Outage Rates (%)		
	EFOR	EFOR′	Service Factor=(service hours) ÷ (period hours) (%)
All fossil units	5.7	4.1	60.5
All gas turbines	55.5	3.4	2.6

The reported, standard equivalent forced outage rate for gas turbines has been omitted since the low duty cycle of gas turbines in peaking service biases the value of effective forced outage rate (EFOR). Using the standard definition earlier, the reported EFOR for all sizes of gas turbine units was 58.9%. This compares with 8.4% for all fossil-fired units. Instead of the aforementioned definition of EFOR, let us use a different rate (call it the EFOR′) that includes reserve shutdown hours and neglects all derated hours to simplify the comparison with the standard definition:

$$\text{EFOR} = \text{forced outage hours} \div (\text{forced outage hours} + \text{hours in service})$$

or

$$\text{EFOR}' = \text{forced outage hours} \div (\text{forced outage hours} + \text{available hours})$$

where the available hours are the sum of the reserve shutdown and service hours. The effect of the short duty cycle may be illustrated using the NERC data.

The significance is not that the NERC definition is “wrong”; for some analytical models it may not be suitable for the purpose at hand. Further, and much more important, the NERC reports provide sufficient data and detail to adjust the historical statistics for use in many different analytical models.

www.mepcafe.com

REFERENCES FOR GENERATION SYSTEMS

1. El-Wakil, M. M., *Power Plant Technology*, McGraw-Hill, New York, 1984.
2. Steinberg, M. J., Smith, T. H., *Economy Loading of Power Plants and Electric Systems*, John Wiley & Sons, Inc., New York, 1943.
3. Performance Specs, *Gas Turbine World*, Vol. 11, Pequot Publications, Inc., Fairfield, CT, Oct. 1990.
4. Foster-Pegg, R. W., *Cogeneration—Interactions of Gas Turbine, Boiler and Steam Turbine*, ASMS paper 84-JPGC-GT-12, Joint Power Generation Conference, 1984.
5. Gantz, J. M., Optimal Implementation of Energy Storage Systems in Power Distribution Networks, MS Thesis, University of Minnesota, Minneapolis, MN, June 2012.
6. Design and operation of power systems with large amounts of wind power, Final report, IEA WIND Task 25, Phase one 2006–2008, VTT TIEDOTTEITA -RESEARCH NOTES 2493.
7. Rose, K., *The Impact of Fuel Costs on Electric Power Prices*, June 2007, see http://www.kenrose.us

FURTHER READING

The following books are suggested as sources of information for the general area covered by this text. Since the publishing of the first edition of this text there has been a great increase in the number of specialized textbooks aimed at specific topics that are covered in this book. Due to this availability, we have entered only a few references to journals and other articles in each chapter.

Power Generation Plants

Anderson, P. M., *Modeling Thermal Power Plants for Dynamic Stability Studies*, Cyclone Copy Center, Ames, IA, 1974. Cochran, R. G., Tsoulfanidis, N. M. I., *The Nuclear Fuel Cycle: Analysis and Management*, American Nuclear Society, La Grange Park, IL, 1990.

El-Wakil, M. M., *Power Plant Technology*, McGraw-Hill, New York, 1984.

Scherer, C. R., *Estimating Electric Power System Marginal Costs*, North-Holland Publishing Company, Amsterdam, 1977.

Steinberg, M. J., Smith, T. H., *Economy Loading of Power Plants and Electric Systems*, John Wiley & Sons, Inc., New York, 1943.

Vardi, J., Avi-Itzhak, B., *Electric Energy Generation, Economics, Reliability, and Rates*, The MIT Press, Cambridge, MA, 1981.

Power System Economic Operation

El-Hawary, M. E., Christensen, G. S., *Optimal Economic Operation of Electric Power Systems*, Academic Press, New York, 1979.

Hano, I., *Operating Characteristics of Electric Power Systems*, Denki Shoin, Tokyo, 1967.

Kirchmayer, L. K., *Economic Operation of Power Systems*, John Wiley & Sons, Inc., New York, 1958.

Kirchmayer, L. K., *Economic Control of Interconnected Systems*, John Wiley & Sons, Inc., New York, 1959.

www.mepcafe.com

Unit Commitment of Generating Units

Hobbs, Benjamin F., et al., (ed.), *The Next Generation of Electric Power Unit Commitment Models*, Kluwer Academic Publishers, Boston, MA, 2001.

Modeling Power Transmission Systems

Arrillaga, J., Arnold, C. P., Harker, B. J., *Computer Modelling of Electrical Power Systems*, John Wiley & Sons, Inc., New York, 1983.

Kundar, P., *Power System Stability and Control*, McGraw-Hill, New York, 1994.

Ajjarapu, V., *Computational Techniques for Voltage Stability Assessment and Control (Power Electronics and Power Systems)*, Springer, Boston, 2007.

Anderson, P., Fouad A., *Power System Control and Stability*, Second Edition, Wiley-IEEE Press, Piscataway, 2002.

Bergen, A. R., Vittal, V., *Power Systems Analysis*, Second Edition, Prentice Hall, Upper Saddle River, 1999.

Heydt, G. T., *Computer Analysis Methods for Power Systems*, Macmillan Pub Co, New York, 1986.

Saadat, H., *Power System Analysis*, McGraw Hill, Inc., Boston, 2002.

Sauer P., Pai M., *Power System Dynamics and Stability*, Prentice Hall, Upper Saddle River, 1998.

Shahidehpour, M., Wang, Y., *Electric Power Systems Communications and Control*, McGraw Hill Professional Engineering, London, 2002.

Voltage Stability

Taylor, C. W., *Power System Voltage Stability*, McGraw-Hill, Inc., New York, 1994.

Van Cutsem, T., Vournas, C., *Voltage Stability of Electric Power Systems*, Kluwer Academic Publishers, Boston, MA, 1998.

State Estimation

Abur, A., Expósito, A. G., *Power System State Estimation, Theory and Implementation*, CRC Press, Marcel Dekker, Inc., Boca Raton, FL 2004.

Monticelli, A., *State Estimation in Electric Power Systems, A Generalized Approach*, Kluwer Academic Publishers, Boston, MA, 1999.

Control of Generation

Cohn, N., *Control of Generation and Power Flow on Interconnected Systems*, John Wiley & Sons, Inc., New York, 1961.

Cregell, T., *Power System Control Technology*, Prentice Hall International, Englewood Cliffs, NJ, 1986.

Hano, I., *Operating Characteristics of Electric Power Systems*, Denki Shoin Publishing Company, Tokyo, 1967.

www.mepcafe.com

Real Time Control and Energy Management Systems

Debs, A. S., *Modern Power Systems Control and Operation*, Kluwer, Norwell, MA, 1988.

Handschin, E. (ed.), *Real-Time Control of Electric Power Systems*, Elsevier, Amsterdam, 1972.

Handschin, E., Petroianu, A., *Energy Management Systems*, Springer-Verlag, Berlin, 1991.

Miller, R. H., Malinowski, J. H., *Power System Operation*, Third Edition, McGraw-Hill, New York, 1994.

Savulescu, S. C. (ed.), *Computerized Operation of Power Systems*, Elsevier, Amsterdam, 1976.

Sterling, M. J. H., *Power System Control*, Peregrinus, London, 1978.

Power System Planning

Stoll, H. G. (ed.), *Least-Cost Electric Utility Planning*, John Wiley & Sons, Inc., New York, 1989.

Sullivan, Robert L., *Power System Planning*, McGraw-Hill Inc., New York, 1977.

Useful Mathematics and Optimization Texts

Bather, J. A., *Decision Theory: An Introduction to Dynamic Programming and Sequential Decisions*, First Edition, Wiley, Chichester, West Sussex, July 13, 2000.

Bazaraa, M. S., Sherali, H. D., Shetty, C. M., *Nonlinear Programming: Theory and Algorithms*, Third Edition, Wiley-Interscience, Hoboken, NJ, May 5, 2006.

Bazaraa, M. S., Jarvis, J. J., Sherali, H. D., *Linear Programming and Network Flows*, Fourth Edition, Wiley, Hoboken, NJ, December 14, 2009.

Bellman, R. E., Dreyfus, S. E., *Applied Dynamic Programming*, Princeton University Press, Princeton, NJ, 1962.

Castillo, E., et al., *Building and Solving Mathematical Programming Models in Engineering and Science*, John Wiley & Sons, Inc., New York, 2002.

Chow, J. H., Wu, F. F., Momoh, J. A., *Applied Mathematics for Restructured Electric Power Systems, Optimization, Control, and Computational Intelligence*, Springer, 2005.

Hillier, F. S., Lieberman, G. J., *Introduction to Operations Research*, Seventh Edition, McGraw-Hill Science/Engineering/Math, New York, March 22, 2002.

Larson, R. E., Costi, J. L., *Principles of Dynamic Programming*, M. Decker, New York, 1978–1982.

Lasdon, L. S., *Optimization Theory for Large Systems*, Macmillan, New York, 1970.

Luenberger, D. G., Ye, Y., *Linear and Nonlinear Programming (International Series in Operations Research & Management Science)*, Third Edition, 2008 Edition, Springer, New York, November 19, 2010.

Song, Y.-H., *Modern Optimization Techniques in Power Systems*, Kluwer Academic Publishers, Boston, MA, 1999.

Strang, G., *An Introduction to Applied Mathematics*, Wellesley-Cambridge Press, Wellesley, MA, 1986.

Deregulation/Restructuring

Allen, E., Ilić, M., *Price-Based Commitment Decisions in the Electricity Market*, Springer, London, 1999.

www.mepcafe.com

Denny, F. I., Dismukes, D. E., *Power System Operations and Electricity Markets*, CRC Press, Boca Raton, FL, 2002.

Ilić, M., Galiana, F., Fink, L., *Power System Restructuring Engineering and Economics*, Kluwer Academic Publishers, Boston, MA, 1998.

Kirschen, D. S., Strbac, G., *Fundamentals of Power System Economics*, John Wiley & Sons, Inc., Hoboken, 2004.

Schweppe, F. C., Caramanis, M. C., Tabors, R. D., Bohn, R. E., *Spot Pricing of Electricity*, Kluwer Academic Publishers, Boston, MA, 1988.

Sheblé, G. B., *Computational Auction Mechanisms for Restructured Power Industry Operation*, Kluwer Academic Publishers, Boston, MA, 1999.

Stoft, S., *Power System Economics, Designing Markets for Electricity*, IEEE Press & Wiley-Interscience, John Wiley & Sons Inc., New York, 2002.

Thompson, G. L., Thore, S., *Computational Economics: Economic Modeling with Optimization Software*, The Scientific Press, San Francisco, CA, 1991.

Vijay, K., *Auction Theory*, Second Edition, Academic Press, August 26, 2009.

Wangensteen, I., *Power System Economics – The Nordic Electricity Market*, Tapir Academic Press, Trondheim, 2007.

Zaccour, G. (ed.), *Deregulation of Electric Utilities*, Kluwer Academic Publishers, Boston, MA, 1998.

Demand Forecasting

Bowerman, B.L., O'Connell, R.T., Koehler, A.B., *Forecasting, Time Series, and Regression: An Applied Approach*, Fourth Edition, Thomson Brooks/Cole, Belmont, CA, 2005.

Box, G. E. P., Jenkins, G. M., *Time Series Analysis: Forecasting and Control*, Holden-Day, Oakland, CA, 1976.

Janacek, G., Swift, L., *Time Series: Forecasting, Simulation, Applications*, Ellis Horwood Limited, West Sussex, 1993.

Montgomery, D. C., Johnson, L. A., Gardiner, J. S., *Forecasting & Time Series Analysis*, Second Edition, McGraw-Hill, New York, 1990.

Schweppe, F. C., *Uncertain Dynamic Systems*, Prentice Hall, Englewood Cliffs, NJ, 1973.

Auction Theory

Auctions: Theory and Practice (*The Toulouse Lectures in Economics*) by Paul Klemperer, Princeton University Press (March 8, 2004).

Auction Theory, Second Edition by Vijay Krishna, Academic Press, (August 26, 2009).

Computational Auction Mechanisms for Restructured Power Industry Operation (*Power Electronics and Power Systems*) [Hardcover], First Edition by Gerald B. Sheblé, Springer (March 31, 1999).

PSERC 98–03 Report, Markets for Electric Power: Experimental Results for Alternative Auction Institutions, by John Bernard, Robert Ethier, Timothy Mount, William Schulze, Ray D. Zimmerman, Deqiang Gan, Carlos Murillo-Sánchez, Robert J. Thomas, Richard Schuler, http://www.pserc.wisc.edu/documents/publications/papers/1998_general_publications/HICSS98_Mount.pdf. Accessed date July 22, 2013.

www.mepcafe.com

2

INDUSTRIAL ORGANIZATION, MANAGERIAL ECONOMICS, AND FINANCE

2.1 INTRODUCTION

The operation of any industry depends on a basic understanding of managerial economics. This chapter gives an introduction to the fundamental concepts of economics for engineers. The reader is introduced to industrial organization to understand the price drivers and the political framework collectively referred to as the business environment. The development takes the reader through a linear model of a regulated company and through the linear models of companies in a competitive market environment. The interactions of companies exchanging products, the transportation of intermediate and final products, and the customer models are developed.

The first section outlines the structure of a traditionally regulated vertically integrated electric and gas utilities and then the segmentation for a competitive framework. The second section outlines the general economic theory of the firm including production costing based on costs and resources. The third section outlines the segmentation of a vertically integrated utility into horizontally organized supply chains and industry input–output models (dynamic Leontief). This is one manner to restructure the division of the previously vertically structured electric power industry into a horizontally structured industry to foster competition. However, the rise of auctions leads to the use of game theory including the presence of market power. The fourth section outlines how the electric products are segmented for operational monitoring, control, and maintenance. The fifth section outlines different competitive economic conditions as the industry reregulates.

Power Generation, Operation, and Control, Third Edition. Allen J. Wood, Bruce F. Wollenberg, and Gerald B. Sheblé.
© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

www.mepcafe.com

The sixth section defines how to calculate the production cost for short-term operation as a regulated and a competitive company for owners of generation, transmission, and distribution. The seventh section discusses opportunity costs and demonstrates some deterministic approximations commonly used. The eight section reviews the basic techniques for market operation based on traditional auction methods. The ninth section introduces the concept of uncertainty and how to more accurately account for the risks of operation and of capital expansion.

The ninth section defines the supply chains involved in electric power production. The tenth section outlines the dynamic Leontief or input–output model used to emulate the competitive energy markets.

2.2 BUSINESS ENVIRONMENTS

There are two fundamental organizational models used in this work: microregulated and macroregulated.

Microregulated is the traditional regulated utility that is managed by the state and federal governments at a microdecision level. The utility proposes new projects, changes to rate classifications, and tariff changes to the state for approval. A key approval is the return on investment allowed by the utility to pay investors to borrow money. Once the state approves any projects or price changes, then the federal government can intervene when energy products cross state boundaries or federal laws, especially environmental protection laws. The operation and planning of such a utility is transparent if the government regulators technically review all engineering and financial details. It is expected that managers and engineers would give testimony and answer questions raised by any of the regulators. Such a utility has to answer to the customers through the elected officials (governors and presidents) via public forums to explain why key decisions were made and why the costs have been incurred. This work refers to microregulated as regulated to conform with the common use of the term.

Macroregulated is an environment where the intercompany contracts such as purchasing, shipping, and buying are the key regulated items. Regulators typically approve all commodity contracts, shipping contracts, joint ventures, and other intercompany arrangements to mitigate market monopolies, shipping congestion, and corporate pricing controls. Regulators are primarily concerned with market power, monopolistic operation or planning, and auction gaming for profit. Macroregulated is often referred to as competitive market structure. However, this work acknowledges that such competition may be lacking due to the supply chain structures, the auction mechanisms, or the market clearing rules. Competitive markets have to be regulated by governments to achieve efficiency, to eliminate contract confusion for gaming, to provide a naturally clearing market without monopolistic actions. This

www.mepcafe.com

work refers to macroregulated as competitive markets even though such markets may not be competitive as established and governed.

2.2.1 Regulated Environment

Utility operation in a regulated environment acknowledged that these companies were monopolies within a geographic area. Each utility was awarded a franchise for a selected area, normally due to historical growth. The vertically integrated utility environment is shown in Figure 2.1, in which the framework of the regulation commissions at the federal and the state level within the United States is given. Other countries with such a layered structure follow this same organization.

The federal government regulation is the top tier of the structure. There are many such government commissions in the United States. The federal level is assigned to the Federal Energy Regulatory Commission (FERC). Each state has a corresponding commission typically called State Public Utilities Commission (SPUC). The National Reliability Standards are defined and standardized by the North American Electric Reliability Organization (NAERO) at a national level (formerly National Electric Reliability Council, NERC) under FERC's approval and oversight.

The utilities in such an environment are microregulated by the federal and state governments at the level of project authorization, tariffs for each type of customer, profit margin allowed, and even profit to be distributed to shareholders, if any.

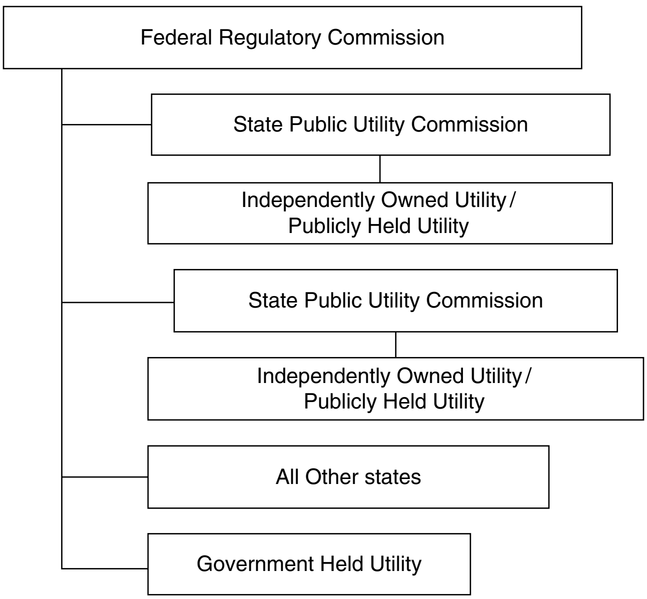


FIGURE 2.1 Regulated industry structure.

www.mepcafe.com

Regulators have complete transparency of all financial contracts, all accounting procedures, data bases, all planning and operating decision processes and risk analysis and management processes. The tariffs are typically reviewed once a year or when an event causes a change in the revenue requirements for the company. It is very seldom that tariffs are reduced unless some accounting discrepancy or mismanagement has occurred. It is common for tariffs to be reduced when fuel prices decline unexpectedly or deflation occurs. The setting of tariffs is beyond the scope of this work.

The independently owned utilities (IOUs) are the traditional for-profit companies that are granted a monopoly franchise for a geographic region. It should be noted that in the United States, most of these companies had the name Edison as part of their names. Government-held utilities are owned by the federal or state governments, such as Bonneville Power Administration/Army Corp of Engineers, Tennessee Valley Power Authority, Western Area Power Authority, and Nebraska Public Power Authority. Publicly held utilities are owned by smaller government entities such as counties and cities. Municipal cooperatives or farm cooperatives are categorized in this last grouping.

2.2.2 Competitive Market Environment

Federal and state legislation called for competition in the power industry from the wholesale level to the retail level through an evolving sequence of laws and regulations. New market structures are sought to search for a good business environment that would ultimately satisfy the regulatory bodies, customers, and suppliers. One approach is the application of brokerage systems to the power industry to promote competition. To accomplish competitive markets, the vertically integrated utilities have been broken up. Figure 2.2 shows the framework of the competitive business environment [1].

The government regulation is the top tier of the structure. There are many such government commissions in the United States. The traditional federal commissions include the Stock Exchange Commission (SEC), the Commodities and Futures Trading Commission (CFTC), and the FERC. Each state has a corresponding commission typically called State Public Utilities Commission (SPUC) for utility oversight. The National Reliability Standards are now more critical to be defined and standardized by the NAERO at a national level (formerly National Electric Reliability Council, NERC).

The utilities in such an environment are macroregulated by the federal and state governments at the level of business protocols, market standards, contract standards, and transparency to mitigate monopoly power. If the retail markets were also reregulated, then the tariffs for each type of customer would be defined at a more general level such as most retailers face in the standard business code of law. Regulators have limited transparency of financial contracts, accounting procedures, data bases, planning and operating decision processes, risk analysis and management processes. The regulators do reserve the right to review any and all records if insider trading or monopoly power is exercised in a fashion not allowed by the laws and

www.mepcafe.com

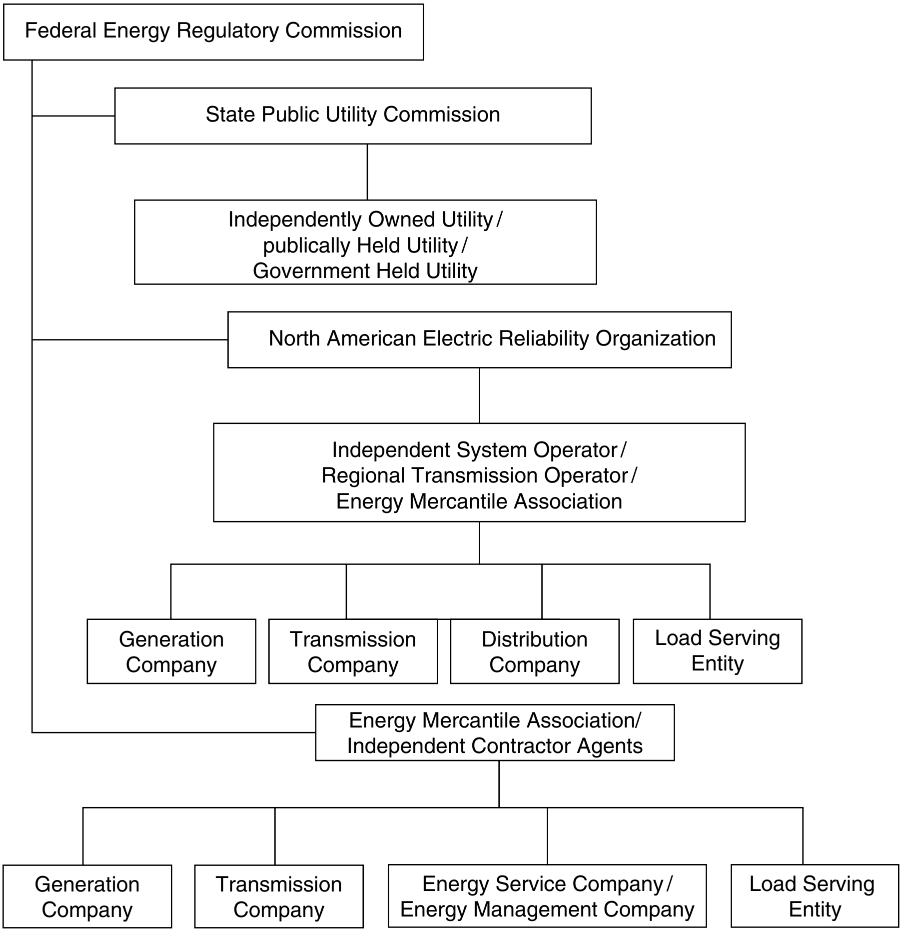


FIGURE 2.2 Competitive market structure.

regulations. Presently, the retail markets have not been reregulated in most countries due to the apparent market power held by very few companies. The design of regulations for contracts, markets, and general business interactions is beyond the scope of this work.

The generation companies primarily produce electric energy (GENCO). The transmission companies (TRANSCOs) own and operate transmission lines. The distribution companies (DISTCOs) own and operate distribution lines. As of this writing, TRANSCOs and DISTCOs may not be forced to divest from each other, as they are both regulated companies. The energy mercantile association (EMA) is a generic financial domain name for a power exchange. The brokers (BROCOs) buy and sell electricity for profit. The energy service companies (ESCOs) purchase electricity acting as agent for consumers. Alternative names for consumer pooling include energy management companies (EMCOs) and load serving entities (LSEs).

www.mepcafe.com

This work uses the aforementioned structure as the presently implemented structures are more diverse. It is noted that other structures are equally valid to achieve the same business environment.

The primary entity that will have direct dealings with the end users under the new market structure is the energy service company (ESCO). The ESCO collects its revenue from the distribution customers for the energy and ancillary services it has provided. It can also act as a wholesaler, purchasing the electric energy through the auction market and reselling it to the other ESCOs, GENCOs, etc. To purchase the desired electric energy to serve its purpose, the ESCO may purchase through an auction market or through direct contracts or utilize the energy reserves that it has accumulated, through the load management programs or the ownership of generation units. Additionally, all supportive services (e.g., load following and regulation) have to be procured to enable delivery of the energy across the transmission grid.

In the reregulated environment, every customer is free to choose any ESCO to serve the energy demand. In addition, the energy purchased from the auction market bears the risk of market price fluctuation. These, from the demand factors to the supply factors, are the risks that the ESCO has to undertake in the new market structure. Since deregulation will render the governmental protection (financially) obsolete, risk management and assessment tools should be considered and applied.

Each GENCO and each ESCO have a fixed and a variable operating cost based on the selected mix of capital expenditures and operating expenditures. The GENCO sees a demand curve by the buyers (ESCOs). The GENCO is a demand curve seen by the fuel suppliers. The ESCOs see the GENCOs through the supply curve. The ESCOs are a supply curve to all customers. This section works from the market model to exemplify the solution of the supply–demand curves. Then the basic models of the supply and demand curves are generated from basic cost components. The supply curve of the marketplace is the sum of the individual supply curves of all suppliers (GENCOs). GENCOs pool suppliers. The demand curve of the marketplace is the sum of the individual demand curves of the buyers. ESCOs pool purchasers and other consumers representatives. These curves are assumed in this work to be the actual cost curves, including profit to pay for capital contracts (stocks, bonds, mortgages, etc.). These curves may be altered if the supplier or the buyer can implant a perceived differentiation of product or if there is a differentiation of product. Such details are beyond the scope of this work.

This work introduces basic operational and economic models for assessment tools to assist any company with emphasis on applications to GENCO or ESCO operation.

2.3 THEORY OF THE FIRM

Each company has a multifaceted optimization problem of using scarce capital resources, scarce operating resources, and scarce labor resources to make and deliver a product. As shown in Figure 2.3, the vertically integrated utility gathers the raw

www.mepcafe.com

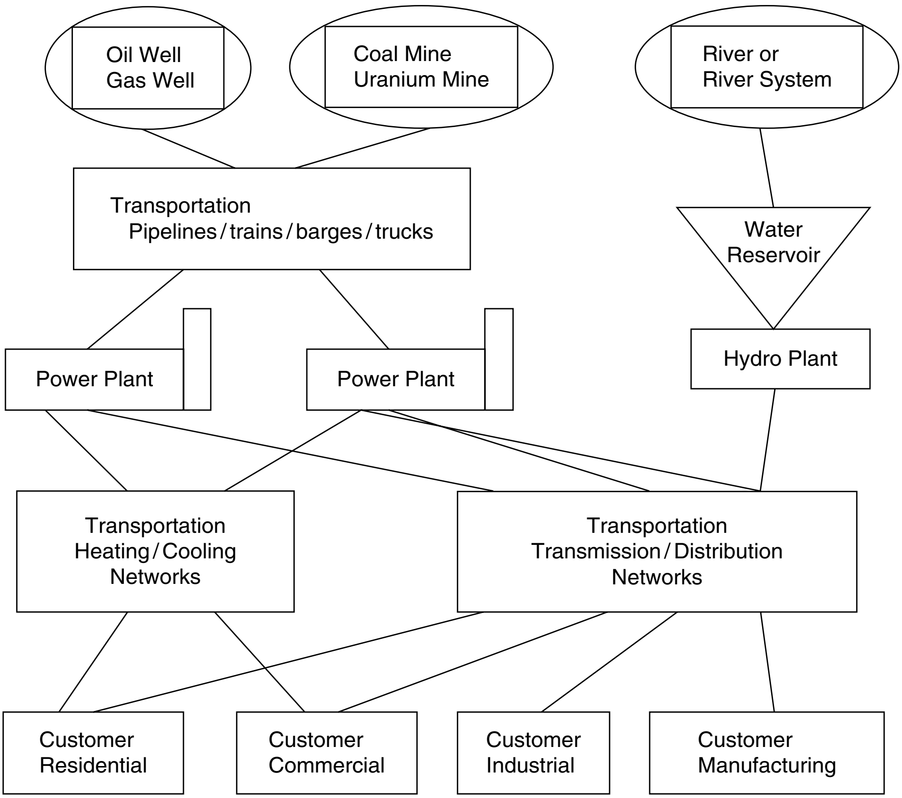


FIGURE 2.3 Vertically integrated utility supply chain.

resources (fuels), transports the fuels to processing plants (distilleries), and transports the raw fuel to the power plant (manufacturing); the produced electricity is then transported and distributed by the transmission and distribution systems to the customers. The customers are categorized in this work as residential (homes and apartment), as commercial (retail shops), industrial (business services and maintenance), and manufacturing (steel mills, auto assembly, etc.).

Within the United States, a vertically integrated utility did not own the oil wells, gas wells, or coal mines. There was a market between the utilities and these resources as oil is also demanded by the petroleum industry, gas is also demanded for direct consumption as a heat resource or process steam resource, and coal was used as a heat resource and as a process resource for metal fabrication and milling. These uses existed before the advent of electricity as an energy infrastructure. Natural gas was the fuel for heating and light. Coal was used for heating. Recently, the use of power plant steam or hot water as another resource has resurfaced as a combined heat and power unit. This returns to the previous use of power plant steam as found in various cities as New York, New York, and Ames, Iowa.

www.mepcafe.com

This chapter notes that many land-grant universities within the United States were independently run microgrids that were self-sufficient utilities due to the remote locations originally selected.

The economic equations to model this subsystem start at the raw resources, the fuels. This will start as a single production facility for simplicity.

It is noted that the price of most products are determined by the Porter five forces, especially the value of the product to the consumer. It is noted that the economic solution of the price being equal to the dual variable is generally not how a price is set. Product prices are primarily set based on demand by the consumer. Product prices have to recover all costs of production including capital expenses, taxes, etc. Such exogenous factors are implicitly included within this work. It is noted that most products are sold at an average price to recover all expenses and profits stated for shareholders at a minimum. It is opined that real option analysis is the superior method for pricing to include volatility or uncertainity.

2.4 COMPETITIVE MARKET SOLUTIONS

A horizontally regulated environment separates the layers from the vertical organization and places intermediate markets between the layers. Markets in this work include all contracts traded on an organized trading floor to contracts traded over the telephone or internet between companies, also known as over-the-counter trading.

There are many possible segmentations. The segment adopted depends on the historical development to date and the deregulation procedures adopted by the governments.

The natural gas industry was the first energy industry to offer lighting and heating to customers. Thus, the harvesting companies were separated historically due to geographic locations and transportation (pipeline) limitations. This segmentation occurred historically within the United States as shown in Figure 2.4. The development of hydro required extensive capital budgeting that could only be covered by the federal government for multiple purposes. Hydro systems were developed not only to provide electric energy in rural parts of the country but also to control flooding during years of high rainfalls. The United States Army Corp of Engineers developed the Mississippi and Columbia rivers for navigation (transportation) as well as for flood control and electric energy production.

The original wellhead companies served local cities directly until more distant cities and states offered higher revenues, causing the splitting of the wellhead companies from the local cities. Additionally, the wellhead companies were traditionally high-risk drilling companies that no one city could underwrite.

Once Edison and Westinghouse developed the electric generator, cities were converted to electric energy from gas energy on a block-by-block basis. Some gas utilities fought the new entrant to the energy industry. Some purchased the competitors to maintain a monopoly status. This is evident in the names of several United States utilities. One such example is Pacific Gas and Electric (PG&E) serving the northern half of California.

www.mepcafe.com

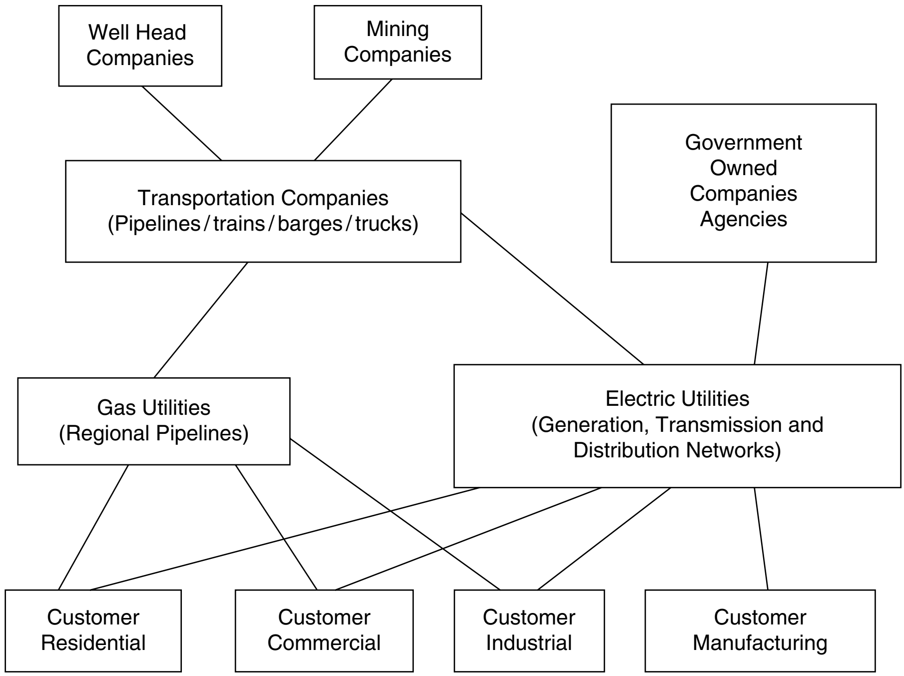


FIGURE 2.4 Initial vertical organization segmentation.

It is noted that many entered into the electricity business as publicly owned municipal utilities, especially when the profitability of service was questionable or uncertain. This was especially true for remote cities, such as Ames, Iowa.

Rural areas, primarily farm and ranch regions of the country, were electrified by federal laws forming rural cooperatives. The Nebraska Electric Power District is one prime example where a complete state formed a cooperative under state law. Such isolated utilities would be called microgrids in more recent analysis.

As there are 50 states in the United States, there are other segmentations that exist based on the history of that state and the political agendas at that time. Such a history is available at the website of the United States Energy Information Administration (EIA).

The resulting segmentation can be viewed as a supply chain consisting of the following layers: resource harvesting, long-distance transportation, regional transportation and distribution, and customers as shown in Figure 2.5. Notice that each input flow shows a link to a supplier and that each output flow identifies a buyer for each company.

It is noted that the original gas and electric utilities were not regulated. Only after the lobbying of Samuel Insul and others were the utilities granted a monopoly franchise for a given geographic area in the United States. It is noted that many countries formed electric companies as part of their federal government due to the view that utilities are a natural monopoly. Given the large capital investment, this can be argued even today. The majority of a utility investment is in the distribution network and the staff to interface with the customers and the customer equipment.

www.mepcafe.com

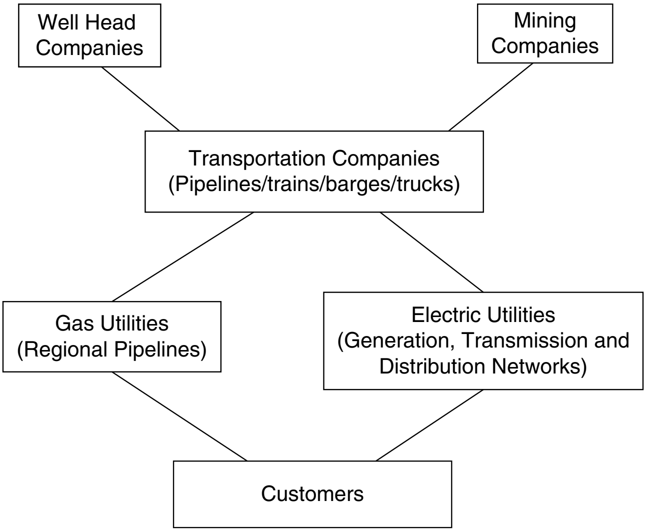


FIGURE 2.5 Energy supply chains.

Reregulation segmented the vertically integrated utility into three categories: generation, transmission, and distribution. This process is neither complete nor stable in organization as of this writing. The companies defined in this work include the following: GENCO, TRANSCO, DISTCO, ESCO, EMCO, LSE, ISO, and RTO as discussed earlier. The analysis of a company does not depend on the business environment within which it operates (regulated or competitive). The business environment only alters the pricing of the product by a government agency or by a competitive market. The basic tools of analysis are the same with a few alterations in data generation.

There are several interesting views to study supply chains. The first is to isolate the suppliers and identify the costs and benefits of working with that supplier. This is normally done as a discrete auction even if a commodity exchange is not used. This work uses the term market to describe any means for suppliers and buyers to exchange information leading to a contract for the exchange of raw resources needed for production. Such means may include organized exchanges, such as NYMEX and CYME in the United States. Such means may be strictly over the counter, such as an internet-based exchange (eBay) or strictly through trade magazines. Discrete auction analysis is introduced in the interchange analysis chapter (Section 11.12). A partial equilibrium system solution is an alternative analysis. These analyses are static solutions and do not show the intricacies of repeat solutions of multiple markets or of time domain analysis of markets.

It is noted that forecasting of the market prices is a key fundamental component of these approaches. Market price forecasting is covered in the subsequent volume to this work.

www.mepcafe.com

Another approach is to identify if a supplier should be used or if the resources should be made within the company. This is the make or buy problem that exists for all companies considering outsourcing. The example for the make or buy decision is the key model. The decision to make or buy is extensively presented in the section on interchange analysis (Chapter 11).

A company in a competitive market has to identify strengths and weaknesses for the various products that are produced. Limited availability of production factors and the production frontier are key analyses that are covered in the economic dispatch (Chapter 3) and unit commitment (Chapter 4).

Supply chain analysis focuses on the analysis of the various networks providing the raw materials from suppliers and the delivery of the products to the buyers (customers). The transportation of commodities through transportation networks, such as oil, natural gas, and coal, is presented in the Chapter 5 on limited energy resource scheduling. Commodity flows in networks are analogous to the flow of electricity in the transmission system with the added benefit of inexpensive storage. Inventory management with storage is a key concept for operation of such systems in parallel with the electric system.

Production in a vertically integrated company was presented earlier in this section to identify the basic calculations for a regulated company. This is generalized to industry modeling with one production location to demonstrate the concepts of a regulated company not connected to other electric companies. This was then generalized once again to industrial modeling with spatial networks to demonstrate the need for an integrated power system solution on a regional basis with other utilities or IPPs. The chapter on interchange continues this development for various contracts under regulated and competitive business environments.

This chapter provides the basic activity analysis and energy modeling problems to introduce the tools used by regulatory bodies to benchmark the operation and planning of regulated or competitive companies.

Another model is the input–output or Leontief network. The static problem is given in a later section.

2.5 SUPPLIER SOLUTIONS

The calculations for a company are the same independent of the activities of that company. This section outlines the financial and economic calculations for a company that is either regulated or reregulated.

Assume that the demand curve is known in a perfectly competitive environment. Thus, the marketplace fixes the price (P) since any individual company cannot change the price. Once the quantity (Q) is selected (known), the production cost is easily found:

$$PC = f(Q)\,|_{\text{static equilibrium}} \tag{2.1}$$

Note that the evaluation assumes that all other factors are constant as this is a static equilibrium. Almost all markets are dynamic. Such dynamic analysis techniques are

www.mepcafe.com

beyond the scope of this work. Fortunately, most markets may be sufficiently explained from a static analysis.

The revenue (R) for the company to use for capital expenses and operating expenses is easily found:

$$R = P \times Q \tag{2.2}$$

The profit for the company is simply the difference between the production cost and the revenue:

$$\Pi = R - PC \tag{2.3}$$

where ∏ is used to symbolize profit.

2.5.1 Supplier Costs

Costs can be segmented into the following categories: capital, revenue, and residual. Capital costs (K) consist of equipment design and procurement. Such costs include equipment specifications, design, development, manufacturing or procurement, installation, training for operations and for maintenance, as well as spare parts for repair. Once these expenditures have been committed, the firm can produce a product for consumers. However, these expenditures are balanced against their impact on asset reliability, asset maintainability, and asset availability as an impact on profit. The production factors include output quantity, product quality, material usage, labor usage, and asset utilization. Revenue costs (C_r) include operating and maintenance costs. Operating costs include cost for direct materials, direct labor, and overheads (direct expenses). Operating costs also include the indirect costs of materials, labor, and establishment overheads. However, these expenditures are balanced against their impact on profit. Maintenance costs includes spare inventory costs, labor costs, facilities and equipment costs, establishment overhead, and revenue lost due to downtime. However, these expenditures (preventative repair) are balanced against their impact on profit. The residual costs are the disposal value and the disposal costs. However, these expenditures are balanced against their impact on the production factors. A detailed discussion of each cost component is beyond the scope of this work. Residual costs are assumed to be absorbed into the capital costs.

2.5.2 Individual Supplier Curves

This work uses the following model for the production cost (C_p) given the capital cost (K) and the operating costs as a function of the production amount P.

$$C_p = g(K) + h\left(P_g\right) \tag{2.4}$$

Most introductory economic texts use linear approximations to this function. This work assumes that the function can be segmented into a piece-wise linear representation as additional accuracy is required.

www.mepcafe.com

The marginal cost of production (MC_p) is the first derivative of the cost of production including all cost components:

$$MC_p = \frac{dC_p}{dQ} \tag{2.5}$$

The average cost of production (AC_p) is the cost of production divided by the quantity sold:

$$AC_p = \frac{P}{Q} \tag{2.6}$$

The total production cost includes indirect costs (e.g., capital costs, as discussed later) as well as the direct production cost discussed earlier. When the total cost is referenced in this work, the subscript p is removed.
The firm's profit Π is the revenue minus the cost of production:

$$\Pi = R - C_p. \tag{2.7}$$

Marginal costs are separated into short run and into long run components. Average costs are separated into short run and into long run components. Short run does not normally include the possibility of capital expenditures. Long run does include the possibility of capital expenditures as well as other major efficiency improvement projects.

2.5.3 Competitive Environments

Note that the cost of production does not set the price in a perfectly competitive environment. Thus, it is natural to understand why all firms endeavor to not engage in a perfectly competitive environment. Many buyers and sellers typify a perfectly competitive market, when each firm produces a homogenous product, buyers and sellers have perfect information, there are no transaction costs, and there is free entry and exit. Under perfect competition, the optimization conditions require that the first derivative of the revenue must be equal to the first derivative of the cost of production:

$$MR = MC = P \tag{2.8}$$

This is most easily visualized through supply versus demand curves. Under the assumption of perfect competition, the curves do not interact. The producer surplus is that area between the supply curve and the perfectly competitive price line bounded by the actual demand. The consumer surplus is that area between the supply curve and the perfectly competitive price line bounded by the actual demand. These surpluses, the actual price, as well as the competitive price are shown in Figure 2.8. As

www.mepcafe.com

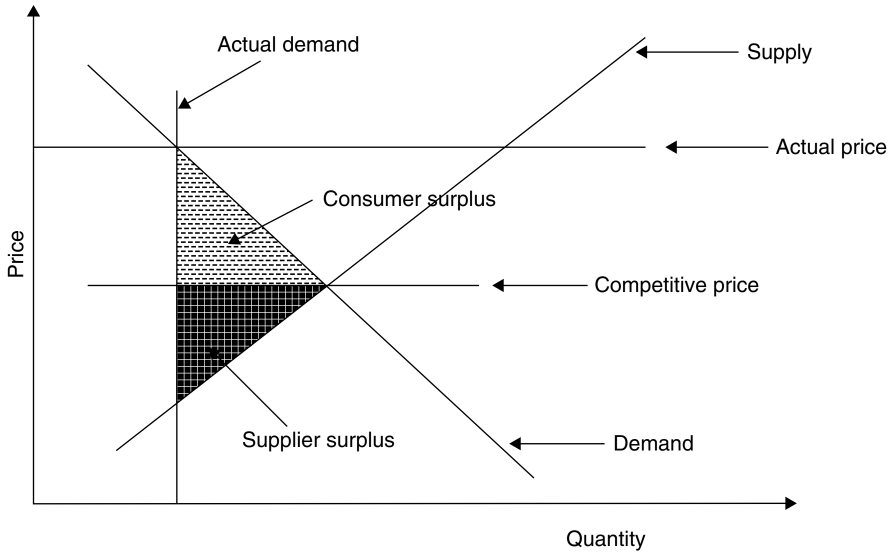


FIGURE 2.6 Supply and demand curves.

the actual price approaches the competitive price, note that the surpluses decrease to 0. When the price is above the competitive price, the supplier is collecting the consumer surplus as well as the producer surplus. Also, the quantity produced is less than the quantity ideally desired by society.

Consumers may alter the demand curve by substitution with alternative products through multimarkets. Use of natural gas instead of electricity is such a substitution. Such substitutions do not necessarily eliminate the consumption of a product as electric energy. However, the slope of the curve may dramatically change. Also, the demand curve may change based on budget constraints (Figure 2.6).

The elasticity of demand is due to budget constraints or due to substitutes. The elasticity is the normalized derivative of the demand curve expressed as a percentage change when the change in P and Q is small. Elasticity of demand is defined as follows:

$$E_{\mathrm{d}} = -\frac{\mathrm{d}Q/Q}{\mathrm{d}P/P} \tag{2.9}$$

When $E_{\mathrm{d}} > 1$, the demand is elastic. When $E_{\mathrm{d}} = 1$, the demand is constant. When $E_{\mathrm{d}} < 1$, the demand is inelastic.

Instead of elasticity, indifference curves may be used to model the use of substitutes. Note that gas is a substitute for electricity in the vertically organized utility in the previous section. Renewables are also substitutes, especially when considered in a distributed environment.

www.mepcafe.com

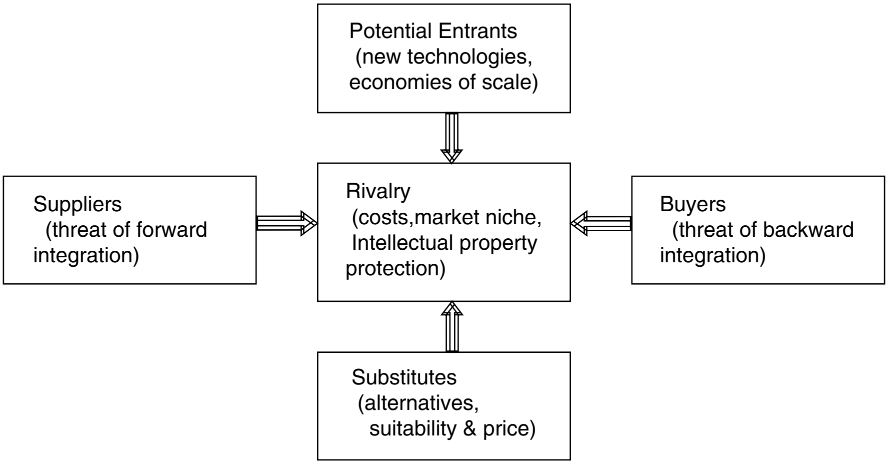


FIGURE 2.7 Porter five forces.

The solution of supply–demand curve is simply the intersection of the two curves under perfect competition.

The business environment structure of a general competitive company was previously shown. The common strategic model for all companies is commonly called the Porter five forces model (Figure 2.7).

It is noted by the authors that the prices of most products are determined by the Porter five forces, especially the value of the product to the consumer. It is noted that the economic solution of the price being equal to the dual variable is generally not how a price is set. Product prices are primarily set based on demand by the consumer. Product prices have to recover all costs of production including capital expenses, taxes, etc. Such exogenous factors are implicitly included within this work. It is noted that most products are sold at an average price to recover all expenses and profits stated for shareholders at a minimum. The Porter five forces encompass the major strategic directions that have to be managed on a daily basis.

The competitive force of potential entry is based on barriers to entry. These are related to economies of scale, the existence of learning and experience curve effects, brand preferences and customer loyalty, capital requirements, cost disadvantages independent of size, access to distribution channels, and government actions and policies.

The competitive force of substitute products is centered on the price and availability of acceptable substitutes for each product and places a ceiling on the prices that the producers of the product can charge. Unless the sellers of each product can upgrade quality, reduce prices via cost reduction, or otherwise differentiate their product from its substitutes, they risk a low growth rate in sales and profits because of the inroads substitutes may make. The competition from substitutes is affected by the ease with which buyers can change over to a substitute. A key consideration is usually the buyers switching costs—the one-time costs facing the buyer in switching from use of a product over to a substitute.

www.mepcafe.com

The economic power of suppliers is based on the ability for a group of supplier firms to exercise more bargaining power:

- When the input is, in one way or another, important to the buyer.
- When the supplier industry is dominated by a few large producers who enjoy reasonably secure market positions and who are not beleaguered by intensely competitive conditions.
- When suppliers' respective products are differentiated to such an extent that it is difficult or costly for buyers to switch from one supplier to another.
- When the buying firms are not important customers of the suppliers.
- When one or more suppliers pose a credible threat of forward integration.

The economic power of customers is based on the leverage and bargaining power of customers and tends to be relatively greater:

- When customers are few in number and when they purchase in large quantities
- When customers' purchases represent a sizable percentage of the selling industry's total sales.
- When the supplying industry is comprised of large numbers of relatively small sellers.
- When the item being purchased is sufficiently standardized among sellers that customers can not only find alternative sellers but they can also switch suppliers at virtually zero cost.
- When customers pose a credible threat of backward integration.
- When the item being bought is not an important input.
- When it is economically feasible for customers to purchase the input from several suppliers rather than one.

The competitive force of rivalry in the existing industry is centered on efficiencies and economies of scale. Rivalry tends to intensify as the number of competitors increases and as they become more equal in size and capacity. Rivalry is usually stronger when demand for the product is growing slowly. Rivalry is more intense when competitors are tempted by industry conditions to use price cuts or other competitive weapons to boost unit volume. Rivalry is stronger when the products and services of competitors are so weakly differentiated that customers incur low costs in switching from one brand to another. Rivalry increases in proportion to the size of the payoff from a successful strategic move. Rivalry tends to be more vigorous when it costs more to get out of a business than to stay in and compete. Rivalry becomes more volatile and unpredictable the more diverse the competitors are in terms of their strategies, personalities, corporate priorities, resources, and countries of origin. Rivalry increases when strong companies outside the industry acquire weak firms in the industry and launch aggressive well-funded moves to transform the newly acquired competitor into a major market contender.

www.mepcafe.com

2.5.4 Imperfect Competition

Imperfectly competitive market environments are more common than perfect competition. A single firm serving an entire market for products that have no close substitutes typifies a monopolized market. Alternatively, monopoly power is achieved by economies of scale, economies of scope, cost complementarities, patents, and other legal barriers, such as import tariffs. Since the price is set by the market conditions and equating marginal revenue to marginal cost sets quantity, a supply curve does not technically exist. However, drawing a perfectly competitive supply curve and then calculating the monopoly rents to increase the marginal cost to the price can form the basic analysis of the monopoly power impact.

Monopolistic competition is a very common condition when there are many buyers and sellers, each firm produces a differentiated product, and there is free entry and exit. Differentiated products may be established in reality or in perception. Customer selection may be based on time of day, reliability, and quality to differentiate electric energy products from different providers. Firms engage in comparative advertising, brand equity, niche marketing, and green marketing to virtually differentiate products.

Other market types are defined under separate categories, such as oligopoly. Only a few firms supplying the product denote an oligopoly market, and each firm is large with respect to the total industry. Several types of oligopoly markets are found in the literature: Cournot, Bertrand, Stackelberg, Sweezy, etc. The analysis of such market structures is beyond this work.

Oligopolies also spawn most of the research due to tacit and explicit collusion. Collusion reduces the competitive market to a monopolistic market. Tacit collusion occurs when there is sufficient public disclosed information for the firms to faithfully predict the competitive actions or reactions. Explicit collusion is when two or more firms engage in price fixing.

Other factors such as the utility of the product instead of the cost are beyond the scope of this work.

Time domain solution is needed for accurate managing of a firm as the time-dependent cash flow is the most critical component of operational management. Such time analyses can be performed by the application of state-variable modern control models. Such applications are beyond the scope of this work.

Supply curves are inelastic in the short run as production and consumption cannot change quickly. If demand increases, then price increases until a new equilibrium is found. As the price increases, more fuel can be procured to meet the increased demand as drilling costs increase. Continuous pressure on supplies with increased price will entice new entries into the market to meet the new demand. Such new entries would include renewable energies, such as biofuels, solar, wind, and hydro resources. Short-run response assumes that equipment is not improved, expanded, or augmented. Long-run response assumes that new equipment may be added or production efficiencies can be implemented. Technology changes are not normally included in the long-run response, as technology changes are fundamental production changes that are hard to predict.

www.mepcafe.com

A perfectly competitive market reaches the following equilibrium under economic theory:

$$\text{SRMC} = \text{SRAC} = \text{LRMC} = \text{LRAC} \tag{2.10}$$

It is interesting to note that under such conditions, the following observations are found. There is no need for marketing as all products are the same at the same price. All products have the same quality, thus no firm can differentiate its products. There is no need for research, as profit is fixed. Such conditions are ideal for the consumer in the short run. However, the lack of product improvement would be one of the constant social concerns. Popcorn was such a commodity until marketing and genetic altering of the kernel enabled branding of the product. It is noted that popcorn is sold at a tremendous markup today and has kept most movie theaters profitable.

2.5.5 Other Factors

There are other exogenous components that can significantly change the supply–demand solution. Electric production processes create costs for a society that does not include buyers or sellers of the product. Such external costs are negative externalities. Pollution is one such negative externality. The proper economic handling of such externalities requires government intervention either directly or indirectly. Direct intervention implies taxes to provide cost reimbursements or regulations. Indirect intervention implies some type of market solution such as the exchange of pollution permits (SO_2, NO_x, carbon, etc.). Then the firms causing pollution as part of their business strategy can efficiently seek the cost of reimbursement. However, there have to be well-defined property rights for such markets to be complete and effective.

The presence of regulations to fix the price provides a ceiling on the solution. A ceiling is a maximum price that can be charged for a product. The presence of a ceiling increases the scarcity of the product. The price increases as a result. Note that price ceilings must be lower than the competitive price to be effective. Alternatively, floors can be introduced through regulation. A floor is a minimum price that can be charged for a product. Floors must be greater than the competitive price to be effective. Floors increase the quantity as the price increases beyond the quantity that would have been produced at the competitive price. The excess quantity has to be purchased by a regulatory body (government) and stored, given away, or destroyed.

Import quotas can lower the supply for a consuming nation, thus raising prices. Export quotas can lower the supply for a supplying nation, thus raising prices. It is impossible today to decouple any energy market from the global economy except under a dictatorship with closed borders. Even then high prices lead to a black market showing the difficulty of isolating a country from the global economy.

As always, indirect taxes increase the cost of production. Such taxes lower the quantity of the product as the supply curve shifts upward. Lump sum or per-unit (excise) tariffs limit foreign competition and thus increase the cost of production. Public goods are often provided from electrical production for goods that are nonrival and nonexclusionary in consumption. Streetlights are a common example. Another recent example is the failed Australian Carbon Tax to reduce emissions.

www.mepcafe.com

The threat of regulatory changes, based on customer perceptions and then requirements, is a constant problem for the electric power industry. Deregulation is actually reregulation. The industry is being forced into a competitive market. The establishment of such markets has taken centuries in other industries. Many government approaches have been taken to best unify the goal of competition with the unique delivery and structure of the electric energy industry. Customers have demanded the reregulation of the electric utility industry at all levels from industrial through residential. The concept of setting prices by a state commission has led to highly varying prices in neighboring states. The competition between states to encourage industrial expansion has resulted in cross-subsidies and a variety of regulations to entice the next best company into their state instead of other states. The number of tariffs is very large in many states, especially the micro-regulated states. Over 200 tariffs exist in the state of Oregon as of this writing.

There are many other markets that significantly influence the electric energy market. All fuel markets (coal, natural gas, oil, uranium) determine the major component of operational expenses. The labor market for engineers and other skilled workers has recently become a major component of corporate operations and planning. The financial markets have always been the significant component for planning and for operations to invest into new equipment and systems. The interaction of all of these markets establishes the price for electricity both directly and indirectly. Such analyses are commonly covered in system planning and tariff texts.

2.6 COST OF ELECTRIC ENERGY PRODUCTION

The cost of production is one of the key components of power system operation. While the accurate cost is not covered in this work, a summary to give the flavor of the calculations and the data required is found in the following.

Find the cost of electricity (CE) from a steam turbine generator plant for the parameters in Table 2.1.

TABLE 2.1 Data for Cost of Electricity Calculation

Plant cost	$10,001 kW
Fixed charge rate	20%
Fixed operation and maintenance costs	$1 51 kW/year
Variable operation and maintenance costs	8 mills/kWh
Heat rate	10,000 kJ/kWh (9479 Btu/kWh)
Fuel cost	$1.30/GJ ($1.37/MBtu)
Capacity factor	0.70 p.u.
Assumed inflation	8%/year
Assumed discount rate	12%/year
Study period	30 year

www.mepcafe.com

Calculation Procedure:

1. *Find Plant Cost*

   Use the given values: annual plant cost (1000)(0.20) = \$200/kW per year and annual operating hours (8760)(0.70) = 6132 h/year. Therefore, plant cost $(200 \times 10^3)/6132 = 32.6$ mills/kWh.

2. *Find Fixed Operation and Maintenance Costs*

   Fixed operation and maintenance costs = $(15 \times 10^3)/6132 = 2.4$ *mills/kWh.*

3. *Compute Levelized Fuel Cost*

   The levelizing factor is found from $P = S/(1+i)^N$ *and* $CRF = 0.12414$. Hence, fuel cost = $(10{,}000)(1.3)(2.06 \times 10^{-3}) = 26.8$ mills/kWh.

4. *Find Levelized Variable Operation and Maintenance Costs*

   Operation and maintenance costs = (*8*) (*2.06*) = *16.5 mills/kWh.*

5. *Determine Cost of Electricity from Plant*

   The cost is found by adding the cost components found in Steps 1, 2, 3, and 4: cost of electricity *32.6 + 2.4 + 26.8 + 16.5 = 78.3 mills/kWh.*

Related Calculations: It should be noted that the result is a levelized value and, as such, should only be used for comparison with other similar generation alternatives. This value is not directly comparable to the cost of electricity in any one year. Note that risk factors (e.g., forced outage rates for each unit) are not included properly. It is appropriate for quick evaluation of the ideal situation from a generation perspective when reliability does not include a financial penalty.

Life-cycle cost calculations may be made by estimating the changing operational costs as system economic conditions change during the estimated lifetime of the unit. The life-cycle cost should be expressed in \$/kW/year and may include differences in unit rating and reliability, as well as the changes in unit capacity factor over the lifetime of the alternatives. Because the total lifetime must be included in the calculation, it is common to use levelized values for the variable components.

The effective capacities of the alternatives, as well as their capacity factors, may be different. It is therefore necessary to assume values for system replacement capacity and system replacement energy costs.

2.7 EVOLVING MARKETS

The original business environment outlined earlier was a single, monopolistic, vertically integrated utility providing gas and electricity. Only in a few countries have a vertically integrated utility from mine or gas well to customers. As natural gas was the original energy industry, those companies are typically separate from the electric companies. Natural gas was the light and power company before the DC dynamotor was invented. Some natural gas companies were visionary to be aware of the impact of electricity and became the gas and electric company. Since natural gas heating is an economic substitute for electric heating, natural gas is still a viable alternative in many locations.

www.mepcafe.com

The natural gas companies within the United States were connected radially to gas wells owned by suppliers. Thus, the term horizontally integrated industry has been used for this structure. It is noted that natural gas has been sold through markets for some time, even before the natural gas industry was reregulated.

The rise of markets in horizontally integrated industries naturally arises as companies find economies of scale and reliability opportunities to interconnect. A market is simply the interface between a buyer and a seller for this book. Such markets are negotiated between companies individually or equivalently through an over-the-counter process. It is noted that many over-the-counter markets are simply bulletin boards, such as newspaper ads, or electronic internet sites (eBay is a prime example) where buyer and seller can "meet" to transact business. When markets are volatile, additional players enter for profit as speculators. Such players provide information on inefficiencies of markets, operation, or planning using better information than the traditional players. Uncertainties in government regulations also give rise to speculators who understand the macroeconomic implications of the global economy. It is noted that imperfect markets also give rise to speculators due to corruption, national frictions, etc.

The original markets were established several thousand years ago. The first recorded market is around 600 BCE. These markets dealt with agricultural products, such as wheat and olive oil. Many of us have been to fish markets, farmer markets, flower markets, etc. This is the definition we will use for markets in this work.

Markets became more organized as the number of the participants grew. Markets standardized on the auction mechanisms to facilitate trade and to regulate the business conducted under the market organization. Auction mechanisms are one of the key interests of economic studies. There are five mechanisms that will be presented within this work. The structure of the offerings is also a key definition of the market mechanism as is the role of an auctioneer, if there is one.

Each market is defined by a contract by the participants. The contract specifies the auction mechanisms, the rules for submitting buy quotes (bids) or sell quotes (asks), the matching of each bid and ask, the rules for recording the transaction, the rules for settlement if discrepancies arise, the quality of the product transferred, the time of delivery, the quantity of product delivered, etc. Actual contracts are available in full at the New York Mercantile Exchange (NYMEX) website.

This work will assume that one-sided auctions dominate most industries. One-sided auctions have an auctioneer representing the buyers or the sellers. The counter party is represented by individuals representing the sellers or the buyers, respectively. We will first assume that the auctioneer is representing the buyers, as is the case for most electric spot markets.

This work also recognizes the dominance of two-sided markets for most commodities. Two-sided markets are structured as a traditional fish or farmers' market with buyers and sellers in a central location, called a trading floor. The London Metals Exchange was one such market structure. The commodity exchanges in the United States are another model of such markets (as shown in the movie "Trading Places."). EBay is a two-sided market as buyers or sellers may offer their products for sellers or buyers to offer a trade. The traditional stock markets use a middle trader to buy from sellers and to sell to buyers who specializes within that industry. This middle

www.mepcafe.com

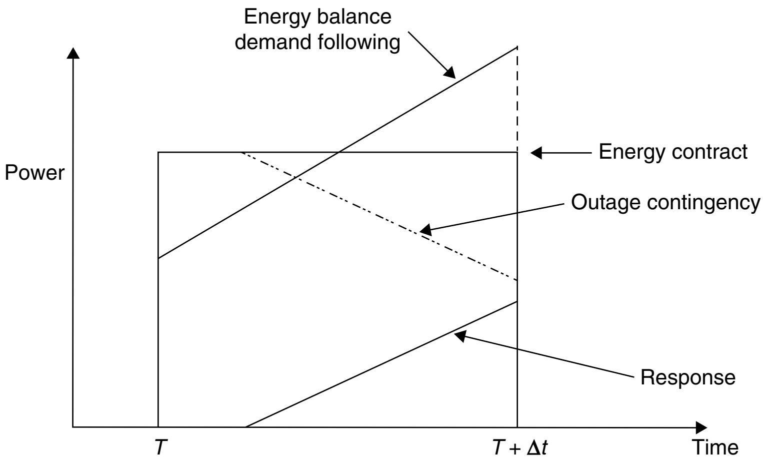


FIGURE 2.8 Segmentation of energy into commodity packages.

trader can hold an inventory to make profit as the market price rises or declines. If the market maker (middleperson) does not forecast the future prices, loses incur instead of profits.

The delivery of the product can occur at different point in time. Exchanges that immediately trade the product are called spot markets. The fish and farmers' markets are traditionally spot markets as the produce is traded immediately.

Thus, the first step is to identify the commodities to be traded. Oil, coal, natural gas, hydrogen, water, and biofuels are easier to identify as they have physical substance and are transported as solids, liquids, or gases.

Electricity is not as easily defined. The first step is to identify the duration of the electric energy to be traded. The traditional hour (dt) is segmented as shown in Figure 2.8.

The basic energy contract is the energy for the hour as if the hour is block loaded. This is the traditional interchange view as discussed later. The energy balance or demand following a contract is the change of the demand from the start of the hour to the end of the hour, shown by the rising straight line over the hour. The possibility of an outage contingency, shown as a decreasing trend is offset by a contingent contract, the first of which is known as spinning reserve. It is noted that the inertia of the system is the first response but is not yet traded as of this writing. Also the governor response is the second response as the steam or water valve is opened to provide more mechanical power to the turbine. The frequency response is the next action implemented by automatic generation control. All of these ancillary services are needed to provide the energy balance that is instantaneously required to satisfy the laws of physics. They should all be traded in a competitive market to establish the inherent response of the power system to changes in demand, transmission availability, and generation resources. The deconstruction of the demand curve into these commodities is needed to understand why these markets are needed to maintain system integrity.

The advent of wind and solar generation has given rise to a need for "inertia markets" as power electronics isolate spinning inertia from the electric power system.

www.mepcafe.com

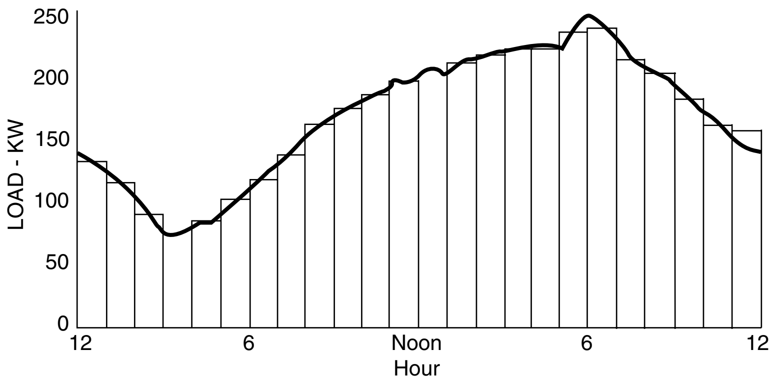


FIGURE 2.9 Demand curve [reference 1].

Contingent contracts are needed for reliability (security) such as spinning reserve and ready reserve. Contingent contracts often offer demand resources as well as static capacitor or dynamic equipment resources. Batteries and flywheels offer inertia response as a prime example. Thermal energy storage offers balancing resources. There are many other systems being implemented for these "ancillary markets."

2.7.1 Energy Flow Diagram

As a switch is turned on at home, this increase in demand is served by the conversion of the system rotating mass. This causes the frequency to decrease. This decrease is sensed at the plant by the governor and then the steam valve is opened to increase the plant output. The governor is comparing the present frequency with the actual frequency to detect these changes. Once the valve is opened sufficiently, the production is once again equal to demand. Now that the power system is at an increased demand operating point, a system-wide controller (automatic generation control) adjusts the governor setting to restore the frequency to the desired base value (60 Hz). The governor settings are found by finding the most efficient, cost-effective allocation of generation based on fuel costs, costs of conversion, and costs of transportation across the transmission grid. The slow response of the governors is masked by the fast response of the rotating inertia (mass). The slow response of the power plant boilers is masked by the faster response of the governors. The slow response of the boilers is masked by the number of boilers committed to operation by optimal scheduling of power plants given the daily demand cycle.

The daily demand cycle is based on the human work cycle. We sleep about 8 h, we work 8 h, we play 4 h, and we move from place to place over 4 h. A generic hourly demand cycle is shown in the graph (Figure 2.9).

Let us segment the demand into small, almost instantaneous, pieces representing one hour of operation, and remember that electric energy is produced and consumed almost instantaneously. The problem, once again, is to level the demand curve.

www.mepcafe.com

1. The first task is to determine the total energy for the hour to schedule the commitment of resources. The commitment of resources is subject to the capability of the generation to respond as the demand increases, decreases, or stays constant. Note that the generation has to additionally respond to the cyclic changes during the hour. The cyclic demand occurs faster than a pumped hydro unit can be switched from pumping to generation. Thus, another resource has to be used.
2. Historically, this additional resource is the fossil-fueled units. However, batteries and flywheels are also used to reduce the cost of operating the fossil-fueled power plants. Pumped hydro storage cannot be changed quickly without concerns for water dynamics. Pumped hydro is normally operated on an hourly basis as a generator or pump. Changes in pumping or generation can provide response capability in a limited fashion, just as fossil-fueled plants can respond.
3. Batteries cannot be changed quickly without loss of expected life from generation to storage. Since batteries are not as fast as flywheels, batteries are often used for the cyclic component and not the random component.
4. Thus, the demand is finally separated into two components: cyclic and random. Flywheels provide energy by adjusting their speed just as the rotating mass of each generator provides energy through adjustment of speed. Flywheels then respond only to quick and random changes so the generation can be operated with minimal governor response resulting in almost constant generator speed. Synchronous condensers used for voltage control also offer inertia response and frequency control. Thus, there is incentive to never retire an old generation unit.

Each of the aforementioned is solved as a competitive market with bids from each resource to establish inherently stable system at all future points of time within the operation planning horizon.

2.8 MULTIPLE COMPANY ENVIRONMENTS

2.8.1 Leontief Model: Input–Output Economics

The input–output concept was devised by Leontief as a representation of the interdependencies among various productive sectors of an economy. A sector is an industry or group of industries. The input–output model includes the inputs to each sector needed to make its own product. The desired result is the gross output needed to cover final and intermediate demands arising from other sectors given the final demand for the outputs of all sectors.

The Leontief system is often used to compute the economic impact of a given change in final demand of a government program. The initial conditions and the stimulus are propagated through the economy as each producer places orders for changes in inputs. Note that the total production must equal the sum of the final demands plus all intermediate stage demands. The Energy Information Administration (EIA) is the government body in the United States that performs such analysis. There are many available case studies at their website.

www.mepcafe.com

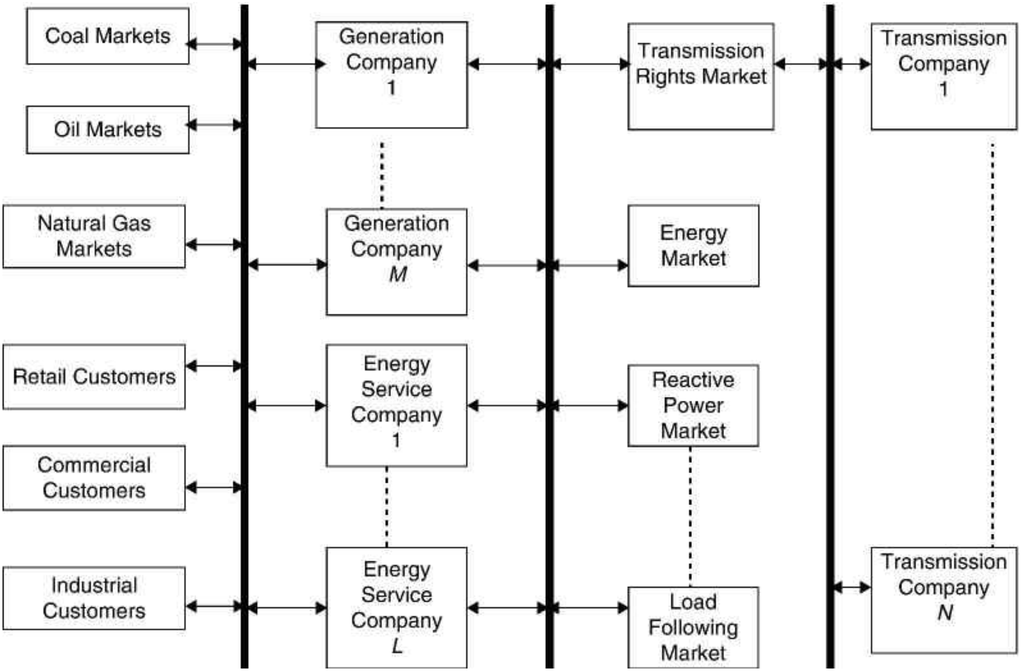


FIGURE 2.10 Leontief model with multiple markets.

www.mepcafe.com

2.8.2 Scarce Fuel Resources

Classical Leontief model calculations include assumptions of constant fuel availability and constant prices. There is always a limit to the fuel supply both in the short term due to developed harvesting operations and in the long term with the amount of fuel still interred within the earth. Thus, eventually, the fuel is exhausted. Fuel rates start rising and prices will be pushed upward as the cost rises. Such studies imbed the Leontief model into a setting or scarce resources with given and known supply curves of fuel and of other production factors. Increased final demand will increase the demands on fuel and will drive some prices upward. An extreme case is the exhaustion of fuel with no increase in output. The scarcity of the final product will drive prices considerably higher. In Figure 2.10, we see a Leontief model showing some of the multiple markets that we now deal with.

In Figure 2.11, Leontief model from Figure 2.10 has been separated and folded so that the top of the figure shows fuel to generation to transmission and the bottom shows transmission to distribution to energy service companies to retail, commercial, and industrial customers. At the top of Figure 2.11, we see two arrows, product flow

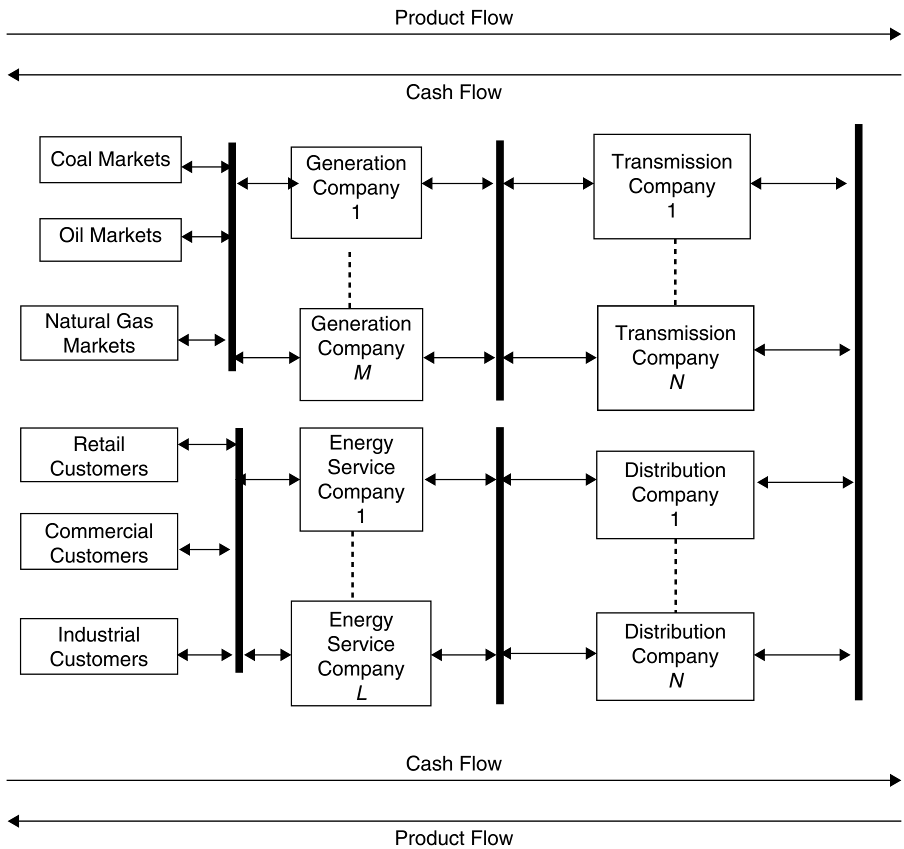


FIGURE 2.11 Cash and product flows from supplier to consumer as a supply chain.

www.mepcafe.com

and cash flow, which apply only to the top half of this folded diagram, while at the bottom the same two arrows appear but reversed in direction to apply to the bottom half of the figure.

The supply chain is central to resource management of resources. Supply chain management optimizes the profit at risk by optimal selection of contracts given the present and future costs of resources. The analysis of the Leontief model as a dynamic system into the future is central to this analysis. This analysis is based on the operation optimization as detailed in this text. The operation is the input to the Leontief dynamic model. The forecasting of fuel resources and market prices, especially fuel, is another major source of input data for energy policy and supply chain analysis. Market pricing is beyond the scope of this text and is covered in the sequel.

2.9 UNCERTAINTY AND RELIABILITY

One of the key performance indices for an electric utility is reliability. It is also the least understood at a customer level. The reliability is based on the percentage of time service is available, adjusted for those events that cannot be mitigated as "acts of god." The definition of reliability and the level of mitigation for contingent events are presently under review as of this writing. Reliability is compounded as markets are introduced since the various competitive and interindustrial relationships introduce a portfolio of risk assessment indices that require various risk management tools.

Reliability is a key issue that will be covered in the sequel to this work. However, it is a key concept that guides not only the system planning but also the operational planning of all power systems. Most utilities operate under an Obligation to Serve (OTS) rule. OTS requires that all demands have to be served but does not fix the price under such regulations. Since it is assumed that units are independent, the availability of the power system is simply the product of the availability of all units as if operated in parallel. This is a simplification that is normally accepted. The availability of a unit is determined by the forced outage rate of the unit based on historical data. The availability is one minus the forced outage rate. Reliability is an active research topic that is beyond the scope of this text. It is covered in the sequel.

PROBLEMS

2.1 The interested student should rewrite the equations for the vertically integrated utility to show that the total revenue received from all activities is equal to the sum of all input costs, plus the sum of all capacity charges. Note that the total revenue equals total cost.

2.2 The interested student should identify the regulations for his country and state (province). How many regulatory bodies are overlooking the utilities? How many interconnections exist? Which fuels are used? What storage is available? How many markets exist? How are they related? How many

www.mepcafe.com

contracts exist for the utility (GENCO, etc.) and for the customer? Are there option contracts?

2.3 What government agency analyzes and/or provides the impact of future fuel prices or energy policies?

2.4 The text by Thompson and Thore is the classic reference to model the material in this chapter using static analysis with Linear Programming. There are many excellent models in that text for energy modeling as a separate course.

REFERENCE

1. Lee Willis, H., *Power Distribution Planning Reference Book*, Second Edition Revised and Expanded, Marcel Dekker, Inc., 2004, p. 57.

www.mepcafe.com

3

ECONOMIC DISPATCH OF THERMAL UNITS AND METHODS OF SOLUTION

This chapter introduces techniques of power system optimization. For a complete understanding of how optimization problems are carried out, first read the appendix to this chapter where the concepts of the Lagrange multiplier and the Kuhn–Tucker conditions are introduced.

3.1 THE ECONOMIC DISPATCH PROBLEM

Figure 3.1 shows the configuration that will be studied in this section. This system consists of N thermal-generating units connected to a single bus bar serving a received electrical load P_{load}. The input to each unit, shown as F_i, represents the cost rate[1] of the unit. The output of each unit, P_i, is the electrical power generated by that particular unit. The total cost rate of this system is, of course, the sum of the costs of each of the individual units. The essential constraint on the operation of this system is that the sum of the output powers must equal the load demand.

Mathematically speaking, the problem may be stated very concisely. That is, an objective function, F_{T}, is equal to the total cost for supplying the indicated load. The problem is to minimize F_{T} subject to the constraint that the sum of the powers generated must equal the received load. Note that any transmission losses are neglected and any operating limits are not explicitly stated when formulating this problem. That is,

[1] Generating units consume fuel at a specific rate (e.g., MBtu/h), which as noted in Chapter 1 can be converted to \$/h, which represents a cost rate. Starting in this chapter and throughout the remainder of the text, we will simply use the term generating unit "cost" to refer to \$/h.

Power Generation, Operation, and Control, Third Edition. Allen J. Wood, Bruce F. Wollenberg, and Gerald B. Sheblé.
© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

www.mepcafe.com

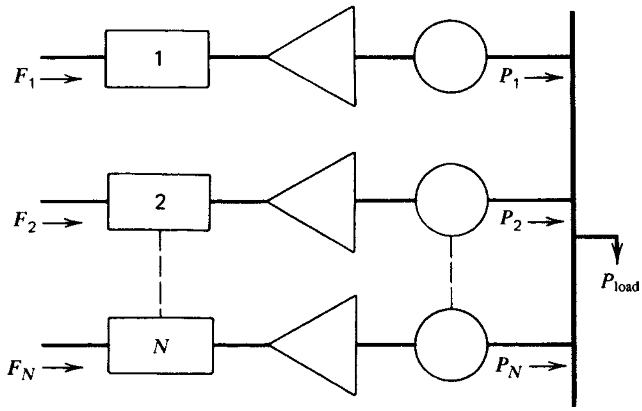


FIGURE 3.1 N_{gen} thermal units committed to serve a load of P_{load}.

$$F_T = F_1 + F_2 + F_3 + \cdots + F_{N_{\text{gen}}} \tag{3.1}$$

$$F_T = \sum_{i=1}^{N_{\text{gen}}} F_i(P_i) \tag{3.2}$$

$$\phi = 0 = P_{\text{load}} - \sum_{i=1}^{N_{\text{gen}}} P_i \tag{3.3}$$

This is a constrained optimization problem that may be attacked formally using advanced calculus methods that involve the Lagrange function.

In order to establish the necessary conditions for an extreme value of the objective function, add the constraint function to the objective function after the constraint function has been multiplied by an undetermined multiplier. This is known as the *Lagrange function* and is shown in Equation 3.4:

$$\mathcal{L} = F_T + \lambda\phi \tag{3.4}$$

The necessary conditions for an extreme value of the objective function result when we take the first derivative of the Lagrange function with respect to each of the independent variables and set the derivatives equal to 0. In this case, there are $N+1$ variables, the N values of power output, P_i, plus the undetermined Lagrange multiplier, λ. The derivative of the Lagrange function with respect to the undetermined multiplier merely gives back the constraint equation. On the other hand, the N equations that result when we take the partial derivative of the Lagrange function with respect to the power output values one at a time give the set of equations shown as Equation 3.5:

$$\frac{\partial \mathcal{L}}{\partial P_i} = \frac{\mathrm{d}F_i(P_i)}{\mathrm{d}P_i} - \lambda = 0$$

www.mepcafe.com

or

$$0 = \frac{dF_i}{dP_i} - \lambda \tag{3.5}$$

That is, the necessary condition for the existence of a minimum cost operating condition for the thermal power system is that the incremental cost rates of all the units be equal to some undetermined value, λ. Of course, to this necessary condition, we must add the constraint equation that the sum of the power outputs must be equal to the power demanded by the load. In addition, there are two inequalities that must be satisfied for each of the units. That is, the power output of each unit must be greater than or equal to the minimum power permitted and must also be less than or equal to the maximum power permitted on that particular unit.

These conditions and inequalities may be summarized as shown in the set of equations making up Equation 3.6:

$$\begin{aligned} &\frac{dF_i}{dP_i} = \lambda \quad N_{gen} \text{ equations} \\ &P_{i,\min} \le P_i \le P_{i,\max} \quad 2N_{gen} \text{ inequalities} \\ &\sum_{i=1}^{N} P_i = P_{load} \quad 1 \text{ constraint} \end{aligned} \tag{3.6}$$

When we recognize the inequality constraints, then the necessary conditions may be expanded slightly as shown in the set of equations making up Equation 3.7:

$$\begin{aligned} &\frac{dF_i}{dP_i} = \lambda \quad \text{for} \quad P_{i,\min} < P_i < P_{i,\max} \\ &\frac{dF_i}{dP_i} \le \lambda \quad \text{for} \quad P_i = P_{i,\max} \\ &\frac{dF_i}{dP_i} \ge \lambda \quad \text{for} \quad P_i = P_{i,\min} \end{aligned} \tag{3.7}$$

Several of the examples in this chapter use the following three generator units.

Unit 1: Coal-fired steam unit: Max output = 600 MW
Min output = 150 MW

Input–output curve:

$$H_1\left(\frac{\text{MBtu}}{\text{h}}\right) = 510.0 + 7.2P_1 + 0.00142P_1^2$$

Unit 2: Oil-fired steam unit: Max output = 400
MW Min output = 100 MW

www.mepcafe.com

Input–output curve:

$$H_2\left(\frac{\text{MBtu}}{\text{h}}\right) = 310.0 + 7.85P_2 + 0.00194P_2^2$$

$$\textbf{Unit 3: Oil-fired steam unit}: \text{ Max output} = 200 \text{ MW}$$
$$\text{Min output} = 50 \text{ MW}$$

Input–output curve:

$$H_3\left(\frac{\text{MBtu}}{\text{h}}\right) = 78.0 + 7.97P_3 + 0.00482P_3^2$$

Example 3A: Suppose that we wish to determine the economic operating point for these three units when delivering a total of 850 MW. Before this problem can be solved, the fuel cost of each unit must be specified. Let the following fuel costs be in effect.

$$\textbf{Unit1}: \quad \text{fuel cost} = 1.1 \text{ \$/MBtu}$$
$$\textbf{Unit2}: \quad \text{fuel cost} = 1.0 \text{ \$/MBtu}$$
$$\textbf{Unit3}: \quad \text{fuel cost} = 1.0 \text{ \$/MBtu}$$

Then

$$F_1(P_1) = H_1(P_1) \times 1.1 = 561 + 7.92P_1 + 0.001562P_1^2 \text{ \$/h}$$
$$F_2(P_2) = H_2(P_2) \times 1.0 = 310 + 7.85P_2 + 0.00194P_2^2 \text{ \$/h}$$
$$F_3(P_3) = H_3(P_3) \times 1.0 = 78 + 7.97P_3 + 0.00482P_3^2 \text{ \$/h}$$

Using Equation 3.5, the conditions for an optimum dispatch are

$$\frac{\mathrm{d}F_1}{\mathrm{d}P_1} = 7.92 + 0.003124P_1 = \lambda$$
$$\frac{\mathrm{d}F_2}{\mathrm{d}P_2} = 7.85 + 0.00388P_2 = \lambda$$
$$\frac{\mathrm{d}F_3}{\mathrm{d}P_3} = 7.97 + 0.00964P_3 = \lambda$$

and

$$P_1 + P_2 + P_3 = 850 \text{ MW}$$

Solving for λ, one obtains

$$\lambda = 9.148 \text{ \$/MWh}$$

and then solving for P_1, P_2, and P_3,

www.mepcafe.com

$$P_1 = 393.2 \text{ MW}$$
$$P_2 = 334.6 \text{ MW}$$
$$P_3 = 122.2 \text{ MW}$$

Note that all constraints are met; that is, each unit is within its high and low limit, and the total output when summed over all three units meets the desired 850 MW total.

Example 3B: Suppose the price of coal decreased to 0.9 $/MBtu. The fuel cost function for unit 1 becomes

$$F_1(P_1) = 459 + 6.48P_1 + 0.00128P_1^2$$

If one goes about the solution exactly as done here, the results are

$$\lambda = 8.284 \text{ \$/MWh}$$

and

$$P_1 = 704.6 \text{ MW}$$
$$P_2 = 111.8 \text{ MW}$$
$$P_3 = 32.6 \text{ MW}$$

This solution meets the constraint requiring total generation to equal 850 MW, but units 1 and 3 are not within limit. To solve for the most economic dispatch while meeting unit limits, use Equation 3.6:

Suppose unit 1 is set to its maximum output and unit 3 to its minimum output. The dispatch becomes

$$P_1 = 600 \text{ MW}$$
$$P_2 = 200 \text{ MW}$$
$$P_3 = 50 \text{ MW}$$

From Equation 3.6, we see that λ must equal the incremental cost of unit 2 since it is not at either limit. Then

$$\lambda = \left.\frac{dF_2}{dP_2}\right|_{P_2=200} = 8.626 \text{ \$/MWh}$$

Next, calculate the incremental cost for units 1 and 3 to see if they meet the conditions of Equation 3.6.

$$\left.\frac{dF_1}{dP_1}\right|_{P_1=600} = 8.016 \text{ \$/MWh}$$

$$\left.\frac{dF_3}{dP_3}\right|_{P_3=50} = 8.452 \text{ \$/MWh}$$

www.mepcafe.com

Note that the incremental cost for unit 1 is less than λ, so unit 1 should be at its maximum. However, the incremental cost for unit 3 is not greater than λ, so unit 3 should not be forced to its minimum. Thus, to find the optimal dispatch, allow the incremental cost at units 2 and 3 to equal λ as follows:

$$P_1 = 600 \text{ MW}$$
$$\frac{dF_2}{dP_2} = 7.85 + 0.00388P_2 = \lambda$$
$$\frac{dF_3}{dP_3} = 7.97 + 0.00964P_3 = \lambda$$
$$P_2 + P_3 = 850 - P_1 = 250 \text{ MW}$$

which results in

$$\lambda = 8.576 \text{ \$/MWh}$$

and

$$P_2 = 187.1 \text{ MW}$$
$$P_3 = 62.9 \text{ MW}$$

Note that this dispatch meets the conditions of Equation 3.6 since

$$\left.\frac{dF_1}{dP_1}\right|_{P_1 = 600\text{MW}} = 8.016 \text{ \$/MWh}$$

which is less than λ, while (dF_2/dP_2) and (dF_3/dP_3) both equal λ.

3.2 ECONOMIC DISPATCH WITH PIECEWISE LINEAR COST FUNCTIONS

Many electric utilities prefer to represent their generator cost functions as single- or multiple-segment linear cost functions. The curves shown in Figure 3.2 are representative of such functions. Note that if we were to attempt to use the lambda iteration search method on the single-segment cost function, we would always land on P_{min} or P_{max} unless λ exactly matched the incremental cost at which point the value of P would be undetermined. To resolve this problem, we perform the dispatch differently.

For all units running, we start with all of them at P_{min}, then begin to raise the output of the unit with the lowest incremental cost segment. If this unit hits the right-hand end of a segment, or if it hits P_{max}, we then find the unit with the next lowest incremental cost segment and raise its output. Eventually, we will reach a

www.mepcafe.com

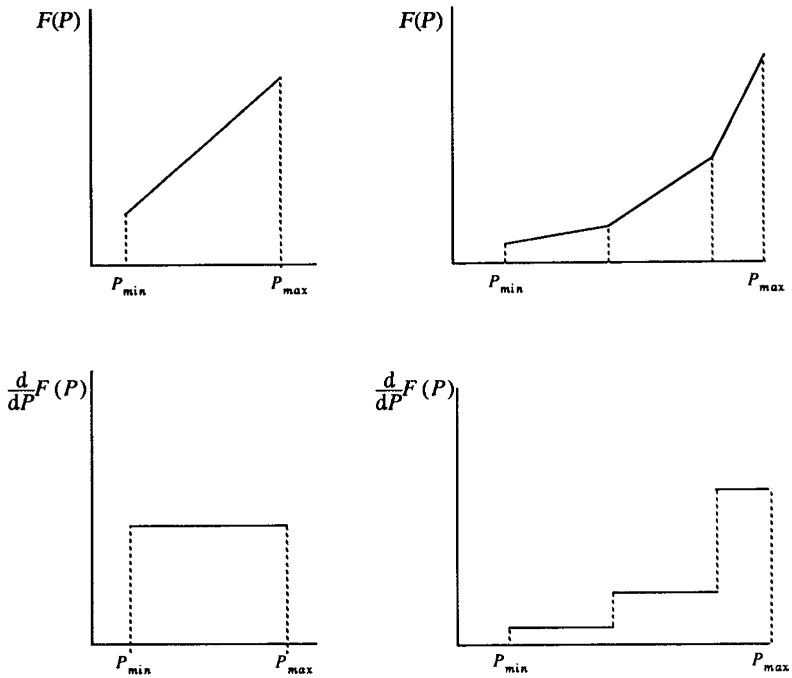


FIGURE 3.2 Piecewise linear cost functions.

point where a unit's output is being raised and the total of all unit outputs equals the total load, or load plus losses. At that point, we assign the last unit being adjusted to have a generation that is partially loaded for one segment. Note that if there are two units with exactly the same incremental cost, we could simply load them equally, although any allocation of generation to such units is arbitrary.

To make this procedure very fast, we can create a table giving each segment of each unit its MW contribution (the right-hand end MW minus the left-hand end MW). Then we order this table by ascending order of incremental cost. By searching from the top down in this table, we do not have to go and look for the next segment each time a new segment is to be chosen. This is an extremely fast form of economic dispatch. In the next section, we show how piecewise linear cost functions can be used in a linear programming (LP) solution.

3.3 LP METHOD

3.3.1 Piecewise Linear Cost Functions

LP is very adept at handling inequality constraints as long as the problem to be solved is such that it can be linearized without loss of accuracy.

In the formulation that follows, we show how the economic dispatch problem can be structured as an LP. First, we tackle the problem of expressing the nonlinear input–output or cost functions as a set of linear functions.

www.mepcafe.com

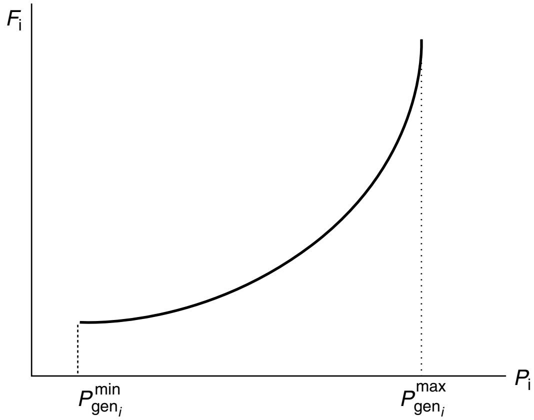


FIGURE 3.3 Nonlinear cost function characteristic.

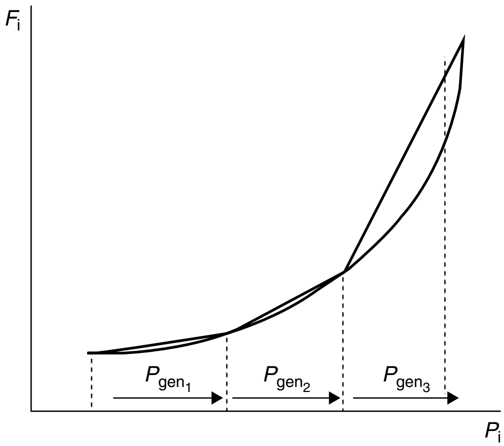


FIGURE 3.4 Nonlinear cost function approximated by straight-line segments.

We start with a nonlinear cost function shown in Figure 3.3.

We can approximate this nonlinear function as a series of straight-line segments as shown in Figure 3.4.

The three segments for generator i shown will be represented as $i1$, $i2$, and $i3$. The P_i variable is replaced with three new variables $P_{gen_{i1}}$, $P_{gen_{i2}}$, and $P_{gen_{i3}}$. Each segment will have a slope designated s_{i1}, s_{i2}, s_{i3} (where $s_{i1} < s_{i2} < s_{i3}$); then the cost function itself is now represented as the sum of the cost at P_i^{min} plus the sum of the linear cost for each segment which is simply its slope times the P_{ij} variable. Then

$$F_i\left(P_{gen_i}\right) = F_i\left(P_{gen_i}^{min}\right) + s_{i1}P_{gen_{i1}} + s_{i2}P_{gen_{i2}} + s_{i3}P_{gen_{i3}}$$

where

www.mepcafe.com

$$0 \le P_{\text{gen}_{ik}} \le P_{\text{gen}_{ik}}^{\min} \quad \text{for} \quad k = 1,2,3$$

and finally

$$P_{\text{gen}_i} = P_{\text{gen}_i}^{\min} + P_{\text{gen}_i} + P_{\text{gen}_{i1}} + P_{\text{gen}_{i2}} + P_{\text{gen}_{i3}}$$

and

$$S_{ik} = \frac{F_i\left(P_{\text{gen}_{ik+1}}\right) - F_i\left(P_{\text{gen}_{ik}}\right)}{\left(P_{\text{gen}_{ik+1}}\right) - \left(P_{\text{gen}_{ik}}\right)}$$

The cost function is now made up of a linear expression in the three variables $P_{\text{gen}_{i1}}, P_{\text{gen}_{i2}}, P_{\text{gen}_{i3}}$.

Because the slopes increase in value, the linear program will cause $P_{\text{gen}_{ik}}$ to be at its limit $P_{\text{gen}_{ik}}^{\max}$ before $P_{\text{gen}_{i(k+1)}}$ increases beyond 0.

3.3.2 Economic Dispatch with LP

Now the LP solution of the economic dispatch can be written as

$$\text{Minimize} \sum_{i=1}^{N_{\text{gen}}} \left(F_i\left(P_{\text{gen}_i}^{\min}\right) + s_{i1} P_{\text{gen}_{i1}} + s_{i2} P_{\text{gen}_{i2}} + s_{i3} P_{\text{gen}_{i3}} \right)$$

$$0 \le P_{\text{gen}_{ik}} \le_{ik} P_{\text{gen}_{ik}}^{\min} \text{ for } k = 1, 2, 3, \ldots \text{ for all generators } i = 1 \ldots N_{\text{gen}}$$

and finally

$$P_i = P_i^{\min} + P_{\text{gen}_{i1}} + P_{\text{gen}_{i2}} + P_{\text{gen}_{i3}} \text{ for all generators } i = 1 \ldots N_{\text{gen}}$$

subject to

$$\sum_{i=1}^{N_{\text{gen}}} P_i = P_{\text{load}}$$

We shall now show how LP solves the same three-generator economic dispatch given in Example 3A. Before we investigate the solutions the LP gives, it is interesting to see what variables the LP will be dealing with. Here is a list of all variables for the three generators, problems with three segments used for each generator:

Generator variables:

$$P_{\text{gen}_1}, P_{\text{gen}_{11}}, P_{\text{gen}_{12}}, P_{\text{gen}_{13}}, P_{\text{gen}_2}, P_{\text{gen}_{21}}, P_{\text{gen}_{22}}, P_{\text{gen}_{23}}, P_{\text{gen}_3}, P_{\text{gen}_{31}}, P_{\text{gen}_{32}}, P_{\text{gen}_{33}}$$

Together with the LP linear cost function, there will be three generation equations to capture the value of the generator as a function of the segment powers, that is, three equations of the type $P_{\text{gen}_i} = P_{\text{gen}_i}^{\min} + P_{\text{gen}_1} + P_{\text{gen}_{i2}} + P_{\text{gen}_{i3}}$, and there will be one equation to force the total generation to equal the load.

www.mepcafe.com

In Chapter 8, we continue this LP example and show how it can easily be used to incorporate transmission line flow constraints.

Example 3C:

Solution Using LP

If we take the three-generator example from the beginning of Example 3A, repeated here:

$$F_1\left(P_{\text{gen}_1}\right) = 561 + 7.92 P_{\text{gen}_1} + 0.00562 P_{\text{gen}_1}^2$$

$$F_2\left(P_{\text{gen}_2}\right) = 310 + 7.85 P_{\text{gen}_2} + 0.00194 P_{\text{gen}1}^2$$

$$F_3\left(P_{\text{gen}_3}\right) = 78 + 7.97 P_{\text{gen}_3} + 0.00482 P_{\text{gen}_3^2}$$

with generator min and max power limits of

Generator	$P_{\text{gen}}^{\text{min}}$	$P_{\text{gen}}^{\text{max}}$
1	150	600
2	100	400
3	50	200

and with load as totaling 850 MW, we will be able to calculate the optimal power flow (OPF) using LP. The only difference is that depending on the number of segments used, the solution will be different from that obtained using the standard method. In the succeeding table, we use 1, 3, 5, 10, and 50 segments for each cost function, and one can observe that the solution closes in on the same solution as the number of segments increases.

Number of Segments	Generator 1 MW	Generator 2 MW	Generator 3 MW	Total Cost ($/h)
1	400	400	50	8227.870
2	375	350	125	8195.369
3	450	300	100	8204.105
5	400	340	110	8195.206
10	385	340	125	8194.554
50	393	335	122	8194.357
Standard solution with lambda search	393.2	334.6	122.2	8194.356

Note that increasing the number of segments does not necessarily bring the solution closer to the exact solution. When going from two segments to three segments, the solution actually gets slightly worse in terms of total cost. This is simply because the breakpoints with three segments fall further from the true solution than the two-segment case. As the number of segments is increased to five, ten, and even fifty the solution comes very close to the exact solution.

www.mepcafe.com

3.4 THE LAMBDA ITERATION METHOD

Figure 3.5 is a block diagram of the lambda iteration method of solution for the all-thermal, dispatching problem-neglecting losses. We can approach the solution to this problem by considering a graphical technique for solving the problem and then extending this into the area of computer algorithms.

Suppose we have a three-machine system and wish to find the optimum economic operating point. One approach would be to plot the incremental cost characteristics for each of these three units on the same graph, such as sketched in Figure 3.6. In order to establish the operating points of each of these three units such that we have minimum cost and at the same time satisfy the specified demand, we could use this sketch and a ruler to find the solution. That is, we could assume an incremental cost rate (λ) and find the power outputs of each of the three units for this value of incremental cost.

Of course, our first estimate will be incorrect. If we have assumed the value of incremental cost such that the total power output is too low, we must increase the λ value and try another solution. With two solutions, we can extrapolate (or interpolate) the two solutions to get closer to the desired value of total received power (see Figure 3.7).

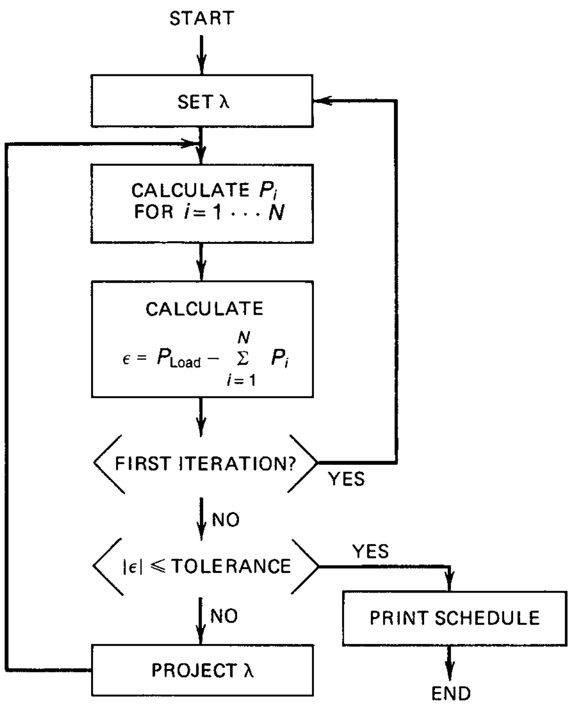


FIGURE 3.5 Economic dispatch by the lambda iteration method.

www.mepcafe.com

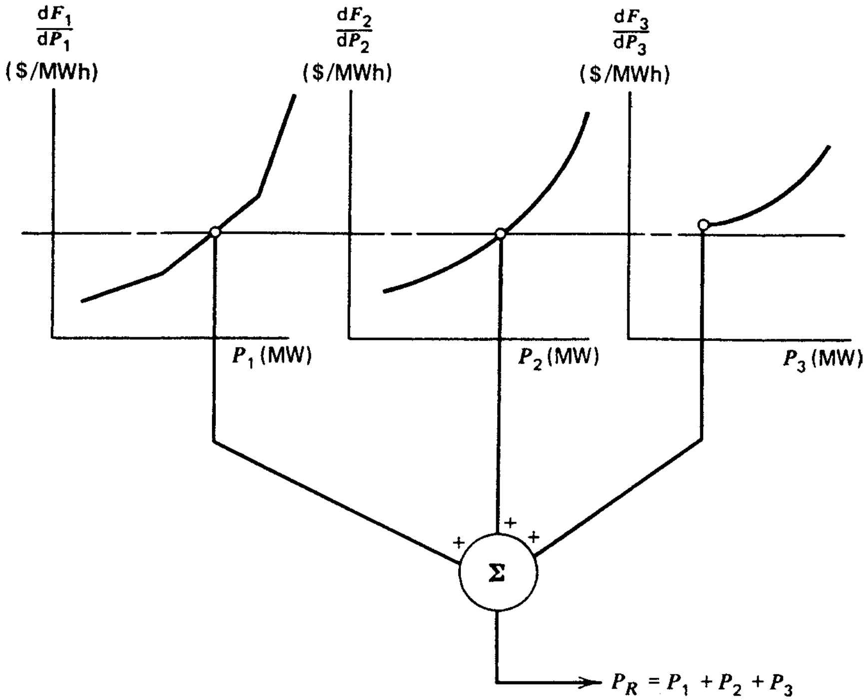


FIGURE 3.6 Graphical solution to economic dispatch.

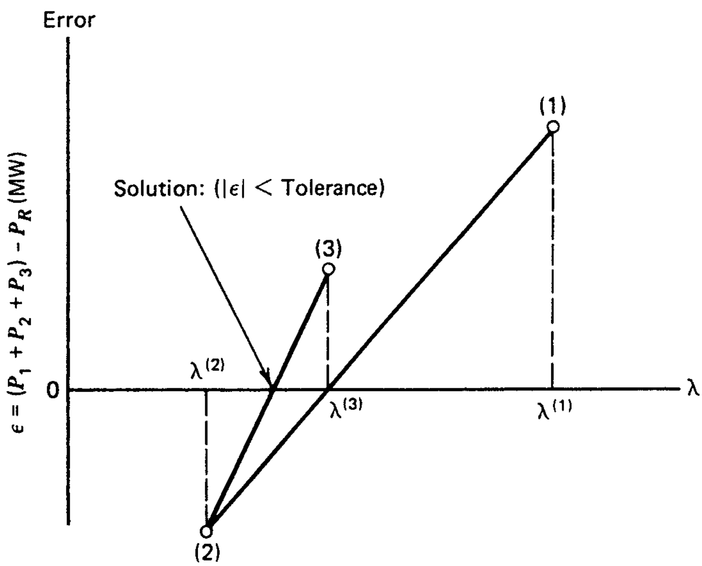


FIGURE 3.7 Lambda projections.

www.mepcafe.com

By keeping track of the total demand versus the incremental cost, we can rapidly find the desired operating point. If we wished, we could manufacture a whole series of tables that would show the total power supplied for different incremental cost levels and combinations of units.

This same procedure can be adopted for a computer implementation as shown in Figure 3.5. That is, we will now establish a set of logical rules that would enable us to accomplish the same objective as we have just done with ruler and graph paper. The actual details of how the power output is established as a function of the incremental cost rate are of very little importance. We could, for example, store tables of data within the computer and interpolate between the stored power points to find exact power output for a specified value of incremental cost rate. Another approach would be to develop an analytical function for the power output as a function of the incremental cost rate, store this function (or its coefficients) in the computer, and use this to establish the output of each of the individual units.

This procedure is an iterative type of computation, and we must establish stopping rules. Two general forms of stopping rules seem appropriate for this application. The first is shown in Figure 3.5 and is essentially a rule based on finding the proper operating point within a specified tolerance. The other, not shown in Figure 3.5, involves counting the number of times through the iterative loop and stopping when a maximum number is exceeded.

The lambda iteration procedure converges very rapidly for this particular type of optimization problem. The actual computational procedure is slightly more complex than that indicated in Figure 3.5, since it is necessary to observe the operating limits on each of the units during the course of the computation. The well-known Newton–Raphson method may be used to project the incremental cost value to drive the error between the computed and desired generation to 0.

Example 3D: Assume that one wishes to use cubic functions to represent the input–output characteristics of generating plants as follows:

$$H(\text{MBtu}/\text{h}) = A + BP + CP^2 + DP^3 \quad (P \text{ in MW})$$

For the three units, find the optimum schedule using the lambda iteration method.

	A	B	C	D
Unit 1	749.55	6.95	9.68×10^{-4}	1.27×10^{-7}
Unit 2	1285.0	7.051	7.375×10^{-4}	6.453×10^{-8}
Unit 3	1531.0	6.531	1.04×10^{-3}	9.98×10^{-8}

Assume the fuel cost to be 1.0 \$/MBtu for each unit and unit limits as follows:

$$320 \text{ MW} \le P_1 \le 800 \text{ MW}$$
$$300 \text{ MW} \le P_2 \le 1200 \text{ MW}$$
$$275 \text{ MW} \le P_3 \le 1100 \text{ MW}$$

www.mepcafe.com

Two sample calculations are shown, both using the flowchart in Figure 3.5. In this calculation, the value for λ on the second iteration is always set at 10% above or below the starting value depending on the sign of the error; for the remaining iterations, lambda is projected as in Figure 3.7.

The first example shows the advantage of starting λ near the optimum value.

$$P_{\text{load}} = 2500 \text{ MW}$$
$$\lambda_{\text{start}} = 8.0 \text{ \$/MWh}$$

Iteration	λ	Total Generation (MW)	P_1	P_2	P_3
1	8.0000	1731.6	494.3	596.7	640.6
2	8.8000	2795.0	800.0	1043.0	952.0
3	8.5781	2526.0	734.7	923.4	867.9
4	8.5566	2497.5	726.1	911.7	859.7
5	8.5586	2500.0	726.9	912.7	860.4

The second example shows the oscillatory problems that can be encountered with a lambda iteration approach.

$$P_{\text{load}} = 2500 \text{ MW}$$
$$\lambda_{\text{start}} = 10.0 \text{ \$/MWh}$$

Iteration	λ	Total Generation (MW)	P_1	P_2	P_3
1	10.0000	3100.0	800.0	1200.0	1100.0
2	9.0000	2974.8	800.0	1148.3	1026.5
3	5.2068	895.0	320.0	300.0	275.0
4	8.1340	1920.6	551.7	674.5	694.4
5	9.7878	3100.0	800.0	1200.0	1100.0
6	8.9465	2927.0	800.0	1120.3	1006.7
7	6.8692	895.0	320.0	300.0	275.0
8	8.5099	2435.0	707.3	886.1	841.7
9	8.5791	2527.4	735.1	924.0	868.3
10	8.5586	2500.1	726.9	912.8	860.4

3.5 ECONOMIC DISPATCH VIA BINARY SEARCH

A very useful lambda iteration algorithm sometimes called the “binary search”[2] avoids oscillations and always succeeds in finding the optimum economic dispatch. First, the user must calculate the incremental cost at the generation

[2] The name “binary search” and development of the method for economic dispatch are attributed to Walter O. Stadlin.

www.mepcafe.com

maximum and minimum output for each generator. Then set λ_{min} to the smallest value among the incremental costs at generator P_{min} values and then set λ_{max} as the largest value among the incremental costs at generator P_{max} values. If $\lambda = \lambda_{min}$, then the lambda search algorithm would set all generators to $P = P_{min}$, and if $\lambda = \lambda_{max}$, then the lambda search algorithm would set all generators to $P = P_{max}$. Binary search starts with

$$\Delta\lambda = \frac{(\lambda_{max} - \lambda_{min})}{2}$$

$$\lambda_i = \lambda_{min} + \Delta\lambda$$

which is halfway between each extreme.

Now calculate the generator output that corresponds to each generator having this incremental cost. If the value λ_1 is less than the incremental cost at P_{min}, then just set the generator output to P_{min}, and if the value λ_1 is greater than the incremental cost at P_{max}, then just set the generator output to P_{max}; otherwise, calculate the P value for the generator from the incremental cost function. Now add all generator outputs:

- If $\sum_{i=1}^{N_{gen}} P_i > P_{load}$, we must reduce lambda, so then

$$\Delta\lambda = \frac{\Delta\lambda}{2}$$

$$\lambda_{i+1} = \lambda_i - \Delta\lambda$$

- If $\sum_{i=1}^{N_{gen}} P_i < P_{load}$, we must increase lambda, so then

$$\Delta\lambda = \frac{\Delta\lambda}{2}$$

$$\lambda_{i+1} = \lambda_i + \Delta\lambda$$

- If $\text{abs}\left(\sum_{i=1}^{N_{gen}} P_i - P_{load}\right) \le \text{tolerance}$, algorithm is done

The delta lambda value changes by one half each iteration. Binary search will work with piecewise linear incremental cost functions and even piecewise linear incremental cost functions that have jumps between linear segments.

www.mepcafe.com

3.6 ECONOMIC DISPATCH USING DYNAMIC PROGRAMMING

As we saw in Chapter 2, when we considered the valve points in the input–output curve (e.g., Figure 1.6), the possibility of nonconvex curves must be accounted for if extreme accuracy is desired. If nonconvex input–output curves are to be used, we cannot use an equal incremental cost methodology since there are multiple values of MW output for any given value of incremental cost.

Under such circumstances, there is a way to find an optimum dispatch that uses dynamic programming (DP). If the reader has no background in DP, Appendix 3D of this chapter should be read at this time.

The DP solution to economic dispatch is done as an allocation problem, as given in Appendix 3D. Using this approach, we do not calculate a single optimum set of generator MW outputs for a specific total load supplied—rather, we generate a set of outputs, at discrete points, for an entire set of load values.

Example 3E: Assume there are three units in the system; all are on-line. Their input–output characteristics are *not* smooth *nor* convex. Data are as follows.

Power Levels (MW)	Costs ($/h)		
$P_1 = P_2 = P_3$	F_1	F_2	F_3
0	∞	∞	∞
50	810	750	806
75	1355	1155	1108.5
100	1460	1360	1411
125	1772.5	1655	1704.5
150	2085	1950	1998
175	2427.5	∞	2358
200	2760	∞	∞
225	∞	∞	∞

The total demand is $D = 310$ MW. This does not fit the data exactly, so that we need to interpolate between the closest values that are available from the data, 300 and 325 MW.

Scheduling units 1 and 2, we find the minimum cost for the function

$$f_2 = F_1(D - P_2) + F_2(P_2)$$

over the allowable range of P_2 and for $100 \leq D \leq 350$ MW. The search data are given in the succeeding table. We need to save the cost for serving each value of D that is minimal and the load level on unit 2 for each demand level.

www.mepcafe.com

		$P_2 = 0$	50	75	100	125	150 (MW)		
		$F_2(P_2) = \infty$	750	1155	1360	1655	1950 ($/h)		
D (MW)	$F_1(D)$ ($/h)							f_2 ($/h)	P_2^* (MW)
0	∞	∞	∞	∞	∞	∞	∞	∞	
50	810	∞	∞	∞	∞	∞	∞	∞	
75	1355	∞	∞	∞	∞	∞	∞	∞	
100	1460	∞	1560	∞	∞	∞	∞	1560	50
125	1772.5	∞	2105	1965	∞	∞	∞	1965	75
150	2085	∞	2210	2510	2170	∞	∞	2170	100
175	2427.5	∞	3177.5	2615	2715	2465	∞	2465	125
200	2760	∞	2834	2927.5	2820	3010	2760	2760	150
225	∞	∞	3177.5	3240	3125	3115	3305	3115	125
250	∞	∞	3510	3582.5	3445	3427	3410	3410	150
275	∞	∞	∞	3915	3787.5	3740	3722.5	3722.5	150
300	∞	∞	∞	∞	4120	4082.5	4025	4035	150
325	∞	∞	∞	∞	∞	4415	4377.5	4377.5	150
350	∞	∞	∞	∞	∞	∞	4710	4710	150

This results in

D	f_2	P_2^a
50	∞	
100	1560	50
125	1965	75
150	2170	100
175	2465	125
200	2760	150
225	3115	125
250	3410	150
275	3722.5	150
300	4035	150
325	4377.5	150
350	4710	150
375	∞	

[a]Loading of unit 2 at minimal cost level.

Next, we minimize

$$f_3 = f_2(D - P_3) + F_3(P_3)$$

for $50 \le P_3 \le 175$ MW and $D = 300$ and 325 MW. Scheduling the third unit for the two different demand levels only requires two rows of the next table.

www.mepcafe.com

		$P_3 = 0$	50	75	100	125	150	175	(MW)	
		$F_3(P_3) = \infty$	806	1108.5	1411	1704.5	1998	2358	($/h)	
D (MW)	f_2 ($/h)								f_3	P_3^*
.			.							
.				.						
.					.					
300	4035	∞	4216	4223.5	4171	4169.5	4168	4323	4168	150
325	4377.5	∞	4528.5	4518.5	4526	4464	4463	4528	4463	150
.			.							
.				.						
.					.					

The results show

D	Cost	P_3^*	P_2^*	P_1^*
300	4168	150	100	50
325	4463	150	125	50

so that between the 300- and 325-MW demand levels, the marginal unit is unit 2. (That is, it is picking up all of the additional demand increase between 300 and 325 MW.) We can, therefore, interpolate to find the cost at a load level of 310 MW or an output level on unit 2 of 110 MW. The results for a demand level of 310 MW are

$$P_1 = 50,\ P_2 = 110, \quad \text{and} \quad P_3 = 150 \quad \text{for a total cost of } 4286 \text{ \$/h}$$

One problem that is common to economic dispatch with DP is the poor control performance of the generators. We shall deal with the control of generators in Chapter 10 when we discuss automatic generation control (AGC). When a generator is under AGC and a small increment of load is added to the power system, the AGC must raise the output of the appropriate units so that the new generation output meets the load and the generators are at economic dispatch. In addition, the generators must be able to move to the new generation value within a short period of time. However, if the generators are large steam generator units, they will not be allowed to change generation output above a prescribed "maximum rate limit" of so many megawatts per minute. When this is the case, the AGC must allocate the change in generation to many other units, so that the load change can be accommodated quickly enough.

When the economic dispatch is to be done with DP and the cost curves are non-convex, we encounter a difficult problem whenever a small increment in load results in a new dispatch that calls for one or more generators to drop their output a great deal and others to increase a large amount. The resulting dispatch may be at the most

www.mepcafe.com

economic values as determined by the DP, but the control action is not acceptable and will probably violate the ramp rates for several of the units.

The only way to produce a dispatch that is acceptable to the control system, as well as being the optimum economically, is to add the ramp rate limits to the economic dispatch formulation itself. This requires a short-range load forecast to determine the most likely load and load-ramping requirements of the units. This problem can be stated as follows.

Given a load to be supplied at time increments $t = 1 \ldots t_{\text{max}}$, with load levels of P^t_{load}, and N_{gen} generators on-line to supply the load,

$$\sum_{i=1}^{N_{\text{gen}}} P_i^t = P^t_{\text{load}}$$

Each unit must obey a rate limit such that

$$P_i^{t+1} = P_i^t + \Delta P_i$$

and

$$-\Delta P_i^{\max} \le \Delta P_i \le \Delta P_i^{\max}$$

Then we must schedule the units to minimize the cost to deliver power over the time period as

$$F^{\text{total}} = \sum_{t=1}^{T_{\max}} \sum_{i=1}^{N_{\text{gen}}} F_i(P_i^t)$$

subject to

$$\sum_{i=1}^{N_{\text{gen}}} P_i^t = P^t_{\text{load}} \quad \text{for} \quad t = 1 \ldots t_{\max}$$

and

$$P_i^{t+1} = P_i^t + \Delta P_i$$

with

$$-\Delta P_i^{\max} \le \Delta P_i \le \Delta P_i^{\max}$$

This optimization problem can be solved with DP, and the "control performance" of the dispatch will be considerably better than that using DP and no ramp limit constraints.

3.7 COMPOSITE GENERATION PRODUCTION COST FUNCTION

A useful technique to facilitate the study of large numbers of generators is to develop a composite generation production cost curve. For example, suppose there were N_{gen} generator units to be scheduled against a generator T that has a fuel schedule constraint (see Chapter 5) as shown in Figure 3.8.

www.mepcafe.com

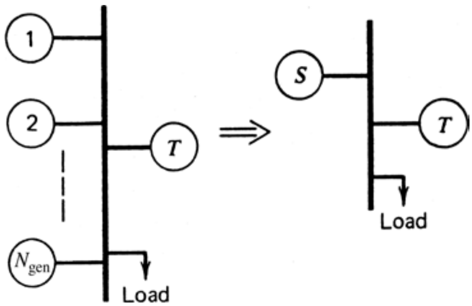


FIGURE 3.8 Composite generator unit.

Then a composite cost curve for units $1, 2, \ldots, N_{\text{gen}}$ can be developed:

$$F_s(P_s) = F_1(P_1) + \cdots + F_{N_{\text{gen}}}(P_{N_{\text{gen}}})$$

where

$$P_s = P_1 + \cdots + P_{N_{\text{gen}}}$$

and

$$\frac{dF_1}{dP_1} = \frac{dF_2}{dP_2} = \cdots = \frac{dF_{N_{\text{gen}}}}{dP_{N_{\text{gen}}}} = \lambda$$

If one of the units hits a limit, its output is held constant.

A simple procedure to allow one to generate $F_s(P_s)$ consists of adjusting λ from $\lambda^{\min}$ to $\lambda^{\max}$ in specified increments, where

$$\lambda^{\min} = \min\left(\frac{dF_i}{dP_i}, \quad i = 1 \ldots N_{\text{gen}}\right)$$

$$\lambda^{\max} = \max\left(\frac{dF_i}{dP_i}, \quad i = 1 \ldots N_{\text{gen}}\right)$$

At each increment, calculate the total fuel consumption and the total power output for all the units. These points represent points on the $F_s(P_s)$ curve. The points may be used directly by assuming that $F_s(P_s)$ consists of straight-line segments between the points, or a smooth curve may be fit to the points using a least-squares fitting program. Be aware, however, that such smooth curves may have undesirable properties such as nonconvexity (e.g., the first derivative is not monotonically increasing). The procedure to generate the points on $F_s(P_s)$ is shown in Figure 3.9.

www.mepcafe.com

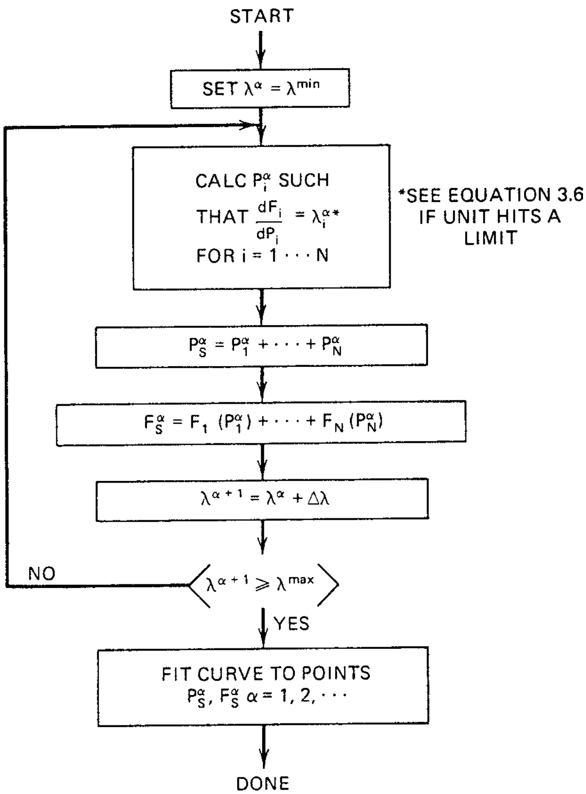


FIGURE 3.9 Procedure for obtaining composite cost curve.

Example 3F: The three generating units from Example 3A are to be combined into a composite generating unit. The fuel costs assigned to these units will be

$$\text{Fuel cost for unit1} = 1.1 \ \$/\text{MBtu}$$
$$\text{Fuel cost for unit2} = 1.4 \ \$/\text{MBtu}$$
$$\text{Fuel cost for unit3} = 1.5 \ \$/\text{MBtu}$$

Figure 3.10 shows the individual unit incremental costs, which range from 8.3886 to 14.847 \$/MWh. A program was written based on Figure 3.9, and λ was stepped from 8.3886 to 14.847.

At each increment, the three units are dispatched to the same λ and then outputs and generating costs are added as shown in Figure 3.9. The results are given in Table 3.1. The result, called F_s Approx in Table 3.1 and shown in Figure 3.11, was calculated by fitting a second-order polynomial to the and F_s points using a least-squares fitting program. The equivalent unit function is

www.mepcafe.com

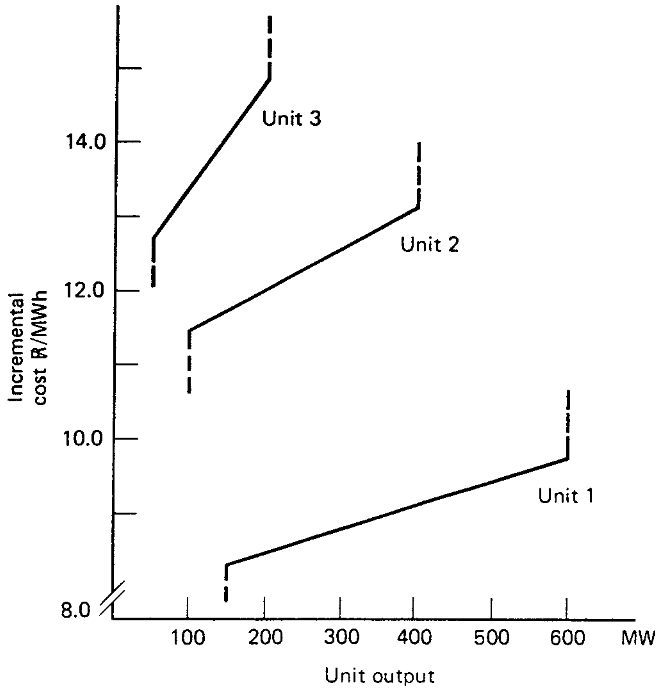


FIGURE 3.10 Unit incremental costs.

TABLE 3.1 Lambda Steps Used in Constructing a Composite Cost Curve for Example 3F

Step	λ	P_s	F_s	F_s Approx
1	8.3886	300.0	4,077.12	4,137.69
2	8.7115	403.4	4,960.92	4,924.39
3	9.0344	506.7	5,878.10	5,799.07
4	9.3574	610.1	6,828.66	6,761.72
5	9.6803	713.5	7,812.59	7,812.35
6	10.0032	750.0	8,168.30	8,204.68
7	11.6178	765.6	8,348.58	8,375.29
8	11.9407	825.0	9,048.83	9,044.86
9	12.2636	884.5	9,768.28	9,743.54
10	12.5866	943.9	10,506.92	10,471.31
11	12.9095	1019.4	11,469.56	11,436.96
12	13.2324	1088.4	12,369.40	12,360.58
13	13.5553	1110.67	12,668.51	12,668.05
14	13.8782	1133.00	12,974.84	12,979.63
15	14.2012	1155.34	13,288.37	13,295.30
16	14.5241	1177.67	13,609.12	13,615.09
17	14.8470	1200.00	13,937.00	13,938.98

www.mepcafe.com

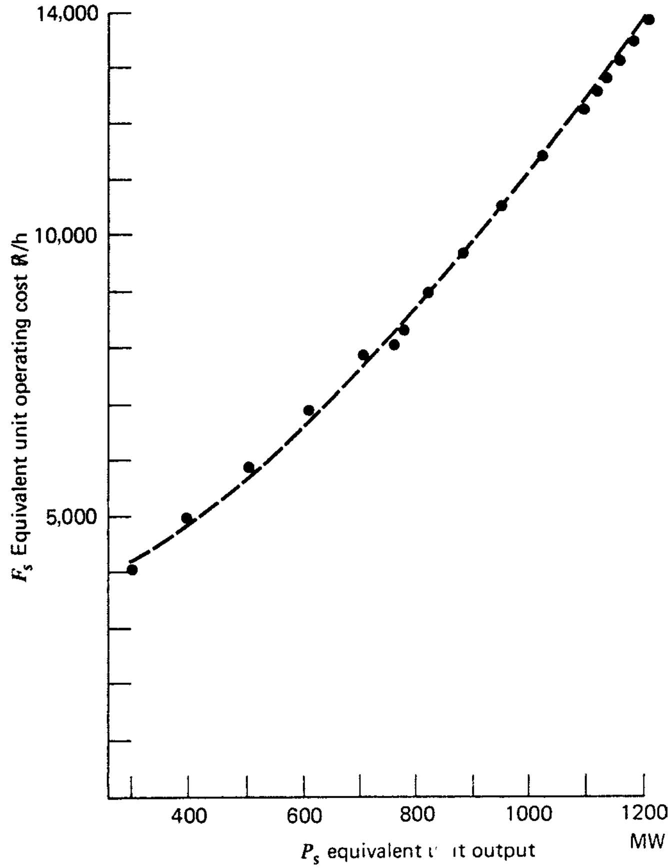


FIGURE 3.11 Equivalent unit input–output curve.

$$F_s \text{ approx}(P_s) = 2352.65 + 4.7151P_s + 0.0041168P_s^2$$
$$(\$/\text{h}) \qquad 300 \text{ MW} \le P_s \le 1200 \text{ MW}$$

The reader should be aware that when fitting a polynomial to a set of points, many choices can be made. The preceding function is a good fit to the total operating cost of the three units, but it is not that good at approximating the incremental cost. More advanced fitting methods should be used if one desires to match total operating cost as well as incremental cost. See Problem 6.2 for an alternative procedure.

3.8 BASE POINT AND PARTICIPATION FACTORS

This method assumes that the economic dispatch problem has to be solved repeatedly by moving the generators from one economically optimum schedule to another as the load changes by a reasonably small amount. We start from a

www.mepcafe.com

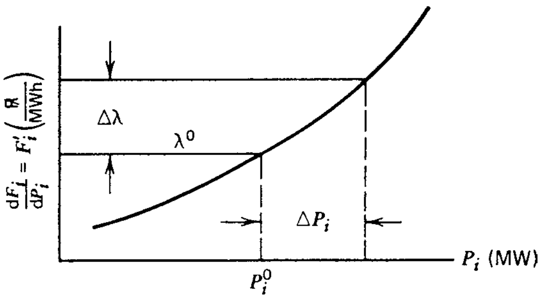


FIGURE 3.12 Relationship of $\Delta\lambda$ and ΔP_i.

given schedule—the base point. Next, the scheduler assumes a load change and investigates how much each generating unit needs to be moved (i.e., "participate" in the load change) in order that the new load be served at the most economic operating point.

Assume that both the first and second derivatives in the cost versus power output function are available (i.e., both F'_i and F''_i exist). The incremental cost curve of the ith unit is given in Figure 3.12. As the unit load is changed by an amount ΔP_i, the system incremental cost moves from λ^0 to $\lambda^0 + \Delta\lambda$. For a small change in power output on this single unit,

$$\Delta\lambda_i = \Delta\lambda \cong F''_i \ \Delta P_i$$

This is true for each of the N units on the system, so that

$$\Delta P_1 = \frac{\Delta\lambda}{F''_1}$$

$$\Delta P_2 = \frac{\Delta\lambda}{F''_2}$$

$$\vdots$$

$$\Delta P_N = \frac{\Delta\lambda}{F''_N}$$

The total change in generation (= change in total system demand) is, of course, the sum of the individual unit changes. Let P_D be the total demand on the generators (where $P_D = P_{load} + P_{loss}$), then

$$\Delta P_D = \Delta P_1 + \Delta P_2 + \cdots + \Delta P_N$$

Then the *participation factor* for each unit is

$$\left(\frac{\Delta P_i}{\Delta P_D}\right) = \frac{(1/F''_i)}{\sum_i (1/F''_i)} \tag{3.8}$$

www.mepcafe.com

The computer implementation of such a scheme of economic dispatch is straightforward. It might be done by provision of tables of the values of F_i'' as a function of the load levels and devising a simple scheme to take the existing load plus the projected increase to look up these data and compute the factors.

A somewhat less elegant scheme to provide participation factors would involve a repeat economic dispatch calculation at $P_D^0 + \Delta P_D$. The base-point economic generation values are then subtracted from the new economic generation values and the difference divided by ΔP_D to provide the participation factors. This scheme works well in computer implementations where the execution time for the economic dispatch is short and will always give consistent answers when units reach limits, pass through breakpoints on piecewise linear incremental cost functions, or have nonconvex cost curves.

Example 3G: Starting from the optimal economic solution found in Example 3A, use the participation factor method to calculate the dispatch for a total load of 900 MW.

Using Equation 3.8,

$$\frac{\Delta P_1}{\Delta P_D} = \frac{(0.003124)^{-1}}{(0.003124)^{-1} + (0.00388)^{-1} + (0.00964)^{-1}} = \frac{320.10}{681.57} = 0.47$$

Similarly,

$$\frac{\Delta P_2}{\Delta P_D} = \frac{(0.00388)^{-1}}{681.57} = 0.38$$

$$\frac{\Delta P_3}{\Delta P_D} = \frac{103.73}{681.57} = 0.15$$

$$\Delta P_D = 900 - 850 = 50$$

The new value of generation is calculated using

$$P_{\text{new}_i} = P_{\text{base}_i} + \left(\frac{\Delta P_i}{\Delta P_D}\right)\Delta P_D \quad \text{for} \quad i = 1,2,3$$

Then for each unit

$$P_{\text{new}_1} = 393.2 + (0.47)(50) = 416.7$$

$$P_{\text{new}_2} = 334.6 + (0.38)(50) = 353.6$$

$$P_{\text{new}_3} = 122.2 + (0.15)(50) = 129.7$$

www.mepcafe.com

3.9 THERMAL SYSTEM DISPATCHING WITH NETWORK LOSSES CONSIDERED

Figure 3.13 shows symbolically an all-thermal power generation system connected to an equivalent load bus through a transmission network. The economic dispatch problem associated with this particular configuration is slightly more complicated to set up than the previous case. This is because the constraint equation is now one that must include the network losses. The objective function, F_T, is the same as that defined for Equation 3.1. However, the constraint equation previously shown in Equation 3.2 must now be expanded to the one shown in Equation 3.9:

$$P_{\text{load}} + P_{\text{loss}} - \sum_{i=1}^{N_{\text{gen}}} P_i = \phi = 0 \tag{3.9}$$

The same procedure is followed in the formal sense to establish the necessary conditions for a minimum cost operating solution. The Lagrange function is shown in Equation 3.10. In taking the derivative of the Lagrange function with respect to each of the individual power outputs, P_i, it must be recognized that the loss in the transmission network, P_{loss}, is a function of the network impedances and the currents flowing in the network. For our purposes, the currents will be considered only as a function of the independent variables P_i and the load P_{load}. Taking the derivative of the Lagrange function with respect to any one of the N_{gen} values of P_i results in Equation 3.11. There are N_{gen} equations of this type to be satisfied along with the constraint equation shown in Equation 3.9. This set of equations is known collectively as the *coordination equations*:

$$\mathcal{L} = F_T + \lambda\phi \tag{3.10}$$

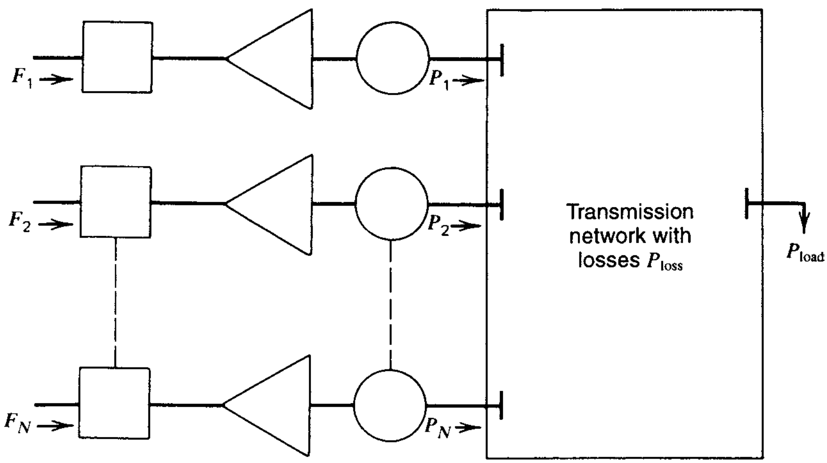


FIGURE 3.13 N_{gen} thermal units serving load through transmission network.

www.mepcafe.com

$$\frac{\partial \mathcal{L}}{\partial P_i} = \frac{dF_i}{dP_i} - \lambda\left(1 - \frac{\partial P_{\text{loss}}}{\partial P_i}\right) = 0 \tag{3.11}$$

or

$$\frac{dF_i}{dP_i} + \lambda \frac{\partial P_{\text{loss}}}{\partial P_i} = \lambda$$

$$P_{\text{load}} + P_{\text{loss}} - \sum_{i=1}^{N_{\text{gen}}} P_i = 0$$

It is much more difficult to solve this set of equations than the previous set with no losses since this second set involves the computation of the network loss in order to establish the validity of the solution in satisfying the constraint equation. There have been two general approaches to the solution of this problem. The first is the development of a mathematical expression for the losses in the network solely as a function of the power output of each of the units. This is the loss-formula method discussed at some length in Kirchmayer's *Economic Operation of Power Systems* (see Chapter 1). The other basic approach to the solution of this problem is to incorporate the power flow equations as essential constraints in the formal establishment of the optimization problem. This general approach is known as the *optimal power flow*.

Example 3H: Starting with the same units and fuel costs as in Example 3A, we will include a simplified loss expression:

$$P_{\text{loss}} = 0.00003P_1^2 + 0.00009P_2^2 + 0.00012P_3^2$$

This simplified loss formula will suffice to show the difficulties in calculating a dispatch for which losses are accounted. Note that real-world loss formulas are more complicated than the one used in this example.

Applying the coordination equations,

$$\frac{dF_1}{dP_1} = \lambda\left(1 - \frac{\partial P_{\text{loss}}}{\partial P_i}\right)$$

becomes

$$7.92 + 0.003124P_1 = \lambda\left[1 - 2(0.00003)P_1\right]$$

Similarly for P_2 and P_3,

$$7.85 + 0.00388P_2 = \lambda[1 - 2(0.00009)P_2]$$
$$7.97 + 0.00964P_3 = \lambda[1 - 2(0.00012)P_3]$$

www.mepcafe.com

and

$$P_1 + P_2 + P_3 - 850 - P_{\text{loss}} = 0$$

We no longer have a set of linear equations as in Example 3A. This necessitates a more complex solution procedure as follows.

Step 1 Pick a set of starting values for P_1, P_2, and P_3 that sum to the load.

Step 2 Calculate the incremental losses $\partial P_{\text{loss}}/\partial P_i$ as well as the total losses P_{loss}. The incremental losses and total losses will be considered constant until we return to step 2.

Step 3 Calculate the value of λ that causes P_1, P_2, and P_3 to sum to the total load plus losses. This is now as simple as the calculations in Example 3A since the equations are again linear.

Step 4 Compare the P_1, P_2, and P_3 from step 3 to the values used at the start of step 2. If there is no significant change in any one of the values, go to step 5; otherwise, go back to step 2.

Step 5 Done.

Using this procedure, we obtain:

Step 1 Pick the P_1, P_2, and P_3 starting values as

$$P_1 = 400.0 \text{ MW}$$
$$P_2 = 300.0 \text{ MW}$$
$$P_3 = 150.0 \text{ MW}$$

Step 2 Incremental losses are

$$\frac{\partial P_{\text{loss}}}{\partial P_1} = 2(0.00003)400 = 0.0240$$
$$\frac{\partial P_{\text{loss}}}{\partial P_2} = 2(0.00009)300 = 0.0540$$
$$\frac{\partial P_{\text{loss}}}{\partial P_3} = 2(0.00012)150 = 0.0360$$

Total losses are 15.6 MW.

Step 3 We can now solve for λ using the following:

$$7.92 + 0.003124P_1 = \lambda(1 - 0.0240) = \lambda(0.9760)$$
$$7.85 + 0.00388P_2 = \lambda(1 - 0.0540) = \lambda(0.9460)$$
$$7.97 + 0.00964P_3 = \lambda(1 - 0.0360) = \lambda(0.9640)$$

www.mepcafe.com

and

$$P_1 + P_2 + P_3 - 850 - 15.6 = P_1 + P_2 + P_3 - 865.6 = 0$$

These equations are now linear, so we can solve for λ directly. The results are

$$\lambda = 9.5252 \ \$/\text{MWh}$$

and the resulting generator outputs are

$$P_1 = 440.68$$
$$P_2 = 299.12$$
$$P_3 = 125.77$$

Step 4 Since these values for P_1, P_2, and P_3 are quite different from the starting values, we will return to step 2.

Step 2 The incremental losses are recalculated with the new generation values:

$$\frac{\partial P_{\text{loss}}}{\partial P_1} = 2(0.00003)440.68 = 0.0264$$
$$\frac{\partial P_{\text{loss}}}{\partial P_2} = 2(0.00009)299.12 = 0.0538$$
$$\frac{\partial P_{\text{loss}}}{\partial P_3} = 2(0.00012)125.77 = 0.0301$$

Total losses are 15.78 MW.

Step 3 The new incremental losses and total losses are incorporated into the equations, and a new value of λ and P_1, P_2, and P_3 is solved for

$$7.92 + 0.003124P_1 = \lambda(1 - 0.0264) = \lambda(0.9736)$$
$$7.85 + 0.00388P_2 = \lambda(1 - 0.0538) = \lambda(0.9462)$$
$$7.97 + 0.00964P_2 = \lambda(1 - 0.0301) = \lambda(0.9699)$$
$$P_1 + P_2 + P_3 - 850 - 15.78 = P_1 + P_2 + P_3 - 865.78 = 0$$

resulting in $\lambda = 9.5275$ \$/MWh and

$$P_1 = 433.94 \text{ MW}$$
$$P_2 = 300.11 \text{ MW}$$
$$P_3 = 131.74 \text{ MW}$$
$$\vdots$$

Table 3.2 summarizes the iterative process used to solve this problem.

www.mepcafe.com

TABLE 3.2 Iterative Process Used to Solve Iteration Example 3H

Interation	P_1(MW)	P_2 (MW)	P_3 (MW)	Losses (MW)	λ($/MWh)
Start	400.00	300.00	150.00	15.60	9.5252
1	440.68	299.12	125.77	15.78	9.5275
2	433.94	300.11	131.74	15.84	9.5285
3	435.87	299.94	130.42	15.83	9.5283
4	435.13	299.99	130.71	15.83	9.5284

3.10 THE CONCEPT OF LOCATIONAL MARGINAL PRICE (LMP)

This section demonstrates what is meant by the locational marginal price (LMP). To begin, we will assume the two-bus system shown next (Figure 3.14):

The generator cost functions used here will be piecewise linear for ease in calculating the economic dispatch.

The cost function data is shown in Table 3.3.

The cost functions are shown in Figure 3.15.

As shown in Section 3.2, when performing an economic dispatch on generators with linear segment cost functions, the calculation proceeds simply by increasing the generation on the segment with the lowest marginal price (slope) until a limit is hit or the segment MW limit is hit. We start therefore by raising the generation on generator 1, segment A, with a marginal cost of 5.00 until we exhaust its 400 MW; next we start consuming the segment with the next highest cost, which is segment C at 6.50, and we consume all of its 200 MW. We now are generating only 600 MW and need 300 MW more. The next highest cost segment is segment B at 7.50, and we only use 300 MW out of its 400 MW.

The resulting dispatch is:

Generation dispatch:

Segment	MW	Price
A	400	5.00
C	200	6.50
B	300	7.50

FIGURE 3.14 Two-bus system.

www.mepcafe.com

TABLE 3.3 Generator Cost Function Segment Data

Generator 1	MW	Marginal Cost ($/MWh)	Generator 2	Bid MW	Marginal Cost ($/MWh)
Segment A	400	5.00	Segment C	200	6.50
Segment B	800	7.50	Segment D	400	8.00

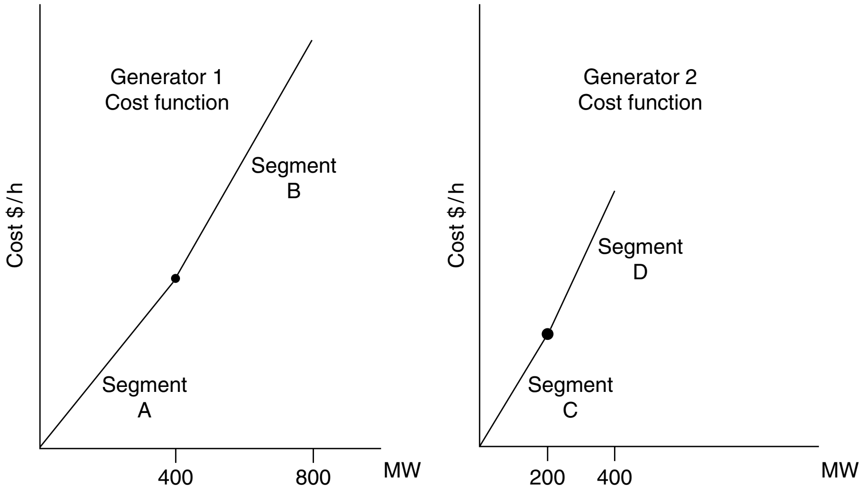


FIGURE 3.15 Generator linear segment cost functions.

The final value of 7.50 is called the "clearing price" since at that cost all the load is satisfied. The resulting power flow is shown in Figure 3.16.

A very important observation at this point is that if we were to add 1 MW to the load at either bus, that additional load would be supplied from segment B at a cost of 7.50 $/MWh. Thus, we can say that the cost of additional load at either bus is 7.50 or that the marginal price of power at either bus is 7.50. The electric power industry has settled on the term locational marginal price or LMP for this concept.

Where the concept of LMP becomes very important is when the transmission system cannot carry the power flow that might result from a simple application of economic dispatch with regard to transmission limits. We will use the two-bus system to illustrate this.

Suppose now that the transmission line between buses 1 and 2 is limited to a maximum of 100-MW flow in either direction. The economic dispatch now must be aware of the flow limit, and we can do this easily for this simple system. Again we start with the lowest cost segment, segment A of generator 1, and bring generator 1 up to 400 MW. We will assume for simplicity that the load at bus 1 is satisfied first, so that at this point the 400 MW of segment A flows into the load at bus 1 leaving another 100 MW to go. Now the

www.mepcafe.com

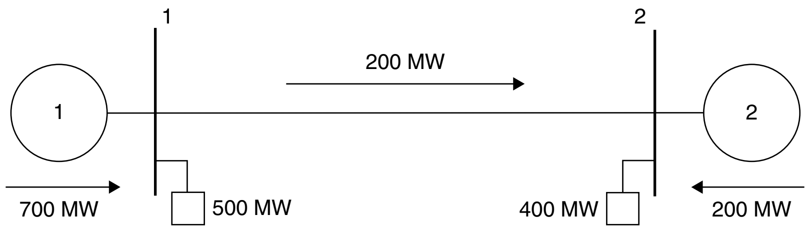


FIGURE 3.16 Resulting power flow for base dispatch.

second highest cost segment is dispatched which means we begin to use segment C of generator 2, and we dispatch all 200 MW of segment C and this 200 MW flows into the load at bus 2. Now we begin to increase the generation on segment B of generation 1. Assume the first 100 MW of this segment satisfy the load at bus 1 and then power flows over the transmission line. But this time we can only send 100 MW over the line since it is limited to 100 MW maximum. Segment B has contributed only 200 MW. We have now satisfied the load at bus 1 and are sending 100 MW to bus 2. The load at bus 2 is receiving 200 MW from segment C of generator 2 and 100 MW over the transmission line. To satisfy the remaining 100 MW needed by the load on bus 2 we must generate using segment D of generator 2. The resulting dispatch is:

Generation dispatch:

Segment	MW	Price
A	400	5.00
C	200	6.50
B	200	7.50
D	100	8.00

The flows are shown in Figure 3.17.
What we need to focus on here is the LMP for bus 1 and 2.

Bus 1: If we add 1 MW of load to the load on bus 1, it must come from segment B of generator 1 at a cost of 7.50. The LMP for bus 1 then is 7.50.
Bus 2: If we add 1 MW to the load at bus 2, it must come from segment D of generator 2 since no more power can come over the transmission line. Thus, the LMP at bus 2 is 8.0.

Whenever we ignore transmission losses (such as in the linear or DC power flow shown in Chapter 6), we can make the following observation:

- When all LMP values are the same, none of the transmission lines are at maximum flow.
- When the LMP values at different buses differ, the cause is a limiting transmission line.

www.mepcafe.com

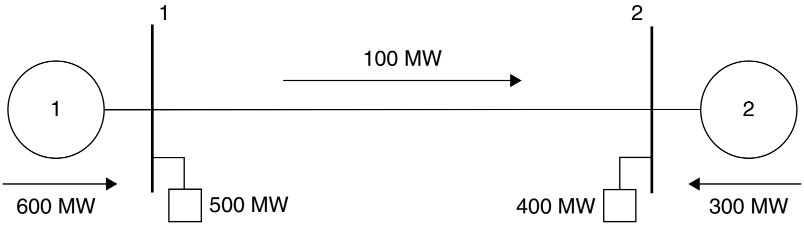


FIGURE 3.17 Flows and generation dispatch with transmission limit.

In an AC power system with transmission losses, the LMPs reflect the cost of generation, the incremental losses of the transmission system, and the limiting transmission lines. For a derivation of this, see Chapter 8.

3.11 AUCTION MECHANISMS

3.11.1 PJM Incremental Price Auction as a Graphical Solution

Recall the first Karush–Kuhn–Tucker (KKT) condition when applied to the general system (Eq. 3.9 shown again here for convenience):

$$\frac{\partial F_i}{\partial P_i} = 0 \Rightarrow \frac{\partial F_i(P_i)}{\partial P_i} - \lambda + \mu_i^- - \mu_i^+ = 0 \quad \forall \quad i = 1, N_{\text{gen}}$$

If we assume that all binding inequality constraints have been converted to equality constraints, so that the μ's are 0, then the previous equation reduces to

$$\frac{\partial F_i}{\partial P_i} = 0 \Rightarrow \frac{\partial F_i(P_i)}{\partial P_i} - \lambda = 0 \quad \forall \quad i = 1, N_{\text{gen}}$$
$$i \notin B$$

where B is the set of all generators without binding constraints. This equation implies that for all regulating generators (i.e., units not at their limits), each generator's incremental costs are the same and are equal to λ:

$$\frac{\partial F_1(P_1)}{\partial P_1} = \frac{\partial F_2(P_2)}{\partial P_2} = \cdots = \frac{\partial F_i(P_i)}{\partial P_i} = \lambda \quad \forall \quad i = 1, n \quad i \notin B$$

This very important principle provides the basis on which to apply the graphical solution method. The graphical solution is illustrated in Figure 3.6 (note that "ICC" means incremental cost curve). The unit's data are simply plotted adjacent to each other. Then, a value for λ is chosen (judiciously), a "ruler" is placed horizontally across the graphs at the value of λ, and the generations are added. If the total generation is equal to the total demand P_T, then the optimal solution has been found.

www.mepcafe.com

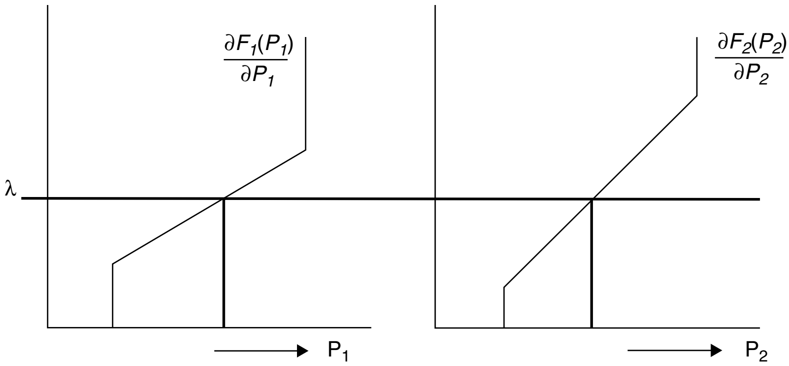


FIGURE 3.18 Graphical solution of EDC.

Otherwise, a new value for λ is chosen and the process repeated. The limitations of each unit are included as vertical lines since the solution must not include generation beyond unit capabilities (Figure 3.18).

Example 3I: Our two-unit problem can be solved using a graphical approach as shown in Figure 3.19 but converted for direct reading of the solution. The λ axis is on the far right and is used for all units since all units must have the same value for λ at the optimum solution. The ruler yields the generation for each unit at the given value of λ and is shown as a line with solid dots at each end. The ruler can then be used to find the generation for each unit for a given function of λ by moving it up and down. These generation values are then added to find the total generation. If the total generation is the generation to be dispatched, then the placement of the ruler is optimal. Otherwise, the ruler has to be moved up if the total generation is too low and down if the total generation is too high. To simplify the operation, note that the total generation for each value of λ is shown on the far right. Also the λ axis is provided at both the left- and right-hand sides for convenience. A similar production cost curve is shown in Figure 3.20 with a ruler that would move in parallel with the aforementioned ruler. The solution indicated in Figure 3.19 corresponds to a loading level of about $P_T = 410$ MW, $\lambda = 9.30$ \$/MWh, $P_1 = 223$ MW, and $P_2 = 187$ MW. See if you can find solutions for $P_T = 400$ MW and $P_T = 550$ MW.

This method was patented in the early part of the twentieth century as noted by John Zaborsky. Both authors of this chapter have seen the mechanical implementation of this method in several energy control centers, in operation as late as 1976 and more often just a showpiece as of this writing. It was used at the emergency backup control room at Pacific Gas and Electric (PG&E). This was the method offered on several energy management systems. It was also the original method for dispatch of the PJM power pool. The process was implemented to minimize the communication requirements between the various pool members. Only the status of each unit was needed that could be given only upon change of status.

www.mepcafe.com

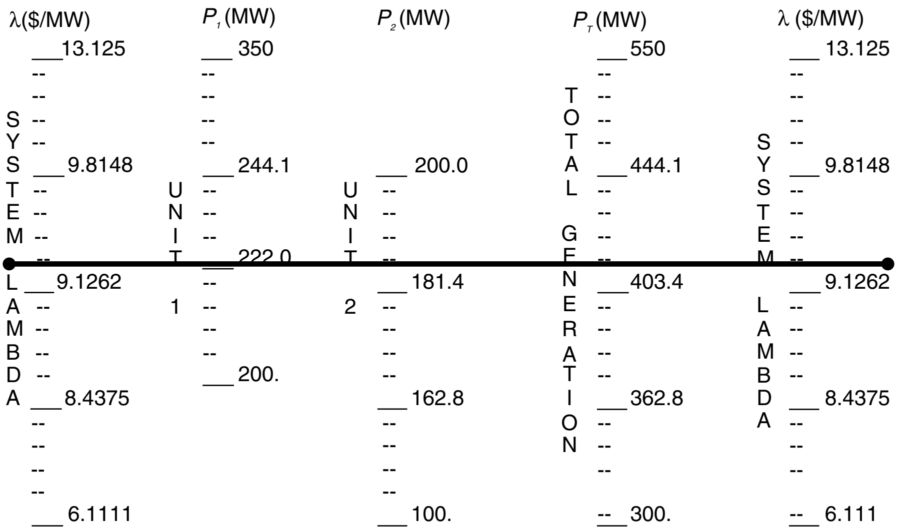


FIGURE 3.19 Economic dispatch graphical solution.

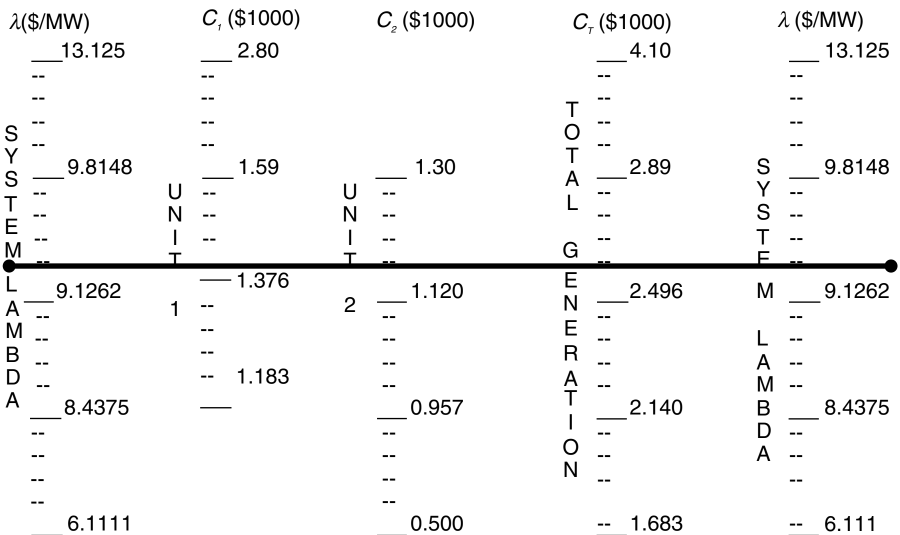


FIGURE 3.20 Production costing graph.

Procedure

PJM constructed the composite incremental cost curve for each member of the pool. Then, the complete composite curve was constructed for the pool. Originally, inter-area flow limitations were not included.

www.mepcafe.com

The total pool demand is the sum of all generation plus the interchange leaving the pool. The total pool demand was then used to find the system incremental cost from the pool composite data. This incremental cost was modified for intermember flow constraints to find the incremental cost to send to each member. Each member used their incremental cost to find the generation dispatch for each unit. The generation dispatch for each unit was entered into the AGC program for each member. The resulting economic dispatch was the reduced cost of operation as if all of the companies had merged into one company.

The difference between the dispatch as a separate company and the dispatch as a pool was used to price the benefit or cost for each member. This difference was summed and netted each month as an interchange contract and billed or paid accordingly.

3.11.2 Auction Theory Introduction

Auctions are the most widely studied economic mechanism given the rise of spectrum sales, Internet sales, and electric industry reregulation. Auctions are often referred to as arbitrary resource allocation problems with self-motivated participants. Suppliers and buyers are motivated to maximize profit. Auctions refer to procurement mechanisms as well as to sales mechanisms for profit maximization by individuals. Market-based auctions are a more efficient mechanism for suppliers and buyers to meet to exchange commodities. There are two types of auctions. An auctioneer and bidders are the most common. Another type of auction is the open-pit outcry with suppliers and buyers in the same arena, such as a market pit, a farmers market, or a fish market. Either forum can be implemented on the Internet.

Auctioneer-based auctions sell commodities where there is a pooling of suppliers to one seller with multiple buyers competing for the commodity. As an example, selling items on eBay yields bids from thousands of buyers for an individual seller. Reverse auction refers to a single buyer, now a seller, reselling a commodity to multiple sellers who are now buyers. Overstock commodities are often disposed of by reverse auctions. Note that in reverse auctions, buyers from the original auction may also be buyers in this auction. This is very widely used by procurement departments. Auctions and reverse auctions are the same theory from a different perspective. Auctions sell down the supply chain; reverse auctions sell up the supply chain.

Historically, auctions have been used for thousands of years; documents report that auctions were used in Babylon around 500 B.C. Commerce was established between Greece and Russia, wheat for olive oil, at the same time [Ancient Mariners]. The Praetorian Guard sold the Roman Empire by means of an auction in 193 A.D. after having killed Emperor Pertinax. This is the first example of the winner's curse as the buyer was later killed by the Praetorian Guard.
Auctions are used for many commodities:

- Gold
- Silver
- Platinum
- Oil

www.mepcafe.com

- Natural gas
- Pork bellies
- Orange juice

Auctions are also used at various levels of society:

- Liquidation of assets upon the death of a person
- Homes in Sydney, Australia
- Government Treasury auctions (bill, notes, bonds, securities)
- Transfer assets from public to private sector
- Transfer assets from monopolistic utilities to competitive generation companies
- Land Use Rights to drill oil, offshore oil lease
- Use of the electromagnetic spectrum
- Government and private corporations that solicit delivery of supplies
- Private firms selling products (flowers, fish, agriculture, herds, precious stones, and metals)
- Internet auctions like eBay, QuiBids, and Overstock.com.

The major question for a player in any market is how to solve the value information problem. The seller has incomplete information about buyers' valuation(s); otherwise, the seller would just set the price as the maximum valuation of the buyer. The problem for each side, suppliers and buyers, is to generate a pricing strategy to perform well in an incomplete information setting. Is an auction better suited for the goal of profit maximization for the supplier? Which type of auction yields greater revenue and thus greater profits? What are good bidding strategies for the buyer? How does a player discover collusion? Should a player enter into collusion if allowed?

Main terminology for auctions is a set of bids and asks. A bid is the price a buyer is willing to pay. An ask is the price the supplier is willing to accept for the commodity. If there is a middle player, such as in many stock markets, they make profit based on the difference between the bid and the ask or the bid–ask spread.

The comparison criterion between types of auctions is the revenue received, the efficiency of obtaining resources, the private value, the pure common value, the correlated value, and the utility. Revenue is the expected selling price. Efficiency is achieved if the object ends up in the hands of the person who values it the most. Resale does not improve efficiency when resources are not properly spread across society. The private value is when no bidder knows with certainty the valuation of the other bidders, and knowledge of the other bidders' valuation would not affect the value of the particular bidder. The pure common value is the actual value which is the same for all bidders, but bidders have different private information what that value actually is for them. An example is the auction of an oil field; as the amount of oil is unknown, different bidders have different geological signals, and acquiring another signal would change the valuation for that bidder. Different bidders have different means to market the product resulting in profit differences. The correlated value is the agent's value of a commodity,

www.mepcafe.com

which depends partly on its own preferences and partly on others' values for that commodity.

This section outlines the following:

- Single-sided auctions (auctioneer)
- Double-sided auctions

Multi-unit auctions are not included within this work where the commodities are associated with the same resource.

As noted in the introductory chapter to economics, there are several different markets based on the contracts exchanged in each ring.

Spot contracts are for immediate exchange of the commodity, such as a farm harvest just finished. Spot contracts deal with the next hour within this work. Forward contracts are for eventual exchange of the commodity or for financial settlement at some future date. Forward contracts in this work deal with the future production beyond the next hour, as is typical of a unit commitment-based solution.

Futures contracts are for monthly energy over the next 18 months as defined originally by the New York Mercantile Exchange. These contracts are beyond the scope of this work.

Options are contracts that are derived from one of the aforementioned contracts. An option gives the buyer the right but not the obligation to exercise the contract at some future point in time. These contracts are beyond the scope of this work.

3.11.3 Auction Mechanisms

Auctions are divided into one-sided and two-sided markets. One-sided markets represent either the supplier or the buyer by a common auctioneer. A two-sided market consists of buyers and sellers in a regulated ring environment. The movie ***Trading Places***[1] demonstrated a two-sided market as commonly used for commodities. The London Metal Exchange was another two-sided market that was used for centuries. The energy industry within the United States made extensive use of the NYMEX, CBOT, CME, and COMEX exchanges. These exchanges are now operated by an umbrella company (CME).

One-sided markets are used extensively in society for the majority of products. Grocery stores are a common example of a one-sided market. The suppliers contract for space to show their products to the public on a given shelf. The customers browse the store and select those products in quantity determined by the price, alternative products, and the demand for that product.

One-sided auctions have been implemented as independent system operators (ISOs) represent the customers. One-sided auctions are discussed first in this segment. ISOs use extended unit commitment programs to select suppliers based on price and quantity bids.

The first two mechanisms are the English and the Dutch mechanisms. The second pair is the first-price sealed bid and the second-price sealed bid. The structure of the bid and of the ask is also part of the mechanism. Each will be discussed within the following.

[1] Cinema Group Ventures, Eddie Murphy Productions, 1983.

www.mepcafe.com

First-price and second-price sealed bids are variants of the English method except that all bidders submit one bid for the auctioneer to accept. The bids are sealed so that other bidders do not know what is offered until the auctioneer opens them for the selection. The first-price sealed bid is awarded to the bid of the highest value. The second-price sealed bid is awarded to the highest bidder with the second highest value. Most of the electric market auctions are first-price sealed bid.

Double-sided auctions are common for most commodities in the energy industry. The **bid–offer spread** (also known as *bid–ask* or *buy–sell spread* and their equivalents using slashes in place of the dashes) for securities (such as stocks, futures contracts, options, or currency pairs) is the difference between the prices quoted (either by a single market maker or in a limit order book) for an immediate sale (offer) and an immediate purchase (bid). The size of the bid–offer spread in a security is one measure of the liquidity of the market and of the size of the transaction cost. If the spread is 0, then it is a frictionless asset.

The double-sided auction is the most interesting to watch. The London Metal Exchange is an interesting use of a trading floor for multiple commodities. One of the authors studied this market as part of his educational plan. Each metal is traded for a fixed period (15 min). At the beginning of the trading period, the traders for that metal will assemble on the floor, normally seated in a round circle as a means of achieving a civil initial posture. As the clock counts down the time remaining, the activity increases as buyers who do not get a good deal lose an opportunity. The same is true for the sellers. Thus, the initial bids and asks are given calmly but grow to a crescendo as the clock time decreases. The last two minutes are frantic as buyers need the commodities for production and the sellers need to deliver the commodities to gain income for the producers. The same frantic closings can be seen on Internet-based trading sites for the same reasons.

The auctions in this work are single-item auctions. Specifically only one product is offered per auction. Combinatorial auctions are beyond the scope of this book.

3.11.4 English (First-Price Open-Cry = Ascending)

Algorithm: Each bidder is free to raise his bid. When no bidder is willing to raise, the auction ends, and the highest bidder wins the item at the price of his bid.

Strategy: Series of bids as a function of agent's private value, his prior estimates of others' valuations, and past bids. The Porter five forces model assists with identifying the various prices possible.

Best strategy: In private value auctions, bidder's dominant strategy is to always bid a small amount more than current highest bid, and stop when his private value price is reached.

Variations:

Correlated value auctions have auctioneer increasing price at a constant rate or as he thinks is appropriate (Japanese auction).

www.mepcafe.com

Open-exit variation requires the bidder to openly declare exit without reentering possibility. This gives more information to other bidders about the agent's valuation.

ED algorithm: Any algorithm that raises the price until there is one remaining bidder and the quantity desired is purchased is of this variety.

Example 3J: Given the four generating units next,

Unit	Minimum (MW)	Maximum (MW)
1	100	500
2	50	300
3	100	400
4	100	500
Total	350	1700

the generator cost functions are

Input/Output Parameters	a ($)	b ($/MW)	c ($/MW2)	Minimum Lambda ($/MW)	Maximum Lambda ($/MW)
1	300	1.8	0.01	3.8	11.8
2	210	2.24	0.012	3.44	9.44
3	290	2.35	0.006	3.55	7.15
4	340	2.5	0.008	4.1	10.5

The English auction starts with a value of lambda and calculates the resulting P_i for each generator using the derivative curve as follows:

$$F_i(P_i) = a_i + b_i P_i + c_i P_i^2$$

$$\frac{\partial F_i(P_i)}{\partial P_i} = b_i + 2c_i P_i = \lambda$$

$$P_i = \frac{\lambda - b_i}{2c_i}$$

If P_i is below its low limit, it is set to the low limit value; similarly, if P_i is above its high limit, then it is set to the high limit value.

This is done for each generator and $\sum_{i=1}^{N_{gen}} P_i$ is subtracted from the demand to obtain an error. The error is multiplied by a "step size" and added to the initial lambda. The step size is kept as is unless the sign of the error changes from the previous step in

www.mepcafe.com

which case the step size is divided by 10. (See Appendix 3E of this chapter for the lambda update algorithm.) The steps for our three units using an English auction would be

	400			
Demand (MW)	Start	Step 1	Step 2	Step 3
English Auction				
Lambda	3.3000	3.8000	3.9417	3.8885
P_1	100.0000	100.0000	107.0833	104.4271
P_2	50.0000	65.0000	70.9028	68.6892
P_3	100.0000	120.8333	132.6389	128.2118
P_4	100.0000	100.0000	100.0000	100.0000
Sum P_i	350.0000	385.8333	410.6250	401.3281
Error	50.0000	14.1667	−10.6250	−1.3281
Step size	0.0100	0.0100	0.0050	
Delta lambda	0.5000	0.1417	−0.0531	

The English auction results in a convergence to the economic dispatch.

3.11.5 Dutch (Descending)

Algorithm: Auctioneer continuously lowers the price until a bidder takes the item at the current price. Strategically equivalent to first-price sealed-bid algorithm in all auction settings.

Strategy: Bid as a function of agent's private value and his prior estimates of others' valuations.

Best strategy: No dominant strategy exists in general. Lying (down-biasing bids) and counter speculation can be found from game theory.

It is possible to determine Nash equilibrium strategies via common knowledge assumptions regarding the probability distributions of others' values. Again, the Porter five forces can be used to estimate the other bids.

This mechanism requires multiple rounds of posting current price for the initial value to be determined as a higher price than any bidder valuation is needed.

Examples: Dutch flower market, Ontario tobacco auction, Filene's basement, Waldenbooks

ED algorithm: Any algorithm that started with all units generating and then undispatched units would mimic the Dutch algorithm. The decommitment algorithm for unit commitment by Tseng is of this generic type.

Example 3K: Using the same generator data as in we can start with a value of lambda which is too high and then converge to the solution mentioned earlier as shown in the following table:

www.mepcafe.com

	400			
Demand (MW)	Start	Step 1	Step 2	Step 3
Dutch				
Lambda	4.0000	3.7917	3.8677	3.8793
P_1	110.0000	100.0000	103.3854	103.9648
P_2	73.3333	64.6528	67.8212	68.3040
P_3	137.5000	120.1389	126.4757	127.4414
P_4	100.0000	100.0000	100.0000	100.0000
Sum P_i	420.8333	384.7917	397.6823	399.7103
Error	−20.8333	15.2083	2.3177	0.2897
Step size	0.0100	0.0050	0.0050	
Delta lambda	−0.2083	0.0760	0.0116	

The descending lambda as in a Dutch auction also converges to the economic dispatch.

3.11.6 First-Price Sealed Bid

Algorithm: Each bidder submits one bid without knowing others' bids. Bids are opened at the same time without alteration. The highest bidder wins the item at the price of his bid. This is a single round of bidding.

Strategy: Bid as a function of agent's private value and his prior estimates of others' valuations. Again, the Porter five forces can be used to estimate the other bids.

Best strategy: There is no dominant strategy in general. Strategic underbidding and counter speculation can be derived from game theory.

It is possible to determine Nash equilibrium strategies via common knowledge assumptions about the probability distributions from which valuations are drawn.

Variant: *k*th price used instead of highest price bid. If values are uniformly distributed on [0,1], then the equilibrium bid is

$$(N-1)\times\frac{x}{N}$$

where

x is the valuation of the bidder
N is the number of bidders
(Proof is left for the interested student to generate or to find.)

ED algorithm: Any algorithm that started with all units generating and submitting a bid curve and then the auctioneer opening the bids and accepting the bids in the order of increased incremental system dispatch (price). The price is the last highest bid accepted to meet the demand. This is the method used by the energy broker presented in Chapter 11. An example is given in that section.

www.mepcafe.com

3.11.7 Vickrey (Second-Price Sealed Bid)

Algorithm: Each bidder submits one bid without knowing others' bids as in first-price sealed bid. Highest bidder wins the item at 2nd highest price.

Strategy: Bid as a function of agent's private value and prior estimates of others' valuations. Again, the Porter five forces can be used to estimate the other bids.

Best strategy: In a private value auction with risk neutral bidders, Vickrey is strategically equivalent to English. In such settings, the dominant strategy is to bid one's true valuation.

No counter speculation can be found for this mechanism. It is independent of others' bidding plans, operating environments, capabilities, etc. It is often used with a single round of bidding as it is advocated for computational multiagent systems.

The old algorithm is often used for adaptive agents in agent-based computation economics such as genetic algorithm or evolutionary programming but not widely used among human traders.

Revelation principle—proxy bidder agents on www.ebay.com, www.webauction.com, www.onsale.com

ED algorithm: Any algorithm that started with all units generating and submitting a bid curve and then the auctioneer opening the bids and accepting the bids in the order of increased incremental system dispatch (price). The price is the second to last highest bid accepted to meet the demand. This is the method used by the energy broker presented in Chapter 11.

3.11.8 All Pay (e.g., Lobbying Activity)

Algorithm: Each bidder is free to raise his bid. When no bidder is willing to raise the bid, the auction ends, and the highest bidder wins the item. All bidders have to pay their last bid.

Strategy: Series of bids as a function of agent's private value, his prior estimates of others' valuations, and past bids. Again, the Porter five forces can be used to estimate the other bids.

Best strategy: In private value settings it can be computed (low bids) by game theory. It is potentially a long bidding process.

Variations: Each agent pays only part of his highest bid since all bidders are paying. Possibly, each agent's payment is a function of the highest bid of all agents. The objective is to pay almost the same amount as the English mechanism.

This bidding is associated with political action funds in government.

ED algorithm: Any algorithm that started with all units generating and submitting a bid curve and then the auctioneer opening the bids and accepting the bids in the order of increased incremental system dispatch (price). The price is the second to

www.mepcafe.com

last highest bid accepted to meet the demand. The repeated energy broker algorithm, presented in the Interchange chapter, is this type of mechanism.

APPENDIX 3A
Optimization Within Constraints

Suppose you are trying to maximize or minimize a function of several variables. It is relatively straightforward to find the maximum or minimum using rules of calculus. First, of course, you must find a set of values for the variables where the first derivative of the function with respect to each variable is 0. In addition, the second derivatives should be used to determine whether the solution found is a maximum, minimum, or a saddle point.

In optimizing a real-life problem, one is usually confronted with a function to be maximized or minimized, as well as numerous constraints that must be met. The constraints, sometimes called *side conditions*, can be other functions with conditions that must be met, or they can be simple conditions such as limits on the variables themselves.

Before we begin this discussion on constrained optimization, we will put down some definitions. Since the objective is to maximize or minimize a mathematical function, we will call this function the *objective function*. The constraint functions and simple variable limits will be lumped under the term *constraints*. The region defined by the constraints is said to be the *feasible region* for the independent variables. If the constraints are such that no such region exists, that is, there are no values for the independent variables that satisfy all the constraints, then the problem is said to have an *infeasible* solution. When an optimum solution to a constrained optimization problem occurs at the boundary of the feasible region defined by a constraint, we say the constraint is *binding*. If the optimum solution lies away from the boundary, the constraint is *nonbinding*.

To begin, let us look at a simple elliptical objective function:

$$f(x_1, x_2) = 0.25x_1^2 + x_2^2 \tag{3A.1}$$

This is shown in Figure 3.21 for various values of f.

Note that the minimum value f can attain is 0, but that it has no finite maximum value. The following is an example of a constrained optimization problem:

$$\begin{aligned} \textbf{Minimize}: &\quad f(x_1,x_2) = 0.25x_1^2 + x_2^2 \\ \textbf{Subject to the constraint}: &\quad \omega(x_1,x_2) = 0 \\ \textbf{Where}: &\quad \omega(x_1,x_2) = 5 - x_1 - x_2 \end{aligned}$$

This optimization problem can be pictured as in Figure 3.22.

We need to observe that the optimum as pictured gives the minimum value for our objective function, f, while also meeting the constraint function, ω. This optimum point occurs where the function f is exactly tangent to the function ω. Indeed, this

www.mepcafe.com

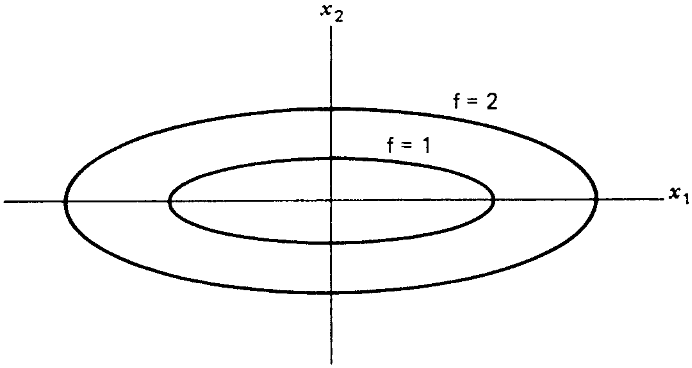


FIGURE 3.21 Elliptical objective function.

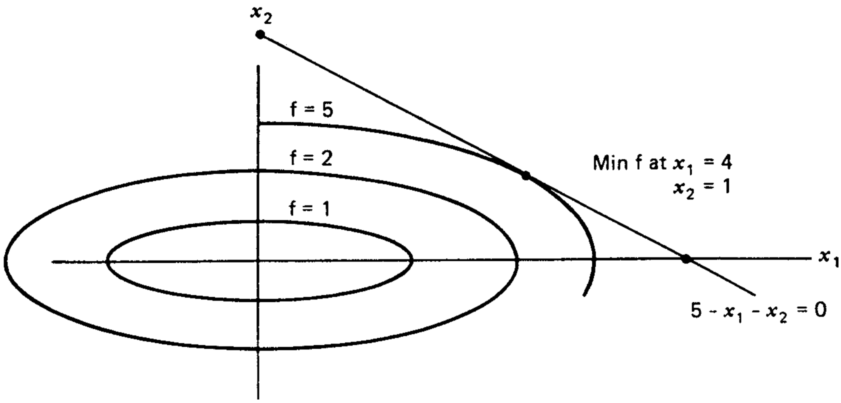


FIGURE 3.22 Elliptical objective function with equality constraint.

observation can be made more rigorous and will form the basis for our development of Lagrange multipliers.

First, redraw the function f for several values of f around the optimum point. At the point (x'_1,x'_2), calculate the gradient vector of f. This is pictured in Figure 3.23 as $\nabla f(x'_1,x'_2)$. Note that the gradient at (x'_1,x'_2) is perpendicular to f but not to ω and therefore has a nonzero component along ω. Similarly, at the point (x''_1,x''_2) the gradient of f has a nonzero component along ω. The nonzero component of the gradient along ω tells us that a small move along ω in the direction of this component will increase the objective function. Therefore, to minimize f, we should go along ω in the opposite direction to the component of the gradient projected onto ω. At the optimum point, the gradient of f is perpendicular (mathematicians say "normal") to ω, and therefore there can be no improvement in f by moving off this point. We can solve for this optimum point mathematically by using this "normal" property at the optimum. To guarantee that the gradient of f (i.e., ∇f) is normal to ω, we simply require that ∇f and the gradient of ω, $\nabla\omega$ be linearly dependent vectors. Vectors that are linearly

www.mepcafe.com

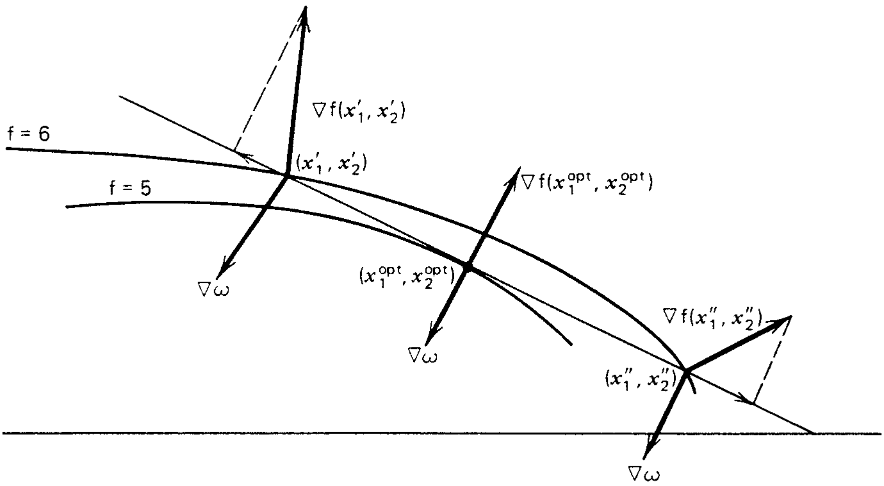


FIGURE 3.23 Gradients near a constrained optimum.

dependent must "line up" with each other (i.e., they point in exactly the same or exactly the opposite direction), although they may be different in magnitude. Mathematically, we can then set up the following equation:

$$\nabla f + \lambda \nabla \omega = 0 \tag{3A.3}$$

That is, the two gradients can be added together in such a way that they cancel each other as long as one of them is scaled. The scaling variable, λ, is called a *Lagrange multiplier*, and instead of using the gradients as shown in Equation 3A.3, we will restate them as

$$\mathcal{L}(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda \omega(x_1, x_2) \tag{3A.4}$$

This equation is called the *Lagrange equation* and consists of three variables, x_1, x_2, and λ. When we solve for the optimum values for x_1 and x_2, we will automatically calculate the correct value for λ. To meet the conditions set down in Equation 3A.3, we simply require that the partial derivative of $\mathcal{L}$ with respect to each of the unknown variables, x_1, x_2, and λ, be equal to 0. That is,

$$\textbf{Al the optimum:} \quad \frac{\partial \mathcal{L}}{\partial x_1} = 0$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = 0 \tag{3A.5}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 0$$

To show how this works, solve for the optimum point for the sample problem using Lagrange's method:

www.mepcafe.com

$$\mathcal{L}(x_1, x_2, \lambda) = 0.25x_1^2 + x_2^2 + \lambda(5 - x_1 - x_2)$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_1} &= 0.5x_1 - \lambda = 0 \\ \frac{\partial \mathcal{L}}{\partial x_2} &= 2x_2 - \lambda = 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= 5 - x_1 - x_2 = 0 \end{aligned} \tag{3A.6}$$

Note that the last equation in (3A.6) is simply the original constraint equation. The solution to Equation 3A.6 is

$$\begin{aligned} x_1 &= 4 \\ x_2 &= 1 \\ \lambda &= 2 \end{aligned} \tag{3A.7}$$

When there is more than one constraint present in the problem, the optimum point can be found in a similar manner to that just used. Suppose there were three constraints to be met, then our problem would be as follows:

$$\begin{aligned} \textbf{Minimize}: &\quad f(x_1, x_2) \\ \textbf{Subject to}: &\quad \omega_1(x_1, x_2) = 0 \\ &\quad \omega_2(x_1, x_2) = 0 \\ &\quad \omega_3(x_1, x_2) = 0 \end{aligned} \tag{3A.8}$$

The optimum point would possess the property that the gradient of f and the gradients of ω_1, ω_2, and ω_3 are linearly dependent. That is,

$$\nabla f + \lambda_1 \nabla \omega_1 + \lambda_2 \nabla \omega_2 + \lambda_3 \nabla \omega_3 = 0 \tag{3A.9}$$

Again, we can set up a Lagrangian equation as before:

$$\mathcal{L} = f(x_1, x_2) + \lambda_1 \omega_1(x_1, x_2) + \lambda_2 \omega_2(x_1, x_2) + \lambda_3 \omega_3(x_1, x_2) \tag{3A.10}$$

whose optimum occurs at

$$\begin{aligned} &\frac{\partial \mathcal{L}}{\partial x_1} = 0 \quad \frac{\partial \mathcal{L}}{\partial x_2} = 0 \\ &\frac{\partial \mathcal{L}}{\partial \lambda_1} = 0 \quad \frac{\partial \mathcal{L}}{\partial \lambda_2} = 0 \quad \frac{\partial \mathcal{L}}{\partial \lambda_3} = 0 \end{aligned} \tag{3A.11}$$

Up until now, we have assumed that all the constraints in the problem were equality constraints; that is, $\omega(x_1, x_2, \ldots) = 0$. In general, however, optimization

www.mepcafe.com

problems involve inequality constraints, that is, $g(x_1,x_2,\ldots)\le 0$, as well as equality constraints. The optimal solution to such problems will not necessarily require all the inequality constraints to be binding. Those that are binding will result in $g(x_1,x_2,\ldots)=0$ at the optimum.

The fundamental rule that tells when the optimum has been reached is presented in the paper by Kuhn and Tucker (reference 1) and in the Karush M.Sc. thesis (reference 2) in. *The* KKT conditions are presented here.

$$\begin{aligned}&\textbf{Minimize}: f(\mathbf{x})\\ &\textbf{Subject to}: \omega_i(\mathbf{x})=0 \quad i=1,2,\ldots,N\omega\\ &g_i(\mathbf{x})\le 0 \quad i=1,2,\ldots,Ng\\ &\mathbf{x}=\text{vector of real numbers, dimension}=N\end{aligned}$$

Then, forming the Lagrange function,

$$\mathcal{L}(\mathbf{x},\lambda,\mu)=f(\mathbf{x})+\sum_{i=1}^{N\omega}\lambda_i\omega_i(\mathbf{x})+\sum_{i=1}^{Ng}\mu_i g_i(\mathbf{x})$$

The conditions for an optimum for the point $\mathbf{x}^0,\lambda^0,\mu^0$ are

1. $\dfrac{\partial\mathcal{L}}{\partial x_i}\left(\mathbf{x}^0,\lambda^0,\mu^0\right)=0 \quad \text{for} \quad i=1\ldots N$

2. $\omega_i\left(\mathbf{x}^0\right)=0 \quad \text{for} \quad i=1\ldots N\omega$

3. $g_i\left(\mathbf{x}^0\right)\le 0 \quad \text{for} \quad i=1\ldots Ng$

4. $\left.\begin{aligned}\mu_i^0 g_i\left(\mathbf{x}^0\right)=0\\ \mu_i^0\ge 0\end{aligned}\right\} \quad \text{for} \quad i=1\ldots Ng$

The first condition is simply the familiar set of partial derivatives of the Lagrange function that must equal 0 at the optimum. The second and third conditions are simply a restatement of the constraint conditions on the problem. The fourth condition, often referred to *as the complimentary slackness condition*, provides a concise mathematical way to handle the problem of binding and nonbinding constraints. Since the product $\mu_i^0 g_i(\mathbf{x}^0)$ equals 0, either μ_i^0 is equal to 0 or $g_i(\mathbf{x}^0)$ is equal to 0, or both are equal to 0. If μ_i^0 is equal to 0, $g_i(\mathbf{x}^0)$ is free to be nonbinding; if μ_i^0 is positive, then $g_i(\mathbf{x}^0)$ must be 0. Thus, we have a clear indication of whether the constraint is binding or not by looking at μ_i^0.

To illustrate how the Kuhn–Tucker equations are used, we will add an inequality constraint to the sample problem used earlier in this appendix. The problem we will solve is as follows:

www.mepcafe.com

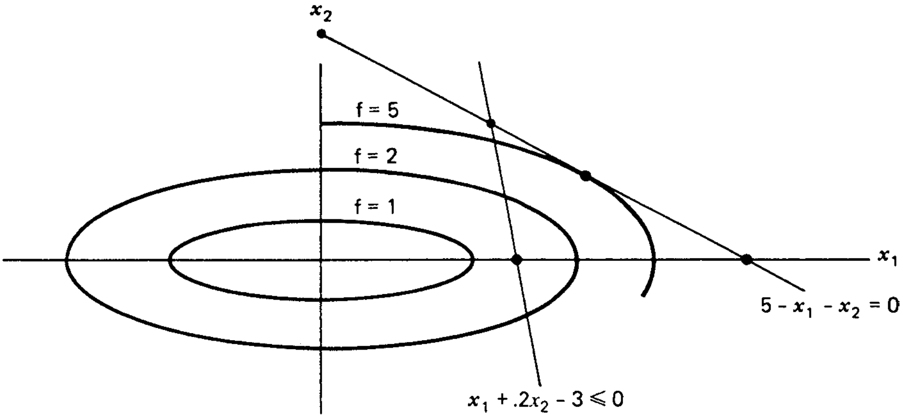


FIGURE 3.24 Elliptical objective function with equality and inequality constraints.

$$\textbf{Minimize}: \quad f(x_1,x_2) = 0.25x_1^2 + x_2^2$$
$$\textbf{Subject to}: \quad \omega(x_1,x_2) = 5 - x_1 - x_2 = 0$$
$$g(x_1,x_2) = x_1 + 0.2x_2 - 3 \le 0$$

which can be illustrated as in Figure 3.24.

First, set up the Lagrange equation for the problem:

$$\begin{aligned}\mathcal{L} &= f(x_1,x_2) + \lambda[\omega(x_1,x_2)] + \mu[g(x_1,x_2)]\\ &= 0.25x_1^2 + x_2^2 + \lambda(5 - x_1 - x_2) + \mu(x_1 + 0.2x_2 - 3)\end{aligned}$$

The first condition gives

$$\frac{\partial \mathcal{L}}{\partial x_1} = 0.5x_1 - \lambda + \mu = 0$$
$$\frac{\partial \mathcal{L}}{\partial x_2} = 2x_2 - \lambda + 0.2\mu = 0$$

The second condition gives

$$5 - x_1 - x_2 = 0$$

The third condition gives

$$x_1 + 0.2x_2 - 3 \le 0$$

The fourth condition gives

$$\mu(x_1 + 0.2x_2 - 3) = 0$$
$$\mu \ge 0$$

www.mepcafe.com

At this point, we are confronted with the fact that the Kuhn–Tucker conditions only give necessary conditions for a minimum, not a precise, procedure as to how that minimum is to be found. To solve the problem just presented, we must literally experiment with various solutions until we can verify that one of the solutions meets all four conditions. First, let $\mu = 0$, which implies that $g(x_1,x_2)$ can be less than or equal to 0. However, if $\mu = 0$, we can see that the first and second conditions give the same solution as we had previously, without the inequality constraint. But the previous solution violates our inequality constraint; and therefore the four Kuhn–Tucker conditions do not hold with $\mu = 0$. In summary,

> If $\mu = 0$, then by conditions 1 and 2
>
> $$x_1 = 4$$
> $$x_2 = 1$$
> $$\lambda = 2$$
>
> but
>
> $$g(x_1,x_2)\Big|_{\substack{x_1=4 \\ x_2=1}} = 4 + 0.2(1) - 3 = 1.2 \not\leq 0$$

Now we will try a solution in which $\mu > 0$. In this case, $g(x_1,x_2)$ must be exactly 0 and our solution can be found by solving for the intersection of $g(x_1,x_2)$ and $\omega(x_1,x_2)$, which occurs at $x_1 = 2.5$, $x_2 = 2.5$. Further, condition 1 gives $\lambda = 5.9375$ and $\mu = 4.6875$, and all four of the Kuhn–Tucker conditions are met. In summary,

> If $\mu > 0$, then by conditions 2 and 3
>
> $$x_1 = 2.5$$
> $$x_2 = 2.5$$
>
> by condition 1
>
> $$\lambda = 5.9375$$
> $$\mu = 4.6875$$
>
> and
>
> $$g(x_1,x_2)\Big|_{x_1=x_2=2.5} = 2.5 + 0.2(2.5) - 3 = 0$$
>
> All conditions are met

Considerable insight can be gained into the characteristics of optimal solutions through use of the Kuhn–Tucker conditions. One important insight comes from formulating the optimization problem so that it reflects our standard power system economic dispatch problems. Specifically, we will assume that the objective function

www.mepcafe.com

consists of a sum of individual cost functions, each of which is a function of only one variable. For example,

$$f(x_1,x_2) = C_1(x_1) + C_2(x_2)$$

Further, we will restrict this problem to have one equality constraint of the form

$$\omega(x_1,x_2) = L - x_1 - x_2 = 0$$

and a set of inequality constraints that act to restrict the problem variables within an upper and lower limit. That is,

$$x_1^- \le x_1 \le x_1^+ \rightarrow \begin{cases} g_1(x_1) = x_1 - x_1^+ \le 0 \\ g_2(x_1) = x_1^- - x_1 \le 0 \end{cases}$$

$$x_2^- \le x_2 \le x_2^+ \rightarrow \begin{cases} g_3(x_2) = x_2 - x_2^+ \le 0 \\ g_4(x_2) = x_2^- - x_2 \le 0 \end{cases}$$

Then the Lagrange function becomes

$$\begin{aligned} &= f(x_1,x_2) + \lambda\omega(x_1,x_2) + \mu_1 g_1(x_1) + \mu_2 g_2(x_1) + \mu_3 g_3(x_2) + \mu_4 g_4(x_2) \\ &= C_1(x_1) + C_2(x_2) + \lambda(L - x_1 - x_2) + \mu_1(x_1 - x_1^+) + \mu_2(x_1^- - x_1) \\ &\quad + \mu_3(x_2 - x_2^+) + \mu_4(x_2^- - x_2) \end{aligned}$$

Condition 1 gives

$$\begin{aligned} C_1'(x_1) - \lambda + \mu_1 - \mu_2 &= 0 \\ C_2'(x_2) - \lambda + \mu_3 - \mu_4 &= 0 \end{aligned}$$

Condition 2 gives

$$L - x_1 - x_2 = 0$$

Condition 3 gives

$$\begin{aligned} x_1 - x_1^+ &\le 0 \\ x_1^- - x_1 &\le 0 \\ x_2 - x_2^+ &\le 0 \\ x_2^- - x_2 &\le 0 \end{aligned}$$

Condition 4 gives

$$\begin{aligned} \mu_1(x_1 - x_1^+) &= 0 \quad \mu_1 \ge 0 \\ \mu_2(x_1^- - x_1) &= 0 \quad \mu_2 \ge 0 \\ \mu_3(x_2 - x_2^+) &= 0 \quad \mu_3 \ge 0 \\ \mu_4(x_2^- - x_2) &= 0 \quad \mu_4 \ge 0 \end{aligned}$$

www.mepcafe.com

Case 1

If the optimum solution occurs at values for x_1 and x_2 that are not at either an upper or a lower limit, then all μ values are equal to 0 and

$$C_1'(x_1) = C_2'(x_2) = \lambda$$

That is, the incremental costs associated with each variable are equal and this value is exactly the λ we are interested in.

Case 2

Now suppose that the optimum solution requires that x_1 be at its upper limit (i.e., $x_1 - x_1^+ = 0$) and that x_2 is not at its upper or lower limit. Then, $\mu_1 \geq 0$ and μ_2, μ_3, and μ_4 will each equal 0. Then, from condition 1,

$$C_1'(x_1) = \lambda - \mu_1 \rightarrow C_1'(x_1) \leq \lambda$$
$$C_2(x_2) = \lambda$$

Therefore, the incremental cost associated with the variable that is at its upper limit will always be less than or equal to λ, whereas the incremental cost associated with the variable that is not at limit will exactly equal λ.

Case 3

Now suppose the opposite of Case 2 obtains; that is, let the optimum solution require x_1 to be at its lower limit (i.e., $x_1^- - x_1 = 0$) and again assume that x_2 is not at its upper or lower limit. Then

$$\mu_2 \geq 0$$

and μ_1, μ_3, and μ_4 will each equal 0. Then from condition 1,

$$C_1'(x_1) = \lambda + \mu_2 \Rightarrow C_1'(x_1) \geq \lambda$$
$$C_2'(x_2) = \lambda$$

Therefore, the incremental cost associated with a variable at its lower limit will be greater than or equal to λ, whereas, again, the incremental cost associated with the variable that is not at limit will equal λ.

Case 4

If the optimum solution requires that both x_1, x_2 are at limit and the equality constraint can be met, then λ and the nonzero μ values are indeterminate. For example, suppose the optimum required that

$$x_1 - x_1^+ = 0$$

www.mepcafe.com

and

$$x_2 - x_2^+ = 0$$

Then

$$\mu_1 \geq 0 \quad \mu_3 \geq 0 \quad \mu_2 = \mu_4 = 0$$

Condition 1 would give

$$C_1'(x_1) = \lambda - \mu_1$$
$$C_2'(x_2) = \lambda - \mu_3$$

and the specific values for λ, μ_1, and μ_3 would be undetermined. In summary, for the general problem of N variables:

$$\textbf{Minimize}: \quad C_1(x_1) + C_2(x_2) + \cdots + C_N(x_N)$$
$$\textbf{Subject to}: \quad L - x_1 - x_2 - \cdots - x_N = 0$$

$$\textbf{And}: \quad \left.\begin{aligned} x_i - x_i^+ &\leq 0 \\ x_i^- - x_i &\leq 0 \end{aligned}\right\} \quad \text{for} \quad i = 1 \ldots N$$

Let the optimum lie at $x_i = x_i^{\text{opt}} i = 1 \ldots N$ and assume that at least one x_i is not at limit. Then,

$$\begin{array}{llll} \text{If } x_i^{\text{opt}} < x_i^+ & \text{and} \quad x_i^{\text{opt}} > x_i^-, & \text{then} & C_i(x_i^{\text{opt}}) = \lambda \\ \text{If } x_i^{\text{opt}} = x_i^+ & & & C_i'(x_i^{\text{opt}}) \leq \lambda \\ \text{If } x_i^{\text{opt}} = x_i^- & & & C_i'(x_i^{\text{opt}}) \geq \lambda \end{array}$$

Slack Variable Formulation. An alternate approach to the optimization problem with inequality constraints requires that all inequality constraints be made into equality constraints. This is done by adding slack variables in the following way:

$$\textbf{If}: g(x_1) = x_1 - x_1^+ \leq 0$$
$$\textbf{Then}: g(x_1, S_1) = x_1 - x_1^+ + S_1^2 = 0$$

We add S_1^2 rather than S_1 so that S_1 need not be limited in sign.

Making all inequality constraints into equality constraints eliminates the need for conditions 3 and 4 of the Kuhn–Tucker conditions. However, as we will see shortly, the result is essentially the same. Let us use our two-variable problem again.

www.mepcafe.com

$$\textbf{Minimize}: f(x_1,x_2) = C_1(x_1) + C_2(x_2)$$
$$\textbf{Subject to}: \omega(x_1,x_2) = L - x_1 - x_2 = 0$$
$$\textbf{And}: \begin{array}{lll} g_1(x_1) = x_1 - x_1^+ \le 0 & \text{or} & g_1(x_1,S_1) = x_1 - x_1^+ + S_1^2 = 0 \\ g_2(x_1) = x_1^- - x_1 \le 0 & & g_2(x_1,S_2) = x_1^- - x_1 + S_2^2 = 0 \\ g_3(x_2) = x_2 - x_2^+ \le 0 & & g_3(x_2,S_3) = x_2 - x_2^+ + S_3^2 = 0 \\ g_4(x_2) = x_2^- - x_2 \le 0 & & g_4(x_2,S_4) = x_2^- - x_2 + S_4^2 = 0 \end{array}$$

The resulting Lagrange function is

$$\begin{aligned} \mathcal{L} = {} & f(x_1,x_2) + \lambda_0 \omega(x_1,x_2) + \lambda_1 g_1(x_1,S_1) \\ & + \lambda_2 g_2(x_1,S_2) + \lambda_3 g_3(x_2,S_3) + \lambda_4 g_4(x_2,S_4) \end{aligned}$$

Note that all constraints are now equality constraints, so we have used only λ values as Lagrange multipliers.

Conditional 1 gives:
$$\frac{\partial \mathcal{L}}{\partial x_1} = C_1'(x_1) - \lambda_0 + \lambda_1 - \lambda_2 = 0$$
$$\frac{\partial \mathcal{L}}{\partial x_2} = C_2'(x_2) - \lambda_0 + \lambda_3 - \lambda_4 = 0$$
$$\frac{\partial \mathcal{L}}{\partial S_1} = 2\lambda_1 S_1 = 0$$
$$\frac{\partial \mathcal{L}}{\partial S_2} = 2\lambda_2 S_2 = 0$$
$$\frac{\partial \mathcal{L}}{\partial S_3} = 2\lambda_3 S_3 = 0$$
$$\frac{\partial \mathcal{L}}{\partial S_4} = 2\lambda_4 S_4 = 0$$

Condition 2 gives:
$$\mathcal{L} - x_1 - x_2 = 0$$
$$\left(x_1 - x_1^+ + S_1^2\right) = 0$$
$$\left(x_1^- - x_1 + S_2^2\right) = 0$$
$$\left(x_2 - x_2^+ + S_3^2\right) = 0$$
$$\left(x_2^- - x_2 + S_4^2\right) = 0$$

We can see that the derivatives of the Lagrange function with respect to the slack variables provide us once again with a complimentary slackness rule. For example, if $2\lambda_1 S_1 = 0$, then either $\lambda_1 = 0$ and S_1 is free to be any value or $S_1 = 0$ and λ_1 is free (or λ_1 and S_1 can both be 0). Since there are as many problem variables whether one uses

www.mepcafe.com

the slack variable form or the inequality constraint form, there is little advantage to either, other than perhaps a conceptual advantage to the student.

APPENDIX 3B
LINEAR PROGRAMMING (LP)

LP is without doubt one of the most powerful optimization techniques ever developed (see references 3–6). LP can be used to solve many problems whether they are purely linear (linear equations in the objective function and the constraints) or even nonlinear objective functions and constraints through linearization. The most powerful aspect of LP is the ability to handle equality and inequality constraints. The following topics will be covered in this appendix. Students will see LP used in this book to solve economic dispatch, hydro and fuel scheduling problems, OPF problems, etc.

Students can make use of LP through the excellent package available in the MATLAB Optimization Toolbox as well as through the LP built into Microsoft Excel.

Here is a classic LP problem which we shall use to show how an LP is solved:
Maximize production of 2 plants:
x_1 is the output from plant 1
x_2 is the output from plant 2
Benefit of product plant 1, $c_1 = 3$ [\$/MWh]
Benefit of product plant 2, $c_2 = 5$ [\$/MWh]
The "objective function" expresses the overall benefit to be optimized; our objective function is

$$\text{Maximize}: 3x_1 + 5x_2$$

Plant maximum output constraints:
Maximum production of plant 1, $b_1 = 4$
Maximum production of plant 2, $b_2 = 6$
Plants must meet an additional inequality constraint:

$$3x_1 + 2x_2 \le 18$$

Mathematically this can be stated as follows:

$$\begin{aligned} \text{Maximize}: \quad & 3x_1 + 5x_2 \\ \text{Subject to}: \quad & x_1 \le 4 \\ & x_2 \le 6 \\ & 3x_1 + 2x_2 \le 18 \\ & x_1 \ge 0 \\ & x_2 \ge 0 \end{aligned}$$

A picture of the x_1, x_2 plane shows that we are optimizing over an area bounded by linear constraints. The shaded region is said to be feasible because any point in that region meets all the constraint conditions (Figure 3.25).

www.mepcafe.com

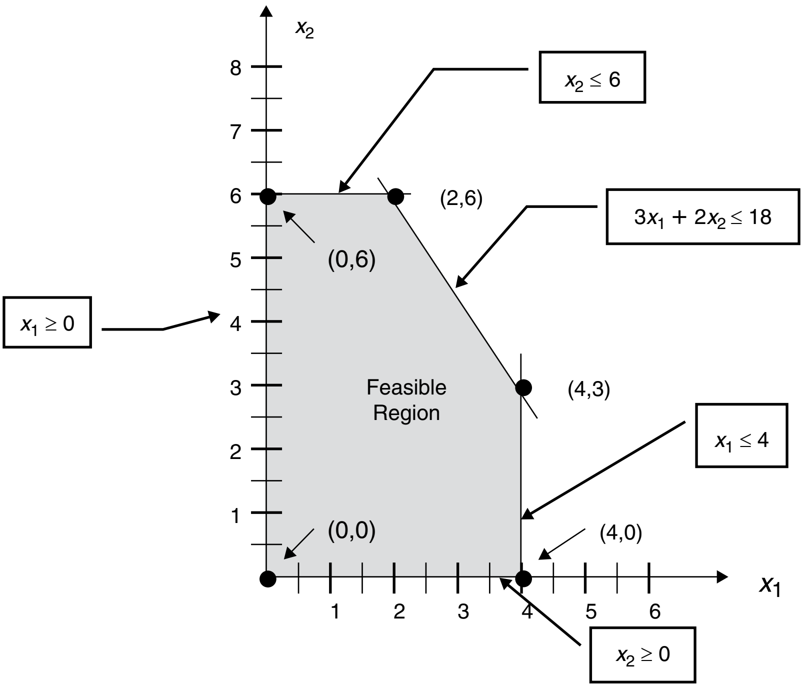
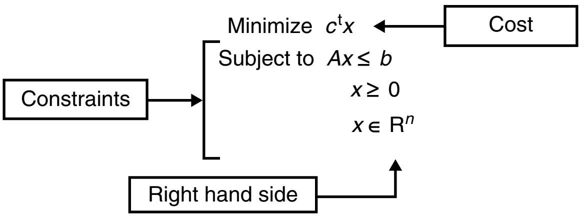


FIGURE 3.25 x_1, x_2 plane showing constraints and the feasible region.

The LP can now be summarized mathematically as



where

x is an unknown $n \times 1$ vector
c is the $n \times 1$ vector of cost coefficients
A is the $m \times n$ matrix of constraint coefficients
b is the right hand side $m \times 1$ vector

There are n variables in the x vector.

There are m constraint equations in the A matrix.

Now we can include the cost function in our x_1, x_2 plane drawing. In this figure the dotted lines represent the linear function:

www.mepcafe.com

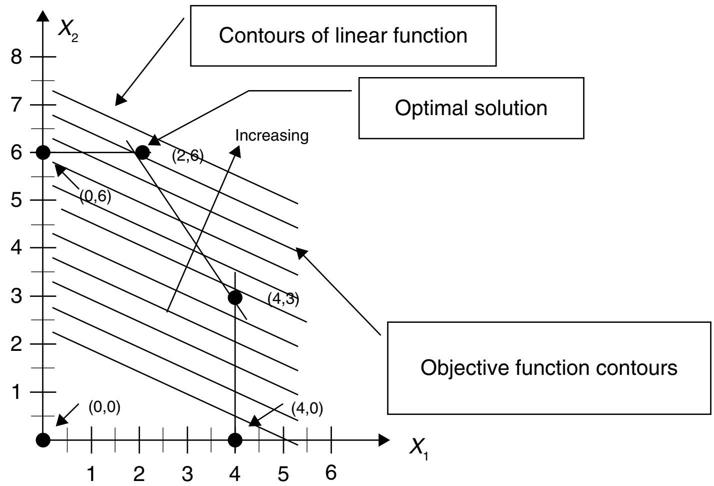


FIGURE 3.26 x_1,x_2 plane with cost contours and the optimal solution shown.

$$z = 3x_1 + 5x_2$$

where z is the total benefit

The individual lines are for different values of z. The line that passes through (4,0) would be for $z = 12$, the line through (4,3) would be for $z = 27$, and the line passing through (2,6) would be for $z = 36$ which is the maximum value of the cost function within the feasible space. The optimum in an LP will always be at the edge of the feasible space, either at the intersection of two or more constraints or along a constraint if it has the same coefficients as the cost function (more on this later). Note that the cost contours (i.e., the value of z) increase as x_1 and x_2 increase in value. Here is the x_1,x_2 plane with the cost contours (Figure 3.26):

We now manipulate the equations to place the problem into what we call "standard form." The standard form makes possible an easy algebraic manipulation of the equations.

Step 1: The standard form starts with noting that LP algorithms are built to minimize, not maximize, an objective function. This can be done by simply minimizing the negative of the objective function. The mathematical form now becomes

$$\begin{aligned}
\text{Minimize:}\quad & -3x_1 - 5x_2 \\
\text{Subject to:}\quad & x_1 \le 4 \\
& x_2 \le 6 \\
& 3x_1 + 2x_2 \le 18 \\
& x_1 \ge 0 \\
& x_2 \ge 0
\end{aligned}$$

www.mepcafe.com

Step 2: We convert all "greater than equal to" constraints (see a that follows) to equality constraints (see b that follows) by subtracting new variables called "slack variables":

$$\text{(a)} \quad \sum_j a_{ij} x_j \ge b_i$$
$$\text{(b)} \quad \sum_j a_{ij} x_j - y_i = b_i ; y_i \ge 0$$

Here the new variables y_i are the slack variables. As long as all y_i are positive, the original greater than equal to constraints are still in force.

Similarly, we convert all "less than equal to" constraints to equality constraints by adding slack variables y_i, and again if all y_i are positive, the less than equal to constraints are still in force:

$$\text{(a)} \quad \sum_j a_{ij} x_j \le b_i$$
$$\text{(b)} \quad \sum_j a_{ij} x_j + y_i = b_i ; y_i \ge 0$$

Step 3: If some or all the variables x are to have infinite positive and negative values allowed, then we add slack variables like this:

$$\text{(a)} \quad -\infty < x_i < +\infty$$
$$\text{(b)} \quad x_i = y_i - z_i ; y_i \ge 0,\ z_i \ge 0$$

We now rewrite our mathematical form with all the slack variables added:

$$\begin{array}{lcrlrlrlrlrlcr}
\text{Minimize} & z = & -3 & x_1 & -5 & x_2 & & & & & & & & \\
\text{Subject to:} & & & x_1 & & & + & x_3 & & & & & = & 4 \\
 & & & & & x_2 & & & + & x_4 & & & = & 6 \\
 & & 3 & x_1 & +2 & x_2 & & & & & + & x_5 & = & 18 \\
 & & & x_1 & & & & & & & & & \ge & 0 \\
 & & & & & x_2 & & & & & & & \ge & 0 \\
 & & & & & & & x_3 & & & & & \ge & 0 \\
 & & & & & & & & & x_4 & & & \ge & 0 \\
 & & & & & & & & & & & x_5 & \ge & 0
\end{array}$$

where x_3, x_4, and x_5 are the new slack variables added for our problem.

www.mepcafe.com

We now partition the x and c vectors into "basic" and "non-basic" parts. We similarly partition the A matrix as well. The result is

$$A \equiv [B \;\vdots\; N]$$

$$x \equiv \begin{bmatrix} x_B \\ x_N \end{bmatrix}$$

$$c \equiv \begin{bmatrix} c_B \\ c_N \end{bmatrix}$$

where

B is a nonsingular square $n \times n$ submatrix called the basis
N is the matrix containing the remaining columns of A
x_B is a vector of basic variables; basic bariables are in the basis
x_N is a vector of nonbasic variables; nonbasic variables are not in the basis
c_B is a vector of cost coefficients corresponding to the basic variables
c_N is a vector of cost coefficients corresponding to nonbasic variables

Solutions are determined using these equations:

$$\{x : Ax = b \quad \text{Solution}$$

$$\begin{cases} x : Ax = b & \text{Feasible solution} \\ x \geq 0 \end{cases}$$

$$\begin{cases} x : Ax = b & \text{Basic feasible solution} \\ x \geq 0 \\ x_N = 0 \end{cases}$$

Note here that we have not stated yet how to find the optimal solution, only a basic feasible solution. Any vector x that meets the basic feasible solution conditions shown earlier might be an optimal solution, but we will need another part of the algorithm to find our way to that optimum.

In the effort to get to the optimal solution, we note that by partitioning A we can write the steps discussed earlier as

$$\begin{cases} [B\,N]\begin{bmatrix} x_B \\ x_N \end{bmatrix} = b \\ x_N = 0 \end{cases}$$

$$\begin{cases} Bx_B = b \\ x_N = 0 \end{cases}$$

$$\begin{cases} x_B = B^{-1}b \\ x_N = 0 \end{cases}$$

www.mepcafe.com

We will now look at the set of all basic feasible solutions, starting from the A matrix and its associated b vector:

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 3 & 2 & 0 & 0 & 1 \end{bmatrix} \quad b = \begin{bmatrix} 4 \\ 6 \\ 18 \end{bmatrix}$$

Set 1: $$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 3 & 2 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 6 \\ 18 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \\ 2 \end{bmatrix} \Rightarrow z = -36(*) \quad \begin{matrix} x_4 = 0 \\ x_5 = 0 \end{matrix}$$

Set 2: $$\begin{bmatrix} x_1 \\ x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 3 & 2 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 6 \\ 18 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 3 \end{bmatrix} \Rightarrow z = -27 \quad \begin{matrix} x_3 = 0 \\ x_5 = 0 \end{matrix}$$

Set 3: $$\begin{bmatrix} x_1 \\ x_2 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 6 \\ 18 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ -6 \end{bmatrix} \Rightarrow \text{OUT} \quad \begin{matrix} x_3 = 0 \\ x_4 = 0 \end{matrix}$$

fails condition that all x are ≥ 0

Set 4: $$\begin{bmatrix} x_1 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 3 & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 6 \\ 18 \end{bmatrix} = \begin{bmatrix} 6 \\ -2 \\ 6 \end{bmatrix} \Rightarrow \text{OUT} \quad \begin{matrix} x_2 = 0 \\ x_5 = 0 \end{matrix}$$

fails condition that all x are ≥ 0

Set 5: $$\begin{bmatrix} x_1 \\ x_3 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 3 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 6 \\ 18 \end{bmatrix} = \begin{bmatrix} \infty \\ \infty \\ \infty \end{bmatrix} \Rightarrow \text{OUT} \quad \begin{matrix} x_2 = 0 \\ x_4 = 0 \end{matrix}$$

A matrix singular

Set 6: $$\begin{bmatrix} x_1 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 6 \\ 18 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ 6 \end{bmatrix} \Rightarrow z = -12 \quad \begin{matrix} x_2 = 0 \\ x_3 = 0 \end{matrix}$$

www.mepcafe.com

Simplex Method

Starting with our basic LP problem,

$$\begin{array}{llcl} \text{Minimize} & z & = & c^{\mathrm{T}} x \\ \text{Subject to} & & & A x = b \\ & & & x \geq 0 \end{array}$$

We partition the A matrix and the x and c vectors into basic and non-basic parts:

$$\begin{array}{lllllll} \text{Minimize} & z = & c_{\mathrm{B}}^{\mathrm{T}} x_{\mathrm{B}} & + & c_{\mathrm{N}}^{\mathrm{T}} x_{\mathrm{N}} & & \\ \text{Subject to} & & \mathrm{B}\, x_{\mathrm{B}} & + & N x_{\mathrm{N}} & = & b \\ & & x_{\mathrm{B}} & & & \geq & 0 \\ & & & & x_{\mathrm{N}} & \geq & 0 \end{array}$$

Last of all we carry out the following matrix multiplys:

$$x_{\mathrm{B}} = B^{-1}b - B^{-1}Nx_{\mathrm{N}}$$
$$z = c_{\mathrm{B}}^{\mathrm{T}}B^{-1}b - c_{\mathrm{B}}^{\mathrm{T}}B^{-1}Nx_{\mathrm{N}} + c_{\mathrm{N}}^{\mathrm{T}}x_{\mathrm{N}}$$

Now we will observe:

$$\tilde{b} \equiv B^{-1}b$$
$$\lambda^{\mathrm{T}} \equiv c_{\mathrm{B}}^{\mathrm{T}}B^{-1}$$
$$Y \equiv B^{-1}N$$

where λ is a vector called the "dual" variable vector.

This leads to a reformulation of our standard form of the LP:

$$\begin{array}{lllll} \text{Minimize} & z & = & c_{\mathrm{B}}^{\mathrm{T}} \tilde{b} - \left[\lambda^{\mathrm{T}}N - c_{\mathrm{N}}^{\mathrm{T}}\right] x_{\mathrm{N}} & \\ \text{Subject to} & x_{\mathrm{B}} & = & \tilde{b} - Y x_{\mathrm{N}} & \\ & x_{\mathrm{B}} & & & \geq 0 \\ & & & x_{\mathrm{N}} & \geq 0 \end{array}$$

We will now define $d^{\mathrm{T}} \equiv \lambda^{\mathrm{T}}N - c_{\mathrm{N}}^{\mathrm{T}}$ as the reduced cost vector, and this allows us to write the new standard form as

$$\begin{array}{lllll} \text{Minimize} & z & = & c_{\mathrm{B}}^{\mathrm{T}} \tilde{b} - d^{\mathrm{T}} x_{\mathrm{N}} & \\ \text{Subject to} & x_{\mathrm{B}} & = & \tilde{b} - Y x_{\mathrm{N}} & \\ & x_{\mathrm{B}} & & & \geq 0 \\ & & & x_{\mathrm{N}} & \geq 0 \end{array}$$

www.mepcafe.com

Set 7: $\begin{bmatrix} x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 6 \\ 18 \end{bmatrix} = \begin{bmatrix} 9 \\ 4 \\ -3 \end{bmatrix} \Rightarrow \text{NO} \quad \begin{matrix} x_1 = 0 \\ x_5 = 0 \end{matrix}$

fails condition that all x are ≥ 0

Set 8: $\begin{bmatrix} x_2 \\ x_3 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 2 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 6 \\ 18 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ 6 \end{bmatrix} \Rightarrow z = -30 \quad \begin{matrix} x_1 = 0 \\ x_4 = 0 \end{matrix}$

Set 9: $\begin{bmatrix} x_2 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 6 \\ 18 \end{bmatrix} = \begin{bmatrix} \infty \\ \infty \\ \infty \end{bmatrix} \Rightarrow \text{OUT} \quad \begin{matrix} x_1 = 0 \\ x_3 = 0 \end{matrix}$ A matrix singular

Set 10: $\begin{bmatrix} x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 6 \\ 18 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ 18 \end{bmatrix} \Rightarrow z = 0 \quad \begin{matrix} x_1 = 0 \\ x_2 = 0 \end{matrix}$

We could simply use this as a means to find the optimum:

- If LP has an optimal solution, then this solution is a basic feasible solution:
- Examine all basic feasible solutions
- Evaluate the objective function for every basic feasible solution
- Choose the basic feasible solution with the lowest objective function value.

However, the upper bound on the number of basic feasible solutions is $\binom{n}{m} = \frac{n!}{m!(n-m)!}$, so just evaluating all basic feasible solutions is not a good means to the optimum.

We shall instead use a fundamental theorem:

If LP problem has an optimal solution,
then
That solution is a basic feasible solution,

which leads to the simplex method:

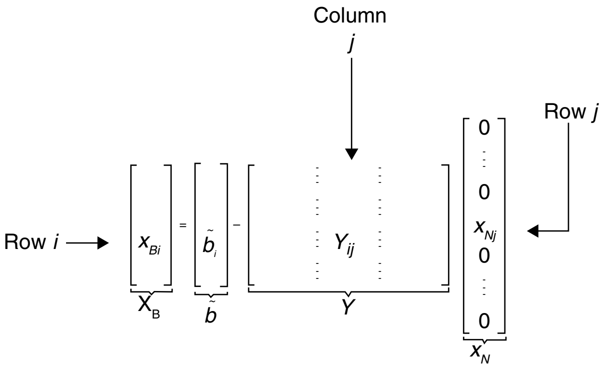
We move from corner to corner by always reducing the value of the objective function until no further reduction is possible, at that point we are at the minimum.

www.mepcafe.com

In this form, the objective function z is now shown to be a function of the nonbasic variables, and the basic variables x_B are also a function of the nonbasic variables.

Pivoting

Pivoting makes use of one element of the Y matrix as shown in the following:



The simplex mechanism can now be stated using these steps and noting the relationship of the variables in the equation shown previously.

Simplex Mechanism

- If d_j is positive, the objective function decreases if the nonbasic variable x_{Nj} increases.
- If the nonbasic variable x_{Nj} increases, the basic variable x_{Bi} decreases if Y_{ij} is positive.
- If some Y_{ij}'s are positive, the corresponding basic variables x_{Bi}'s decrease if the nonbasic variable x_{Nj} increases.
- The nonbasic variable x_{Nj} can be increased until the first basic variable becomes 0, that is, until $\tilde{b}_i - Y_{ij}x_{Nj}$ becomes 0 for the first i.
- The value of x_{Nj} becomes $\underset{1 \le i \le m}{\text{Minimum}} \left\{ \dfrac{\tilde{b}_i}{Y_{ij}} : Y_{ij} > 0 \right\}$.
- The value of x_{Bi} becomes 0.

Simplex Algorithm Pseudo Code

1. Get a basic feasible solution (see succeeding text).
2. Find out if the current solution is the minimizer: if all d_j<0 stop, the current solution is the minimizer; otherwise, go on.

www.mepcafe.com

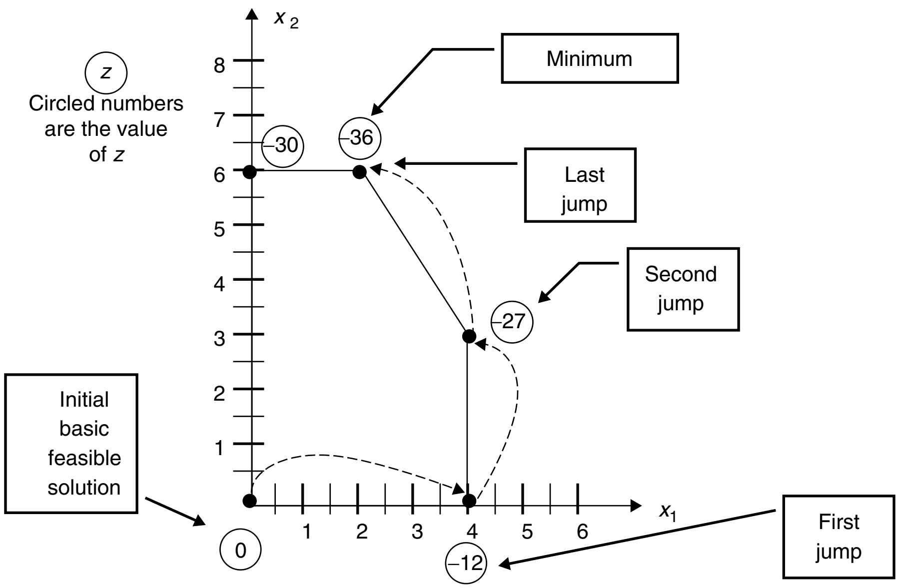


FIGURE 3.27 Simplex algorithms steps from initial basic feasible solution to minimum z.

3. Find out which nonbasic variable enters the basis, x_{Nj} (most positive d_j).
4. Find out which basic variable x_{Bi} leaves the basis, $\underset{1 \le i \le m}{\text{Minimum}} \left\{ \frac{\tilde{b}_i}{Y_{ij}} : Y_{ij} > 0 \right\}$.
5. Build the new basis.
6. Get a new basic feasible solution.
7. Go to 2

For our original simple problem these steps look like this (Figure 3.27):
Simplex algorithms steps from initial basic feasible solution to minimum z
Initial Basic Feasible Solution

$$
\begin{array}{llll}
\text{Minimize} & c^{T}x & & \\
\text{Subject to} & Ax & = & b \\
 & x & \ge & 0 \\
 & x & \in & \mathfrak{R}^{n}
\end{array}
$$

The initial basic feasible solution can be obtained by solving the LP problem which follows:

www.mepcafe.com

Minimize $\sum_{i=1}^{m} y_i$

Subject to

$$\begin{aligned} A\,x + y &= b \\ x &\geq 0 \\ y &\geq 0 \\ x &\in \Re^n \\ y &\in \Re^m \end{aligned}$$

This is called the "phase 1" method

where $y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$ is a column of "artificial variables."

It should be noted that this second problem has a trivial basic feasible solution, $y = b$.

Sensitivity and LP Calculation of Lagrange Variables

If B is the optimal basis

$$x_B^* = B^{-1}b \tag{3.1}$$

$$z = c_B^T x_B^* \tag{3.2}$$

then we can calculate the vector λ as

$$\lambda^{*T} = c_B^T B^{-1} \tag{3.3}$$

Let Δb be an increment in b so that the basis does not change $b \rightarrow b + \Delta b$.

Then $\begin{matrix} x_B^* & \rightarrow & x_B^* & + & \Delta x_B \\ z^* & \rightarrow & z^* & + & \Delta z \end{matrix}$,and from (1) and (2) mentioned earlier

$\Delta x_B = B^{-1}\Delta b$ and $\Delta z = c_B^T \Delta x_B$. Using these last two equations and (3), we get $\Delta z = c_B^T \Delta x_B = c_B^T B^{-1} \Delta b = \lambda^{*T} \Delta b$ and finally $\Delta z = \lambda^{*T} \Delta b$. This can be written as

$$\lambda_j^* = \frac{\Delta z}{\Delta b_j} \quad \forall j = 1, 2, \ldots, m$$

λ_j^* is the marginal change in the objective function as a result of a marginal change in the right-hand side constraint j limit (as long as the basis remains

www.mepcafe.com

unchanged). Note that the Lagrange multiplier λ which we developed in Appendix A has this same definition. In pure economic dispatch λ is the ratio of change in total cost to a change in the load in the generation = load balance constraint.

Some Useful Features of Available LP Codes

Upper and lower bounding: A feature of many LP algorithms available today (e.g., the LP in the MATLAB Optimization Toolbox) is to directly give vectors of upper and lower bounds on each LP variable instead of explicitly building constraints for each limit. Most of the LP problems we have encountered in power system operation have upper and lower bounds on variables.

Sensitivity analysis: The MATLAB LP allows students to see the lambda values for all constraints including all variable upper bound constraints, all variable lower bound constraints, all linear equality constraints, and all linear inequality constraints.

Initial value of the variables: For some problems this may make a difference in run time; if not specified, the algorithm will find an initial basis by itself and then proceed to the optimum.

APPENDIX 3C
Non-Linear Programming

Nonlinear programming refers to optimization problems that have either nonlinear objective functions or nonlinear constraints or both. An example of a method that solves a nonlinear objective function with linear constraints is quadratic programming. An example of a nonlinear programming method that can solve both nonlinear objective functions and nonlinear constraints is the "interior point" method. Both quadratic programming and the interior point method are dealt with in Chapter 8.

APPENDIX 3D
Dynamic Programming (DP)

The application of digital methods to solve a wide variety of control and dynamics optimization problems in the late 1950s led Dr. Richard Bellman and his associates to the development of DP. These techniques are useful in solving a variety of problems and can greatly reduce the computational effort in finding optimal trajectories or control policies.

The theoretical mathematical background, based on the calculus of variations, is somewhat difficult. The applications are not, however, since they depend on a

www.mepcafe.com

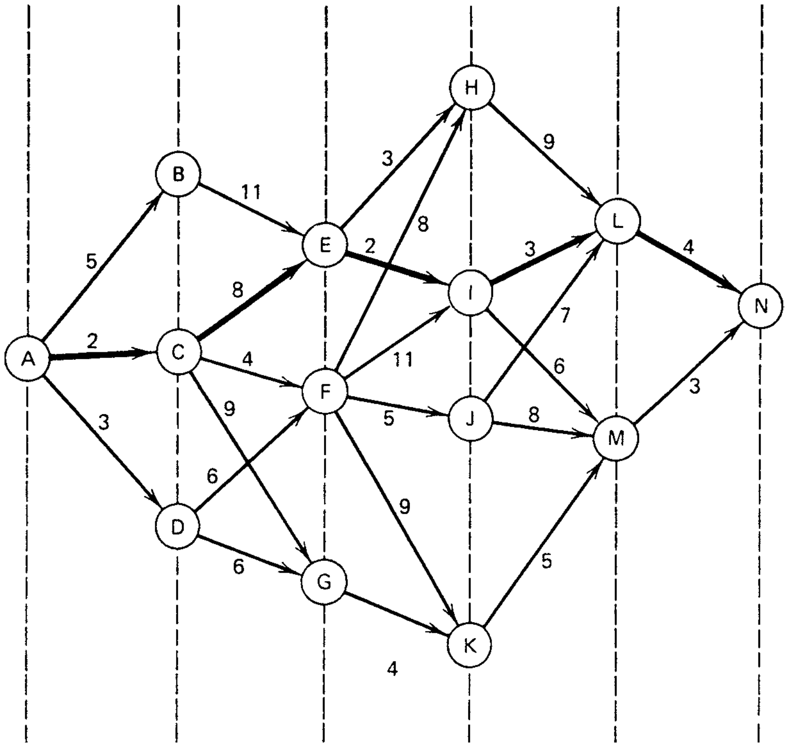


FIGURE 3.28 DP example.

willingness to express the particular optimization problem in terms appropriate for a DP formulation.

In the scheduling of power generation systems, DP techniques have been developed for the following:

- The economic dispatch of thermal systems
- The solution of hydrothermal economic-scheduling problems
- The practical solution of the unit commitment problem

This chapter will touch on all three areas.

First, however, it will be as well to introduce some of the notions of DP by means of some one-dimensional examples. Figure 3.28 represents the cost of transporting a unit shipment from node *A* to node *N*. The values on the arcs are the costs, or values, of shipping the unit from the originating node to the terminating node of the arc. The problem is to find the minimum cost route from *A* to *N*. The method to be illustrated is that of DP. The first two examples are from reference 8 and are used by permission.

www.mepcafe.com

Starting at A, the minimum cost path to N is ACEILN.
Starting at C, the least cost path to N is CEILN.
Starting at E, the least cost path to N is EILN
Starting at I, the least cost path to N is ILN.
Starting at L, the least cost path to N is LN.

The same type of statements could be made about the maximum cost path from A to N (ABEHLN). That is, the maximum cost to N, starting from any node on the original maximal path, is contained in that original path.

The choice of route is made in sequence. There are various stages traversed. The optimum sequence is called the *optimal policy*; any subsequence is a *subpolicy*. From this it may be seen that the optimal policy (i.e., the minimum cost route) contains only optimal subpolicies. This is the *theorem of optimality*.

An optimal policy must contain only optimal subpolicies.

Bellman and Dreyfus (reference 7) call it the "principle of optimality" and state it as

A policy is optimal if, at a stated stage, whatever the preceding decisions may have been, the decisions still to be taken constitute an optimal policy when the result of the previous decisions is included.

We continue with the same example; only now let us find the minimum cost path. Figure 3.29 identifies the stages (I, II, III, IV, V). At the terminus of each stage, there is a set of choices of nodes $\{X_i\}$ to be chosen $[\{X_3\} = \{\text{H, I, J, K}\}]$. The symbol $V_a(X_i, X_i + 1)$ represents the "cost" of traversing stage a(=I, …, V) and depends on the variables selected from the sets $\{X_i\}$ and $\{X_i + 1\}$. That is, the cost, V_a, depends on the starting and terminating nodes. Finally, $f_a(X_i)$ is the minimum cost for stages I through a to arrive at some particular node X_i at the end of that stage, starting from A. The numbers in the node circles in Figure 3.29 represent this minimum cost.

$$\begin{array}{lll} \{X_0\}: \text{A} & \{X_2\}: \text{E, F, G} & \{X_4\}: \text{L, M} \\ \{X_1\}: \text{B, C, D} & \{X_3\}: \text{H, I, J, K} & \{X_5\}: \text{N} \end{array}$$

$f_{\text{I}}(X_1)$: Minimum cost for the first stage is obvious:

$$f_{\text{I}}(\text{B}) = V_1(\text{A, B}) = 5$$
$$f_{\text{I}}(\text{C}) = V_1(\text{A, C}) = 2$$
$$f_{\text{I}}(\text{D}) = V_1(\text{A, D}) = 3$$

www.mepcafe.com

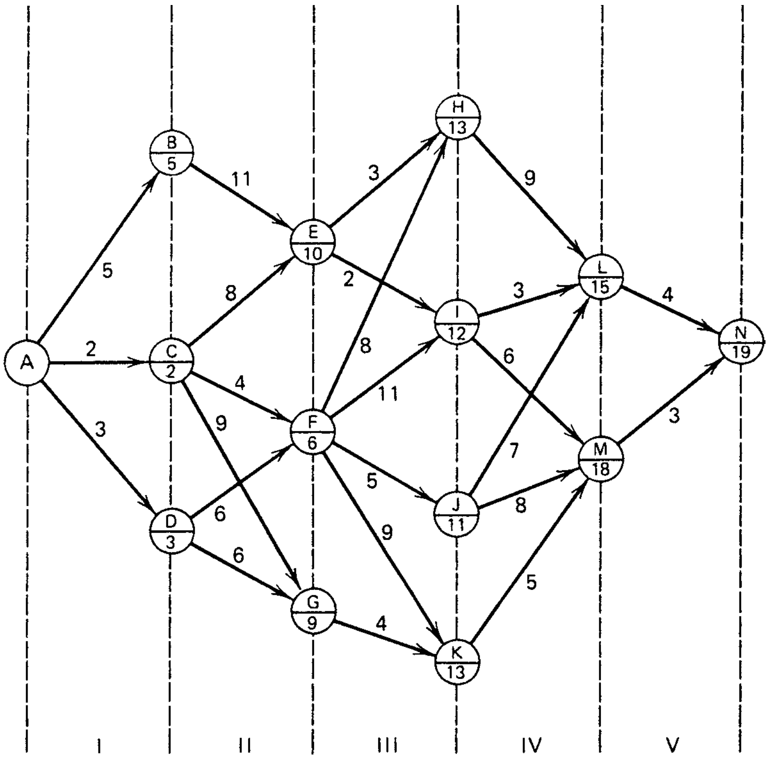


FIGURE 3.29 DP example showing minimum cost at each node.

$f_{\mathrm{II}}(X_2)$: Minimum cost for stages I and II as a function of X_2:

$$f_{\mathrm{II}}(\mathrm{E}) = \min_{\{X_1\}}[f_{\mathrm{I}}(X_1) + V_{\mathrm{II}}(X_1, \mathrm{E})]$$

$$= \min[5+11, \quad 2+8, \quad 3+\infty] = 10$$

$$X_1 = D \qquad = C \qquad = D \qquad X_1 = C$$

The cost is infinite for node D since there is no path from D to E:

$$f_{\mathrm{II}}(f) = \min_{\{X_1\}}[f_{\mathrm{I}}(X_1) + V_{\mathrm{II}}(X_1, f)] = \min[\infty, 6, 9] \ = 6, X_1 = \mathrm{C}$$

$$f_{\mathrm{II}}(\mathrm{G}) = \min_{\{X_1\}}[f_{\mathrm{I}}(X_1) + V_{\mathrm{II}}(X_1, \mathrm{G})] = \min[\infty, 11, 9] = 9, X_1 = \mathrm{D}$$

Thus, at each stage we should record the minimum cost and the termination starting the stage in order to achieve the minimum cost path for each of the nodes terminating the current stage.

www.mepcafe.com

$(\mathbf{X_2})$	E	F	G
$f_{II}(X_2)$	10	6	9
Path X_0X_1	AC	AC	AD

$f_{III}(X_3)$: Minimum cost of stages I, II, and III as a function of X_3:

$$f_{III}(\mathrm{H}) = \min_{\{X_2\}}\left[f_{II}(X_2) + V_{III}(X_2,\mathrm{H})\right] = \min\left[13,14,\infty\right] = 13 \quad \text{with} \quad X_2 = \mathrm{E}$$

In general,

$$f_{III}(X_3) = \min_{\{X_2\}}\left[f_{II}(X_2) + V_{III}(X_2, X_3)\right]$$

giving

$\mathbf{X_3}$	H	I	J	K
$f_{III}(X_3)$	13	12	11	13
Path $X_0X_1X_2$	ACE	ACE	ACF	ADG

f_{IV}: Minimum cost of stages I through IV as a function of X_4:

$$f_{IV}(X_4) = \min_{\{X_3\}}[f_{III}(X_3) + V_{IV}(X_3, V_4)]$$
$$f_{IV}(\mathrm{L}) = \min[13+9, 12+3, 11+7, 13+\infty] = 15, \quad X_3 = \mathrm{I}$$
$$X_3 = \mathrm{H} = \mathrm{I} = \mathrm{J} = \mathrm{K}$$
$$f_{IV}(\mathrm{M}) = [13+\infty, 12+6, 11+8, 13+5] = 18 \quad X_3 = \mathrm{I} \quad \text{or} \quad \mathrm{K}$$
$$X_3 = \mathrm{H} = \mathrm{I} = \mathrm{J} = \mathrm{K}$$

f_V: Minimum cost of I through V as a function of X_5:

$$f_V(N) = \min_{\{X_4\}}[f_{IV}(X_4) + V_V(X_4, X_5)]$$
$$= \min\left[15+4, 18+3\right] = 19 \quad X_4 = \mathrm{L}$$
$$X_4 = \mathrm{L} = \mathrm{M}$$

Tracing back, the path of minimum cost is found as follows:

Stage 1	$\{\mathbf{X_i}\}$	f_i
1	B, Ⓒ, D	5, ②, 3
2	Ⓔ, F, G	⑩, 6, 9
3	H, Ⓘ, J, K	13, ⑫, 11, 13
4	Ⓛ, M	⑮, 18
5	Ⓝ	⑲

www.mepcafe.com

TABLE 3.4 Profit Versus Investment

Investment Amount	Profit from Venture I	II	III	IV
0	0	0	0	0
1	0.28	0.25	0.15	0.20
2	0.45	0.41	0.25	0.33
3	0.65	0.55	0.40	0.42
4	0.78	0.65	0.50	0.48
5	0.90	0.75	0.65	0.53
6	1.02	0.80	0.73	0.56
7	1.13	0.85	0.82	0.58
8	1.23	0.88	0.90	0.60
9	1.32	0.90	0.96	0.60
10	1.38	0.90	1.00	0.60

It would be possible to carry out this procedure in the opposite direction just as easily.

An Allocation Problem. Table 3.4 lists the profits to be made in each of four ventures as a function of the investment in the particular venture. Given a limited amount of money to allocate, the problem is to find the optimal investment allocation. The only restriction is that investments must be made in integer amounts. For instance, if one had 10 units to invest and the policy were to put 3 in I, 1 in II, 5 in III, and 1 in IV, then

$$\text{Profit} = 0.65 + 0.25 + 0.65 + 0.20 = 1.75$$

The problem is to find an allocation policy that yields the maximum profit. Let

$$\begin{aligned}&X_1, X_2, X_3, X_4 \text{ be investments in I through IV}\\&V(X_1), V(X_2), V(X_3), V(X_4) \text{ be profits}\\&X_1 + X_2 + X_3 + X_4 = 10 \text{ is the constraint; that is,}\\&10 \text{ units must be invested}\end{aligned}$$

To transform this into a multistage problem, let the stages be

$$X_1, U_1, U_2, A$$

where

$$\begin{aligned}U_1 &= X_1 + X_2 \quad U_1 \le A \quad U_2 \le A\\U_2 &= U_1 + X_3 \quad \{A\} = 0, 1, 2, 3, \ldots, 10\\A &= U_2 + X_4\end{aligned}$$

www.mepcafe.com

The total profit is

$$f(X_1,X_2,X_3,X_4) = V_1(X_1) + V_2(X_2) + V_3(X_3) + V_4(X_4)$$

which can be written

$$f(X_1,U_1,U_2,A) = V_1(X_1) + V_2(U_1 - X_1) + V_3(U_2 - U_1) + V_4(A - U_2)$$

At the second stage, we can compute

$$f_2(U_1) = \max_{X_1=0,1,\ldots,U_1}\left[V_1(X_1) + V_2(U_1 - X_1)\right]$$

X_1, X_2, or U_1	$V_1(X_1)$	$V_2(X_2)$	$f_2(U_1)$	Optimal Subpolicies for I and II
0	0	0	0	0, 0
1	0.28	0.25	0.28	1, 0
2	0.45	0.41	0.53	1, 1
3	0.65	0.55	0.70	2, 1
4	0.78	0.65	0.90	3, 1
5	0.90	0.75	1.06	3, 2
6	1.02	0.80	1.20	3, 3
7	1.13	0.85	1.33	4, 3
8	1.23	0.88	1.45	5, 3
9	1.32	0.90	1.57	6, 3
10	1.38	0.90	1.68	7, 3

Next, at the third stage,

$$f_3(U_2) = \max_{U_1=0,1,2,\ldots,U_2}\left[f_2(U_1) + V_3(U_2 - U_1)\right]$$

				Optimal Subpolicies	
U_1, U_2, or X_3	$f_2(U_1)$	$V_3(X_3)$	$f_3(U_2)$	For I and II	For I, II, and III
0	0	0	0	0, 0	0, 0, 0
1	0.28	0.15	0.28	1, 0	1, 0, 0
2	0.53	0.25	0.53	1, 1	1, 1, 0
3	0.70	0.40	0.70	2, 1	2, 1, 0
4	0.90	0.50	0.90	3, 1	3, 1, 0
5	1.06	0.62	1.06	3, 2	3, 2, 0
6	1.20	0.73	1.21	3, 3	3, 2, 1
7	1.33	0.82	1.35	4, 3	3, 3, 1
8	1.45	0.90	1.48	5, 3	4, 3, 1
9	1.57	0.96	1.60	6, 3	5, 3, 1 or 3, 3, 3
10	1.68	1.00	1.73	7, 3	4, 3, 3

Finally, the last stage is

$$f_4(A) = \max_{\{U_2\}}\left[f_3(U_2) + V_4(A - U_2)\right]$$

www.mepcafe.com

U_2, A, or X_4	$f_3(U_2)$	$V_4(X_4)$	$f_4(A)$	Optimal Subpolicy for I, II, and III	Optimal Policy
0	0	0	0	0, 0, 0	0, 0, 0, 0
1	0.28	0.20	0.28	1, 0, 0	1, 0, 0, 0
2	0.53	0.33	0.53	1, 1, 0	1, 1, 0, 0
3	0.70	0.42	0.73	2, 1, 0	1, 1, 0, 1
4	0.90	0.48	0.90	3, 1, 0	3, 1, 0, 0 or 2, 1, 0, 1
5	1.06	0.53	1.10	3, 2, 0	3, 1, 0, 1
6	1.21	0.56	1.26	3, 2, 1	3, 2, 0, 1
7	1.35	0.58	1.41	3, 3, 1	3, 2, 1, 1
8	1.48	0.60	1.55	4, 3, 1	3, 3, 1, 1
9	1.60	0.60	1.68	5, 3, 1 or 3, 3, 3	4, 3, 1, 1 or 3, 3, 1, 2
10	1.73	0.60	1.81	4, 3, 3	4, 3, 1, 2

Consider the procedure and solutions:

a. It was not necessary to enumerate all possible solutions. Instead, we used an orderly, stage-wise search, the form of which was the same at each stage.
b. The solution was obtained not only for $A = 10$, but for the complete set of A values $\{A\} = 0, 1, 2, \ldots, 10$.
c. The optimal policy contains only optimal subpolicies. For instance, $A = 10, (4,3,1,2)$ is the optimal policy. For stages I, II, and III, $U_2 = 8, (4,3,1)$ is the optimal subpolicy. For stages I and II, $U_1 = 7, (4,3)$ is the optimal subpolicy. For stage I only, $X_1 = 4$ fixes the policy.
d. Notice also that by storing the intermediate results, we could work a number of different variations of the same problem with the data already computed.

APPENDIX 3E
Convex Optimization

Another way to solve an optimization problem is to use a technique that solves for the Lagrange variables directly and then solves for the problem variables themselves. This formulation is known as a "dual solution," and in it the Lagrange multipliers are called "dual variables." We shall use the example solved previously to demonstrate this technique.

The presentation up to now has been concerned with the solution of what is formally called the "primal problem," which was stated in Equation 3A.2 as

$$\textbf{Minimize:} \quad f(x_1, x_2) = 0.25x_1^2 + x_2^2$$
$$\textbf{Subject to:} \quad \omega(x_1, x_2) = 5 - x_1 - x_2$$

and its Lagrangian function is

$$\mathcal{L}(x_1, x_2, \lambda) = 0.25x_1^2 + x_2^2 + \lambda(5 - x_1 - x_2)$$

www.mepcafe.com

If we define a dual function, $q(\lambda)$, as

$$q(\lambda) = \min_{x_1 x_2} \mathcal{L}(x_1, x_2, \lambda) \tag{3E.12}$$

then the "dual problem" is to find

$$q^*(\lambda) = \max_{\lambda \geq 0} q(\lambda) \tag{3E.13}$$

The solution in the case of the dual problem involves two separate optimization problems. The first requires us to take an initial set of values for x_1 and x_2 and then find the value of λ that maximizes $q(\lambda)$. We then take this value of λ and, holding it constant, we find values of x_1 and x_2 that minimize $\mathcal{L}(x_1, x_2, \lambda)$. This process is repeated or iterated until the solution is found.

In the case of convex objective functions, such as the example used in this appendix, this procedure is guaranteed to solve to the same optimum as the primal problem solution presented earlier.

The reader will note that in the case of the functions presented in Equation 3A.2, we can simplify the procedure discussed earlier by eliminating x_1 and x_2 from the problem altogether, in which case we can find the maximum of $q(\lambda)$ directly. If we express the problem variables in terms of the Lagrange multiplier (or dual variable), we obtain

$$x_1 = 2\lambda$$
$$x_2 = \frac{\lambda}{2}$$

We now eliminate the original problem variables from the Lagrangian function:

$$q(\lambda) = -\left(\frac{5}{4}\right)\lambda^2 + 5\lambda$$

We can use the dual variable to solve our problem as follows:

$$\frac{\partial}{\partial \lambda} q(\lambda) = 0 = \left(\frac{5}{2}\right)\lambda - 5$$

or

$$\lambda = 2$$

Therefore, the value of the dual variable is $q^*(\lambda)=5$. The values of the primal variables are $x_1=4$ and $x_2=1$.

In the economic dispatch problem dealt with in this chapter, one cannot eliminate the problem variables since the generating unit cost functions may be piecewise linear or other complex functions. In this case, we must use the dual optimization algorithm described earlier: we first optimize on λ and then on the problem variables, then go back and update λ, etc. Since the dual problem requires that we find

$$q^*(\lambda) = \max_{\lambda \geq 0} q(\lambda)$$

www.mepcafe.com

and we do not have an explicit function in λ (as we did in the previous text), we must adopt a slightly different strategy. In the case of economic dispatch or other problems where we cannot eliminate the problem variables, we find a way to adjust λ so as to move $q(\lambda)$ from its initial value to one which is larger. The simplest way to do this is to use a gradient adjustment so that

$$\lambda^1 = \lambda^0 + \left[\frac{\mathrm{d}}{\mathrm{d}\lambda} q(\lambda)\right]\alpha$$

where α merely causes the gradient to behave well. A more useful way to apply the gradient technique is to let λ be adjusted upwards at one rate and downward at a much slower rate; for example,

$$\alpha = 0.5 \quad \text{when} \quad \frac{\mathrm{d}}{\mathrm{d}\lambda} q(\lambda) \text{ is positive}$$

and

$$\alpha = 0.1 \quad \text{when} \quad \frac{\mathrm{d}}{\mathrm{d}\lambda} q(\lambda) \text{ is negative}$$

The closeness to the final solution in the dual optimization method is measured by noting the relative size of the "gap" between the primal function and the dual function. The primal optimization problem can be solved directly in the case of the problem stated in Equation 3A.2, and the optimal value will be called J^* and it is defined as

$$J^* = \min \mathcal{L} \tag{3E.14}$$

This value will be compared to the optimum value of the dual function, q^*. The difference between them is called the "duality gap." A good measure of the closeness to the optimal solution is the "relative duality gap," defined as

$$\frac{J^* - q^*}{q^*} \tag{3E.15}$$

For a convex problem with continuous variables, the duality gap will become 0 at the final solution. When we again take up the dual optimization method in Chapter 4, we will be dealing with nonconvex problems with noncontinuous variables and the duality gap will never actually go to 0.

Using the dual optimization approach on the problem given in Equation 3A.2 and starting at $\lambda=0$, we obtain the results shown in Table 3.5. As can be seen, this procedure converges to the correct answer.

A special note about lambda search. The reader should note that the dual technique, when applied to economic dispatch, is the same as the lambda search technique we introduced earlier in this chapter to solve the economic dispatch problem.

www.mepcafe.com

TABLE 3.5 Dual Optimization

Iteration	λ	x_1	x_2	ω	J^*	q^*	$\frac{J^*-q^*}{q^*}$
1	0	0	0	5.0	5.0	0	—
2	2.5	5.0	1.25	−1.25	5.0	4.6875	0.0666
3	2.375	4.75	1.1875	−0.9375	5.0	4.8242	0.0364
4	2.2813	4.5625	1.1406	−0.7031	5.0	4.9011	0.0202
5	2.2109	4.4219	1.1055	−0.5273	5.0	4.9444	0.01124
⋮							
20	2.0028	4.0056	1.0014	−0.007	5.0	5.0	0

PROBLEMS

3.1 Assume that the fuel inputs in MBtu per hour for units 1 and 2, which are both on-line, are given by

$$H_1 = 8P_1 + 0.024P_1^2 + 80$$
$$H_2 = 6P_2 + 0.04P_2^2 + 120$$

where

H_n = fuel input to unit n in MBtu per hour (millions of Btu per hour)
P_n = unit output in megawatts

a. Plot the input–output characteristics for each unit expressing input in MBtu per hour and output in megawatts. Assume that the minimum loading of each unit is 20 MW and that the maximum loading is 100 MW.

b. Calculate the net heat rate in Btu per kilowatt-hour, and plot against output in megawatts.

c. Assume that the cost of fuel is 1.5 \$/MBtu. Calculate the incremental production cost in \$/MWh of each unit, and plot against output in megawatts.

3.2 Dispatch with Three-Segment Piecewise Linear Incremental Heat Rate Function

Given: two generating units with incremental heat rate (IHR) curves specified as three connected line segments (four points as shown in Figure 3.30).

Unit 1:

Point	MW	IHR (Btu/kWh)
1	100	7,000
2	200	8,200
3	300	8,900
4	400	11,000

Fuel cost for unit 1 = 1.60\$/MBtu

www.mepcafe.com

Unit 2:

Point	MW	IHR (Btu/kWh)
1	150	7500
2	275	7700
3	390	8100
4	450	8500
Fuel cost for unit 2=2.10$/MBtu		

Both units are running. Calculate the optimum schedule (i.e., the unit megawatt output for each unit) for various total megawatt values to be supplied by the units. Find the schedule for these total megawatt values:

300 MW, 500 MW, 700 MW, 840 MW

Notes: Piecewise linear increment cost curves are quite common in digital computer executions of economic dispatch. The problem is best solved by using a "search" technique. In such a technique, the incremental cost is given a value and the units are scheduled to meet this incremental cost. The megawatt outputs for the units are added together and compared to the desired total. Depending on the difference, and whether the resulting total is above or below the desired total, a new value of incremental cost is "tried." This is repeated until the incremental cost is found that gives the correct desired value. The trick is to search in an efficient manner so that the number of iterations is minimized.

3.3 Assume the system load served by the two units of Problem 3.1 varies from 50 to 200 MW. For the data of Problem 3.1, plot the outputs of units 1 and 2 as a function of total system load when scheduling generation by equal incremental production costs. Assume that both units are operating.

3.4 As an exercise, obtain the optimum loading of the two generating units in Problem 3.1 using the following technique. The two units are to deliver 100 MW. Assume both units are on-line and delivering power. Plot the total fuel cost for 100 MW of delivered power as generation is shifted from one unit to

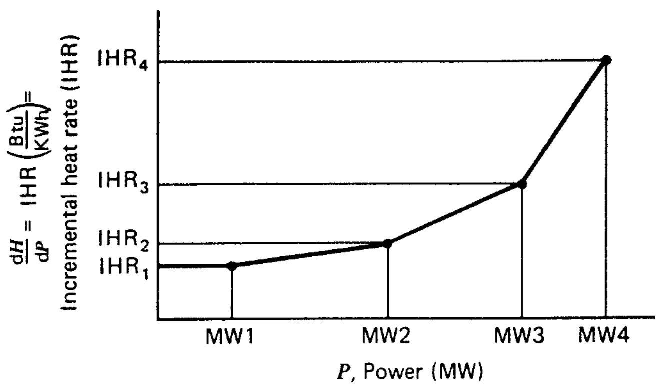


FIGURE 3.30 Piecewise linear IHR curve for Problem 3.2.

www.mepcafe.com

the other. Find the minimum cost. The optimum schedule should check with the schedule obtained by equal incremental production costs.

3.5 This problem demonstrates the complexity involved when we must commit (turn on) generating units, as well as dispatch them economically. This problem is known as the *unit commitment problem* and is the subject of Chapter 4.

Given the two generating units in Problem 3.1, assume that they are both off-line at the start. Also, assume that load starts at 50 MW and increases to 200 MW. The most economic schedule to supply this varying load will require committing one unit first, followed by commitment of the second unit when the load reaches a higher level.

Determine which unit to commit first and at what load the remaining unit should be committed. Assume no "start-up" costs for either unit.

3.6 The system to be studied consists of two units as described in Problem 3.1. Assume a daily load cycle as follows.

Time Band	Load (MW)
0000–0600	50
0600–1800	150
1800–0000	50

Also, assume that a cost of 180$ is incurred in taking either unit off-line and returning it to service after 12 h. Consider the 24-h period from 0600 one morning to 0600 the next morning.

a. Would it be more economical to keep both units in service for this 24-h period or to remove one of the units from service for the 12-h period from 1800 one evening to 0600 the next morning?

b. What is the economic schedule for the period of time from 0600 to 1800 (load = 150 MW)?

c. What is the economic schedule for the period of time from 1800 to 0600 (load = 50 MW)?

3.7 **Economic Dispatch Using Different Methods of Solution**

Assume that all three of the thermal units described next are running. Find the economic dispatch schedules as requested in each part. Use the method and starting conditions given.

Unit Data	Minimum (MW)	Maximum (MW)	Fuel Cost ($/MBtu)
$H_1 = 225 + 8.4P_1 + 0.0025P_1^2$	45	350	0.80
$H_2 = 729 + 6.3P_2 + 0.0081P_2^2$	45	350	1.02
$H_3 = 400 + 7.5P_3 + 0.0025P_3^2$	47.5	450	0.90

www.mepcafe.com

a. Use the lambda iteration method to find the economic dispatch for a total demand of 450 MW.

b. Use the base-point and participation factor method to find the economic schedule for a demand of 495 MW. Start from the solution to part a.

c. Use a gradient method to find the economic schedule for a total demand of 500 MW, assuming the initial conditions (i.e., loadings) on the three units are

$$P_1 = P_3 = 100 \text{ MW} \quad \text{and} \quad P_2 = 300 \text{ MW}$$

Give the individual unit loadings and cost per hour, as well as the total cost per hour to supply each load level: (MBtu = millions of Btu; H_j = heat input in Btu/h;

$$P_i = \text{electric power output in MW}; \quad i = 1, 2, 3).$$

3.8 Thermal Scheduling with Straight-Line Segments for Input–Output Curves

The following data apply to three thermal units. Compute and sketch the input–output characteristics and the IHR characteristics. Assume the unit input–output curves consist of straight-line segments between the given power points.

Unit No.	Power Output (MW)	Net Heat Rate (Btu/kWh)
1	45	13,512.5
	300	9,900.0
	350	9,918.0
2	45	22,764.5
	200	11,465.0
	300	11,060.0
	350	11,117.9
3	47.5	16,039.8
	200	10,000.0
	300	9,583.3
	450	9,513.9

Fuel costs are

Unit No.	Fuel Cost ($/MBtu)
1	0.61
2	0.75
3	0.75

Compute the economic schedule for system demands of 300, 460, 500, and 650 MW, assuming all three units are on-line. Give unit loadings and costs per hour as well as total costs in $ per hour.

www.mepcafe.com

3.9 Environmental Dispatch

Recently, there has been concern that optimum *economic* dispatch was not the best environmentally. The principles of economic dispatch can fairly easily be extended to handle this problem. The following is a problem based on a real situation that occurred in the Midwestern United States in 1973. Other cases have arisen with "NO_x" emission in Los Angeles.Two steam units have input–output curves as follows:

$$H_1 = 400 + 5P_1 + 0.01P_1^2, \quad \text{MBtu/h}, \quad 20 \le P_1 \le 200 \text{ MW}$$
$$H_2 = 600 + 4P_2 + 0.015P_2^2, \quad \text{MBtu/h}, \quad 20 \le P_2 \le 200 \text{ MW}$$

The units each burn coal with a heat content of 11,500 Btu/lb that costs \$13.50t (2,000 lb). The combustion process in each unit results in 11.75% of the coal by weight going up the stack as fly ash.

a. Calculate the net heat rates of both units at 200 MW.

b. Calculate the IHRs; schedule each unit for optimum *economy* to serve a total load of 250 MW with both units on-line.

c. Calculate the cost of supplying that load.

d. Calculate the rate of emission of fly ash for that case in pounds (lb) per hour, assuming no fly ash removal devices are in service.

e. Unit 1 has a precipitator installed that removes 85% of the fly ash; unit 2's precipitator is 89% efficient. Reschedule the two units for the minimum *total fly ash emission rate* with both on-line to serve a 250-MW load.

f. Calculate the rate of emission of ash and the cost for this schedule to serve the 250-MW load. What is the cost penalty?

g. Where does all that fly ash go?

3.10 Take the generation data shown in Example 3A. Ignore the generation limits and solve for the economic dispatch using the gradient method and Newton's method. Solve for a total generation of 900 MW in each case.

3.11 You have been assigned the job of building an oil pipeline from the West Coast of the United States to the East Coast. You are told that any one of the three West Coast sites is satisfactory and any of the three East Coast sites is satisfactory. The numbers in Figure 3.31 represent relative cost in hundreds of millions ($\$ \times 10^8$). Find the cheapest West Coast to East Coast pipeline.

3.12 The Stagecoach Problem

A mythical salesman who had to travel west by stagecoach, through unfriendly country, wished to take the safest route. His starting point and destination were fixed, but he had considerable choice as to which states he would travel through en route. The possible stagecoach routes are shown in Figure 3.32. After some thought, the salesman deduced a clever way of determining his safest route. Life insurance policies were offered to passengers, and since the cost of each policy was based on a careful evaluation of the safety of that run, the safest route should be the one with the cheapest policy. The cost of the standard policy on the stagecoach run from state i to state j, denoted as C_{ij}, is given in Figure 3.33. Find the safest path(s) for the salesman to take.

www.mepcafe.com

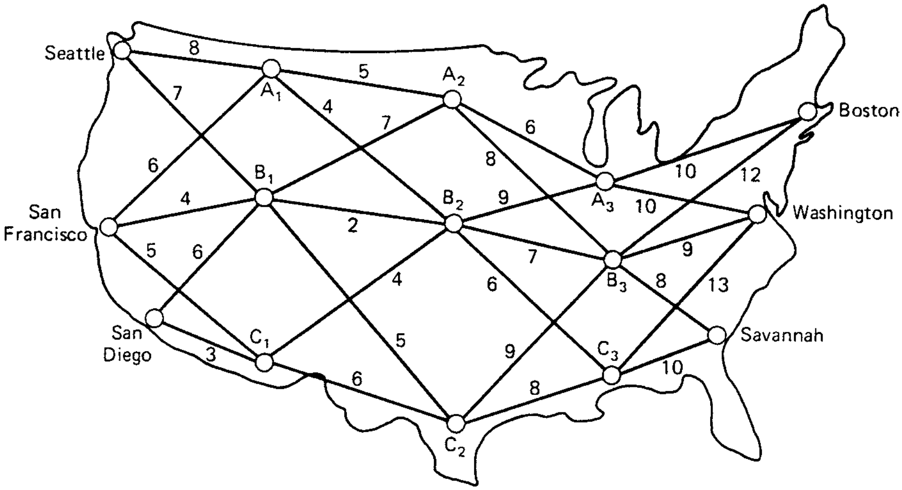


FIGURE 3.31 Possible oil pipeline routes for Problem 3.11.

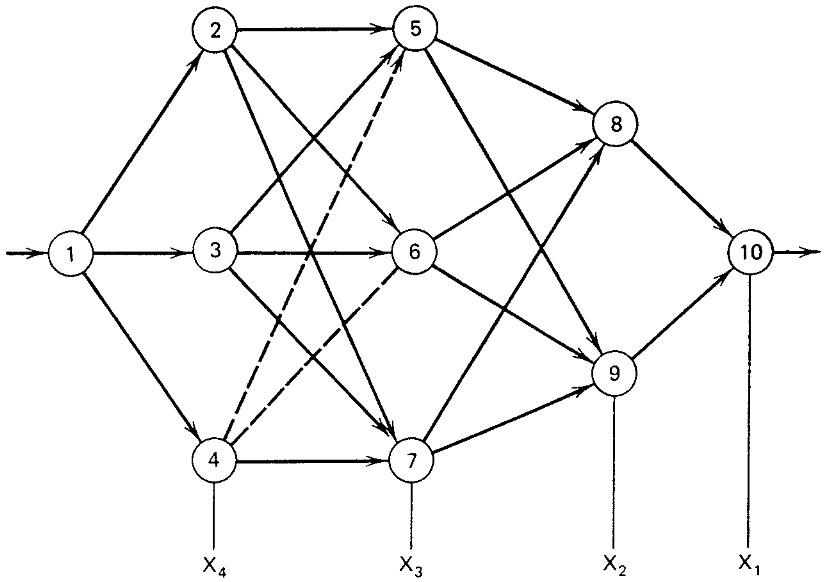


FIGURE 3.32 Possible stagecoach routes for Problem 3.12.

3.13 Economic Dispatch Problem

Consider three generating units that do not have convex input–output functions. (This is the type of problem one encounters when considering valve points in the dispatch problem.)

Unit 1:

$$H_1(P_1) = \begin{cases} 80 + 8P_1 + 0.024P_1^2 & 20 \text{ MW} \le P_1 \le 60 \text{ MW} \\ 196.4 + 3P_1 + 0.075P_1^2 & 60 \text{ MW} \le P_1 \le 100 \text{ MW} \end{cases}$$

www.mepcafe.com

i \ j	1	2	3	4	5	6	7	8	9	10
= 1		2	4	3						
2					7	4	6			
3					3	2	4			
4					4	1	5			
5								1	4	
6								6	3	
7								3	2	
8										3
9										4
10										

FIGURE 3.33 Cost to go from state *i* to state *j* in Problem 3.12. Costs not shown are infinite.

Generation limits are $20\text{ MW} \le P_1 \le 100\text{ MW}$.

Unit 2:

$$H_2(P_2) = \begin{cases} 120 + 6P_2 + 0.04P_2^2 & 20\text{ MW} \le P_2 \le 40\text{ MW} \\ 157.335 + 3.3333P_2 + 0.08333P_2^2 & 40\text{ MW} \le P_2 \le 100\text{ MW} \end{cases}$$

Generation limits are $20\text{ MW} \le P_2 \le 100\text{ MW}$.

Unit 3:

$$H_3(P_3) = \begin{cases} 100 + 4.6666P_3 + 0.13333P_3^2 & 20\text{ MW} \le P_3 \le 50\text{ MW} \\ 316.66 + 2P_3 + 0.1P_3^2 & 50\text{ MW} \le P_3 \le 100\text{ MW} \end{cases}$$

Generation limits are $20\text{ MW} \le P_3 \le 100\text{ MW}$. Fuel costs $= 1.5\ \$/\text{MBtu}$ for all units.

a. Plot the cost function for each unit (see Problem 3.1).

b. Plot the incremental cost function for each unit.

c. Find the most economical dispatch for the following total demands assuming all units are on-line:

$$P_D = 100\text{ MW}$$
$$P_D = 140\text{ MW}$$
$$P_D = 180\text{ MW}$$
$$P_D = 220\text{ MW}$$
$$P_D = 260\text{ MW}$$

www.mepcafe.com

where

$$P_D = P_1 + P_2 + P_3$$

Solve using DP and discrete load steps of 20 MW, starting at 20 MW through 100 MW for each unit.

d. Can you solve these dispatch problems without DP? If you think you know how, try solving for $P_D = 100$ MW.

3.14 Composite Generating Cost Function

Refer to Example 3F, where three generating units are combined into a single composite generating unit. Repeat the example, except develop an equivalent incremental cost characteristic using only the incremental characteristics of the three units. Using this composite incremental characteristic plus the zero-load intercept costs of the three units, develop the total cost characteristic of the composite. (Suggestion: Fit the composite incremental cost data points using a linear approximation and a least-squares fitting algorithm.)

3.15 Refer to Problem 3.8, where three generator units have input–output curves specified as a series of straight-line segments. Can you develop a composite input–output curve for the three units? Assume all three units are on-line and that the composite input–output curve has as many linear segments as needed.

3.16 Given: the two generating units in the following with piecewise linear cost functions $F(P)$ as shown.

Unit 1:

$$P_1^{\min} = 25 \text{ MW} \quad \text{and} \quad P_1^{\max} = 200 \text{ MW}$$

P_1 (MW)	$F_1(P_1)$ (\$/h)
25	289.0
100	971.5
150	1436.5
200	1906.5

Unit 2:

$$P_2^{\min} = 50 \text{ MW} \quad \text{and} \quad P_2^{\max} = 400 \text{ MW}$$

P_2 (MW)	$F_2(P_2)$ (\$/h)
50	3800
100	4230
200	5120
400	6960

www.mepcafe.com

Find the optimum generation schedule for a total power delivery of 350 MW (assume both generators are on-line).

3.17 Given: two generator units with piecewise linear incremental cost functions as shown.

Unit 1:

$$P_1^{\min} = 100 \text{ MW} \quad \text{and} \quad P_1^{\max} = 400 \text{ MW}$$

P_1(MW)	$\frac{d}{dP_1}F_1(P_1)$(\$/MWh)
100	6.5
200	7.0
300	8.0
400	11.0

Unit 2:

$$P_2^{\min} = 120 \text{ MW} \quad \text{and} \quad P_2^{\max} = 300 \text{ MW}$$

P_2(MW)	$\frac{d}{dP_2}F_2(P_2)$(\$/MWh)
120	8.0
150	8.3
200	9.0
300	12.5

Find the optimum schedule for a total power delivery of 500 MW.

REFERENCES

1. Kuhn, H. W., Tucker,, A. W., *Nonlinear Programming in Second Berkeley Symposium on Mathematical Programming Statistics and Probability, 1950*, University of California Press, Berkeley, Berkeley,, CA, 1951.
2. Karush., *Minima of Functions of Several Variables with Inequalities as Side Constraints*, MSc Dissertation, Department of Mathematics, University of Chicago, Chicago, IL, 1939.,
3. Bazaara, M. S., Jarvis, J. J., Sherali, H. D., *Linear Programming and Network Flows*, Second Edition, John Wiley & Sons, New York, 1990.
4. Frederick, S. H., Lieberman, G. J., *(Introduction to) Operations Research*, Holden-Day, San Francisco, CA, c1974, etc. any edition.
5. Luenberger, D. G., *Introduction to Linear and Nonlinear Programming*, Second Edition, Addison-Wesley Publishing Company, Reading, MA.
6. Luenberger, D. G., *Linear and Nonlinear Programming*, Second Edition, Addison-Wesley Publishing Company, Reading, MA.
7. Bellman, R. E., Dreyfus, S. E., *Applied Dynamic Programming*, Princeton University Press, Princeton, NJ, 1962.
8. Kaufmann, A., *Graphs, Dynamic Programming and Markov Processes*, Wiley and Technology Press, New York, 1967.

www.mepcafe.com

4

UNIT COMMITMENT

4.1 INTRODUCTION

Because human activity follows cycles, most systems supplying services to a large population will experience cycles. This includes transportation systems, communication systems, as well as electric power systems. In the case of an electric power system, the total load on the system will generally be higher during the daytime and early evening when industrial loads are high, lights are on, and so forth and lower during the late evening and early morning when most of the population is asleep. In addition, the use of electric power has a weekly cycle, the load being lower over weekend days than weekdays. But why is this a problem in the operation of an electric power system? Why not just simply commit enough units to cover the maximum system load and leave them running? Note that to "commit" a generating unit is to "turn it on," that is, to bring the unit up to speed, synchronize it to the system, and connect it so it can deliver power to the network. The problem with "commit enough units and leave them on-line" is one of economics. As will be shown in Example 4A, it is quite expensive to run too many generating units. A great deal of money can be saved by turning units off (decommitting them) when they are not needed.

4.1.1 Economic Dispatch versus Unit Commitment

At this point, it may be as well to emphasize the essential difference between the unit commitment and economic dispatch problem. The economic dispatch problem

Power Generation, Operation, and Control, Third Edition. Allen J. Wood, Bruce F. Wollenberg, and Gerald B. Sheblé.
© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

www.mepcafe.com

assumes that there are N_{gen} units already connected to the system. The purpose of the economic dispatch problem is to find the optimum operating policy for these N_{gen} units. This is the problem that we have been investigating so far in this text.

The unit commitment problem, on the other hand, is more complex. We may assume that we have N_{gen} units available to us and that we have a forecast of the demand to be served. The question that is asked in the unit commitment problem area is approximately as follows.

> Given that there are a number of subsets of the complete set of N_{gen} generating units that would satisfy the expected demand, which of these subsets should be used in order to provide the minimum operating cost?

This unit commitment problem may be extended over some period of time, such as the 24 h of a day or the 168 h of a week. The unit commitment problem is a much more difficult problem to solve. The solution procedures involve the economic dispatch problem as a subproblem. That is, for each of the subsets of the total number of units that are to be tested, for any given set of them connected to the load, the particular subset should be operated in optimum economic fashion. This will permit finding the minimum operating cost for that subset, but it does not establish which of the subsets is in fact the one that will give minimum cost over a period of time.

A later chapter will consider the unit commitment problem in some detail. The problem is more difficult to solve mathematically since it involves integer variables. That is, generating units must be either all on or all off. (How can you turn a switch half on?)

Example 4A: Suppose one had the three units given here:

$$\textbf{Unit 1}: \quad \text{Min} = 150\,\text{MW}$$
$$\text{Max} = 600\,\text{MW}$$
$$H_1 = 510.0 + 7.2P_1 + 0.00142P_1^2 \text{ MBtu/h}$$

$$\textbf{Unit 2}: \quad \text{Min} = 100\,\text{MW}$$
$$\text{Max} = 400\,\text{MW}$$
$$H_2 = 310.0 + 7.85P_2 + 0.00194P_2^2 \text{ MBtu/h}$$

$$\textbf{Unit 3}: \quad \text{Min} = 50\,\text{MW}$$
$$\text{Max} = 200\,\text{MW}$$
$$H_3 = 78.0 + 7.97P_3 + 0.00482P_3^2 \text{ MBtu/h}$$

with fuel costs:

$$\text{Fuel cost}_1 = 1.1\,\text{R/MBtu}$$
$$\text{Fuel cost}_2 = 1.0\,\text{R/MBtu}$$
$$\text{Fuel cost}_3 = 1.2\,\text{R/MBtu}$$

www.mepcafe.com

If we are to supply a load of 550 MW, what unit or combination of units should be used to supply this load most economically? To solve this problem, simply try all combinations of the three units. Some combinations will be infeasible if the sum of all maximum MW for the units committed is less than the load or if the sum of all minimum MW for the units committed is greater than the load. For each feasible combination, the units will be dispatched using the techniques of Chapter 3. The results are presented in Table 4.1.

Note that the least expensive way to supply the generation is not with all three units running or even any combination involving two units. Rather, the optimum commitment is to only run unit 1, the most economic unit. By only running the most economic unit, the load can be supplied by that unit operating closer to its best efficiency. If another unit is committed, both unit 1 and the other unit will be loaded further from their best efficiency points such that the net cost is greater than unit 1 alone.

Suppose the load follows a simple "peak-valley" pattern as shown in Figure 4.1. If the operation of the system is to be optimized, units must be shut down as the load goes down and then recommitted as it goes back up. We would like to know which units to drop and when. As we will show later, this problem is far from trivial when real generating units are considered. One approach to this solution is demonstrated in Example 4B, where a simple priority-list scheme is developed.

Example 4B: Suppose we wish to know which units to drop as a function of system load. Let the units and fuel costs be the same as in Example 4A, with the load varying from a peak of 1200 MW to a valley of 500 MW. To obtain a "shutdown rule," simply use a brute-force technique wherein all combinations of units will be tried (as in Example 4A) for each load value taken in steps of 50 MW from 1200 to 500. The results of applying this brute-force technique are given in Table 4.2. Our shutdown rule is quite simple.

When load is above 1000 MW, run all three units; between 1000 and 600 MW, run units 1 and 2; below 600 MW, run only unit 1.

Figure 4.2 shows the unit commitment schedule derived from this shutdown rule as applied to the load curve of Figure 4.1.

So far, we have only obeyed one simple constraint: *Enough units will be committed to supply the load*. If this were all that was involved in the unit commitment problem—that is, just meeting the load—we could stop here and state that the problem was "solved." Unfortunately, other constraints and other phenomena must be taken into account in order to claim an optimum solution. These constraints will be discussed in the next section, followed by a description of some of the presently used methods of solution.

www.mepcafe.com

TABLE 4.1 Unit Combinations and Dispatch for 550-MW Load of Example 4A

Unit 1	Unit 2	Unit 3	Max Gen	Min Gen	P_1	P_2	P_3	F_1	F_2	F_3	Total Gen Cost $F_1 + F_2 + F_3$
Off	Off	Off	0	0							Infeasible
Off	Off	On	200	50							Infeasible
Off	On	Off	400	100							Infeasible
Off	On	On	600	150	0	400	150	0	3760	1658	5418
On	Off	Off	600	150	550	0	0	5389	0	0	5389
On	Off	On	800	200	500	0	50	4911	0	586	5497
On	On	Off	1000	250	295	255	0	3030	2440	0	5471
On	On	On	1200	300	267	233	50	2787	2244	586	5617

www.mepcafe.com

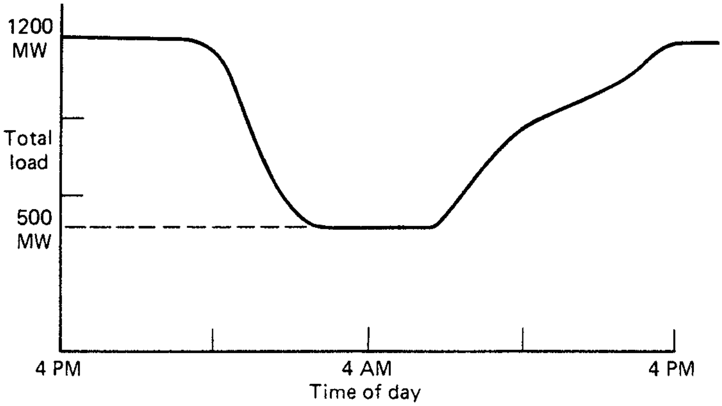


FIGURE 4.1 Simple "peak-valley" load pattern.

TABLE 4.2 "Shutdown Rule" Derivation for Example 4B

	Optimum Combination		
Load	Unit 1	Unit 2	Unit 3
1200	On	On	On
1150	On	On	On
1100	On	On	On
1050	On	On	On
1000	On	On	Off
950	On	On	Off
900	On	On	Off
850	On	On	Off
800	On	On	Off
750	On	On	Off
700	On	On	Off
650	On	On	Off
600	On	Off	Off
550	On	Off	Off
500	On	Off	Off

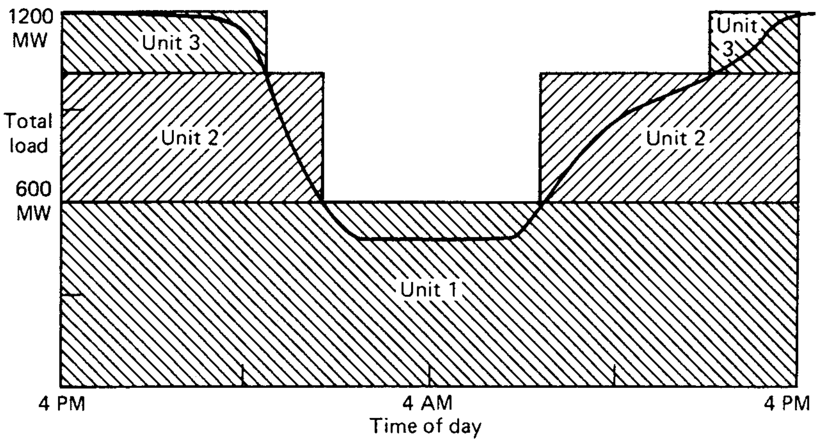


FIGURE 4.2 Unit commitment schedule using the shutdown rule.

www.mepcafe.com

4.1.2 Constraints in Unit Commitment

Many constraints can be placed on the unit commitment problem. The list presented here is by no means exhaustive. Each individual power system, power pool, reliability council, and so forth may impose different rules on the scheduling of units, depending on the generation makeup, load-curve characteristics, and such.

4.1.3 Spinning Reserve

Spinning reserve is the term used to describe the total amount of generation available from all units synchronized (i.e., spinning) on the system, minus the present load and losses being supplied. Spinning reserve must be carried so that the loss of one or more units does not cause too far a drop in system frequency (see Chapter 10). Quite simply, if one unit is lost, there must be ample reserve on the other units to make up for the loss in a specified time period.

Spinning reserve must be allocated to obey certain rules, usually set by regional reliability councils (in the United States) that (specify how the) reserve is to be allocated to various units. Typical rules specify that reserve must be a given percentage of forecasted peak demand or that reserve must be capable of making up the loss of the most heavily loaded unit in a given period of time. Others calculate reserve requirements as a function of the probability of not having sufficient generation to meet the load.

Not only must the reserve be sufficient to make up for a generating unit failure, but the reserves must be allocated among fast-responding units and slow-responding units. This allows the automatic generation control system (see Chapter 10) to restore frequency and interchange quickly in the event of a generating unit outage.

Beyond spinning reserve, the unit commitment problem may involve various classes of "scheduled reserves" or "off-line" reserves. These include quick-start diesel or gas-turbine units as well as most hydro-units and pumped-storage hydro-units that can be brought on-line, synchronized, and brought up to full capacity quickly. As such, these units can be "counted" in the overall reserve assessment, as long as their time to come up to full capacity is taken into account.

Reserves, finally, must be spread around the power system to avoid transmission system limitations (often called "bottling" of reserves) and to allow various parts of the system to run as "islands," should they become electrically disconnected.

Example 4C: Suppose a power system consisted of two isolated regions: a western region and an eastern region. Five units, as shown in Figure 4.3, have been committed to supply 3090 MW. The two regions are separated by transmission tie lines that can together transfer a maximum of 550 MW in either direction. This is also shown in Figure 4.3. What can we say about the allocation of spinning reserve in this system?

The data for the system in Figure 4.3 are given in Table 4.3. With the exception of unit 4, the loss of any unit on this system can be covered by the spinning reserve on the remaining units. Unit 4 presents a problem, however. If unit 4 were to be lost and unit 5 were to be run to its maximum of 600 MW, the eastern region would still need 590 MW to cover the load in that region. The 590 MW would have to be transmitted over the tie lines from the western region, which can easily supply 590 MW from its

www.mepcafe.com

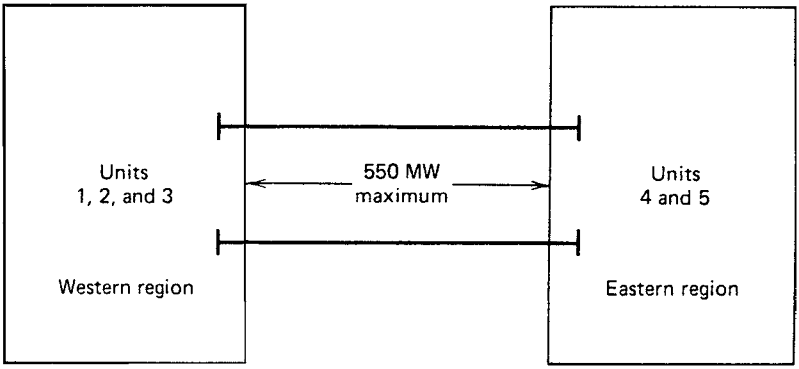


FIGURE 4.3 Two-region system.

TABLE 4.3 Data for the System in Figure 4.3

Region	Unit	Unit Capacity (MW)	Unit Output (MW)	Regional Generation (MW)	Spinning Reserve	Regional Load (MW)	Interchange (MW)
Western	1	1000	900		100		
	2	800	420	1740	380	1900	160 in
	3	800	420		380		
Eastern	4	1200	1040	1350	160	1190	160 out
	5	600	310		290		
Total	1–5	4400	3090	3090	1310	3090	

reserves. However, the tie capacity of only 550 MW limits the transfer. Therefore, the loss of unit 4 cannot be covered even though the entire system has ample reserves. The only solution to this problem is to commit more units to operate in the eastern region.

4.1.4 Thermal Unit Constraints

Thermal units usually require a crew to operate them, especially when turned on and turned off. A thermal unit can undergo only gradual temperature changes, and this translates into a time period of some hours required to bring the unit on-line. As a result of such restrictions in the operation of a thermal plant, various constraints arise, such as:

- **Minimum uptime**: once the unit is running, it should not be turned off immediately.
- **Minimum downtime**: once the unit is decommitted, there is a minimum time before it can be recommitted.
- **Crew constraints**: if a plant consists of two or more units, they cannot both be turned on at the same time since there are not enough crew members to attend both units while starting up.

www.mepcafe.com

In addition, because the temperature and pressure of the thermal unit must be moved slowly, a certain amount of energy must be expended to bring the unit on-line. This energy does not result in any MW generation from the unit and is brought into the unit commitment problem as a *start-up cost*.

The start-up cost can vary from a maximum "cold-start" value to a much smaller value if the unit was only turned off recently and is still relatively close to operating temperature. There are two approaches to treating a thermal unit during its down period. The first allows the unit's boiler to cool down and then heat back up to operating temperature in time for a scheduled turn on. The second (called *banking*) requires that sufficient energy be input to the boiler to just maintain operating temperature. The costs for the two can be compared so that, if possible, the best approach (cooling or banking) can be chosen.

$$\text{Start-up cost when cooling} = C_c\left(1-\varepsilon^{-t/\alpha}\right)\times F + C_f$$

where

$$\begin{aligned}
C_c &= \text{cold-start cost (MBtu)}\\
F &= \text{fuel cost}\\
C_f &= \text{fixed cost (includes crew expense, maintenance expenses) (in \$)}\\
\alpha &= \text{thermal time constant for the unit}\\
t &= \text{time (h) the unit was cooled}
\end{aligned}$$

$$\text{Start-up cost when banking} = C_t \times t \times F + C_f$$

where

$$C_t = \text{cost (MBtu/h) of maintaining unit at operating temperature}$$

Up to a certain number of hours, the cost of banking will be less than the cost of cooling, as is illustrated in Figure 4.4.

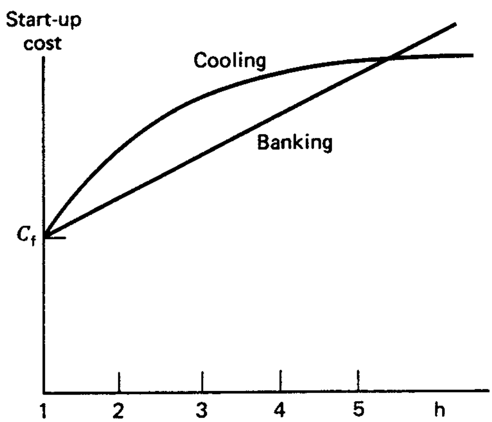


FIGURE 4.4 Time-dependent start-up costs.

www.mepcafe.com

Finally, the capacity limits of thermal units may change frequently, due to maintenance or unscheduled outages of various equipment in the plant; this must also be taken into account in unit commitment.

4.1.5 Other Constraints

4.1.5.1 Must Run. Some units are given a must-run status during certain times of the year for reason of voltage support on the transmission network or for such purposes as supply of steam for uses outside the steam plant itself.

4.1.5.2 Fuel Constraints. We will treat the "fuel scheduling" problem briefly in Chapter 5. A system in which some units have limited fuel, or else have constraints that require them to burn a specified amount of fuel in a given time, presents a most challenging unit commitment problem.

4.1.5.3 Hydro-Constraints. Unit commitment cannot be completely separated from the scheduling of hydro-units. In this chapter, we will assume that the hydrothermal scheduling (or "coordination") problem can be separated from the unit commitment problem. We, of course, cannot assert flatly that our treatment in this fashion will always result in an optimal solution.

4.2 UNIT COMMITMENT SOLUTION METHODS

The commitment problem can be very difficult. As a theoretical exercise, let us postulate the following situation.

- We must establish a loading pattern for M periods.
- We have N_{gen} units to commit and dispatch.
- The M load levels and operating limits on the N_{gen} units are such that any one unit can supply the individual loads and that any combination of units can also supply the loads.

Next, assume we are going to establish the commitment by enumeration (brute force). The total number of combinations we need to try each hour is

$$C\left(N_{gen},1\right)+C\left(N_{gen},2\right)+\cdots+C\left(N_{gen},N_{gen}-1\right)+C\left(N_{gen},N_{gen}\right)=2^{N_{gen}}-1$$

where $C(N_{gen},j)$ is the combination of N_{gen} items taken j at a time. That is,

$$C(N_{gen},j)=\left[\frac{N_{gen}!}{(N_{gen}-j)!\,j!}\right]$$

$$j!=1\times2\times3\times\cdots\times j$$

For the total period of M intervals, the maximum number of possible combinations is $(2^{N_{gen}}-1)^M$, which can become a horrid number to think about.

www.mepcafe.com

For example, take a 24-h period (e.g., 24 one-hour intervals) and consider systems with 5, 10, 20, and 40 units. The value of $(2^{N_{gen}} - 1)^{24}$ becomes the following.

N_{gen}	$\left(2^{N_{gen}} - 1\right)^{24}$
5	6.2×10^{35}
10	1.73×10^{72}
20	3.12×10^{144}
40	(Too big)

These very large numbers are the upper bounds for the number of enumerations required. Fortunately, the constraints on the units and the load–capacity relationships of typical utility systems are such that we do not approach these large numbers. Nevertheless, the real practical barrier in the optimized unit commitment problem is the high dimensionality of the possible solution space.

The most talked-about techniques for the solution of the unit commitment problem are:

- Priority-list schemes,
- Dynamic programming (DP),
- Lagrange relaxation (LR).
- Mixed integer linear programming (MILP)

4.2.1 Priority-List Methods

The simplest unit commitment solution method consists of creating a priority list of units. As we saw in Example 4B, a simple shutdown rule or priority-list scheme could be obtained after an exhaustive enumeration of all unit combinations at each load level. The priority list of Example 4B could be obtained in a much simpler manner by noting the full-load average production cost of each unit, where the full-load average production cost is simply the net heat rate at full load multiplied by the fuel cost.

Example 4D: Construct a priority list for the units of Example 4A. (Use the same fuel costs as in Example 4A.) First, the full-load average production cost will be calculated:

Unit	**Full Load Average Production Cost** (\$/MWh)
1	9.79
2	9.48
3	11.188

A strict priority order for these units, based on the average production cost, would order them as follows:

www.mepcafe.com

Unit	$/ MWh	Min MW	Max MW
2	9.401	100	400
1	9.79	150	600
3	11.188	50	200

And the commitment scheme would (ignoring min up-/downtime, start-up costs, etc.) simply use only the following combinations.

Combination	Min MW from Combination	Max MW Combination
2+1+3	300	1200
2+1	250	1000
2	100	400

Note that such a scheme would not completely parallel the shutdown sequence described in Example 4B, where unit 2 was shut down at 600 MW leaving unit 1. With the priority-list scheme, both units would be held on until load reached 400 MW, then unit 1 would be dropped.

Most priority-list schemes are built around a simple shutdown algorithm that might operate as follows:

- At each hour when load is dropping, determine whether dropping the next unit on the priority list will leave sufficient generation to supply the load plus spinning-reserve requirements. If not, continue operating as is; if yes, go on to the next step.
- Determine the number of hours, *H*, before the unit will be needed again, that is, assuming that the load is dropping and will then go back up some hours later.
- If *H* is less than the minimum shutdown time for the unit, keep the commitment as is and go to the last step; if not, go to the next step.
- Calculate two costs. The first is the sum of the hourly production costs for the next *H* hours with the unit up. Then recalculate the same sum for the unit down and add in the start-up cost for either cooling the unit or banking it, whichever is less expensive. If there is sufficient savings from shutting down the unit, it should be shut down; otherwise, keep it on.
- Repeat this entire procedure for the next unit on the priority list. If it is also dropped, go to the next and so forth.

Various enhancements to the priority-list scheme can be made by grouping of units to ensure that various constraints are met. We will note later that DP methods usually create the same type of priority list for use in the DP search.

4.2.2 Lagrange Relaxation Solution

The DP method of solution of the unit commitment problem has many disadvantages for large power systems with many generating units. This is because of the necessity

www.mepcafe.com

of forcing the DP solution to search over a small number of commitment states to reduce the number of combinations that must be tested in each time period.

In the Lagrange relaxation technique, these disadvantages disappear (although other technical problems arise and must be addressed, as we shall see). This method is based on a dual optimization approach as introduced in Appendix 3E and further expanded in Appendix 4A to this chapter. (The reader should be familiar with both of these appendices before proceeding further.)

We start by defining the variable U_i^t as

$$U_i^t = 0 \text{ if unit } i \text{ is off-line during period } t$$
$$U_i^t = 1 \text{ if unit } i \text{ is on-line during period } t \tag{4.1}$$

We shall now define several constraints and the objective function of the unit commitment problem:

1. Loading constraints:

$$P_{\text{load}}^t - \sum_{i=1}^{N_{\text{gen}}} P_i^t U_i^t = 0 \quad \text{for} \quad t = 1 \ldots T \tag{4.2}$$

2. Unit limits:

$$U_i^t P_i^{\min} \le P_i^t \le U_i^t P_i^{\max} \quad \text{for} \quad i = 1 \ldots N_{\text{gen}} \quad \text{and} \quad t = 1 \ldots T \tag{4.3}$$

3. Unit minimum up- and downtime constraints. Note that other constraints can easily be formulated and added to the unit commitment problem. These include transmission security constraints (see Chapter 7), generator fuel limit constraints, and system air quality constraints in the form of limits on emissions from fossil-fired plants, spinning-reserve constraints, etc.
4. The objective function is

$$\sum_{t=1}^{T} \sum_{i=1}^{N_{\text{gen}}} [F_i\left(P_i^t\right) + \text{start-up cost}_{i,t}] U_i^t = F\left(P_i^t, U_i^t\right) \tag{4.4}$$

We can then form the Lagrange function similar to the way we did in the economic dispatch problem:

$$\mathcal{L}(P, U, \lambda) = F\left(P_i^t, U_i^t\right) + \sum_{t=1}^{T} \lambda^t \left(P_{\text{load}}^t - \sum_{i=1}^{N_{\text{gen}}} P_i^t U_i^t \right) \tag{4.5}$$

The unit commitment problem requires that we minimize this Lagrange function, subject to the local unit constraints 2 and 3, which can be applied to each unit separately. Note:

1. The cost function, $F\left(P_i^t, U_i^t\right)$, together with constraints 2 and 3 are each *separable over units*. That is, what is done with one unit does not affect the

www.mepcafe.com

cost of running another unit, as far as the cost function and the unit limits (constraint 2) and the unit up- and downtime (constraint 3) are concerned.

2. Constraints 1 are *coupling constraints* across the units so that what we do to one unit affects what will happen on other units if the coupling constraints are to be met.

The Lagrange relaxation procedure solves the unit commitment problem by "relaxing" or temporarily ignoring the coupling constraints and solving the problem as if they did not exist. This is done through the dual optimization procedure as explained in Appendix 4A of this chapter. The dual procedure attempts to reach the constrained optimum by maximizing the Lagrangian with respect to the Lagrange multipliers while minimizing with respect to the other variables in the problem; that is,

$$q^*(\lambda) = \max_{\lambda^t} q(\lambda) \tag{4.6}$$

where

$$q(\lambda) = \min_{P_i^t, U_i^t} \mathcal{L}(P, U, \lambda) \tag{4.7}$$

This is done in two basic steps:

Step 1. Find a value for each λ^t which moves $q(\lambda)$ toward a larger value.

Step 2. Assuming that the λ^t found in step 1 are now fixed, find the minimum of $\mathcal{L}$ by adjusting the values of P^t and U^t.

The adjustment of the λ^t values will be dealt with at a later time in this section; assume then that a value has been chosen for all the λ^t and that they are now to be treated as fixed numbers. We shall minimize the Lagrangian as follows.

First, we rewrite the Lagrangian as

$$\mathcal{L} = \sum_{t=1}^{T}\sum_{i=1}^{N_{\text{gen}}}\left[F_i\left(P_i^t\right) + \text{start-up cost}_{i,t}\right]U_i^t + \sum_{t=1}^{T}\lambda^t\left(P_{\text{load}}^t - \sum_{i=1}^{N_{\text{gen}}} P_i^t U_i^t\right) \tag{4.8}$$

This is now rewritten as

$$\mathcal{L} = \sum_{t=1}^{T}\sum_{i=1}^{N_{\text{gen}}}\left[F_i\left(P_i^t\right) + \text{start-up cost}_{i,t}\right]U_i^t + \sum_{t=1}^{T}\lambda^t P_{\text{load}}^t - \sum_{t=1}^{T}\sum_{i=1}^{N_{\text{gen}}}\lambda^t P_i^t U_i^t \tag{4.9}$$

The second term in the preceding equation is constant and can be dropped (since the λ^t are fixed). Finally, we write the Lagrange function as

$$\mathcal{L} = \sum_{i=1}^{N_{\text{gen}}}\left(\sum_{t=1}^{T}\left\{\left[F_i\left(P_i^t\right) + \text{start-up cost}_{i,t}\right]U_i^t - \lambda^t P_i^t U_i^t\right\}\right) \tag{4.10}$$

Here, we have achieved our goal of separating the units from one another. The term inside the outer brackets, that is,

www.mepcafe.com

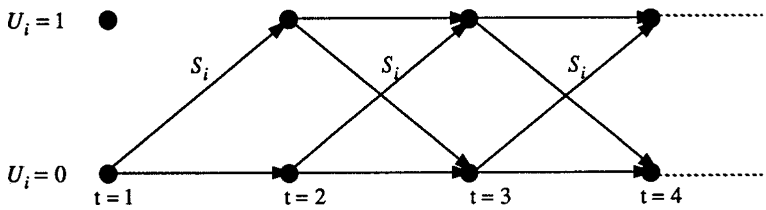


FIGURE 4.5 DP steps for a single unit.

$$\sum_{t=1}^{T}\left\{\left[F_i\left(P_i^t\right)+\text{start-up cost}_{i,t}\right]U_i^t-\lambda^t P_i^t U_i^t\right\}$$

can be solved separately for each generating unit, without regard for what is happening on the other generating units. The minimum of the Lagrangian is found by solving for the minimum for each generating unit over all time periods; that is,

$$\min q(\lambda)=\sum_{i=1}^{N_{\text{gen}}}\min\sum_{t=1}^{T}\left\{\left[F_i\left(P_i^t\right)+\text{start-up cost}_{i,t}\right]U_i^t-\lambda^t P_i^t U_i^t\right\} \tag{4.11}$$

subject to

$$U_i^t P_i^{\min}\le P_i^t\le U_i^t P_i^{\max}\quad\text{for}\quad t=1\ldots T$$

and the up- and downtime constraints. This is easily solved as a DP problem in one variable. This can be visualized in Figure 4.5, which shows the only two possible states for unit i (i.e., $U_i^t = 0$ or 1):

where S_i is the start-up cost for unit i.

At the $U_i^t = 0$ state, the value of the function to be minimized is trivial (i.e., it equals 0); at the state where $U_i^t = 1$, the function to be minimized is (the start-up cost is dropped here since the minimization is with respect to P_i^t)

$$\min\left[F_i\left(P_i\right)-\lambda^t P_i^t\right] \tag{4.12}$$

The minimum of this function is found by taking the first derivative:

$$\frac{\mathrm{d}}{\mathrm{d}P_i^t}\left[F_i\left(P_i\right)-\lambda^t P_i^t\right]=\frac{\mathrm{d}}{\mathrm{d}P_i^t}F_i\left(P_i^t\right)-\lambda^t=0 \tag{4.13}$$

The solution to this equation is

$$\frac{\mathrm{d}}{\mathrm{d}P_i^t}F_i\left(P_i^{\text{opt}}\right)=\lambda^t \tag{4.14}$$

There are three cases to be concerned with depending on the relation of P_i^{opt} and the unit limits:

1. If $P_i^{\text{opt}}\le P_i^{\min}$, then

www.mepcafe.com

$$\min\left[F_i\left(P_i\right)-\lambda^t P_i^t\right]=F_i\left(P_i^{\min}\right)-\lambda^t P_i^{\min} \tag{4.14a}$$

2. If $P_i^{\min} \le P_i^{\text{opt}} \le P_i^{\max}$, then

$$\min\left[F_i\left(P_i\right)-\lambda^t P_i^t\right]=F_i\left(P_i^{\text{opt}}\right)-\lambda^t P_i^{\text{opt}} \tag{4.14b}$$

3. If $P_i^{\text{opt}} \ge P_i^{\max}$, then

$$\min\left[F_i\left(P_i\right)-\lambda^t P_i^t\right]=F_i\left(P_i^{\max}\right)-\lambda^t P_i^{\max} \tag{4.14c}$$

The solution of the two-state dynamic program for each unit proceeds in the normal manner as was done for the forward DP solution of the unit commitment problem itself. Note that since we seek to minimize $\left[F_i\left(P_i\right)-\lambda^t P_i^t\right]$ at each stage and that when $U_i^t=0$ this value goes to 0, then the only way to get a lower value is to have

$$\left[F_i\left(P_i\right)-\lambda^t P_i^t\right]<0$$

The dynamic program should take into account all the start-up costs, S_i, for each unit, as well as the minimum up- and downtime for the generator. Since we are solving for each generator independently, however, we have avoided the dimensionality problems that affect the DP solution.

4.2.2.1 Adjusting Lambda. So far, we have shown how to schedule generating units with fixed values of λ^t for each time period. As shown in Appendix 4A to this chapter, the adjustment of λ^t must be done carefully so as to maximize $q(\lambda)$. Most references that work on the Lagrange relaxation procedure use a combination of gradient search and various heuristics to achieve a rapid solution. Note that unlike in Appendix 4A, the λ here is a vector of values, each of which must be adjusted. Much research in recent years has been aimed at ways to speed the search for the correct values of λ for each hour. In Example 4E, we shall use the same technique of adjusting λ for each hour that is used in the appendix. For the unit commitment problem solved in Example 4E, however, the λ adjustment factors are different:

$$\lambda^t=\lambda^t+\left[\frac{\mathrm{d}}{\mathrm{d}\lambda}q(\lambda)\right]\alpha \tag{4.16}$$

where

$$\alpha=0.01 \quad \text{when} \quad \frac{\mathrm{d}}{\mathrm{d}\lambda}q(\lambda) \text{ is positive} \tag{4.17}$$

$$\alpha=0.002 \quad \text{when} \quad \frac{\mathrm{d}}{\mathrm{d}\lambda}q(\lambda) \text{ is negative} \tag{4.18}$$

Each λ^t is treated separately. The overall Lagrange relaxation unit commitment algorithm is shown in Figure 4.6.

A measure of the closeness to the solution, called the "relative duality gap," is defined as $(J^*-q^*)/q^*$. The relative duality gap is used in Example 4E as a measure

www.mepcafe.com

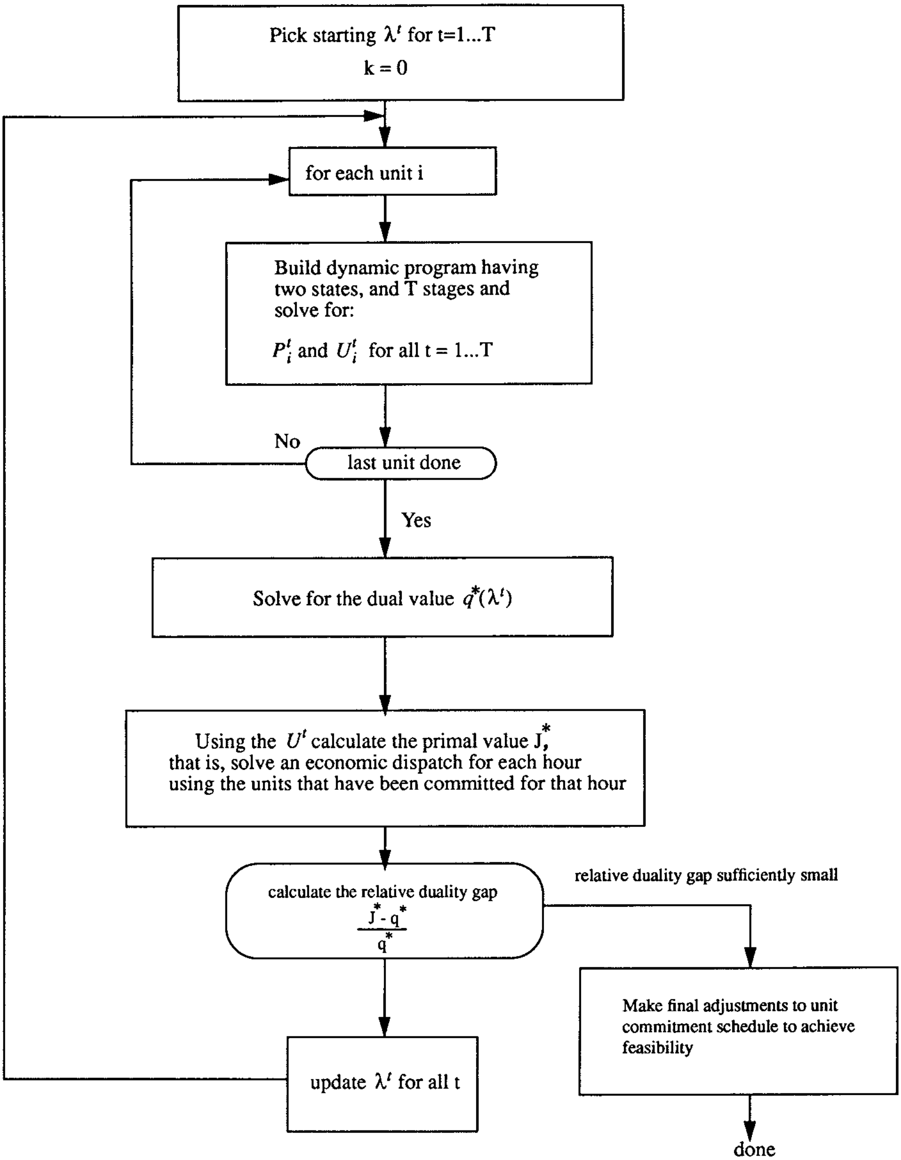


FIGURE 4.6 Lagrange relaxation procedure for unit commitment.

of the closeness to the solution. The paper in the footnote[1] points out several useful things about dual optimization applied to the unit commitment problem.

[1] Bertsekas, D., Lauer, G. S., Sandell, N. R., Posbergh, T. A., "Optimal Short-Term Scheduling of Large-Scale Power Systems," IEEE Transactions on Automatic Control, Vol. AC-28, No. 1, January 1983, pp. 1–11.

www.mepcafe.com

1. For large, real-sized, power system unit commitment calculations, the duality gap does become quite small as the dual optimization proceeds, and its size can be used as a stopping criterion. The larger the problem (larger number of generating units), the smaller the gap.
2. The convergence is unstable at the end, meaning that some units are being switched in and out, and the process never comes to a definite end.
3. There is no guarantee that when the dual solution is stopped, it will be at a feasible solution.

All of these are demonstrated in Example 4E. The duality gap is large at the beginning and becomes progressively smaller as the iterations progress. The solution reaches a commitment schedule when at least enough generation is committed so that an economic dispatch can be run, and further iterations only result in switching marginal units on and off. Finally, the loading constraints are not met by the dual solution when the iterations are stopped.

Many of the Lagrange relaxation unit commitment programs use a few iterations of a DP algorithm to get a good starting point, then run the dual optimization iterations, and finally, at the end, use heuristic logic or restricted DP to get to a final solution. The result is a solution that is not limited to search windows, such as what had to be done in strict application of DP.

EXAMPLE 4E In this example, a three-generator, four-hour unit commitment problem will be solved. The data for this problem are as follows. Given the three generating units in the following,

$$F_1(P_1) = 500 + 10P_1 + 0.002P_1^2 \quad \text{and} \quad 100 \le P_1 \le 600$$
$$F_2(P_2) = 300 + 8P_2 + 0.0025P_2^2 \quad \text{and} \quad 100 \le P_2 \le 400$$
$$F_3(P_3) = 100 + 6P_3 + 0.005P_3^2 \quad \text{and} \quad 50 \le P_3 \le 200$$

Load:

t	$\mathbf{P}^{\mathbf{t}}_{\text{load}}$ (MW)
1	170
2	520
3	1100
4	330

no start-up costs, no minimum up- or downtime constraints.

This example is solved using the Lagrange relaxation technique. Shown in the following text are the results of several iterations, starting from an initial condition where all the λ^t values are set to 0. An economic dispatch is run for each hour, provided there is sufficient generation committed that hour. If there is not enough generation committed, the total cost for that hour is set arbitrarily to 10,000. Once each hour has enough generation committed, the primal value J^* simply represents the total generation cost summed over all hours as calculated by the economic dispatch.

www.mepcafe.com

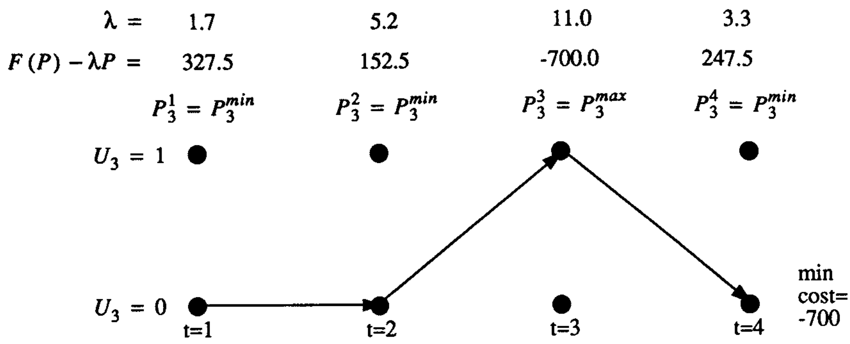


FIGURE 4.7 DP step for unit 3.

The dynamic program for each unit with a $\lambda^t = 0$ for each hour will always result in all generating units off-line.

Iteration 1

Hour	λ	u_1	u_2	u_3	P_1	P_2	P_3	$P_{\text{load}}^t - \sum_{i=1}^{N} P_i^t U_i^t$	P_1^{edc}	P_2^{edc}	P_3^{edc}
1	0	0	0	0	0	0	0	170	0	0	0
2	0	0	0	0	0	0	0	520	0	0	0
3	0	0	0	0	0	0	0	1100	0	0	0
4	0	0	0	0	0	0	0	330	0	0	0

$$q(\lambda) = 0.0, \quad J^* = 40{,}000, \quad \text{and} \quad \frac{J^* - q^*}{q^*} = \text{undefined}$$

In the next iteration, the λ^t values have been increased. To illustrate the use of DP to schedule each generator, we will detail the DP steps for unit 3 in Figure 4.7.

The result is to schedule unit 3 off during hours 1, 2, and 4 and on during hour 3. Further, unit 3 is scheduled to be at its maximum of 200 MW during hour 3. The results, after all the units have been scheduled by DP, are as follows.

Iteration 2

Hour	λ	u_1	u_2	u_3	P_1	P_2	P_3	$P_{\text{load}}^t - \sum_{i=1}^{N_{\text{gen}}} P_i^t U_i^t$	P_1^{edc}	P_2^{edc}	P_3^{edc}
1	1.7	0	0	0	0	0	0	170	0	0	0
2	5.2	0	0	0		0	0	520	0	0	0
3	11.0	0	1	1	0	400	200	500	0	0	0
4	3.3	0	0	0	0	0	0	330	0	0	0

www.mepcafe.com

$$q(\lambda) = 14{,}982, \quad J^* = 40{,}000, \quad \text{and} \quad \frac{J^* - q^*}{q^*} = 1.67$$

Iteration 3

Hour	λ	u_1	u_2	u_3	P_1	P_2	P_3	$P_{\text{load}}^t - \sum_{i=1}^{N} P_i^t U_i^t$	P_1^{edc}	P_2^{edc}	P_3^{edc}
1	3.4	0	0	0	0	0	0	170	0	0	0
2	10.4	0	1	1	0	400	200	−80	0	320	200
3	16.0	1	1	1	600	400	200	−100	500	400	200
4	6.6	0	0	0	0	0	0	330	0	0	0

$$q(\lambda) = 18{,}344, \quad J^* = 36{,}024, \quad \text{and} \quad \frac{J^* - q^*}{q^*} = 0.965$$

Iteration 4

Hour	λ	u_1	u_2	u_3	P_1	P_2	P_3	$P_{\text{load}}^t - \sum_{i=1}^{N} P_i^t U_i^t$	P_1^{edc}	P_2^{edc}	P_3^{edc}
1	5.1	0	0	0	0	0	0	170	0	0	0
2	10.24	0	1	1	0	400	200	−80	0	320	200
3	15.8	1	1	1	600	400	200	−100	500	400	200
4	9.9	0	1	1	0	380	200	−250	0	130	200

$$q(\lambda) = 19{,}214, \quad J^* = 28{,}906, \quad \text{and} \quad \frac{J^* - q^*}{q^*} = 0.502$$

Iteration 5

Hour	λ	u_1	u_2	u_3	P_1	P_2	P_3	$P_{\text{load}}^t - \sum_{i=1}^{N} P_i^t U_i^t$	P_1^{edc}	P_2^{edc}	P_3^{edc}
1	6.8	0	0	0	0	0	0	170	0	0	0
2	10.08	0	1	1	0	400	200	−80	0	320	200
3	15.6	1	1	1	600	400	200	−100	500	400	200
4	9.4	0	0	1	0	0	200	130	0	0	0

$$q(\lambda) = 19.532, \quad J^* = 36{,}024, \quad \text{and} \quad \frac{J^* - q^*}{q^*} = 0.844$$

www.mepcafe.com

Iteration 6

Hour	λ	u_1	u_2	u_3	P_1	P_2	P_3	$P_{\text{load}}^t - \sum_{i=1}^{N} P_i^t U_i^t$	P_1^{edc}	P_2^{edc}	P_3^{edc}
1	8.5	0	0	1	0	0	200	−30	0	0	170
2	9.92	0	1	1	0	384	200	−64	0	320	200
3	15.4	1	1	1	600	400	200	−100	500	400	200
4	10.7	0	1	1	0	400	200	−270	0	130	200

$$q(\lambda) = 19{,}442, \quad J^* = 20{,}170, \quad \text{and} \quad \frac{J^* - q^*}{q^*} = 0.037$$

The commitment schedule does not change significantly with further iterations, although it is not by any means stable. Further iterations do reduce the duality gap somewhat, but these iterations are unstable in that unit 2 is on the borderline between being committed and not being committed and is switched in and out with no final convergence. After 10 iterations, $q(\lambda) = 19{,}485$, $J^* = 20{,}017$, and $(J^* - q^*)/q^* = 0.027$. This latter value will not go to 0, nor will the solution settle down to a final value; therefore, the algorithm must stop when $(J^* - q^*)/q^*$ is sufficiently small (e.g., less than 0.05 in this case).

4.2.3 Mixed Integer Linear Programming

If we go back and reexamine the equations developed at the beginning of Section 4.2.2 on Lagrange relaxation, we note that there are variables that take a value of only 1 or 0.

Define U_i^t as $U_i^t = 0$ if unit i is off-line during period t

$U_i^t = 1$ if unit i is on-line during period t

The objective function is

$$\sum_{t=1}^{T} \sum_{i=1}^{N_{\text{gen}}} \left[F_i\left(P_i^t\right) + \text{start-up cost}_{i,t} \right] U_i^t = F\left(P_i^t, U_i^t\right)$$

We shall now define several constraints and the objective function of the unit commitment problem:

- Loading constraints:

$$P_{\text{load}}^t - \sum_{i=1}^{N_{\text{gen}}} P_i^t U_i^t = 0 \quad \text{for} \quad t = 1 \ldots T \tag{4.2}$$

- Unit limits:

$$U_i^t P_i^{\min} \le P_i^t \le U_i^t P_i^{\max} \quad \text{for} \quad i = 1 \ldots N_{\text{gen}} \quad \text{and} \quad t = 1 \ldots T \tag{4.3}$$

www.mepcafe.com

Other constraints can also be formulated such as:

- Unit minimum up- and downtime constraints
- Transmission security constraints (see Chapter 7)
- Spinning-reserve constraints
- Generator fuel limit constraints and system air quality constraints in the form of limits on emissions from fossil-fired plants

Since the publishing of the second edition of this text, there have been advances in the field of what we shall here simply refer to as integer programming, that is, programming methods that directly manipulate the integer variables and reach optimal solutions that are better than those reached using Lagrange relaxation. Many of the large market operators in the United States now use these integer programming codes to do the unit commitment calculation.

4.3 SECURITY-CONSTRAINED UNIT COMMITMENT (SCUC)

If one is going to operate the power system according to the schedule of unit commitment output by the UC program, then it is necessary that the economic dispatch within the UC include "security constraints." These are additional constraints that force the dispatch of generation so that no transmission lines are overloaded and that force the dispatch so that no transmission lines are overloaded even when another line is lost. The constraints themselves are dealt with in Chapter 7 on System Security and the incorporation of these constraints into a dispatch algorithm is dealt with in Chapter 8. It is commonplace today to use a security-constrained unit commitment (SCUC) algorithm for all unit commitment calculations—especially for markets where the geographic territory of the market may encompass a large set of transmission lines.

4.4 DAILY AUCTIONS USING A UNIT COMMITMENT

The markets described in Chapter 11 use a unit commitment, usually a SCUC in which bids are placed for each hour, to supply generation from units. The bids may contain multiple segments (see Section 3.2) and the unit commitment is run for 24 h using these bids as input.

APPENDIX 4A
Dual Optimization on a Nonconvex Problem

We introduced the concept of dual optimization in Appendix 3E and pointed out that when the function to be optimized is convex, and the variables are continuous, then the maximization of the dual function gives the identical result as minimizing

www.mepcafe.com

the primal function. Dual optimization is also used in solving the unit commitment problem. However, in the unit commitment problem, there are variables that must be restricted to two values: 1 or 0. These 1–0 variables cause a great deal of trouble and are the reason for the difficulty in solving the unit commitment problem.

The application of the dual optimization technique to the unit commitment problem has been given the name "Lagrange relaxation" and the formulation of the unit commitment problem using this method is shown in the chapter in Section 4.2.2. In this appendix, we illustrate this technique with a simple geometric problem. The problem is structured with 1–0 variables, which makes it clearly nonconvex. Its form is generally similar to the form of the unit commitment problems, but that is incidental for now.

The sample problem to be solved is given next. It illustrates the ability of the dual optimization technique to solve the unit commitment problem. Given:

$$J(x_1, x_2, u_1, u_2) = (0.25x_1^2 + 15)u_1 + (0.255x_2^2 + 15)u_2 \tag{4A.1}$$

subject to

$$\omega = 5 - x_1u_1 - x_2u_2 \tag{4A.2}$$

and

$$0 \le x_1 \le 10 \tag{4A.3}$$

$$0 \le x_2 \le 10 \tag{4A.4}$$

where x_1 and x_2 are continuous real numbers, and

$$u_1 = 1 \text{ or } 0$$
$$u_2 = 1 \text{ or } 0$$

Note that in this problem we have two functions, one in x_1 and the other in x_2. The functions were chosen to demonstrate certain phenomena in a dual optimization. Note that the functions are numerically close and only differ by a small, constant amount. Each of these functions is multiplied by a 1–0 variable and combined into the overall objective function. There is also a constraint that combines the x_1 and x_2 variables again with the 1–0 variables. There are four possible solutions.

1. If u_1 and u_2 are both 0, the problem cannot have a solution since the equality constraint cannot be satisfied.
2. If $u_1=1$ and $u_2=0$, we have the trivial solution that $x_1=5$ and x_2 does not enter into the problem anymore. The objective function is 21.25.
3. If $u_1=0$ and $u_2=1$, then we have the trivial result that $x_2=5$ and x_1 does not enter into the problem. The objective function is 21.375.
4. If $u_1=1$ and $u_2=1$, we have a simple Lagrange function of

$$\mathcal{L}(x_1, x_2, \lambda) = (0.25x_1^2 + 15) + (0.255x_2^2 + 15) + \lambda(5 - x_1 - x_2) \tag{4A.5}$$

www.mepcafe.com

The resulting optimum is at $x_1 = 2.5248$, $x_2 = 2.4752$, and $\lambda = 1.2624$, with an objective function value of 33.1559. Therefore, we know the optimum value for this problem; namely, $u_1 = 1$, $u_2 = 0$, and $x_1 = 5$.

What we have done, of course, is to enumerate all possible combinations of the 1–0 variables and then optimize over the continuous variables. When there are more than a few 1–0 variables, this cannot be done because of the large number of possible combinations. However, there is a systematic way to solve this problem using the dual formulation.

The Lagrange relaxation method solves problems such as the previous one as follows. Define the Lagrange function as

$$\mathcal{L}(x_1, x_2, u_1, u_2, \lambda) = (0.25x_1^2 + 15)u_1 + (0.255x_2^2 + 15)u_2 + \lambda(5 - x_1u_1 - x_2u_2) \quad (4A.6)$$

As shown in Appendix 3E, we define $q(\lambda)$ as

$$q(\lambda) = \min_{x_1, x_2, u_1, u_2} L \quad (4A.7)$$

where x_1, x_2, u_1, u_2 obey the limits and the 1–0 conditions as before. The dual problem is then to find

$$q^*(\lambda) = \max_{\lambda \geq 0} q(\lambda) \quad (4A.8)$$

This is different from the dual optimization approach used in Appendix 3E because of the presence of the 1–0 variables. Because of the presence of the 1–0 variables, we cannot eliminate variables; therefore, we keep all the variables in the problem and proceed in alternating steps as shown in Appendix 3E.

Step 1. Pick a value for λ^k and consider it fixed. Now the Lagrangian function can be minimized. This is much simpler than the situation we had before since we are trying to minimize

$$\left(0.25x_1^2 + 15\right)u_1 + \left(0.255x_2^2 + 15\right)u_2 + \lambda^k\left(5 - x_1u_1 - x_2u_2\right)$$

where the value of λ^k is fixed.

We can then rearrange the preceding equation as

$$(0.25x_1^2 + 15 - x_1\lambda^k)u_1 + (0.255x_2^2 + 15 - x_2\lambda^k)u_2 + \lambda^k 5$$

The last term in the preceding equation is fixed and we can ignore it. The other terms are now given in such a way that the minimization of this function is relatively easy. Note that the minimization is now over two terms, each being multiplied by a 1–0 variable. Since these two terms are summed in the Lagrangian, we can minimize the entire function by minimizing each term separately. Since each term is the product of a function in x and λ (which is fixed) and these are all multiplied by the 1–0 variable u, then the minimum will be 0 (i.e., with $u = 0$) or it will be negative, with $u = 1$ and

www.mepcafe.com

the value of x set so that the term inside the parentheses is negative. Looking at the first term, the optimum value of x_1 is found by (ignore u_1 for a moment):

$$\frac{d}{dx_1}\left(0.25x_1^2+15-x_1\lambda^k\right)=0 \tag{4A.9}$$

If the value of x_1 which satisfies the value stated earlier falls outside the limits of 0 and 10 for x_1, we force x_1 to the limit violated. If the term in the first brackets

$$(0.25x_1^2+15-x_1\lambda^k)$$

is positive, then we can minimize the Lagrangian by merely setting $u_1=0$; otherwise, $u_1=1$.

Looking at the second term, the optimum value of x_2 is found by (again, ignore u_2)

$$\frac{d}{dx_2}\left(0.255x_2^2+15-x_2\lambda^k\right)=0 \tag{4A.10}$$

and if the value of x_2 which satisfies the value stated earlier falls outside the 0–10 limits on x_2, we set it to the violated limit. Similarly, the term in the second brackets

$$(0.255x_2^2+15-x_2\lambda^k)$$

is evaluated. If it is positive, then we minimize the Lagrangian by making $u_2=0$; otherwise, $u_2=1$. We have now found the minimum value of L with a specified fixed value of λ^k.

Step 2 Assume that the variables x_1, x_2, u_1, u_2 found in step 1 are fixed and find a value for λ that maximizes the dual function. In this case, we cannot solve for the maximum since $q(\lambda)$ is unbounded with respect to λ. Instead, we form the gradient of $q(\lambda)$ with respect to λ, and we adjust λ so as to move in the direction of increasing $q(\lambda)$. That is, given

$$\nabla q=\frac{dq}{d\lambda} \tag{4A.11}$$

which for our problem is

$$\frac{dq}{d\lambda}=5-x_1u_1-x_2u_2 \tag{4A.12}$$

we adjust λ according to

$$\lambda^{k+1}=\lambda^k+\frac{dq}{d\lambda}\alpha \tag{4A.13}$$

where α is a multiplier chosen to move λ only a short distance. Note also that if both u_1 and u_2 are 0, the gradient will be 5, indicating a positive value telling us to increase λ. Eventually, increasing λ will result in a negative value for

$$(0.25x_1^2+15-x_1\lambda^k)$$

www.mepcafe.com

or for

$$(0.255x_2^2 + 15 - x_2\lambda^k)$$

or for both, and this will cause u_1 or u_2, or both, to be set to 1. Once the value of λ is increased, we go back to step 1 and find the new values for x_1, x_2, u_1, u_2 again.

The real difficulty here is in not increasing λ by too much. In the example presented earlier, the following scheme was imposed on the adjustment of λ:

- If $\frac{dq}{d\lambda}$ is positive, then use $\alpha=0.2$.
- If $\frac{dq}{d\lambda}$ is negative, then use $\alpha=0.005$.

This lets λ approach the solution slowly, and if it overshoots, it backs up very slowly. This is a common technique to make a gradient "behave."

We must also note that, given the few variables we have, and given the fact that two of them are 1–0 variables, the value of λ will not converge to the value needed to minimize the Lagrangian. In fact, it is seldom possible to find a da^k that will make the problem feasible with respect to the equality constraint. However, when we have found the values for u_1 and u_2 at any iteration, we can then calculate the minimum of $J(x_1,x_2,u_1,u_2)$ by solving for the minimum of

$$\left[(0.25x_1^2 + 15)u_1 + (0.255x_2^2 + 15)u_2 + \lambda(5 - x_1u_1 - x_2u_2)\right]$$

using the techniques in Appendix 3E (since the u_1 and u_2 variables are now known).

The solution to this minimum will be at $x_1 = \overline{x_1}, x_2 = \overline{x_2}$ and $\lambda = \overline{\lambda}$. For the case where u_1 and u_2 are both 0, we shall arbitrarily set this value to a large value (here we set it to 50). We shall call this minimum value $J^*(\overline{x_1}, \overline{x_2}, u_1, u_2)$, and we shall observe that it starts out with a large value, and decreases, while the dual value $q^*(\lambda)$ starts out with a value of 0, and increases. Since there are 1–0 variables in this problem, the primal values and the dual values never become equal. The value $J^* - q^*$ is called the duality gap and we shall call the value

$$\frac{J^* - q^*}{q^*}$$

the relative duality gap.

The presence of the 1–0 variables causes the algorithm to oscillate around a solution with one or more of the 1–0 variables jumping from 1 to 0 to 1, etc. In such cases, the user of the Lagrange relaxation algorithm must stop the algorithm, based on the value of the relative duality gap.

The iterations starting from $\lambda=0$ are shown in Table 4.4. The table shows eight iterations and illustrates the slow approach of λ toward the threshold when both of the 1–0 variables flip from 0 to 1. Also note that ω became negative and the value of λ must now be decreased. Eventually, the optimal solution is reached and the relative duality gap becomes small. However, as is typical with the dual optimization on a

www.mepcafe.com

TABLE 4.4 Dual Optimization on a Sample Problem

Iteration	λ	u_1	u_2	x_1	x_2	q^*	ω	$\bar{\lambda}$	$\bar{x}_1$	$\bar{x}_2$	J^*	$\frac{J^*-q^*}{q^*}$
1	0	0	0	0	0	0	5.0	—	—	—	50.0	—
2	1.0	0	0	2.0	1.9608	5.0	5.0	—	—	—	50.0	9.0
3	2.0	0	0	4.0	3.9216	10.0	5.0	—	—	—	50.0	4.0
4	3.0	0	0	6.0	5.8824	15.0	5.0	—	—	—	50.0	2.33
5	4.0	1	1	8.0	7.8431	18.3137	−10.8431	1.2624	2.5248	2.4752	33.1559	0.8104
6	3.9458	1	1	7.8916	7.7368	18.8958	−10.6284	1.2624	2.5248	2.4752	33.1559	0.7546
7	3.8926	1	0	7.7853	7.6326	19.3105	−2.7853	2.5	5.0	—	21.25	0.1004
8	3.8787	1	0	7.7574	7.6053	19.3491	−2.7574	2.5	5.0	—	21.25	0.0982

www.mepcafe.com

problem with 1–0 variables, the solution is not stable and if iterated further it exhibits further changes in the 1–0 variables as λ is adjusted. Both the q^* and J^* values and the relative duality gap are shown in Table 4.4.

APPENDIX 4B
Dynamic-Programming Solution to Unit Commitment

4B.1 INTRODUCTION

Dynamic programming has many advantages over the enumeration scheme, the chief advantage being a reduction in the dimensionality of the problem. Suppose we have found units in a system and any combination of them could serve the (single) load. There would be a maximum of $2^4 - 1 = 15$ combinations to test. However, if a strict priority order is imposed, there are only four combinations to try:

Priority 1 unit
Priority 1 unit + Priority 2 unit
Priority 1 unit + Priority 2 unit + Priority 3 unit
Priority 1 unit + Priority 2 unit + Priority 3 unit + Priority 4 unit 6pt

The imposition of a priority list arranged in order of the full-load average-cost rate would result in a theoretically correct dispatch and commitment only if:

1. No load costs are 0.
2. Unit input–output characteristics are linear between zero output and full load.
3. There are no other restrictions.
4. Start-up costs are a fixed amount.

In the DP approach that follows, we assume that:

1. A *state* consists of an array of units with specified units operating and the rest off-line.
2. The start-up cost of a unit is independent of the time it has been off-line (i.e., it is a fixed amount).
3. There are no costs for shutting down a unit.
4. There is a strict priority order, and in each interval a specified minimum amount of capacity must be operating.

A feasible state is one in which the committed units can supply the required load and that meets the minimum amount of capacity each period.

www.mepcafe.com

4B.2 FORWARD DP APPROACH

One could set up a DP algorithm to run backward in time starting from the final hour to be studied back to the initial hour. Conversely, one could set up the algorithm to run forward in time from the initial hour to the final hour. The forward approach has distinct advantages in solving generating unit commitment. For example, if the start-up cost of a unit is a function of the time it has been off-line (i.e., its temperature), then a forward DP approach is more suitable since the previous history of the unit can be computed at each stage. There are other practical reasons for going forward. The initial conditions are easily specified and the computations can go forward in time as long as required. A forward DP algorithm is shown by the flowchart in Figure 4.8.

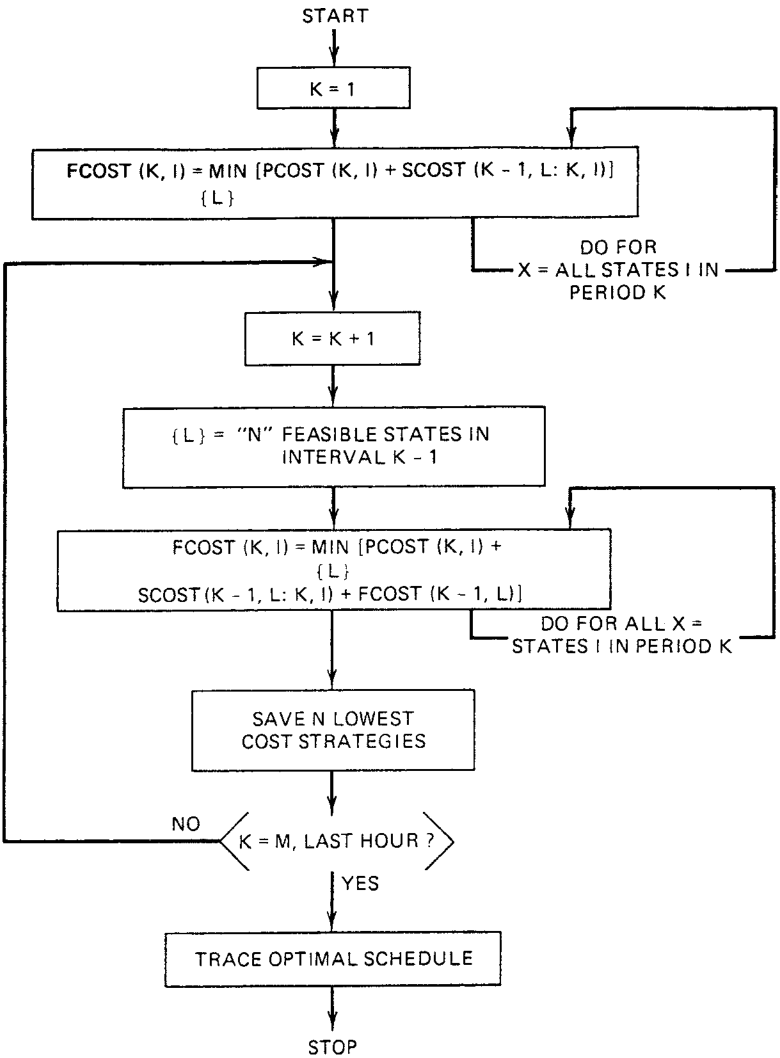


FIGURE 4.8 Unit commitment via forward DP.

www.mepcafe.com

The recursive algorithm to compute the minimum cost in hour K with combination I is

$$F_{\text{cost}}(K,I) = \min_{\{L\}}\left[P_{\text{cost}}(K,I)+S_{\text{cost}}(K-1,L:K,I)+F_{\text{cost}}(K-1,L)\right] \tag{4B.1}$$

where

$$F_{\text{cost}}(K,I) = \text{least total cost to arrive at state } (K,I)$$
$$P_{\text{cost}}(K,I) = \text{production cost for state } (K,I)$$
$$S_{\text{cost}}(K-1,L:K,I) = \text{transition cost from state } (K-1,L) \text{ to state } (K,I)$$

State (K,I) is the Ith combination in hour K. For the forward DP approach, we define a *strategy* as the transition, or path, from one state at a given hour to a state at the next hour.

Note that two new variables, X and N, have been introduced in Figure 4.9:

$$X = \text{number of states to search each period}$$
$$N = \text{number of strategies, or paths, to save at each step}$$

These variables allow control of the computational effort (see Figure 4.9). For complete enumeration, the maximum number of the value of X or N is $2^n - 1$.

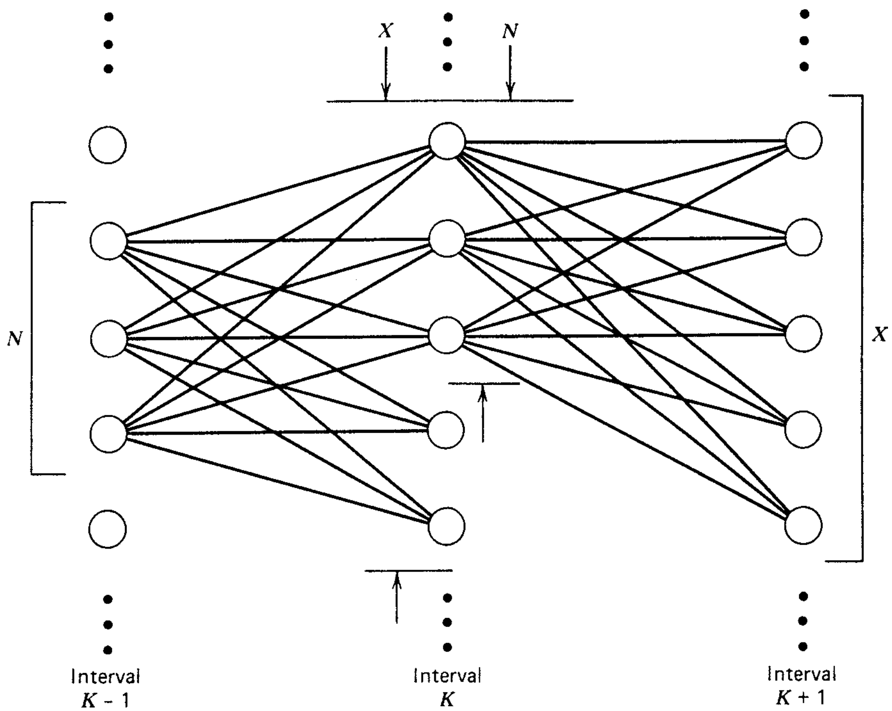


FIGURE 4.9 Restricted search paths in DP algorithm with N=3 and X= 5.

www.mepcafe.com

For example, with a simple priority-list ordering, the upper bound on X is n, the number of units. Reducing the number N means that we are discarding the highest cost schedules at each time interval and saving only the lowest N paths or strategies. There is no assurance that the theoretical optimal schedule will be found using a reduced number of strategies and search range (the X value); only experimentation with a particular program will indicate the potential error associated with limiting the values of X and N below their upper bounds.

Example 4F: For this example, the complete search range will be used and three cases will be studied. The first is a priority-list schedule; the second is the same example with complete enumeration. Both of the first two cases ignore hot-start costs and minimum up- and downtimes. The third case includes the hot-start costs, as well as the minimum up- and downtimes. Four units are to be committed to serve an 8-h load pattern. Data on the units and the load pattern are contained in Table 4.5.

In order to make the required computations more efficiently, a simplified model of the unit characteristics is used. For our example, only a single step between minimum

TABLE 4.5 Unit Characteristics, Load Pattern, and Initial Status for the Cases in Example 4F

Unit	Max (MW)	Min (MW)	Incremental Heat Rate (Btu/kWh)	No-Load Cost ($ / h)	Full-Load Ave. Cost ($/MWh)	Minimum Times (h)	
						Up	Down
1	80	25	10,440	213.00	23.54	4	2
2	250	60	9,000	585.62	20.34	5	3
3	300	75	8,730	684.74	19.74	5	4
4	60	20	11,900	252.00	28.00	1	1

Unit	Initial Conditions	Start-Up Costs	Note: Fuel Cost = 2.0 $/MBtu	
	Hours Off-Line (–) or On-Line (+)	Hot ($)	Cold ($)	Cold Start (h)
1	–5	150	350	4
2	8	170	400	5
3	8	500	1100	5
4	–6	0	0	0

Load Pattern

Hour	**Load (MW)**
1	450
2	530
3	600
4	540
5	400
6	280
7	290
8	500

www.mepcafe.com

and the maximum power points is used as shown in Figure 4.10. In practical applications, two- or three-section stepped incremental curves might be used, as shown in Figure 4.11.

The units in this example have linear $F(P)$ functions:
The $F(P)$ function is

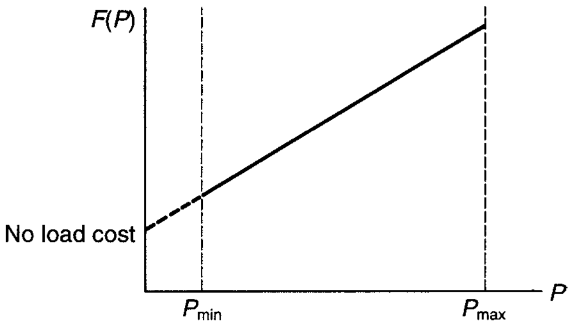


FIGURE 4.10 Single-step cost function used in example.

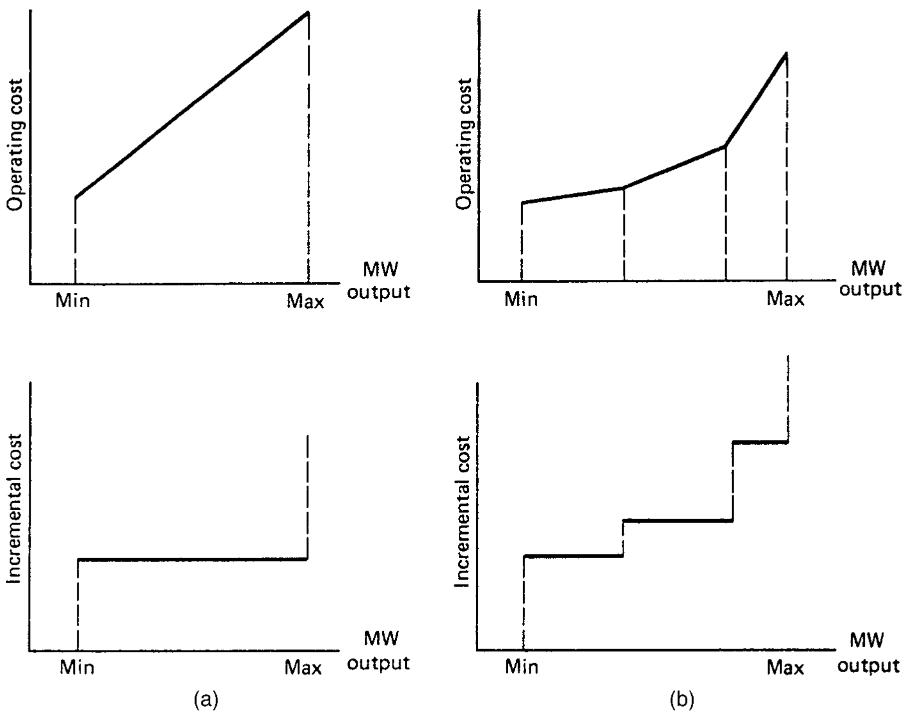


FIGURE 4.11 (**a**) Single-step incremental cost curve and (**b**) multiple-step incremental cost curve.

www.mepcafe.com

TABLE 4.6 Capacity Ordering of the Units

State		Unit Combination[a]				Maximum Net Capacity for Combination
15		1	1	1	1	690
14		1	1	1	0	630
13		0	1	1	1	610
12		0	1	1	0	550
11		1	0	1	1	440
10		1	1	0	1	390
9		1	0	1	0	380
8		0	0	1	1	360
7		1	1	0	0	330
6		0	1	0	1	310
5		0	0	1	0	300
4		0	1	0	0	250
3		1	0	0	1	140
2		1	0	0	0	80
1		0	0	0	1	60
0		0	0	0	0	0
	Unit	1	2	3	4	

[a]1 = Committed (unit operating). 0 = Uncommitted (unit shut down).

$$F(P) = \text{No-load cost} + \text{Inc cost} \times P$$

Note, however, that the unit must operate within its limits. Start-up costs for the first two cases are taken as the cold-start costs. The priority order for the four units in the example is as follows: unit 3, unit 2, unit 1, unit 4. For the first two cases, the minimum up- and downtimes are taken as 1 h for all units.

In all three cases, we will refer to the capacity ordering of the units. This is shown in Table 4.6, where the unit combinations or states are ordered by maximum net capacity for each combination.

Case 1

In Case 1, the units are scheduled according to a strict priority order. That is, units are committed in order until the load is satisfied. The total cost for the interval is the sum of the eight dispatch costs plus the transitional costs for starting any units. In this first case, a maximum of 24 dispatches must be considered.

For Case 1, the only states examined each hour consist of

State No.	Unit Status	Capacity (MW)
5	0 0 1 0	300
12	0 1 1 0	550
14	1 1 1 0	630
15	1 1 1 1	690

www.mepcafe.com

Note that this is the priority order; that is, state 5=unit 3, state 12=units 3+2, state 14=unit 3+2+1, and state 15=units 3+2+1+4. For the first 4 h, only the last three states are of interest. The sample calculations illustrate the technique. All possible commitments start from state 12 since this was given as the initial condition. For hour 1, the minimum cost is state 12. The results for the priority-ordered case are as follows:

Hour	State with Min Total Cost	Pointer for Previous Hour
1	12 (9208)	12
2	12 (19857)	12
3	14 (32472)	12
4	12 (43300)	14
⋮	⋮	⋮

Note that state 13 is not reachable in this strict priority ordering.

Sample Calculations for Case 1

$$F_{\text{cost}}(J,K) = \min_{\{L\}}\left[P_{\text{cost}}(J,K) + S_{\text{cost}}(J-1,L:J,K) + F_{\text{cost}}(J-1,L)\right]$$

Allowable states are

$$\{\,\} = \{0010, 0110, 1110, 1111\} = \{5, 12, 14, 15\}$$

in hour $0\{L\}=\{12\}$, initial condition.

J = 1: **1st hour**

$$\frac{K}{15} \quad F_{\text{cost}}(1,15) = P_{\text{cost}}(1,15) + S_{\text{cost}}(0,12:1,15)$$

$$= 9861 + 350 = 10211$$

$$14 \quad F_{\text{cost}}(1,14) = 9493 + 350 = 9843$$

$$12 \quad F_{\text{cost}}(1,12) = 9208 + 0 = 9208$$

J = 2: **2nd hour**

Feasible states are $\{12,14,15\}=\{K\}$, so X=3. Suppose two strategies are saved at each stage, so N=2, and $\{L\}=\{12,14\}$,

$$\frac{K}{15} \quad F_{\text{cost}}(2,15) = \min_{\{12,14\}}[P_{\text{cost}}(2,15) + S_{\text{cost}}(1,L:2,15) + F_{\text{cost}}(1,L)]$$

$$= 11,301 + \min\begin{bmatrix}(350 + 9,208)\\(0 + 9,843)\end{bmatrix} = 20,859$$

and so on.

www.mepcafe.com

Case 2

In Case 2, complete enumeration is tried with a limit of $(2^4 - 1) = 15$ dispatches each of the 8 h, so that there is a theoretical maximum of $15^8 = 2.56 \times 10^9$ possibilities. Fortunately, most of these are not feasible because they do not supply sufficient capacity and can be discarded with little analysis required.

Figure 4.12 illustrates the computational process for the first 4 h for Case 2. On the figure itself, the circles denote states each hour. The numbers within the circles are the "pointers." That is, they denote the state number in the previous hour that provides the path to that particular state in the current hour. For example, in hour 2, the minimum costs for states 12, 13, 14, and 15 all result from transitions from state 12 in hour 1. Costs shown on the connections are the start-up costs. At each state, the figures shown are the hourly cost/total cost.

In Case 2, the true optimal commitment is found. That is, it is less expensive to turn on the less efficient peaking unit, number 4, for hour 3, than to start up the more efficient unit 1 for that period. By hour 3, the difference in total cost is $l\$165$, or $\$0.104/\text{MWh}$. This is not an insignificant amount when compared with the fuel cost per MWh for an average thermal unit with a net heat rate of 10,000 Btu/kWh and a fuel cost of $\$2.00$ MBtu. A savings of $\$165$ every 3 h is equivalent to $\$481,800$/year.

The total 8-h trajectories for Cases 1 and 2 are shown in Figure 4.13. The neglecting of start-up and shutdown restrictions in these two cases permits the shutting down of all but unit 3 in hours 6 and 7. The only difference in the two trajectories occurs in hour 3, as discussed in the previous paragraph.

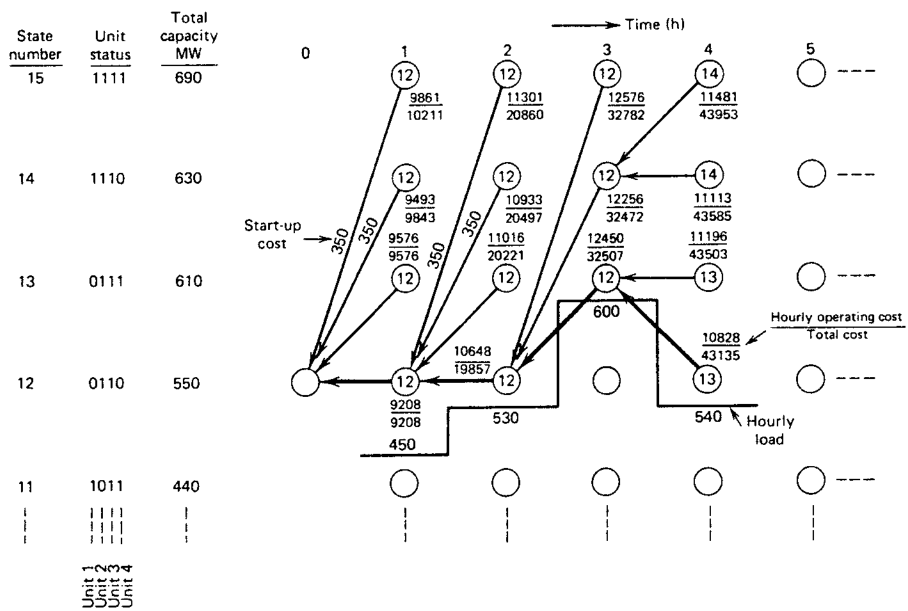


FIGURE 4.12 Example 4F, Cases 1 and 2 (first 4 h).

www.mepcafe.com

Status number	Unit status	Total capacity
15	1111	690
14	1110	630
13	0111	610
12	0110	550
11	1011	440
10	1101	390
9	1010	380
8	0011	360
7	1100	330
6	0101	310
5	0010	300
4	0100	250
3	1001	140
2	1000	80
1	0001	60

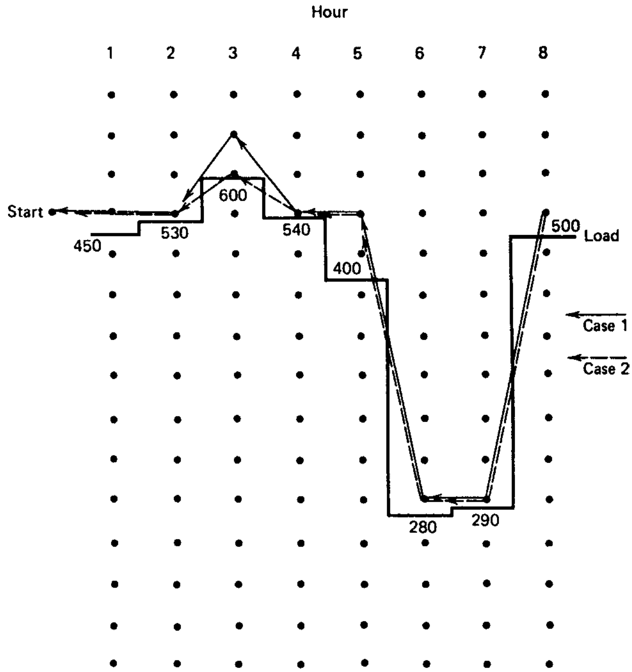


FIGURE 4.13 Example 4F, Cases 1 and 2 (complete solutions).

Case 3

In case 3, the original unit data are used so that the minimum shutdown and operating times are observed. The forward DP algorithm was repeated for the same 8-h period. Complete enumeration was used. That is, the upper bound on X shown in the flowchart was 15. Three different values for N, the number of strategies saved at each stage, were taken as 4, 8, and 10. The same trajectory was found for values of 8 and 10. This trajectory is shown in Figure 4.14. However, when only four strategies were saved, the procedure flounders (i.e., fails to find a feasible path) in hour 8, because the lowest cost strategies in hour 7 have shut down units that cannot be restarted in hour 8 because of minimum unit downtime rules.

The practical remedy for this deficiency in the method shown in Figure 4.13 is to return to a period prior to the low-load hours and temporarily keep more (i.e., higher-cost) strategies. This will permit keeping a nominal number of strategies at each stage. The other alternative is, of course, the method used here: run the entire period with more strategies saved.

These cases can be summarized in terms of the total costs found for the 8-h period, as shown in Table 4.7. These cases illustrate the forward DP method and also point out the problems involved in the practical application of the method.

www.mepcafe.com

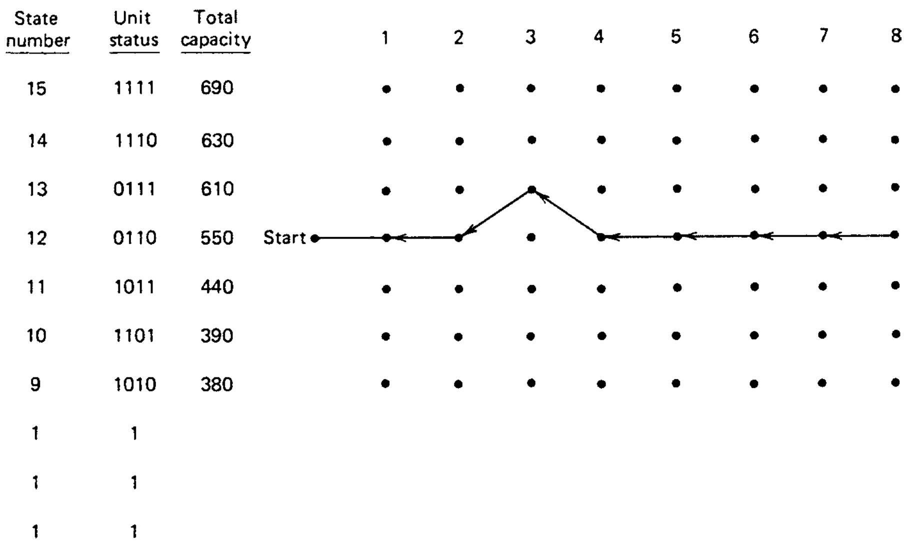


FIGURE 4.14 Example 4F, Case 3.

TABLE 4.7 Summary of Cases 1–3

Case	Conditions	Total Cost ($)
1	Priority order. Up- and downtimes neglected	73,439
2	Enumeration ($X \leq 15$) with 4 strategies (N) saved. Up- and downtimes neglected	73,274
3	$X \leq 15$. Up- and downtimes observed	
	$N = 4$ strategies	No solution
	$N = 8$ strategies	74,110
	$N = 10$ strategies	74,110

PROBLEMS

4.1 Given the unit data in Table 4.8 and Table 4.9, use forward DP to find the optimum unit commitment schedules covering the 8-h period. Table 4.9 gives all the combinations you need, as well as the operating cost for each at the loads in the load data. A "×" indicates that a combination cannot supply the load. The starting conditions are as follows: at the beginning of the first period, units 1 and 2 are up and units 3 and 4 are down and have been down for 8 h.

4.2 Table 4.10 presents the unit characteristics and load pattern for a five-unit, four-time-period problem. Each time period is 2 h long. The input–output characteristics are approximated by a straight line from min to max generation, so that the incremental heat rate is constant. Unit no-load and start-up costs are given in terms of heat energy requirements.

www.mepcafe.com

TABLE 4.8 Unit Commitment Data for Problem 4.1

Unit	Max (MW)	Min (MW)	Incremental Heat Rate (Btu/kWh)	No-Load Energy Input (MBtu/h)	Start-Up Energy (MBtu)
1	500	70	9,950	300	800
2	250	40	10,200	210	380
3	150	30	11,000	120	110
4	150	30	11,000	120	110

Load data (all time periods = 2h):

Time Period	Load (MW)
1	600
2	800
3	700
4	950

Start-up and shutdown rules:

Unit	Minimum Uptime (h)	Minimum Downtime (h)
1	2	2
2	2	2
3	2	4
4	2	4

Fuel cost = 1.00R/MBtu.

www.mepcafe.com

TABLE 4.9 Unit Combinations and Operating Cost for Problem 4.1

					Operating Cost ($ / h)			
					Load	Load	Load	Load
Combination	Unit 1	Unit 2	Unit 3	Unit 4	600 MW	700 MW	800 MW	950 MW
A	1	1	0	0	6,505	7,525	×	×
B	1	1	1	0	6,649	7,669	8,705	×
C	1	1	1	1	6,793	7,813	8,833	10,475

1 = up; 0 = down.

a. Develop the priority list for these units and solve for the optimum unit commitment. Use a strict priority list with a search range of three ($X=3$) and save no more than three strategies ($N=3$). Ignore min uptime/min downtimes for units.

b. Solve the same commitment problem using the strict priority list with $X=3$ and $N=3$ as in part a, but obey the min uptime/min downtime rules.

c. (Optional) Find the optimum unit commitment without use of a strict priority list (i.e., all 32 unit on/off combinations are valid). Restrict the search range to decrease your effort. Obey the min uptime/min downtime rules.

When using a DP method to solve a unit commitment problem with minimum up- and downtime rules, one must save an additional piece of information at each state, each hour. This information simply tells us whether any units are ineligible to be shut down or started up at that state. If such units exist at a particular state, the transition cost, S_{cost}, to a state that violates the start-up/shutdown rules should be given a value of infinity.

4.3 Lagrange Relaxation Problem

Given the three generating units that follows,

$$F_1(P_1) = 30 + 10P_1 + 0.002P_1^2 \quad \text{and} \quad 100 \le P_1 \le 600$$
$$F_2(P_2) = 20 + 8P_2 + 0.0025P_2^2 \quad \text{and} \quad 100 \le P_2 \le 400$$
$$F_3(P_3) = 10 + 6P_3 + 0.005P_3^2 \quad \text{and} \quad 50 \le P_3 \le 200$$

Load:

t	P^t_{load} (MW)
1	300
2	500
3	1100
4	400

No start-up costs; no minimum up- or downtime constraints.

a. Solve for the unit commitment by conventional DP.

b. Set up and carry out four iterations of the Lagrange relaxation method. Let the initial values of λ^t be 0 for $t = 1 \dots 4$.

www.mepcafe.com

TABLE 4.10 The Unit Characteristic and Load Pattern for Problem 4.2

Unit	Max (MW)	Net Full-Load Heat Rate (Btu/kWh)	Incremental Heat Rate (Btu/kWh)	Min (MW)	No-Load Cost (MBtu/h)	Start-up Cost (MBtu)	Min Up/ Downtime (h)
1	200	11,000	9,900	50	220	400	8
2	60	11,433	10,100	15	80	150	8
3	50	12,000	10,800	15	60	105	4
4	40	12,900	11,900	5	40	0	4
5	25	13,500	12,140	5	34	0	4

Load Pattern

Hours	Load (MW)	Conditions
1–2	250	1. Initially (prior to hour 1), only unit 1 is on and has been on for 4 h.
3–4	320	
5–6	110	2. Ignore losses, spinning reserve, etc. The only requirement is that the generation be able to supply the load.
7–8	75	3. Fuel costs for all units may be taken as .1.40 \$/MBtu

www.mepcafe.com

4.4 You are given three generating units and asked to find the optimal unit commitment schedule for the units to supply load over a 4-h time period.

Hour	MW Load
1	400
2	1000
3	1600
4	400

Gen 1: $F_1(P_1) = 2200 + 25P_1 + 0.025 \times P_1^2$ where $220 \le P_1 \le 600\,\text{MW}$

Gen 2: $F_2(P_2) = 1500 + 11P_2 + 0.02 \times P_2^2$ where $350 \le P_2 \le 800\,\text{MW}$

Gen 3: $F_3(P_3) = 1000 + 20P_3 + 0.015 \times P_3^2$ where $150 \le P_3 \le 600$

Each generator has a start-up cost that must be factored in when the unit starts. You should assume that this start-up cost is incurred even if the unit was on then off for only one or two hours and then gets turned back on.

Generating Unit	Start-Up Cost
1	2400
2	3000
3	3400

Find the optimal unit commitment schedule.

www.mepcafe.com

5

GENERATION WITH LIMITED ENERGY SUPPLY

5.1 INTRODUCTION

This chapter removes the assumption of economic dispatch and unit commitment that fuel is not limited. This chapter also deals with water restrictions for hydro plants. Procurement of fuels is through markets with some risk of volume or price liquidity. Other risk factors are also present. Fuel suppliers wish to spread this risk to the buyers as do buyers wish to spread risk to the sellers. Many fuel contracts are based on the seller's need to have a guarantee that resources gathered are sold without storage fees or if there is no storage available. Hydro facilities are often limited to the storage limitations of the reservoir, flow limitations for navigation or for irrigation, or level change limitations for safety restrictions. Wind and solar resources are of a more immediate nature as the resource may not be present due to weather conditions. Since these are more due to weather uncertainties than fuel harvesting restrictions, they are treated in a different fashion. Pumped hydro is a storage resource that is linked to wind and solar generation to provide a continuity of service and/or quality of service.

Of greatest concern when scheduling generation with limited fuel or water are those situations when there is limited fuel or water and the generator cannot be operated at full output for all periods of time. This kind of problem is common to the operation of real power systems, and not only do we need to be concerned with the generator's use of fuel or water but we need to be concerned with things like storage of fuel, fuel delivery by rail or river barge, and in the case of hydroplants, we must be concerned with things like the level of water in rivers and reservoirs and how much water is flowing in downstream facilities. This becomes especially difficult with

Power Generation, Operation, and Control, Third Edition. Allen J. Wood, Bruce F. Wollenberg, and Gerald B. Sheblé.
© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

www.mepcafe.com

hydro facilities that are coupled so that water flowing out of one plant then makes its way into the reservoir of another plant further downstream.

5.2 FUEL SCHEDULING

The economic operation of a power system requires that expenditures for fuel be minimized over all periods of time. When there is no limitation on the fuel supply to any of the plants in the system, the economic dispatch can be carried out with only the present conditions as data in the economic dispatch algorithm. In such a case, the fuel costs are simply the incoming price of fuel with, perhaps, adjustments for fuel handling and maintenance of the plant.

When the energy resource available to a particular plant (be it coal, oil, gas, water, or nuclear fuel) is a limiting factor in the operation of the plant, the entire economic dispatch calculation must be done differently. Each economic dispatch calculation must account for what happened before and what is expected to happen in the future.

This chapter begins the development of solutions to the dispatching problem "over time." The techniques used are an extension of the familiar Lagrange formulation. Concepts involving slack variables and penalty functions are introduced to allow solution under certain conditions.

The example chosen to start with is a fixed fuel supply that must be paid for, whether or not it is consumed. We might have started with a limited fuel supply of natural gas that must be used as boiler fuel because it has been declared as "surplus." The take-or-pay fuel supply contract is probably the simplest of these possibilities.

Alternatively, we might have started directly with the problem of economic scheduling of hydroelectric plants with their stored supply of water or with light-water-moderated nuclear reactors supplying steam to drive turbine generators. Hydroelectric plant scheduling involves the scheduling of water flows, impoundments (storage), and releases into what usually prove to be a rather complicated hydraulic network (viz., the watershed). The treatment of nuclear fuel scheduling requires a longer term horizon than the fuel or hydro scheduling covered in this chapter. Such maintenance scheduling has to include the uncertainties of demand, of fuel prices, and of interchange contracts. Such treatment is beyond the scope of this book.

5.3 TAKE-OR-PAY FUEL SUPPLY CONTRACT

Assume there are N normally fueled thermal plants plus one turbine generator, fueled under a "*take-or-pay*" *agreement*. We will interpret this type of agreement as being one in which the utility agrees to use a contracted amount of fuel during a period (the "take") and it agrees to pay the minimum charge. This last clause is the "pay" part of the "take-or-pay" contract.

While this unit's cumulative fuel consumption is below the minimum, the system excluding this unit should be scheduled to minimize the total fuel cost, subject to the constraint that the total fuel consumption for the period for this particular unit is equal to

www.mepcafe.com

the specified amount. Once the specified amount of fuel has been used, the unit should be scheduled normally. Let us consider a special case where the minimum amount of fuel consumption is also the maximum. The system is shown in Figure 5.1. We will consider the operation of the system over j_{max} time intervals j where $j=1,\ldots,j_{max}$, so that

$$P_{1j}, P_{2j}, \ldots, P_{Tj} \quad \text{(power outputs)}$$
$$F_{1j}, F_{2j}, \ldots, F_{Nj} \quad \text{(fuel cost rate)}$$

and

$$q_{T1}, q_{T2}, \ldots, q_{Tj} \quad \text{(take-or-pay fuel input)}$$

are the power outputs, fuel costs, and take-or-pay fuel inputs, where

$$P_{ij} \triangleq \text{power from } i\text{th unit in the } j\text{th time interval}$$
$$F_{ij} \triangleq \$/\text{h cost for } i\text{th unit during the } j\text{th time interval}$$
$$q_{Tj} \triangleq \text{fuel input for unit T in } j\text{th time interval}$$
$$F_{Tj} \triangleq \$/\text{h cost for unit T in } j\text{th time interval}$$
$$P_{\text{load}j} \triangleq \text{total load in the } j\text{th time interval}$$
$$n_j \triangleq \text{Number of hours in the } j\text{th time interval}$$

Mathematically, the problem is as follows:

$$\min \sum_{j=1}^{j_{max}} \left(n_j \sum_{i=1}^{N} F_{ij} \right) + \sum_{j=1}^{j_{max}} n_j F_{Tj} \tag{5.1}$$

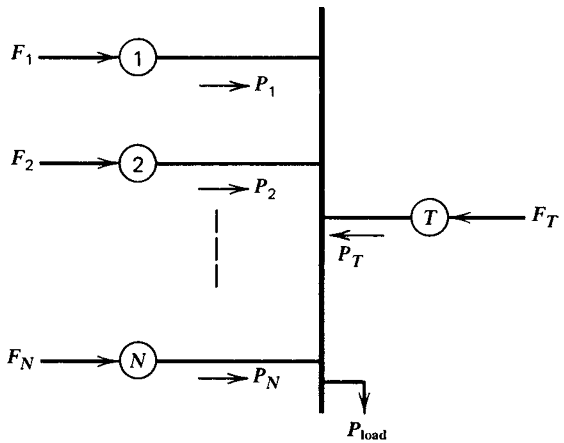


FIGURE 5.1 $N+1$ unit system with take-or-pay fuel supply at unit T.

www.mepcafe.com

subject to

$$\phi = \sum_{j=1}^{j_{\max}} n_j q_{\mathrm{T}j} - q_{\mathrm{TOT}} = 0 \tag{5.2}$$

and

$$\psi_j = P_{\mathrm{load}j} - \sum_{i=1}^{N} P_{ij} - P_{\mathrm{T}j} = 0 \quad \text{for} \quad j = 1 \ldots j_{\max} \tag{5.3}$$

or, in words,

We wish to determine the minimum production cost for units 1 to N subject to constraints that ensure that fuel consumption is as contracted and also subject to the set of constraints to ensure that power supplied is within limits each interval.

Note that (for the present) we are ignoring high and low limits on the units themselves. It should also be noted that the term

$$\sum_{j=1}^{j_{\max}} n_j F_{\mathrm{T}j}$$

is constant because the total fuel to be used in the "T" plant is fixed. Therefore, the total cost of that fuel will be constant, and we can drop this term from the objective function.

The Lagrange function is

$$\mathcal{L} = \sum_{j=1}^{j_{\max}} n_j \sum_{i=1}^{N} F_{ij} + \sum_{j=1}^{j_{\max}} \lambda_j \left(P_{\mathrm{load}j} - \sum_{i=1}^{N} P_{ij} - P_{\mathrm{T}j} \right) + \gamma \left(\sum_{j=1}^{j_{\max}} n_j q_{\mathrm{T}j} - q_{\mathrm{TOT}} \right) \tag{5.4}$$

The independent variables are the powers P_{ij} and $P_{\mathrm{T}j}$, since $F_{ij} = F_i(P_{ij})$ and $q_{\mathrm{T}j} = q_{\mathrm{T}}(P_{\mathrm{T}j})$. For any given time period, $j = k$,

$$\frac{\partial \mathcal{L}}{\partial P_{ik}} = 0 = n_k \frac{\mathrm{d}F_{ik}}{\mathrm{d}P_{ik}} - \lambda_k \quad \text{for} \quad i = 1 \ldots N \tag{5.5}$$

and

$$\frac{\partial \mathcal{L}}{\partial P_{\mathrm{T}k}} = -\lambda_k + \gamma n_k \frac{\mathrm{d}q_{\mathrm{T}k}}{\mathrm{d}P_{\mathrm{T}k}} = 0 \tag{5.6}$$

www.mepcafe.com

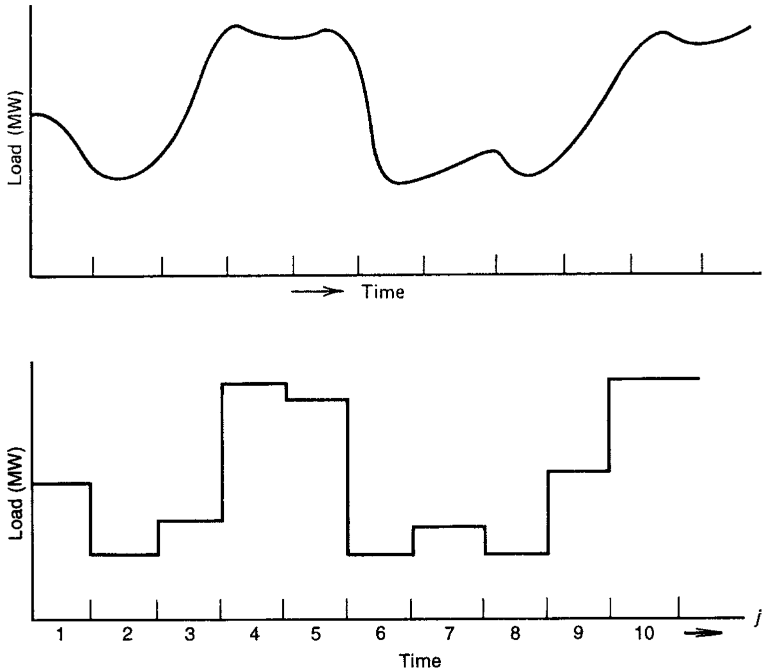


FIGURE 5.2 Conversion of load pattern to discrete loads at each hour.

Note that if one analyzes the dimensions of γ, it would be \$ per unit of q (e.g., \$/ft^3, \$/bbl, \$/t). As such, γ has the units of a "fuel price" expressed in volume units rather than MBtu as we have used up to now. Because of this, γ is often referred to as a "pseudo-price" or "shadow price." In fact, once it is realized what is happening in this analysis, it becomes obvious that we could solve fuel-limited dispatch problems by simply adjusting the price of the limited fuel(s); thus, the terms "pseudo-price" and "shadow price" are quite meaningful.

Since γ appears unsubscripted in Equation 5.6, γ would be expected to be a constant value over all the time periods. This is true unless the fuel-limited machine is constrained by fuel storage limitations. We will encounter such limitations in hydroplant scheduling later in this chapter.

Figure 5.2 (upper part) shows how the load pattern may look. The solution to a fuel-limited dispatching problem will require dividing the load pattern into time intervals, as in Figure 5.2 (lower part), and assuming load to be constant during each interval. Assuming all units are on-line for the period, the optimum dispatch could be done using a simple search procedure for γ, as is shown in Figure 5.3. Note that the procedure shown in Figure 5.3 will only work if the fuel-limited unit does not hit either its high or its low limit in any time interval.

www.mepcafe.com

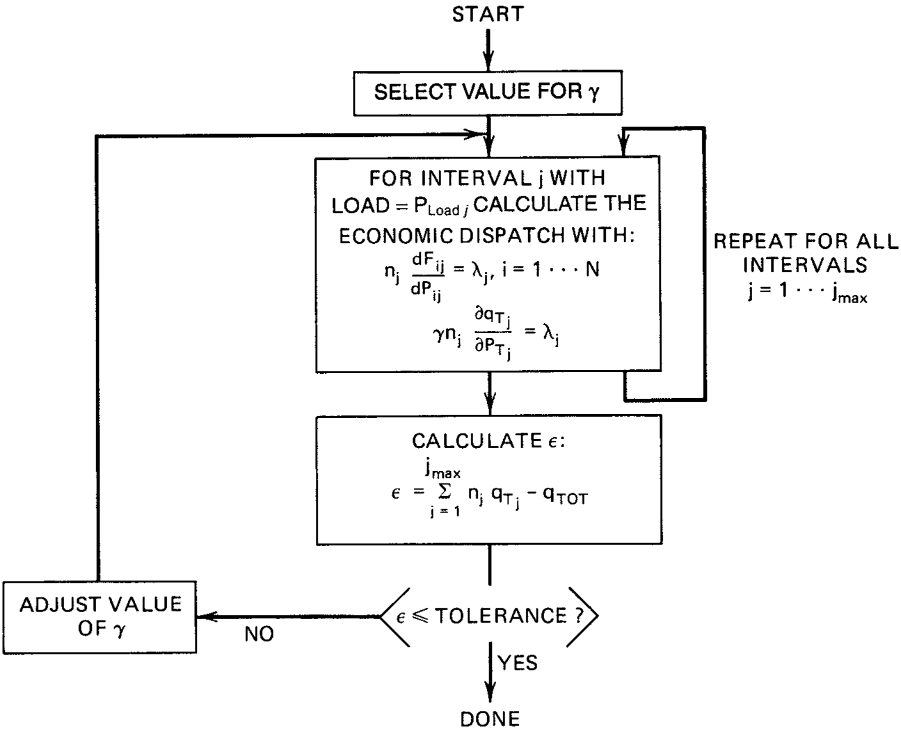


FIGURE 5.3 Gamma search method.

Example 5A: Find the optimal dispatch for a gas-fired steam plant given the following:

Gas-fired plant:

$$\begin{aligned}
&H_T(P_T) = 300 + 6.0P_T + 0.0025P_T^2 \text{ MBtu / h}\\
&\text{Fuel cost for gas} = 2.0\,\$/\text{ccf(where 1 ccf} = 10^3 \text{ ft}^3)\\
&\text{The gas is rated at 1100 Btu / ft}^3\\
&50 \le P_T \le 400
\end{aligned} \tag{5.7}$$

Composite of remaining units:

$$\begin{aligned}
&H_s(P_s) = 200 + 8.5P_s + 0.002P_s^2 \text{ MBtu/h}\\
&\text{Equivalent fuel cost} = 0.6\,\$/\text{MBtu}\\
&50 \le P_s \le 500
\end{aligned} \tag{5.8}$$

The gas-fired plant must burn 40×10^6 ft^3 of gas. The load pattern is shown in Table 5.1. If the gas constraints are ignored, the optimum economic schedule for these two plants appears as is shown in Table 5.2. Operating cost of the composite unit over the

www.mepcafe.com

TABLE 5.1 Load Pattern

Time Period	Load
1. 0000–0400	400 MW
2. 0400–0800	650 MW
3. 0800–1200	800 MW
4. 1200–1600	500 MW
5. 1600–2000	200 MW
6. 2000–2400	300 MW

Where: $n_j = 4, j = 1 \ldots 6$.

TABLE 5.2 Optimum Economic Schedule (Gas Constraints Ignored)

Time Period	P_s	P_T
1	350	50
2	500	150
3	500	300
4	450	50
5	150	50
6	250	50

TABLE 5.3 Optimal Schedule (Gas Constraints Met)

Time Period	P_s	P_T
1	197.3	202.6
2	353.2	296.8
3	446.7	353.3
4	259.7	240.3
5	72.6	127.4
6	135.0	165.0

entire 24-h period is 52,128.03 \$. The total gas consumption is 21.8×10^6 ft^3. Since the gas-fired plant must burn 40×10^6 ft^3 of gas, the cost will be $2.0\ \$/1000\ \text{ft}^3 \times 40.10^6\ \text{ft}^3$, which is 80,000 \$ for the gas. Therefore, the total cost will be 132,128.03 \$. The solution method shown in Figure 5.3 was used with γ values ranging from 0.500 to 0.875. The final value for γ is 0.8742 \$/ccf with an optimal schedule as shown in Table 5.3. This schedule has a fuel cost for the composite unit of 34,937.47 \$. Note that the gas unit is run much harder and that it does not hit either limit in the optimal schedule. Further, note that the total cost is now

$$34{,}937.47\,\$ + 80{,}000\,\$ = 114{,}937.4\,\$$$

so we have lowered the total fuel expense by properly scheduling the gas plant.

www.mepcafe.com

5.4 COMPLEX TAKE-OR-PAY FUEL SUPPLY MODELS

5.4.1 Hard Limits and Slack Variables

This section takes account of hard limits on the take-or-pay generating unit. The limits are

$$P_{\mathrm{T}} \geq P_{\mathrm{T\,min}} \tag{5.9}$$

and

$$P_{\mathrm{T}} \leq P_{\mathrm{T\,max}} \tag{5.10}$$

These may be added to the Lagrangian by the use of two constraint functions and two new variables called *slack variables* (see Appendix 3A). The constraint functions are

$$\psi_{1j} = P_{\mathrm{T}j} - P_{\mathrm{T\,max}} + S_{1j}^2 \tag{5.11}$$

and

$$\psi_{2j} = P_{\mathrm{T\,min}} - P_{\mathrm{T}j} + S_{2j}^2 \tag{5.12}$$

where S_{1j} and S_{2j} are slack variables that may take on any real value including 0.

The new Lagrangian then becomes

$$\begin{aligned} \mathcal{L} = & \sum_{j=1}^{j_{\max}} n_j \sum_{i=1}^{N} F_{ij} + \sum_{j=1}^{j_{\max}} \lambda_j \left(P_{\mathrm{load}j} - \sum_{i=1}^{N} P_{ij} - P_{Tj} \right) + \gamma \left(\sum_{j=1}^{j_{\max}} n_j q_{\mathrm{T}j} - q_{\mathrm{TOT}} \right) r \\ & + \sum_{j=1}^{j_{\max}} \alpha_{1j} \left(P_{\mathrm{T}j} - P_{\mathrm{T\,max}} + S_{1j}^2 \right) + \sum_{j=1}^{j_{\max}} \alpha_{2j} \left(P_{\mathrm{T\,min}} - P_{Tj} + S_{2j}^2 \right) \end{aligned} \tag{5.13}$$

where α_{1j}, α_{2j} are Lagrange multipliers. Now, the first partial derivatives for the kth period are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial P_{ik}} &= 0 = n_k \frac{\mathrm{d}F_{ik}}{\mathrm{d}P_{ik}} - \lambda_k \\ \frac{\partial \mathcal{L}}{\partial P_{\mathrm{T}_k}} &= 0 = -\lambda_k + \alpha_{1k} - \alpha_{2k} + \gamma n_k \frac{\mathrm{d}q_{\mathrm{T}_k}}{\mathrm{d}P_{\mathrm{T}_k}} \\ \frac{\partial \mathcal{L}}{\partial S_{1k}} &= 0 = 2\alpha_{1k} S_{1k} \\ \frac{\partial \mathcal{L}}{\partial S_{2k}} &= 0 = 2\alpha_{2k} S_{2k} \end{aligned} \tag{5.14}$$

As we noted in Appendix 3A, when the constrained variable (P_{T_k} in this case) is within bounds, the new Lagrange multipliers $\alpha_{1k} = \alpha_{2k} = 0$ and S_{1k} and S_{2k} are nonzero.

www.mepcafe.com

TABLE 5.4 Resulting Optimal Schedule with P_{Tmax} = 300 MW

Time Period j	P_{sj}	P_{Tj}	λ_j	$\gamma_{nj}\frac{\partial q_T}{\partial P_{Tj}}$	α_{1j}
1	183.4	216.6	5.54	5.54	0
2	350.0	300.0	5.94	5.86	0.08
3	500.0	300.0	6.3	5.86	0.44
4	245.4	254.6	5.69	5.69	0
5	59.5	140.5	5.24	5.24	0
6	121.4	178.6	5.39	5.39	0

When the variable is limited, one of the slack variables, S_{1k} or S_{2k}, becomes 0, and the associated Lagrange multiplier will take on a nonzero value.

Suppose in some interval k, $P_{Tk} = P_{max}$, then $S_{1k} = 0$ and $\alpha_{1k} \neq 0$. Thus,

$$-\lambda_k + \alpha_{1k} + \gamma n_k \frac{dq_{T_k}}{dP_{T_k}} = 0 \tag{5.15}$$

and if

$$\lambda_k > \gamma n_k \frac{dq_{T_k}}{dP_{T_k}}$$

the value of α_{1k} will take on the value just sufficient to make the equality true.

Example 5B: Repeat Example 5A with the maximum generation on P_T reduced to 300 MW. Note that the optimum schedule in Example 5A gave a $P_T = 353.3$ MW in the third time period. When the limit is reduced to 300 MW, the gas-fired unit will have to burn more fuel in other time periods to meet the 40×10^3 ft^3 gas consumption constraint.

Table 5.4 shows the resulting optimal schedule where $\gamma = 0.8603$ and total cost = 122,984.83 \$.

5.5 FUEL SCHEDULING BY LINEAR PROGRAMMING

Figure 5.4 shows the major elements in the chain making up the delivery system that starts with raw-fuel suppliers and ends up in delivery of electric power to individual customers. The basic elements of the chain are as follows:

The suppliers: These are the coal, oil, and gas companies with which the utility must negotiate contracts to acquire fuel. The contracts are usually written for a long term (10–20 years) and may have stipulations, such as the minimum and maximum limits on the quantity of fuel delivered over a specified time period. The time period may be as long as a year, a month, a week, a day, or even for a period of only a few minutes. Prices may change, subject to the renegotiation provisions of the contracts.

www.mepcafe.com

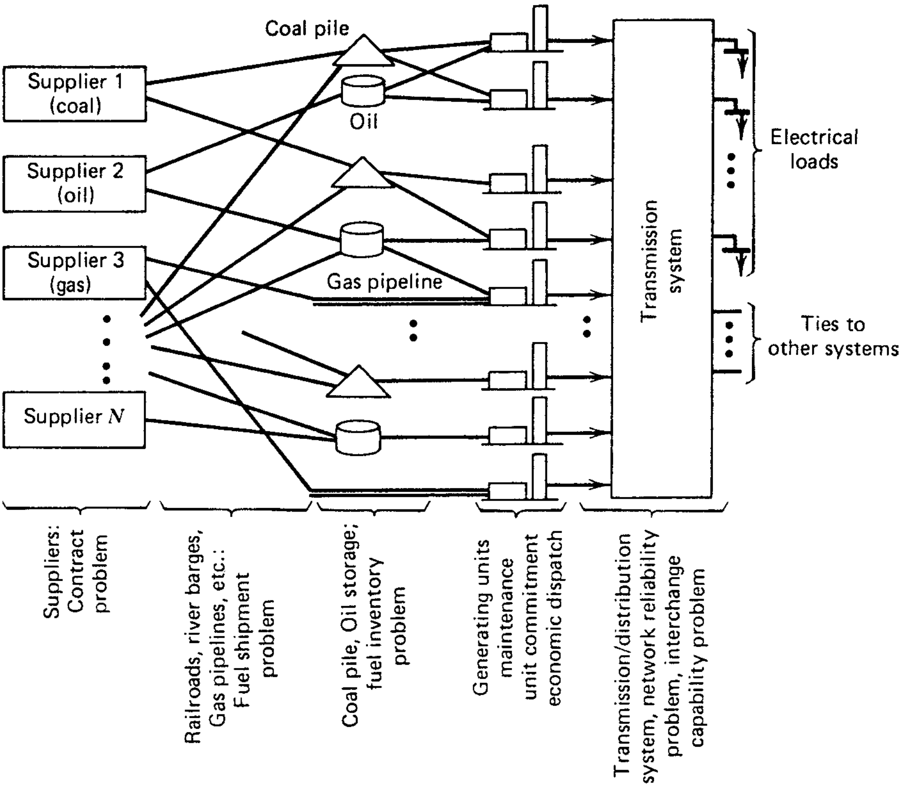


FIGURE 5.4 Energy delivery system.

Transportation: Railroads, unit trains, river barges, gas pipeline companies, and such, all present problems in scheduling of deliveries of fuel.

Inventory: Coal piles, oil storage tanks, and underground gas storage facilities. Inventories must be kept at proper levels to forestall fuel shortages when load levels exceed forecast or suppliers or shippers are unable to deliver. Price fluctuations also complicate the decisions on when and how much to add or subtract from inventories.

The remainder of the system—generators, transmission, and loads—is covered in other chapters.

One of the most useful tools for solving large fuel-scheduling problems is linear programming (LP). If the reader is not familiar with LP, an easily understood algorithm is provided in Appendix 3B.

Linear programming is an optimization procedure that minimizes a linear objective function with variables that are also subject to linear constraints. Because of this limitation, any nonlinear functions either in the objective or in the constraint equations will have to be approximated by linear or piecewise linear functions.

To solve a fuel-scheduling problem with LP, we must break the total time period involved into discrete time increments, as was done in earlier and shown in

www.mepcafe.com

Figure 5.2. The LP solution will then consist of an objective function that is made up of a sum of linear or piecewise linear functions, each of which is a function of one or more variables from only one time step. The constraints will be linear functions of variables from each time step. Some constraints will be made up of variables drawn from onetime step, whereas others will span two or more time steps. The best way to illustrate how to set up an LP to solve a fuel-scheduling problem will be to use an example.

Example 5C: We are given two coal-burning generating units that must both remain on-line for a 3-week period. The combined output from the two units is to supply the following loads (loads are assumed constant for 1 week).

Week	Load (MW)
1	1200
2	1500
3	800

The two units are to be supplied by one coal supplier who is under contract to supply 40,000 t of coal per week to the two plants. The plants have existing coal inventories at the start of the 3-week period. We must solve for the following:

1. How should each plant be operated each week?
2. How should the coal deliveries be made up each week?

The data for the problem are as follows:

Coal: Heat value = 11,500 Btu/lb = 23 MBtu/t (1 t = 2000 lb)

Coal can all be delivered to one plant or the other or it can be split, some going to one plant, some to the other, as long as the total delivery in each week is equal to 40,000 t. The coal costs 30 \$/t or 1.3 \$/MBtu.

Inventories: Plant 1 Has an initial inventory of 70,000 t; its final inventory is not restricted
Plant 2 Has an initial inventory of 70,000 t; its final inventory is not restricted

Both plants have a maximum coal storage capacity of 200,000 t of coal.

Generating units:
The input versus output function will be approximated by a linear function for each unit:

$$H_1(P_1) = 380.0 + 8.267P_1$$
$$H_2(P_2) = 583.3 + 8.167P_2$$

www.mepcafe.com

Unit	Min (MW)	Max (MW)	Heat Input at Min (MBtu/h)	Heat Input at Max (MBtu/h)
1	150	600	1620	5340
2	400	1000	3850	8750

The unit cost curves are

$$F_1(P_1) = 1.3\$/\text{MBtu} \times H_1(P_1) = 495.65 + 10.78P_1 (\$/\text{h})$$
$$F_2(P_2) = 1.3\$/\text{MBtu} \times H_2(P_2) = 760.8 + 10.65P_2 (\$/\text{h})$$

The coal consumption q (t/h) for each unit is

$$q_1(P_1) = \frac{1}{23}\left(\frac{\text{t}}{\text{MBtu}}\right) \times H_1(P_1) = 16.52 + 0.3594P_1 \text{ t/h}$$
$$q_2(P_2) = \frac{1}{23}\left(\frac{\text{t}}{\text{MBtu}}\right) \times H_2(P_2) = 25.36 + 0.3551P_2 \text{ t/h}$$

To solve this problem with LP, assume that the units are to be operated at a constant rate during each week and that the coal deliveries will each take place at the beginning of each week. Therefore, we will set up the problem with 1-week time periods, and the generating unit cost functions and coal consumption functions will be multiplied by 168 h to put them on a “per week” basis; then,

$$\begin{aligned} F_1(P_1) &= 83{,}269.2 + 1811P_1 \ \$/\text{week} \\ F_2(P_2) &= 127{,}814.4 + 1789P_2 \ \$/\text{week} \\ q_1(P_1) &= 2775.4 + 60.4P_1 \ \text{t/week} \\ q_2(P_2) &= 4260.5 + 59.7P_2 \ \text{t/week} \end{aligned} \tag{5.16}$$

We are now ready to set up the objective function and the constraints for our LP solution.

Objective function: To minimize the operating cost over the 3-week period. The objective function is

$$\text{Minimize } Z = F_1[P_1(1)] + F_2[P_2(1)] + F_1[(P_1(2)] + F_2[P_2(2)] + F_1[P_1(3)] + F_2[P_2(3)] \tag{5.17}$$

where $P_i(j)$ is the power output of the ith unit during the jth week, $j = 1 \ldots 3$.

Constraints: During each time period, the total power delivered from the units must equal the scheduled load to be supplied; then

$$\begin{aligned} P_1(1) + P_2(1) &= 1200 \\ P_1(2) + P_2(2) &= 1500 \\ P_1(3) + P_2(3) &= 800 \end{aligned} \tag{5.18}$$

www.mepcafe.com

Similarly, the coal deliveries, D_1 and D_2, made to plant 1 and plant 2, respectively, during each week must sum to 40,000 t; then

$$\begin{aligned} D_1(1) + D_2(1) &= 40{,}000 \\ D_1(2) + D_2(2) &= 40{,}000 \\ D_1(3) + D_2(3) &= 40{,}000 \end{aligned} \tag{5.19}$$

The volume of coal at each plant at the beginning of each week plus the delivery of coal to that plant minus the coal burned at the plant will give the coal remaining at the beginning of the next week. Letting V_1 and V_2 be the volume of coal in each coal pile at the beginning of the week, respectively, we have the following set of equations governing the two coal piles:

$$\begin{aligned} V_1(1) + D_1(1) - q_1(1) &= V_1(2) \\ V_2(1) + D_2(1) - q_2(1) &= V_2(2) \\ V_1(2) + D_1(2) - q_1(2) &= V_1(3) \\ V_2(2) + D_2(2) - q_2(2) &= V_2(3) \\ V_1(3) + D_1(3) - q_1(3) &= V_1(4) \\ V_2(3) + D_2(3) - q_2(3) &= V_2(4) \end{aligned} \tag{5.20}$$

where $V_i(j)$ is the volume of coal in the ith coal pile at the beginning of the jth week.

To set these equations up for the LP solutions, substitute the $q_1(P_1)$ and $q_2(P_2)$ equations from 5.16 into the equations of 5.20. In addition, all constant terms are placed on the right of the equal sign and all variable terms on the left; this leaves the constraints in the standard form for inclusion in the LP. The result is

$$\begin{aligned} D_1(1) - 60.4P_1(1) - V_1(2) &= 2775.4 - V_1(1) \\ D_2(1) - 59.7P_2(1) - V_2(2) &= 4260.5 - V_2(1) \\ V_1(2) + D_1(2) - 60.4P_1(2) - V_1(3) &= 2775.4 \\ V_2(2) + D_2(2) - 59.7P_2(2) - V_2(3) &= 4260.5 \\ V_1(3) + D_1(3) - 60.4P_1(3) - V_1(4) &= 2775.4 \\ V_2(3) + D_2(3) - 59.7P_2(3) - V_2(4) &= 4260.5 \end{aligned} \tag{5.21}$$

Note: $V_1(1)$ and $V_2(1)$ are constants that will be set when we start the problem.

The constraints from Equations 5.18, 5.19, and 5.21 are arranged in a matrix, as shown in Figure 5.5. Each variable is given an upper and lower bound, and we shall assume here that an LP capable of directly accepting vectors of upper and lower variable bounds is in use. The $P_1(t)$ and $P_2(t)$ variables are given the upper and lower bounds corresponding to the upper and lower limits on the generating units. $D_1(t)$ and $D_2(t)$ are given upper and lower bounds of 40,000 and 0. $V_1(t)$ and $V_2(t)$ are given upper and lower bounds of 200,000 and 0.

www.mepcafe.com

Problem Variable	D1(1)	P1(1)	D2(1)	P2(1)	V1(2)	D1(2)	P1(2)	V2(2)	D2(2)	P2(2)	V1(3)	D1(3)	P1(3)	V2(3)	D2(3)	P2(3)	V1(4)	V2(4)	
LP Variable	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	Constraint Units
Constraint																			
1		1		1															1200
2	1		1																40000
3	1	− 60.4			−1														2775.4 − $V_1(1)$
4			1	−59.7				−1											4260.5 − $V_2(1)$
5							1			1									1500
6						1			1										40000
7					1	1	−60.4				−1								2775.4
8								1	1	−59.7				−1					4260.5
9													1			1			800
10												1			1				40000
11											1	1	−60.4				−1		2775.4
12														1	1	−59.7		−1	4260.5
Variable min.	0	150	0	400	0	0	150	0	0	400	0	0	150	0	0	400	0	0	
Variable max.	40000	600	40000	1000	200000	40000	600	200000	40000	1000	200000	40000	600	200000	40000	1000	200000	200000	
	Week 1				Week 2						Week 3						Final Conditions		

FIGURE 5.5 LP constraint matrix for Example 5C.

www.mepcafe.com

Time Period	V_1	D_1	P_1	V_2	D_2	P_2
1	70000.0	0	200	70000.0	40000.0	1000
2	55144.6	0	500	46039.5	40000.0	1000
3	22169.2	19013.5	150	22079.0	20986.5	650
4	29347.3					

$$\text{Optimum cost} = 6{,}913{,}450.8\$.$$

Solution: The solution to this problem was carried out with a computer program capable of incorporating upper and lower bounds on all variables. The first problem solved had coal storage at the beginning of the first week of

$$V_1(1) = 70{,}000 \text{ t}$$
$$V_2(1) = 70{,}000 \text{ t}$$

The solution is:

In this case, there are no constraints on the coal deliveries to either plant, and the system can run in the most economic manner. Since unit 2 has a lower incremental cost, it is run at its maximum when possible. Furthermore, since no restrictions were placed on the coal pile levels at the end of the third week, the coal deliveries could have been shifted a little from unit 2 to unit 1 with no effect on the generation dispatch.

The next case solved was purposely structured to create a fuel shortage at unit 2. The beginning inventory at plant 2 was set to 50,000 t, and a requirement was imposed that at the end of the third week, the coal pile at unit 2 be no less than 8000 t. The solution was made by changing the right-hand side of the fourth constraint from $-65{,}739.5$ (i.e., $4{,}260.5 - 70{,}000$) to $-45{,}739.5$ (i.e., $4{,}260.5 - 50{,}000$) and placing a lower bound on $V_2(4)$ (i.e., variable X_{18}) of 8000. The solution is:

Note that this solution requires unit 2 to drop off its generation in order to meet the end-point constraint on its coal pile. In this case, all the coal must be delivered to plant 2 to minimize the overall cost.

The final case was constructed to show the interaction of the fuel deliveries and the economic dispatch of the generating units. In this case, the initial coal piles were set to 10,000 and 150,000 t, respectively. Furthermore, a restriction of 30,000 t minimum in the coal pile at unit 1 at the end of the third week was imposed.

Time Period	V_1	D_1	P_1	V_2	D_2	P_2
1	70000.0	0	200	50000.0	40000.0	1000
2	55144.6	0	500	26039.5	40000.0	1000
3	22169.2	0	300.5276	2079.0	40000.0	499.4724
4	1241.9307			8000.0		

$$\text{Optimum cost} = 6{,}916{,}762.4\$.$$

www.mepcafe.com

Time Period	V_1	D_1	P_1	V_2	D_2	P_2
1	10000.0	4855.4	200	150000.0	35144.6	1000
2	0.0	40000.0	500	121184.1	0	1000
3	7024.6	40000.0	150	57223.6	0	650
4	35189.2			14158.1		

$$\text{Optimum cost} = 6{,}913{,}450.8\,\$$$

To obtain the most economic operation of the two units over the 3-week period, the coal deliveries will have to be adjusted to insure both plants have sufficient coal. The solution was obtained by setting the right-hand side of the third and fourth constraint equations to −7224.6 and −145739.5, respectively, as well as imposing a lower bound of 30,000 on $V_1(4)$ (i.e., variable X_{17}). The solution is:

The LP was able to find a solution that allowed the most economic operation of the units while still directing enough coal to unit 1 to allow it to meet its end-point coal pile constraint. Note that, in practice, we would probably not wish to let the coal pile at unit 1 go to 0. This could be prevented by placing an appropriate lower bound on all the volume variables (i.e., X_5, X_8, X_{11}, X_{14}, X_{17}, and X_{18}).

This example has shown how a fuel-management problem can be solved with LP. The important factor in being able to solve very large fuel-scheduling problems is to have a LP code capable of solving large problems having perhaps tens of thousands of constraints and as many, or more, problem variables. Using such codes, elaborate fuel-scheduling problems can be optimized out over several years and play a critical role in utility fuel-management decisions.

5.6 INTRODUCTION TO HYDROTHERMAL COORDINATION

The systematic coordination of the operation of a system of hydroelectric generation plants is usually more complex than the scheduling of an all-thermal generation system. The reason is both simple and important. That is, the hydroelectric plants may very well be coupled both electrically (i.e., they all serve the same load) and hydraulically (i.e., the water outflow from one plant may be a very significant portion of the inflow to one or more other, downstream plants).

No two hydroelectric systems in the world are alike. They are all different. The reasons for the differences are the natural differences in the watersheds, the differences in the man-made storage and release elements used to control the water flows, and the very many different types of natural and man-made constraints imposed on the operation of hydroelectric systems. River systems may be simple with relatively few tributaries (e.g., the Connecticut River), with dams in series (hydraulically) along the river. River systems may encompass thousands of acres, extend over vast multinational areas, and include many tributaries and complex arrangements of storage reservoirs (e.g., the Columbia River basin in the Pacific Northwest).

www.mepcafe.com

Reservoirs may be developed with very large storage capacity with a few high-head plants along the river. Alternatively, the river may have been developed with a larger number of dams and reservoirs, each with smaller storage capacity. Water may be intentionally diverted through long raceways that tunnel through an entire mountain range (e.g., the Snowy Mountain scheme in Australia). In European developments, auxiliary reservoirs, control dams, locks, and even separate systems for pumping water back upstream have been added to rivers.

However, the one single aspect of hydroelectric plants that differentiates the coordination of their operation more than any other is the existence of the many, and highly varied, constraints. In many hydrosystems, the generation of power is an adjunct to the control of floodwaters or the regular, scheduled release of water for irrigation. Recreation centers may have developed along the shores of a large reservoir so that only small surface water elevation changes are possible. Water release in a river may well have to be controlled so that the river is navigable at all times. Sudden changes, with high-volume releases of water, may be prohibited because the release could result in a large wave traveling downstream with potentially damaging effects. Fish ladders may be needed. Water releases may be dictated by international treaty.

To repeat: all hydrosystems are different.

5.6.1 Long-Range Hydro-Scheduling

The coordination of the operation of hydroelectric plants involves, of course, the scheduling of water releases. The *long-range hydro-scheduling problem* involves the long-range forecasting of water availability and the scheduling of reservoir water releases (i.e., “drawdown”) for an interval of time that depends on the reservoir capacities.

Typical long-range scheduling goes anywhere from 1 week to 1 year or several years. For hydro schemes with a capacity of impounding water over several seasons, the long-range problem involves meteorological and statistical analyses.

Nearer-term water inflow forecasts might be based on snowmelt expectations and near-term weather forecasts. For the long-term drawdown schedule, a basic policy selection must be made. Should the water be used under the assumption that it will be replaced at a rate based on the statistically expected (i.e., mean value) rate, or should the water be released using a “worst-case” prediction. In the first instance, it may well be possible to save a great deal of electric energy production expense by displacing thermal generation with hydro-generation. If, on the other hand, a worst-case policy was selected, the hydroplants would be run so as to minimize the risk of violating any of the hydrological constraints (e.g., running reservoirs too low, not having enough water to navigate a river). Conceivably, such a schedule would hold back water until it became quite likely that even worst-case lack of rain would still give ample water to meet the constraints.

Long-range scheduling involves optimizing a policy in the context of unknowns such as load, hydraulic inflows, and unit availabilities (steam and hydro). These

www.mepcafe.com

unknowns are treated statistically, and long-range scheduling involves optimization of statistical variables. Useful techniques include:

1. Dynamic programming (DP), where the entire long-range operation time period is simulated (e.g., 1 year) for a given set of conditions.
2. Composite hydraulic simulation models, which can represent several reservoirs.
3. Statistical production cost models.

The problems and techniques of long-range hydro-scheduling are outside the scope of this text, so we will end the discussion at this point and continue with short-range hydro-scheduling.

5.6.2 Short-Range Hydro-Scheduling

Short-range hydro-scheduling (1 day to 1 week) involves the hour-by-hour scheduling of all generation on a system to achieve minimum production cost for the given time period. In such a scheduling problem, the load, hydraulic inflows, and unit availabilities are assumed known. A set of starting conditions (e.g., reservoir levels) is given, and the optimal hourly schedule that minimizes a desired objective, while meeting hydraulic steam, and electric system constraints, is sought. Part of the hydraulic constraints may involve meeting "end-point" conditions at the end of the scheduling interval in order to conform to a long-range, water-release schedule previously established.

5.7 HYDROELECTRIC PLANT MODELS

To understand the requirements for the operation of hydroelectric plants, one must appreciate the limitations imposed on operation of hydro-resources by flood control, navigation, fisheries, recreation, water supply, and other demands on the water bodies and streams, as well as the characteristics of energy conversion from the potential energy of stored water to electric energy. The amount of energy available in a unit of stored water, say a cubic foot, is equal to the product of the weight of the water stored (in this case, 62.4 lb) times the height (in feet) that the water would fall. One thousand cubic feet of water falling a distance of 42.5 ft has the energy equivalent to 1 kWh. Correspondingly, 42.5 ft^3 of water falling 1000 ft also has the energy equivalent to 1 kWh.

Consider the sketch of a reservoir and hydroelectric plant shown in Figure 5.6. Let us consider some overall aspects of the falling water as it travels from the reservoir through the penstock to the inlet gates, through the hydraulic turbine down the draft tube and out the tailrace at the plant exit. The power that the water can produce is equal to the rate of water flow in cubic feet per second times a conversion coefficient that takes into account the net head (the distance through which the water falls, less the losses in head caused by the flow) times the conversion efficiency of the turbine generator. A flow of 1 ft^3/s falling 100 ft has the power equivalent of approximately

www.mepcafe.com

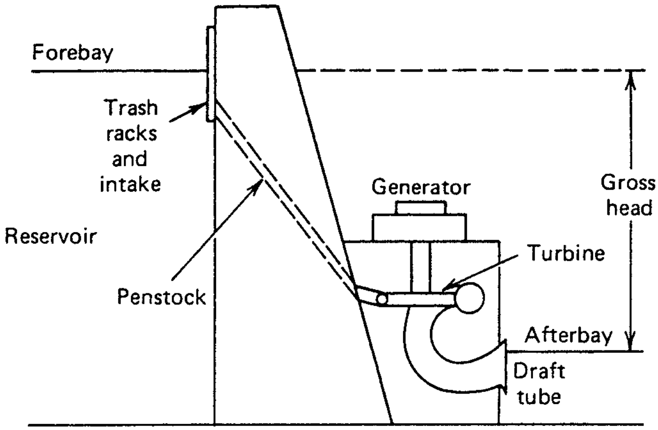


FIGURE 5.6 Hydroplant components.

8.5 kW. If the flow-caused loss in head was 5%, or 5 ft, then the power equivalent for a flow of 1 ft^3 of water per second with the net drop of 100 − 5, or 95 ft, would have the power equivalent of slightly more than 8 kW (8.5 × 95 %). Conversion efficiencies of turbine generators are typically in the range of 85–90% at the best efficiency operating point for the turbine generator, so 1 ft^3/s falling 100 ft would typically develop about 7 kW at most.

Let us return to our description of the hydroelectric plant as illustrated in Figure 5.6. The hydroelectric project consists of a body of water impounded by a dam, the hydroplant, and the exit channel or lower water body. The energy available for conversion to electrical energy of the water impounded by the dam is a function of the gross head, that is, the elevation of the surface of the reservoir less the elevation of the afterbay, or downstream water level below the hydroelectric plant. The head available to the turbine itself is slightly less than the gross head, due to the friction losses in the intake, penstock, and draft tube. This is usually expressed as the *net head* and is equal to the gross head less the flow losses (measured in feet of head). The flow losses can be very significant for low-head (10–60 ft) plants and for plants with long penstocks (several thousand feet). The water level at the afterbay is influenced by the flow out of the reservoir, including plant release and any spilling of water over the top of the dam or through bypass raceways. During flooding conditions such as spring runoff, the rise in afterbay level can have a significant and adverse effect on the energy and capacity or power capacity of the hydroplant.

The type of turbine used in a hydroelectric plant depends primarily on the design head for the plant. By far the largest number of hydroelectric projects use reaction-type turbines. Only two types of reaction turbines are now in common use. For medium heads (i.e., in the range from 60 to 1000 ft), the Francis turbine is used exclusively. For the low-head plants (i.e., for design heads in the range of 10–60 ft), the propeller turbine is used. The more modern propeller turbines have adjustable pitch blading (called *Kaplan turbines*) to improve the operating efficiency over a wide range of plant net head. Typical turbine performance results in an efficiency at full gate loading of

www.mepcafe.com

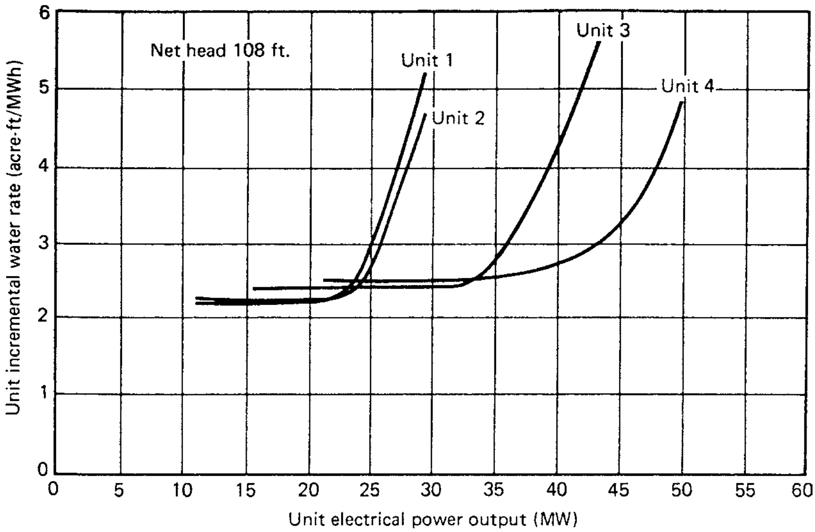


FIGURE 5.7 Incremental water rate versus power output.

between 85 and 90%. The Francis turbine and the adjustable propeller turbine may operate at 65–125% of rated net head as compared to 90–110% for the fixed propeller.

Another factor affecting operating efficiency of hydro-units is the MW loading. At light unit loadings, efficiency may drop below 70% (these ranges are often restricted by vibration and cavitation limits) and at full gate may rise to about 87%. If the best use of the hydro-resource is to be obtained, operation of the hydro-unit near its best efficiency gate position and near the designed head is necessary. This means that unit loading and control of reservoir forebay are necessary to make efficient use of hydro-resources. Unit loading should be near best efficiency gate position, and water-release schedules must be coordinated with reservoir inflows to maintain as high a head on the turbines as the limitations on forebay operations will permit.

Typical plant performance for a medium-head, four-unit plant in South America is illustrated in Figure 5.7. The incremental "water rate" is expressed in acre-feet (AF) per megawatt hour.* The rise in incremental water rate with increasing unit output results primarily from the increased hydraulic losses with the increased flow. A composite curve for multiple unit operation at the plant would reflect the mutual effects of hydraulic losses and rise in afterbay with plant discharge. Very careful attention must be given to the number of units run for a given required output. One unit run at best efficiency will usually use less water than two units run at half that load.

High-head plants (typically over 1000 ft) use impulse or Pelton turbines. In such turbines, the water is directed into spoon-shaped buckets on the wheel by means of one or more water jets located around the outside of the wheel.

*An *acre-foot* is a common unit of water volume. It is the amount of water that will cover 1 acre to a depth of 1 ft (43,560 ft^3). It also happens to be nearly equal to half a cubic foot per second flow for a day (43,200 ft^3). An acre-foot is equal to $1.2335 \cdot 10^3$ m^3.

www.mepcafe.com

In the text that follows, we will assume a characteristic giving the relationship between water flow through the turbine, q, and power output, P (MW), where q is expressed in ft^3/s or acre-ft/h. Furthermore, we will not be concerned with what type of turbine is being used or the characteristics of the reservoir, other than such limits as the reservoir head or volume and various flows.

5.8 SCHEDULING PROBLEMS

5.8.1 Types of Scheduling Problems

In the operation of a hydroelectric power system, three general categories of problems arise. These depend on the balance between the hydroelectric generation, the thermal generation, and the load.

Systems without any thermal generation are fairly rare. The economic scheduling of these systems is really a problem in scheduling water releases to satisfy all the hydraulic constraints and meet the demand for electrical energy. Techniques developed for scheduling hydrothermal systems may be used in some systems by assigning a pseudo-fuel cost to some hydroelectric plant. Then the schedule is developed by minimizing the production "cost" as in a conventional hydrothermal system. In all hydroelectric systems, the scheduling could be done by simulating the water system and developing a schedule that leaves the reservoir levels with a maximum amount of stored energy. In geographically extensive hydroelectric systems, these simulations must recognize water travel times between plants.

Hydrothermal systems where the hydroelectric system is by far the largest component may be scheduled by economically scheduling the system to produce the minimum cost for the thermal system. These are basically problems in scheduling energy. A simple example is illustrated in the next section where the hydroelectric system cannot produce sufficient energy to meet the expected load.

The largest category of hydrothermal systems includes those where there is a closer balance between the hydroelectric and thermal generation resources and those where the hydroelectric system is a small fraction of the total capacity.

In these systems, the schedules are usually developed to minimize thermal-generation production costs, recognizing all the diverse hydraulic constraints that may exist. The main portion of this chapter is concerned with systems of this type.

5.8.2 Scheduling Energy

Suppose, as in Figure 5.8, we have two sources of electrical energy to supply a load, one hydro and another steam. The hydroplant can supply the load by itself for a limited time. That is, for any time period j,

$$P_{Hj}^{max} \geq P_{load\,j} \quad j = 1 \ldots j_{max} \tag{5.22}$$

www.mepcafe.com

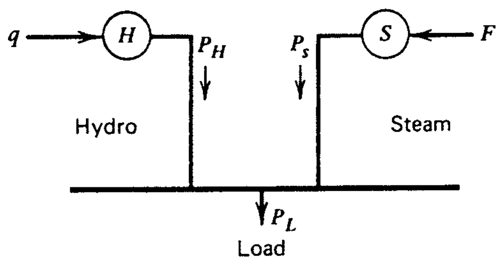


FIGURE 5.8 Two-unit hydrothermal system.

However, the energy available from the hydroplant is insufficient to meet the load:

$$\sum_{j=1}^{j_{\max}} P_{Hj} n_j \le \sum_{j=1}^{j_{\max}} P_{\text{load}\,j} n_j \quad n_j = \text{number of hours in period } j \tag{5.23}$$

$$\sum_{j=1}^{j_{\max}} n_j = T_{\max} = \text{total interval}$$

We would like to use up the entire amount of energy from the hydroplant in such a way that the cost of running the steam plant is minimized. The steam-plant energy required is

$$\underbrace{\sum_{j=1}^{j_{\max}} P_{\text{load}\,j} n_j}_{\text{Load energy}} - \underbrace{\sum_{j=1}^{j_{\max}} P_{Hj} n_j}_{\text{Hydro-energy}} = \underbrace{E}_{\text{Steam energy}} \tag{5.24}$$

We will not require the steam unit to run for the entire interval of $T_{\max}$ hours. Therefore,

$$\sum_{j=1}^{N_s} P_{sj} n_j = E \quad N_s = \text{number of periods the steam plant is run} \tag{5.25}$$

Then

$$\sum_{j=1}^{N_s} n_j \le T_{\max}$$

the scheduling problem becomes

$$\text{Min } F_T = \sum_{j=1}^{N_s} F(P_{sj}) n_j \tag{5.26}$$

subject to

$$\sum_{j=1}^{N_s} P_{sj} n_j - E = 0 \tag{5.27}$$

www.mepcafe.com

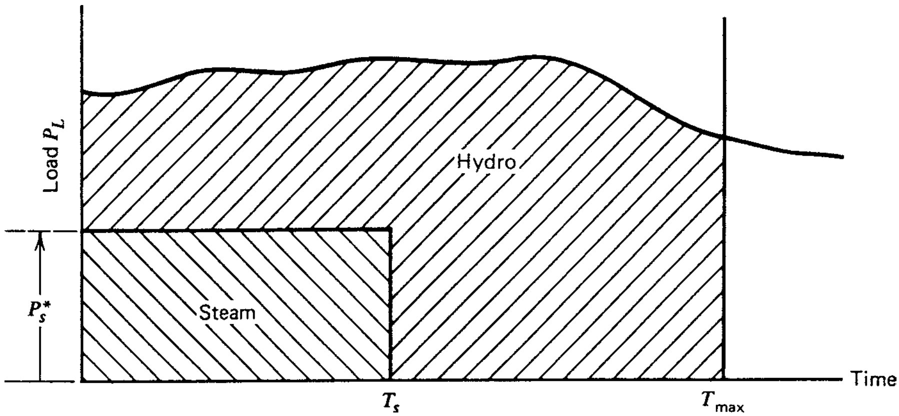


FIGURE 5.9 Resulting optimal hydrothermal schedule.

and the Lagrange function is

$$\mathcal{L} = \sum_{j=1}^{N_s} F(P_{sj})n_j + \alpha\left(E - \sum_{j=1}^{N_s} P_{sj}n_j\right) \tag{5.28}$$

Then

$$\frac{\partial \mathcal{L}}{\partial P_{sj}} = \frac{dF(P_{sj})}{dP_{sj}} - \alpha = 0 \quad \text{for } j = 1 \ldots N_s$$

or (5.29)

$$\frac{dF(P_{sj})}{dP_{sj}} = \alpha \quad \text{for } j = 1 \ldots N_s$$

This means that the steam plant should be run at constant incremental cost for the entire period it is on. Let this optimum value of steam-generated power be P_s^*, which is the same for all time intervals the steam unit is on. This type of schedule is shown in Figure 5.9.

The total cost over the interval is

$$F_T = \sum_{j=1}^{N_s} F(P_s^*)n_j = F(P_s^*)\sum_{j=1}^{N_s} n_j = F(P_s^*)T_s \tag{5.30}$$

where

$$T_s = \sum_{j=1}^{N_s} n_j = \text{the total run time for the steam plant}$$

Let the steam-plant cost be expressed as

$$F(P_s) = A + BP_s + CP_s^2 \tag{5.31}$$

www.mepcafe.com

then

$$F_{\mathrm{T}} = (A + BP_{\mathrm{s}}^{*} + CP_{\mathrm{s}}^{*2})T_{\mathrm{s}} \tag{5.32}$$

also note that

$$\sum_{j=1}^{N_{\mathrm{s}}} P_{\mathrm{s}j} n_j = \sum_{j=1}^{N_{\mathrm{s}}} P_{\mathrm{s}}^{*} n_j = P_{\mathrm{s}}^{*} T_{\mathrm{s}} = E \tag{5.33}$$

Then

$$T_{\mathrm{s}} = \frac{E}{P_{\mathrm{s}}^{*}} \tag{5.34}$$

and

$$F_{\mathrm{T}} = \left(A + BP_{\mathrm{s}}^{*} + CP_{\mathrm{s}}^{*2}\right)\left(\frac{E}{P_{\mathrm{s}}^{*}}\right) \tag{5.35}$$

Now we can establish the value of P_{s}^{*} by minimizing F_{T}:

$$\frac{\mathrm{d}F_{\mathrm{T}}}{\mathrm{d}P_{\mathrm{s}}^{*}} = \frac{-AE}{P_{\mathrm{s}}^{*2}} + CE = 0 \tag{5.36}$$

or

$$P_{\mathrm{s}}^{*} = \sqrt{\frac{A}{C}} \tag{5.37}$$

which means the unit should be operated at its maximum efficiency point long enough to supply the energy needed, E. Note, if

$$F(P_{\mathrm{s}}) = A + BP_{\mathrm{s}} + CP_{\mathrm{s}}^{2} = f_{\mathrm{c}} \times H(P_{\mathrm{s}}) \tag{5.38}$$

where f_{c} is the fuel cost, then the heat rate is

$$\frac{H(P_{\mathrm{s}})}{P_{\mathrm{s}}} = \frac{1}{f_{\mathrm{c}}}\left(\frac{A}{P_{\mathrm{s}}} + B + CP_{\mathrm{s}}\right) \tag{5.39}$$

and the heat rate has a minimum when

$$\frac{\mathrm{d}}{\mathrm{d}P_{\mathrm{s}}}\left[\frac{H(P_{\mathrm{s}})}{P_{\mathrm{s}}}\right] = 0 = \frac{-A}{P_{\mathrm{s}}^{2}} + C \tag{5.40}$$

giving best efficiency at

$$P_{\mathrm{s}} = \sqrt{\frac{A}{C}} = P_{\mathrm{s}}^{*} \tag{5.41}$$

www.mepcafe.com

Example 5D: A hydroplant and a steam plant are to supply a constant load of 90 MW for 1 week (168 h). The unit characteristics are

Hydroplant:

$$q = 300 + 15P_H \text{ acre-ft / h}$$
$$0 \le P_H \le 100 \text{ MW}$$

Steam plant:

$$H_s = 53.25 + 11.27P_s + 0.0213P_s^2$$
$$12.5 \le P_s \le 50 \text{ MW}$$

Example 5D Part 1: Let the hydroplant be limited to 10,000 MWh of energy. Solve for T_s^*, the run time of the steam unit. The load is $90 \times 168 = 15,120$ MWh, requiring 5120 MWh to be generated by the steam plant.

The steam plant's maximum efficiency is at $\sqrt{53.25/0.0213} = 50$ MW. Therefore, the steam plant will need to run for 5120/50 or 102.4 h. The resulting schedule will require the steam plant to run at 50 MW and the hydroplant at 40 MW for the first 102.4 h of the week and the hydroplant at 90 MW for the remainder.

Example 5D Part 2: Instead of specifying the energy limit on the steam plant, let the limit be on the volume of water that can be drawn from the hydroplants' reservoir in 1 week. Suppose the maximum drawdown is 250,000 AF, how long should the steam unit run?

To solve this, we must account for the plant's q versus P characteristic. A different flow will take place when the hydroplant is operated at 40 MW than when it is operated at 90 MW. In this case,

$$q_1 = [300 + 15(40)] \times T_s \text{ acre-ft}$$
$$q_2 = [300 + 15(90)] \times (168 - T_s) \text{ acre-ft}$$

and

$$q_1 + q_2 = 250,000 \text{ acre-ft}$$

Solving for T_s we get 36.27 h.

5.9 THE HYDROTHERMAL SCHEDULING PROBLEM

5.9.1 Hydro-Scheduling with Storage Limitations

A more general and basic short-term hydrothermal scheduling problem requires that a given amount of water be used in such a way as to minimize the cost of running the thermal units. We will use Figure 5.10 in setting up this problem.

The problem we wish to set up is the general, short-term hydrothermal scheduling problem where the thermal system is represented by an equivalent unit, P_s, as was done in Chapter 3. In this case, there is a single hydroelectric plant, P_H. We assume that the hydroplant is not sufficient to supply all the load demands during the period

www.mepcafe.com

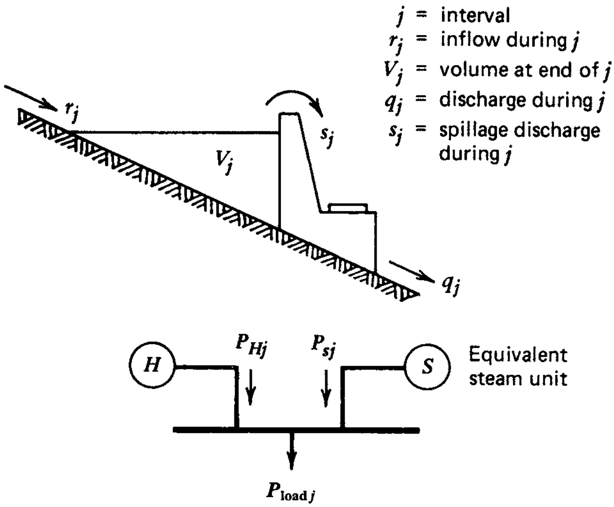


FIGURE 5.10 Hydrothermal system with hydraulic constraints.

and that there is a maximum total volume of water that may be discharged throughout the period of T_{max} hours.

In setting up this problem and the examples that follow, we assume all spillages, s_j, are 0. The only other hydraulic constraint we will impose initially is that the total volume of water discharged must be exactly as defined. Therefore, the mathematical scheduling problem may be set up as follows:

Problem:

$$\text{Min } F_T = \sum_{j=1}^{j_{\max}} n_j F_j \tag{5.42}$$

subject to

$$\sum_{j=1}^{j_{\max}} n_j q_j = q_{\text{TOT}} \quad \text{total water discharge}$$

$$P_{\text{load}\,j} - P_{\text{H}j} - P_{sj} = 0 \quad \text{load balance for } j = 1 \ldots j_{\max}$$

where

$$n_j = \text{length of } j\text{th interval}$$

$$\sum_{j=1}^{j_{\max}} n_j = T_{\max}$$

www.mepcafe.com

FIGURE 5.11 Hydroelectric unit input–output characteristic for constant head.

and the loads are constant in each interval. Other constraints could be imposed, such as

$$V_j \big|_{j=0} = V_s \quad \text{starting volume}$$
$$V_j \big|_{j=j_{\max}} = V_E \quad \text{ending volume}$$
$$q_{\min} \le q_j \le q_{\max} \quad \text{flow limits for } j = 1 \ldots j_{\max}$$
$$q_j = Q_j \quad \text{fixed discharge for a particular hour}$$

Assume constant-head operation and assume a q versus P characteristic is available, as shown in Figure 5.11, so that

$$q = q(P_H) \tag{5.43}$$

We now have a similar problem to the take-or-pay fuel problem. The Lagrange function is

$$\mathcal{L} = \sum_{j=1}^{j_{\max}} \left[n_j F(P_{sj}) + \lambda_j (P_{\text{load}\,j} - P_{Hj} - P_{sj}) \right] + \gamma \left[\sum_{j=1}^{j_{\max}} n_j q_j (P_{Hj}) - q_{\text{TOT}} \right] \tag{5.44}$$

and for a specific interval $j = k$,

$$\frac{\partial \mathcal{L}}{\partial P_{sk}} = 0$$

gives

$$n_k \frac{dF_{sk}}{dP_{sk}} = \lambda_k \tag{5.45}$$

and

$$\frac{\partial \mathcal{L}}{\partial P_{Hk}} = 0$$

gives

$$\gamma n_k \frac{dq_k}{dP_{Hk}} = \lambda_k \tag{5.46}$$

www.mepcafe.com

This is solved using the same techniques shown earlier in this chapter.

Suppose we add the network losses to the problem. Then at each hour,

$$P_{\text{load}j} + P_{\text{loss}j} - P_{\text{H}j} - P_{\text{s}j} = 0 \tag{5.47}$$

and the Lagrange function becomes

$$\mathcal{L} = \sum_{j=1}^{j_{\max}} \left[n_j F(P_{\text{s}j}) + \lambda_j (P_{\text{load}j} + P_{\text{loss}j} - P_{\text{H}j} - P_{\text{s}j}) \right] + \gamma \left[\sum_{j=1}^{j_{\max}} n_j q_j (P_{\text{H}j}) - q_{\text{TOT}} \right] \tag{5.48}$$

with resulting coordination equations (hour k):

$$n_k \frac{\mathrm{d}F(P_{\text{s}k})}{\mathrm{d}P_{\text{s}k}} + \lambda_k \frac{\partial P_{\text{loss}\,k}}{\partial P_{\text{s}k}} = \lambda_k \tag{5.49}$$

$$\gamma n_k \frac{\mathrm{d}q(P_{\text{H}k})}{\mathrm{d}P_{\text{H}k}} + \lambda_k \frac{\partial P_{\text{loss}k}}{\partial P_{\text{H}k}} = \lambda_k \tag{5.50}$$

This gives rise to a more complex scheduling solution requiring three loops, as shown in Figure 5.12. In this solution procedure, ε_1 and ε_2 are the respective tolerances on the load balance and water balance relationships.

Note that this problem ignores volume and hourly discharge rate constraints. As a result, the value of γ will be constant over the entire scheduling period as long as the units remain within their respective scheduling ranges. The value of γ would change if a constraint (i.e., $V_j = V_{max}$) were encountered. This would require that the scheduling logic recognize such constraints and take appropriate steps to adjust γ so that the constrained variable does not go beyond its limit.

Example 5E A load is to be supplied from a hydroplant and a steam system whose characteristics are given here.

Equivalent steam system:

$$H = 500 + 8.0P_{\text{s}} + 0.0016P_{\text{s}}^2 \ \ (\text{MBtu/h})$$
$$\text{Fuel cost} = 1.15 \ \$/\text{MBtu}$$
$$150 \ \text{MW} \le P_{\text{s}} \le 1500 \ \text{MW}$$

Hydroplant:

$$q = 330 + 4.97P_{\text{H}} \ \text{acre-ft/h}$$
$$0 \le P_{\text{H}} \le 1000 \ \text{MW}$$
$$q = 5300 + 12(P_{\text{H}} - 1000) + 0.05(P_{\text{H}} - 1000)^2 \ \text{acre-ft/h}$$
$$1000 < P_{\text{H}} < 1100 \ \text{MW}$$

www.mepcafe.com

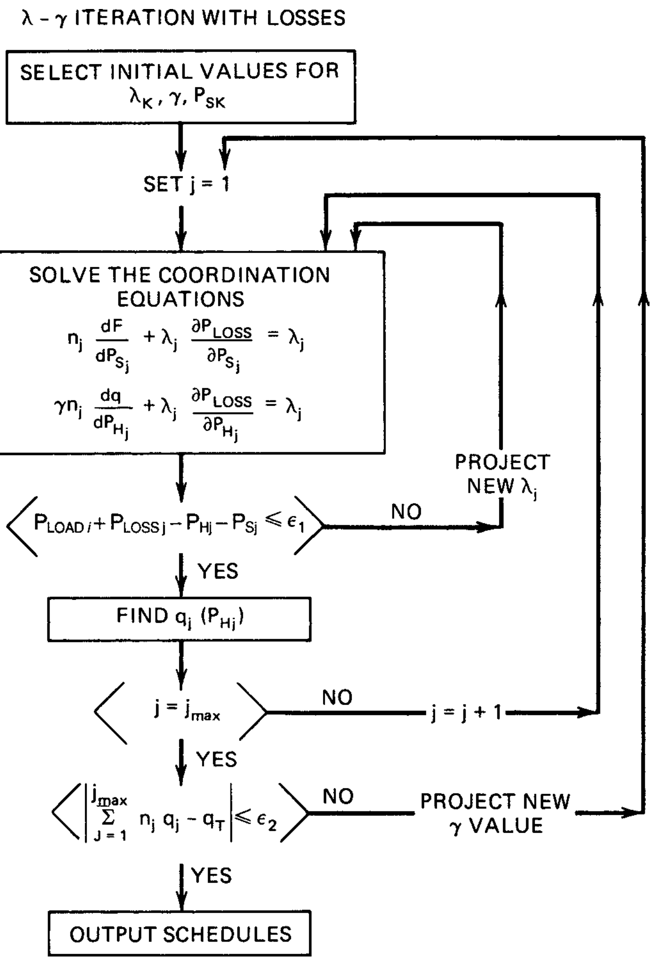


FIGURE 5.12 A $\lambda - \gamma$ iteration scheme for hydrothermal scheduling.

The hydroplant is located a good distance from the load. The electrical losses are

$$P_{\text{loss}} = 0.00008 P_{\text{H}}^2 \text{ MW}$$

The load to be supplied is connected at the steam plant and has the following schedule:

$$2400 - 1200 = 1200 \text{ MW}$$
$$1200 - 2400 = 1500 \text{ MW}$$

The hydro-unit's reservoir is limited to a drawdown of 100,000 AF over the entire 24-h period. Inflow to the reservoir is to be neglected. The optimal

www.mepcafe.com

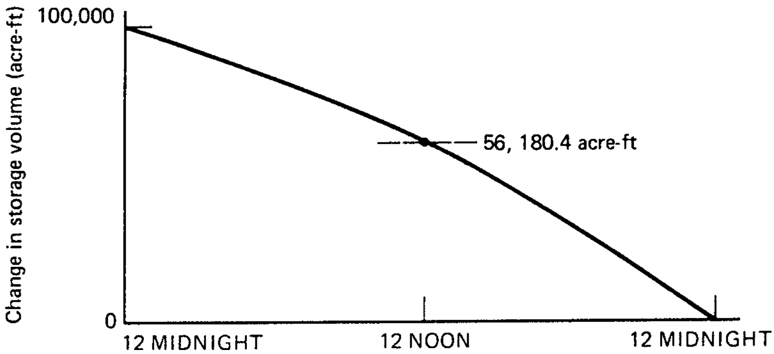


FIGURE 5.13 Change in storage volume versus time for Example 5E.

schedule for this problem was found using a program written using Figure 5.12. The results are:

Time Period	*P* Steam	*P* Hydro	Hydro-Discharge (acre-ft/h)
2400–1200	567.4	668.3	3651.5
1200–2400	685.7	875.6	4681.7

The optimal value for γ is 2.028378 \$/acre-ft. The storage in the hydroplant's reservoir goes down in time as shown in Figure 5.13. No natural inflows or spillage are assumed to occur.

5.9.2 Hydro-Units in Series (Hydraulically Coupled)

Consider now a hydraulically coupled system consisting of three reservoirs in series (see Figure 5.14). The discharge from any upstream reservoir is assumed to flow directly into the succeeding downstream plant with no time lag. The hydraulic continuity equations are

$$V_{1j} = V_{1j-1} + (r_{1j} - s_{1j} - q_{1j})n_j$$
$$V_{2j} = V_{2j-1} + (q_{1j} + s_{1j} - s_{2j} - q_{2j})n_j$$
$$V_{3j} = V_{3j-1} + (q_{2j} + s_{2j} - s_{3j} - q_{3j})n_j$$

where

r_j = inflow

V_j = reservoir volume

s_j = spill rate over the dam´s spillway

q_j = hydroplant discharge

n_j = numbers of hours in each scheduling period

www.mepcafe.com

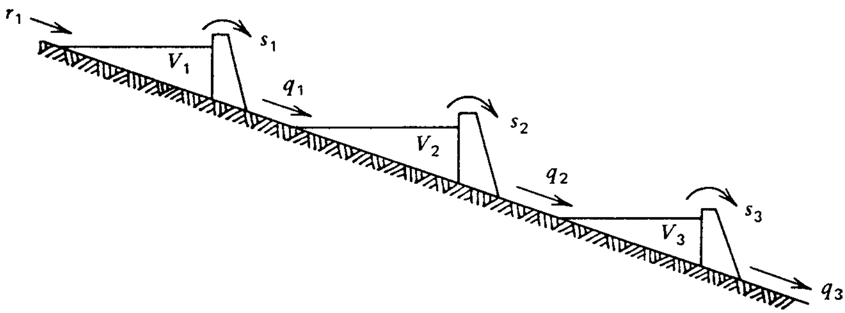


FIGURE 5.14 Hydraulically coupled hydroelectric plants.

The object is to minimize

$$\sum_{j=1}^{j_{\max}} n_j F(P_{sj}) = \text{total cost} \tag{5.51}$$

subject to the following constraints:

$$P_{\text{load}j} - P_{sj} - P_{\text{H}1j} - P_{\text{H}2j} - P_{\text{H}3j} = 0$$

and

$$\begin{aligned} V_{1j} - V_{1j-1} - (r_{1j} - s_{1j} - q_{1j})n_j &= 0 \\ V_{2j} - V_{2j-1} - (q_{1j} + s_{1j} - s_{2j} - q_{2j})n_j &= 0 \\ V_{3j} - V_{3j-1} - (q_{2j} + s_{2j} - s_{3j} - q_{3j})n_j &= 0 \end{aligned} \tag{5.52}$$

All equations in set 5.52 must apply for $j = 1 \ldots j_{max}$.

The Lagrange function would then appear as

$$\begin{aligned} L = \sum_{j=1}^{j_{\max}} \Big\{ & \left[n_j F(P_{sj}) - \lambda_j (P_{\text{load}j} - P_{sj} - P_{\text{H}1j} - P_{\text{H}2j} - P_{\text{H}3j}) \right] \\ & + \gamma_{1j} \left[V_{1j} - V_{1j-1} - (r_{1j} - s_{1j} - q_{1j})n_j \right] \\ & + \gamma_{2j} \left[V_{2j} - V_{2j-1} - (q_{1j} + s_{1j} - s_{2j} - q_{2j})n_j \right] \\ & + \gamma_{3j} \left[V_{3j} - V_{3j-1} - (q_{2j} + s_{2j} - s_{3j} - q_{3j})n_j \right] \Big\} \end{aligned}$$

Note that we could have included more constraints to take care of reservoir volume limits, end-point volume limits, and so forth, which would have necessitated using the Kuhn–Tucker conditions when limits were reached.

Hydro-scheduling with multiple-coupled plants is a formidable task. Lambda–gamma iteration techniques or gradient techniques can be used; in either case, convergence to

www.mepcafe.com

the optimal solution can be slow. For these reasons, hydro-scheduling for such systems is often done with DP (see Appendix 5A) or LP (see Section 5.10).

5.9.3 Pumped-Storage Hydroplants

Pumped-storage hydroplants are designed to save fuel costs by serving the peak load (a high fuel-cost load) with hydro-energy and then pumping the water back up into the reservoir at light load periods (a lower cost load). These plants may involve separate pumps and turbines or, more recently, reversible pump turbines. Their operation is illustrated by the two graphs in Figure 5.15. The first is the composite thermal system input–output characteristic and the second is the load cycle.

The pumped-storage plant is operated until the added pumping cost exceeds the savings in thermal costs due to the *peak shaving operations*. Figure 5.15 illustrates the operation on a daily cycle. If

$$\left.\begin{aligned} e_g &= \text{generation, MWh} \\ e_p &= \text{pumping load, MWh} \end{aligned}\right\} \text{for the same volume of water}$$

then the cycle efficiency is

$$\eta = \frac{e_g}{e_p} \; (\eta \text{ is typically about } 0.67)$$

Storage reservoirs have limited storage capability and typically provide 4–8 or 10 h of continuous operation as a generator. Pumped-storage plants may be operated on a

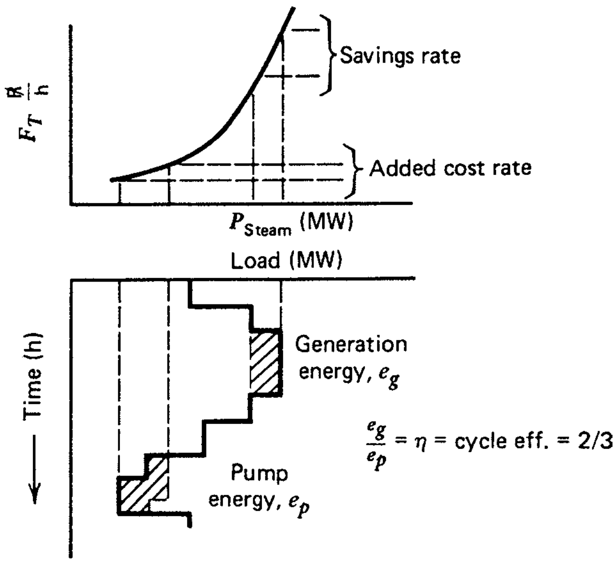


FIGURE 5.15 Thermal input–output characteristic and typical daily load cycle.

www.mepcafe.com

daily or weekly cycle. When operated on a weekly cycle, pumped-storage plants will start the week (say a Monday morning in the United States) with a full reservoir. The plant will then be scheduled over a weekly period to act as a generator during high load hours and to refill the reservoir partially, or completely, during off-peak periods.

Frequently, special interconnection arrangements may facilitate pumping operations if arrangements are made to purchase low-cost, off-peak energy. In some systems, the system operator will require a complete daily refill of the reservoir when there is any concern over the availability of capacity reserves. In those instances, economy is secondary to reliability.

5.9.3.1 Pumped-Storage Hydro-Scheduling with a $\lambda - \gamma$ Iteration

Assume:

1. Constant head hydro-operation.
2. An equivalent steam unit with convex input–output curve.
3. A 24-h operating schedule, each time intervals equals 1 h.
4. In any one interval, the plant is either pumping or generating or idle (idle will be considered as just a limiting case of pumping or generating).
5. Beginning and ending storage reservoir volumes are specified.
6. Pumping can be done continuously over the range of pump capability.
7. Pump and generating ratings are the same.
8. There is a constant cycle efficiency, η.

The problem is set up ignoring reservoir volume constraints to show that the same type of equations can result as those that arose in the conventional hydro-case. Figure 5.16 shows the water flows and equivalent electrical system.

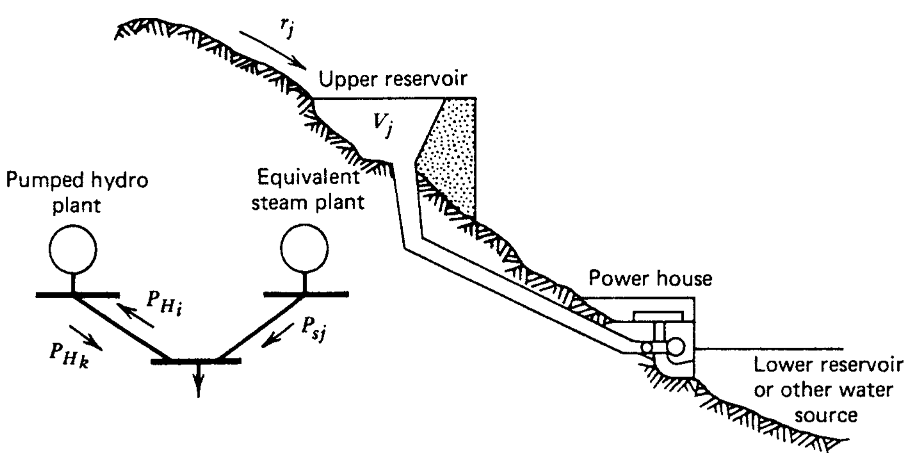


FIGURE 5.16 Pumped-storage hydraulic flows and electric system flows.

www.mepcafe.com

In some interval, j,

$$r_j = \text{inflow (acre-ft / h)}$$
$$V_j = \text{volume at end of interval (acre-ft)}$$
$$q_j = \text{discharge if generating (acre-ft / h)}$$

or

$$w_j = \text{pumping rate if pumping (acre-ft / h)}$$

Intervals during the day are classified into two sets:

$$\{k\} = \text{intervals of generation}$$
$$\{i\} = \text{intervals of pumping}$$

The reservoir constraints are to be monitored in the computational procedure.

The initial and final volumes are

$$V_0 = V_s$$
$$V_{24} = V_e$$

The problem is to minimize the sum of the hourly costs for steam generation over the day while observing the constraints. This total fuel cost for a day is (note that we have dropped n_j here since $n_j = 1$ h)

$$F_T = \sum_{j=1}^{24} F_j(P_{sj})$$

We consider the two sets of time intervals:

Set 1: $\{k\}$: **Generation intervals:** The electrical and hydraulic constraints are

$$P_{\text{load}k} + P_{\text{loss}k} - P_{sk} - P_{Hk} = 0$$
$$V_k - V_{k-1} - r_k + q_k = 0$$

These give rise to a Lagrange function during a generation hour (interval k) of

$$E_k = F_k + \lambda_k(P_{\text{load}k} + P_{\text{loss}k} - P_{sk} - P_{Hk}) + \gamma_k(V_k - V_{k-1} - r_k + q_k) \tag{5.53}$$

Set 2: $\{i\}$: **Pump intervals:** Similarly, for a typical pumping interval, i,

$$\begin{aligned} &P_{\text{load}\,i} + P_{\text{loss}\,i} - P_{si} + P_{Hi} = 0 \\ &V_i - V_{i-1} - r_i - w_i = 0 \\ &E_i = F_i + \lambda_i(P_{\text{load}\,i} + P_{\text{loss}\,i} - P_{si} + P_{Hi}) + \gamma_i(V_i - V_{i-1} - r_i - w_i) \end{aligned} \tag{5.54}$$

www.mepcafe.com

Therefore, the total Lagrange function is

$$E = \sum_{\{k\}} E_k + \sum_{\{i\}} E_i + \varepsilon_s (V_0 - V_s) + \varepsilon_e (V_{24} - V_e) \tag{5.55}$$

where the end-point constraints on the storage have been added.

In this formulation, the hours in which no pumped hydro activity takes place may be considered as pump (or generate) intervals with

$$P_{Hi} = P_{Hk} = 0$$

To find the minimum of $F_T = \Sigma F_j$, we set the first partial derivatives of E to 0.

Set 1. $\{k\}$: **Generation intervals:**

$$\begin{aligned} \frac{\partial E}{\partial P_{sk}} &= 0 = -\lambda_k \left(1 - \frac{\partial P_{loss}}{\partial P_{sk}}\right) + \frac{dF_k}{dP_{sk}} \\ \frac{\partial E}{\partial P_{Hk}} &= 0 = -\lambda_k \left(1 - \frac{\partial P_{loss}}{\partial P_{Hk}}\right) + \gamma_k \frac{dq_k}{dP_{Hk}} \end{aligned} \tag{5.56}$$

Set 2. $\{i\}$: **Pump intervals:**

$$\begin{aligned} \frac{\partial E}{\partial P_{si}} &= 0 = -\lambda_i \left(1 - \frac{\partial P_{loss}}{\partial P_{si}}\right) + \frac{dF_i}{dP_{si}} \\ \frac{\partial E}{\partial P_{Hi}} &= 0 = +\lambda_i \left(1 + \frac{\partial P_{loss}}{\partial P_{Hi}}\right) - \gamma_i \frac{dw_i}{dP_{Hi}} \end{aligned} \tag{5.57}$$

For the $\partial E/\partial V$, we can consider any interval of the entire day—for instance, the ℓth interval—which is not the first or 24th h.

$$\frac{\partial E}{\partial V_\ell} = 0 = \gamma_\ell - \gamma_\ell + 1$$

and for $\ell = 0$ and $= 24$

$$\frac{\partial E}{\partial V_0} = 0 = -\gamma_1 + \varepsilon_s \quad \text{and} \quad \frac{\partial E}{\partial V_{24}} = 0 = \gamma_{24} + \varepsilon_e \tag{5.58}$$

From Equation 5.58, it may be seen that γ is a constant. Therefore, it is possible to solve the pumped-storage scheduling problem by means of a $\lambda - \gamma$ iteration over the time interval chosen. It is necessary to monitor the calculations to prevent a violation of the reservoir constraints, or else to incorporate them in the formulation.

It is also possible to set up the problem of scheduling the pumped-storage hydroplant in a form that uses linear programming.

www.mepcafe.com

5.10 HYDRO-SCHEDULING USING LINEAR PROGRAMMING

One of the more useful ways to solve large hydro-scheduling problems is through the use of LP. Modern LP codes and computers make this an increasingly useful option. In this section, a simple, single reservoir hydroplant operating in conjunction with a single steam plant, as shown in Figure 5.10, will be modeled using LP.

First, we shall show how each of the models needed are expressed as linear models that can be incorporated in an LP. The notation is as follows:

P_{sj} = the steam plant net output at time period j

P_{Hj} = the hydroplant net output at time period j

q_j = the turbine discharge at time period j

s_j = the reservoir spill at time period j

V_j = the reservoir volume at time period j

r_j = the net inflow to the reservoir during time period j

sf_k = the slopes of the piecewise linear steam-plant cost function

sh_k = the slopes of the piecewise linear hydroturbine electrical output versus discharge function

sd_k = the slopes of the piecewise linear spill function

$P_{\text{load}j}$ = the net electrical load at time period j

The steam plant will be modeled with a piecewise linear cost function, $F(P_j)$, as shown in Figure 5.17. The three segments shown will be represented as P_{sj1}, P_{sj2}, P_{sj3} where each segment power, P_{sjk}, is measured from the start of the kth segment. Each segment has a slope designated sf_1, sf_2, sf_3; then, the cost function itself is

$$F(P_{sj}) = F(P_s^{\min}) + sf_1 P_{sj1} + sf_2 P_{sj2} + sf_3 P_{sj3} \tag{5.59}$$

and

$$0 < P_{sjk} < P_{sjk}^{\max} \quad \text{for} \quad k = 1,2,3 \tag{5.60}$$

and finally

$$P_{sj} = P_s^{\min} + P_{sj1} + P_{sj2} + P_{sj3} \tag{5.61}$$

The hydroturbine discharge versus the net electrical output function is designated $P_H(q_j)$ and is also modeled as a piecewise linear curve. The actual characteristic is usually quite nonlinear, as shown by the dotted line in Figure 5.18. Hydroplants are rarely operated close to the low end of this curve, rather they are operated close to their maximum efficiency or full gate flow points. Using the piecewise linear characteristic shown in Figure 5.18, the plant will tend to go to one of these two points.

www.mepcafe.com

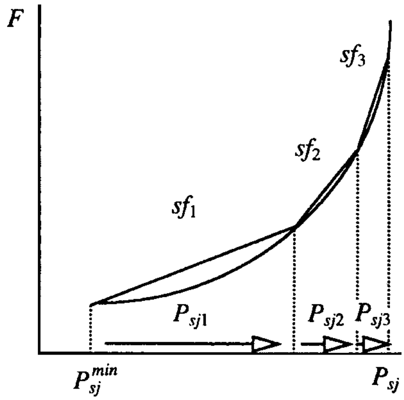


FIGURE 5.17 Steam-plant piecewise linear cost function.

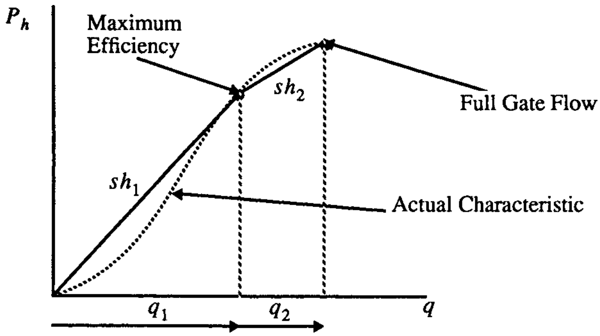


FIGURE 5.18 Hydroturbine characteristic.

In this model, the net electrical output is given as a linear sum:

$$P_{Hj} = sh_1 q_{j1} + sh_2 q_{j2} \tag{5.62}$$

The spill out of the reservoir is modeled as a function of the reservoir volume, and it is assumed that the spill is 0 if the volume of water in the reservoir is less than a given limit. This can easily be modeled by the piecewise linear characteristic in Figure 5.19, where the spill is constrained to be 0 if the volume of water in the reservoir is less than the first volume segment where

$$s_j = sd_1 V_{j1} + sd_2 V_{j2} + sd_3 V_{j3} \tag{5.63}$$

and

$$0 \le V_{jk} \le V_{jk}^{\max} \quad \text{for} \quad k = 1,2,3 \tag{5.64}$$

then

$$V_j = V_{j1} + V_{j2} + V_{j3} \tag{5.65}$$

www.mepcafe.com

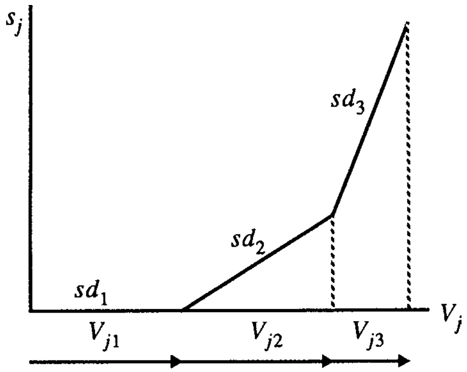


FIGURE 5.19 Spill characteristic.

The hydro-scheduling linear program then consists of the following; minimize

$$\sum_{j=1}^{j_{\max}} F(P_{sj})$$

subject to

$$V_j - V_{j-1} - (r_j - s_j - q_j) = 0 \quad \text{for } j = 1 \ldots j_{\max}$$

where

$$P_{Hj} = P_H(q_j)$$
$$s_j = s(V_j)$$

and

$$P_{sj} + P_{Hj} - P_{\text{load}j} = 0 \quad \text{for } j = 1 \ldots j_{\max}$$

Note that this simple hydro-scheduling problem will generate eight constraints for each time step:

- Two constraints for the steam-plant characteristic.
- Two constraints for the hydroturbine characteristic.
- Two constraints for the spill characteristic.
- One constraint for the volume continuity equation.
- One constraint for the load balance.

In addition, there are 15 variables for each time step. If the linear program were to be run with 1-h time periods for 1 week, it would have to accommodate a model with 1344 constraints and 2520 variables. This may seem quite large but is actually well

www.mepcafe.com

within the capability of modern LP codes. Reference 16 reports on a hydro-scheduling model containing about 10,000 constraints and 35,000 variables.

When multiple reservoir/plant models connected by multiple rivers and channels are modeled, there are many more additional constraints and variables needed. Nonetheless, the use of LP is common and can be relied upon to give excellent solutions.

APPENDIX 5A
Dynamic-Programming Solution to Hydrothermal Scheduling

Dynamic programming may be applied to the solution of the hydrothermal scheduling problem. The multiplant, hydraulically coupled systems offer computational difficulties that make it difficult to use that type of system to illustrate the benefits of applying DP to this problem. Instead we will illustrate the application with a single hydroplant operated in conjunction with a thermal system. Figure 5.20 shows a single, equivalent steam plant, P_s, and a hydroplant with storage, P_H, serving a single series of loads, P_L. Time intervals are denoted by j, where j runs between 1 and j_{max}.

Let

r_j = net inflow rate during period j

V_j = storage volume at the end of period j

q_j = flow rate through the turbine during period j

P_{Hj} = power output during period j

s_j = spillage rate during period j

P_{sj} = steam-plant output

$P_{load\,j}$ = load level

F_j = fuel cost rate for period j

Both starting and ending storage volumes, V_0 and $V_{j_{max}}$, are given, as are the period loads. The steam plant is assumed to be on for the entire period. Its input–output characteristic is

$$F_j = a + bP_{sj} + cP_{sj}^2 \ \$/\text{h} \tag{5A.1}$$

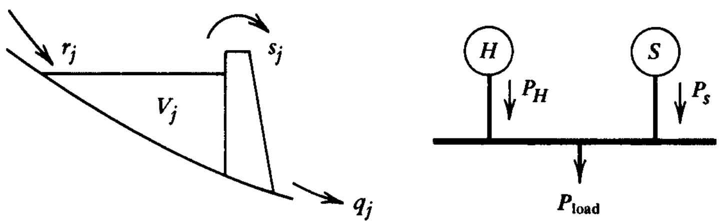


FIGURE 5.20 Hydrothermal system model used in DP illustration.

www.mepcafe.com

The water use rate characteristic of the hydroelectric plant is

and

$$q_j = d + gP_{Hj} + hP_{Hj}^2, \text{acre-ft/h} \quad \text{for } P_{Hj} > 0$$
$$= 0 \quad \text{for } P_{Hj} = 0 \tag{5A.2}$$

The coefficients a through h are constants. We will take the units of water flow rate as acre-ft/h. If each interval, j, is n_j hours long, the volume in storage changes as

$$V_j = V_{j-1} + n_j(r_j - q_j - s_j) \tag{5A.3}$$

Spilling water will not be permitted (i.e., all $s_j = 0$).

If V_i and V_k denote two different volume states, and

$$V_{j-1} = V_i$$
$$V_j = V_k$$

then the rate of flow through the hydro-unit during interval j is

$$q_j = \frac{(V_i - V_k)}{n_j} + r_j$$

where q_j must be nonnegative and is limited to some maximum flow rate, q_{max}, which corresponds to the maximum power output of the hydro-unit. The scheduling problem involves finding the minimum cost trajectory (i.e., the volume at each stage). As indicated in Figure 5.21, numerous feasible trajectories may exist.

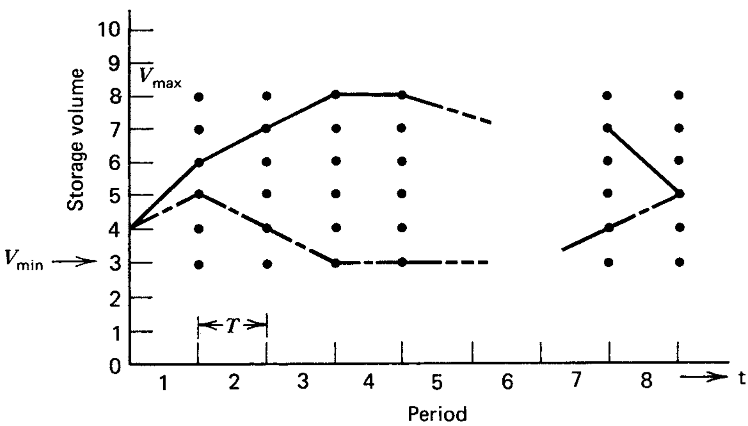


FIGURE 5.21 Trajectories for hydroplant operation.

www.mepcafe.com

The DP algorithm is quite simple. Let

$\{i\}$ = the volume states at the start of the period j

$\{k\}$ = the states at the end of j

$\mathrm{TC}_k(j)$ = the total cost from the start of the scheduling period to the end of period j for the reservoir storage state V_k

$\mathrm{PC}(i,j-1{:}k,j)$ = production cost of the thermal system in period j to go from an initial volume of V_i to an end of period volume V_k

The forward DP algorithm is then

$$\mathrm{TC}_k(0)=0$$

and

$$\mathrm{TC}_k(j)=\min_{\{i\}}\left[\mathrm{TC}_i(j-1)+\mathrm{PC}(i,j-1{:}k,j)\right] \tag{5A.4}$$

We must be given the loads and natural inflows. The discharge rate through the hydro-unit is, of course, fixed by the initial and ending storage levels, and this, in turn, establishes the values of P_{H} and P_{s}. The computation of the thermal production cost follows directly.

There may well be volume states in the set V_k that are unreachable from some of the initial volume states V_i because of the operating limits on the hydroplants. There are many variations on the hydraulic constraints that may be incorporated in the DP computation. For example, the discharge rates may be fixed during certain intervals to allow fish ladders to operate or to provide water for irrigation.

Using the volume levels as state variables restricts the number of hydropower output levels that are considered at each stage, since the discharge rate fixes the value of power. If a variable-head plant is considered, it complicates the calculation of the power level as an average head must be used to establish the value of P_H. This is relatively easy to handle.

5A.1 DYNAMIC-PROGRAMMING EXAMPLE

It is, perhaps, better to use a simple numerical example than to attempt to discuss the DP application generally. Let us consider the two-plant case just described with the steam-plant characteristics as shown in Figure 5.22 with $F = 700 + 4.8P_{\mathrm{s}} + P_{\mathrm{s}}^2/2000$, \$/h, and $\mathrm{d}F/\mathrm{d}P_{\mathrm{s}} = 4.8 + P_{\mathrm{s}}/1000$, \$/MWh, for P_{s} in MW and $200 \le P_{\mathrm{s}} \le 1200$ MW. The hydro-unit is a constant-head plant, shown in Figure 5.23, with

$$q = 260 + 10P_{\mathrm{H}} \quad \text{for } P_{\mathrm{H}} > 0, \quad q = 0 \quad \text{for } P_{\mathrm{H}} = 0$$

where P_{H} is in MW, and $0 \le P_{\mathrm{H}} \le 200$ MW

The discharge rate is in acre-ft/h. There is no spillage, and both initial and final volumes are 10,000 AF. The storage volume limits are 6000 and 18,000 AF. The natural inflow is 1000 acre-ft/h.

www.mepcafe.com

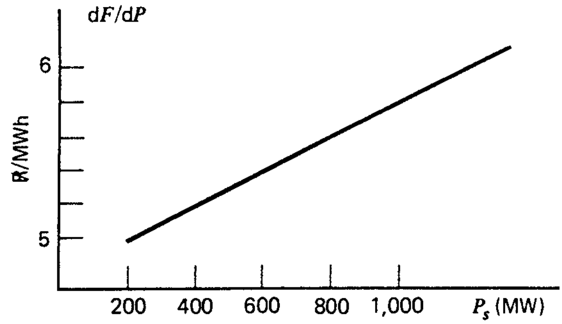


FIGURE 5.22 Steam-plant incremental cost function.

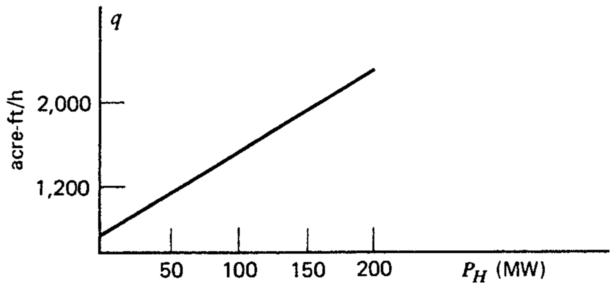


FIGURE 5.23 Hydroplant q versus P_H function.

The scheduling problem to be examined is for a 24-h day with individual periods taken as 4 h each ($n_j = 4.0$ h). The loads and natural inflows into the storage pond are:

Period j	$P_{\text{load}j}$ (MW)	Inflow Rate $r(j)$ (acre-ft/h)
1	600	1000
2	1000	1000
3	900	1000
4	500	1000
5	400	1000
6	300	1000

5A.1.1 Procedure

If this were an actual scheduling problem, we might start the search using a coarse grid on both the time interval and the volume states. This would permit the future refinement of the search for the optimal trajectory after a crude search had established the general neighborhood. Finer grid steps bracketing the range of the coarse steps around the initial optimal trajectory could then be used to establish a better path. The method will work well for problems with convex (concave) functions. For this example, we will limit our efforts to 4-h time steps and storage volume steps that are 2000 AF apart.

www.mepcafe.com

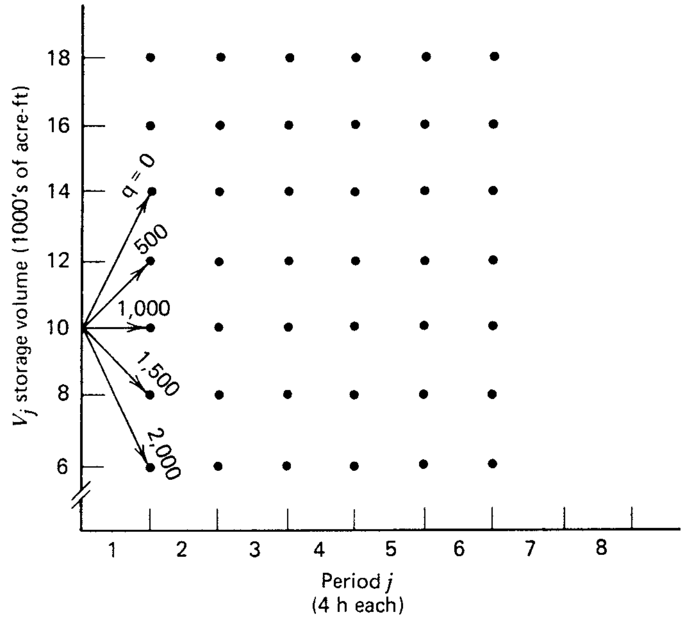


FIGURE 5.24 Initial trajectories for DP example.

During any period, the discharge rate through the hydro-unit is

$$q_j = \frac{(V_{j-1} - V_j)}{4} + 1000 \tag{5A.5}$$

The discharge rate must be nonnegative and not greater than 2260 acre-ft/h. For this problem, we may use the equation that relates P_H, the plant output, to the discharge rate, q. In a more general case, we may have to deal with tables that relate P_H, q, and the net hydraulic head.

The DP procedure may be illustrated for the first two intervals as follows. We take the storage volume steps at 6,000, 8,000, 10,000, …, 18,000 AF. The initial set of volume states is limited to 10,000 AF. (In this example, volumes will be expressed in 1000 AF to save space.) The table here summarizes the calculations for $j = 1$; the graph in Figure 5.24 shows the trajectories. We need not compute the data for greater volume states since it is possible to do no more than shut the unit down and allow the natural inflow to increase the amount of water stored.

	$j = 1$	$P_L(1) = 600$ MW		$\{i\} = 10$
V_k	q	P_H	P_s	$TC_k(j)$($)
14	0	0	600	15,040
12	500	24	576	14,523
10	1000	74	526	13,453
8	1500	124	476	12,392
6	2000	174	426	11,342

www.mepcafe.com

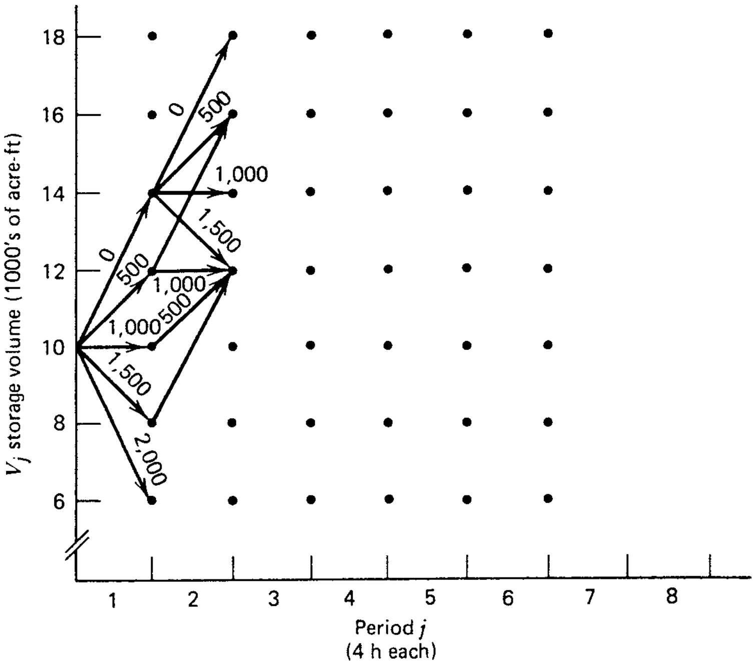


FIGURE 5.25 Second-stage trajectories for DP example.

The tabulation for the second and succeeding intervals is more complex since there are a number of initial volume states to consider. A few are shown in the following table and illustrated in Figure 5.25.

$j = 2$		$P_L = 1000$ MW		$\{i\} = [6, 8, 10, 12, 14]$	
V_k	V_i	q	P_H	P_s	$TC_k(j)$($)
18	14	0	0	1000	39,040[a]
16	14	500	24	976	38,484[a]
16	12	0	0	1000	38,523
14	14	1000	74	926	37,334[a]
14	12	500	24	976	37,967
14	10	0	0	1000	37,453
12	14	1500	124	876	36,194[a]
12	12	1000	74	926	39,818
12	10	500	24	976	36,897
12	8	0	0	1000	36,392
⋮		⋱			⋮
6	10	2000	174	826	33,477[a]
6	8	1500	124	876	33,546
6	6	1000	74	926	33,636

[a]Denotes the minimum cost path.

www.mepcafe.com

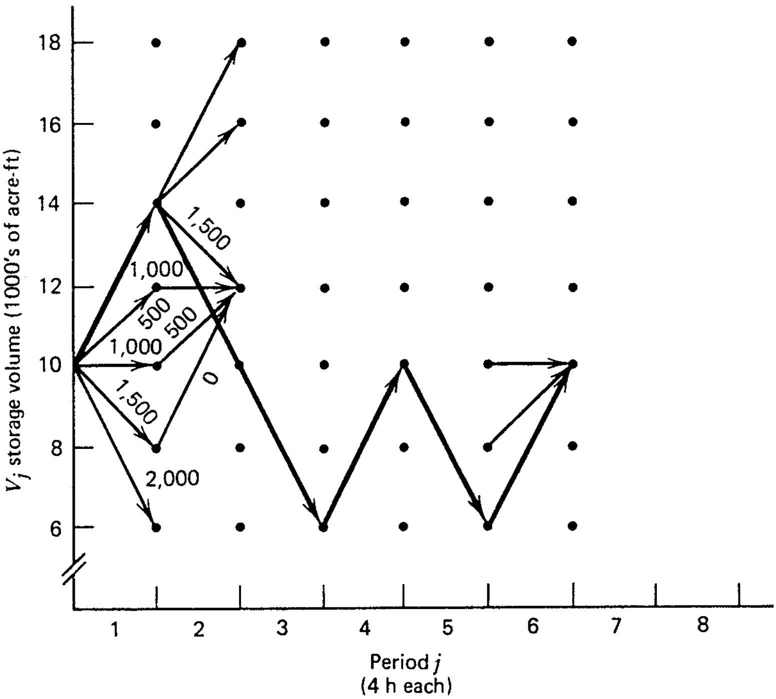


FIGURE 5.26 Final trajectory for hydrothermal scheduling example.

Finally, in the last period, the following combinations

$j = 6$		$P_L = 300$ MW		$\{i\} = [6,8,10,12,14]$	
V_k	V_i	q	P_H	P_s	$TC_k(j)(\$)$
10	10	1000	74	226	82240.61
10	8	500	24	276	82260.21
10	6	0	0	300	81738.46

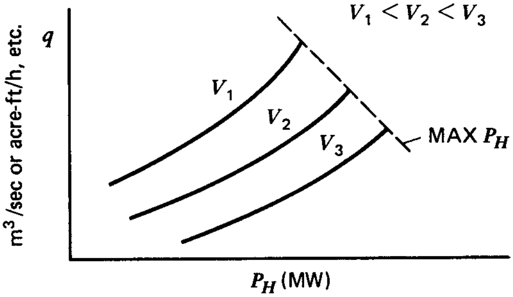
are the only feasible combinations since the end volume is set at 10, and the minimum loading for the thermal plant is 200 MW.

The final, minimum cost trajectory for the storage volume is plotted in Figure 5.26. This path is determined to a rather coarse grid of 2000 AF by 4-h steps in time and could be easily recomputed with finer increments.

5A.1.2 Extension to Other Cases

The DP method is amenable to application in more complex situations. Longer time steps make it useful to compute seasonal *rule curves*, the long-term storage plan for

www.mepcafe.com



Variable head plant
$q = q\ (P_H, \bar{V})$
$\bar{V}$ = average volume used to represent the effect of the hydraulic head

FIGURE 5.27 Input–output characteristic for variable-head hydroelectric plant.

a system of reservoirs. Variable-head cases may be treated. A sketch of the type of characteristics encountered in variable-head plants is shown in Figure 5.27. In this case, the variation in maximum plant output may be as important as the variation in water use rate as the net head varies.

5A.1.3 Dynamic-Programming Solution to Multiple Hydroplant Problem

Suppose we are given the hydrothermal system shown in Figure 5.28. We have the following hydraulic equations when spilling is constrained to 0

$$V_{1j} = V_{1j-1} + r_{1j} - q_{1j}$$
$$V_{2j} = V_{2j-1} + q_{1j} - q_{2j}$$

and the electrical equation

$$P_{H1}(q_{1j}) + P_{H2}(q_{2j}) + P_{sj} - P_{\text{load}\,j} = 0$$

There are a variety of ways to set up the DP solution to this program. Perhaps the most obvious would be to again let the reservoir volumes, V_1 and V_2, be the state variables and then run over all feasible combinations. That is, let V_1 and V_2 both be divided into N volume steps $S_1 \ldots S_2$. Then the DP must consider N^2 steps at each time interval, as shown in Figure 5.29.

www.mepcafe.com

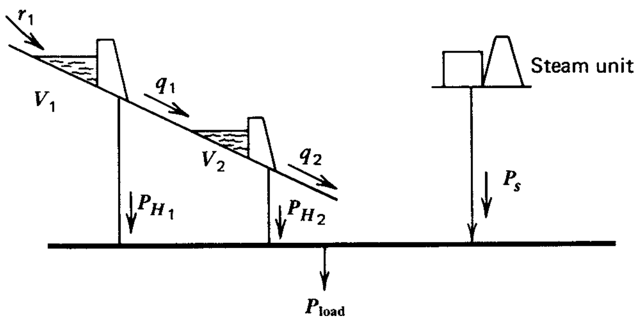


FIGURE 5.28 Hydrothermal system with hydraulically coupled plants.

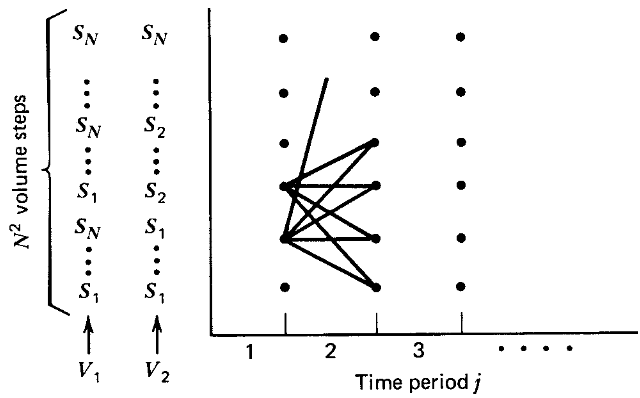


FIGURE 5.29 Trajectory combinations for coupled plants.

This procedure might be a reasonable way to solve the multiple hydroplant scheduling problem if the number of volume steps were kept quite small. However, this is not practical when a realistic schedule is desired. Consider, for example, a reservoir volume that is divided into 10 steps ($N = 10$). If there were only one hydroplant, there would be 10 states at each time period, resulting in a possible 100 paths to be investigated at each stage. If there were two reservoirs with 10 volume steps, there would be 100 states at each time interval with a possibility of 10,000 paths to investigate at each stage.

This dimensionality problem can be overcome through the use of a procedure known as *successive approximation*. This approach starts with a widely spaced grid, finds a solution, and then uses a less widely spaced grid around the previously found solution. This approach is used for each reservoir in sequence until the desired solution tolerance is achieved. The steps taken in a successive approximation method appear in Figure 5.30.

www.mepcafe.com

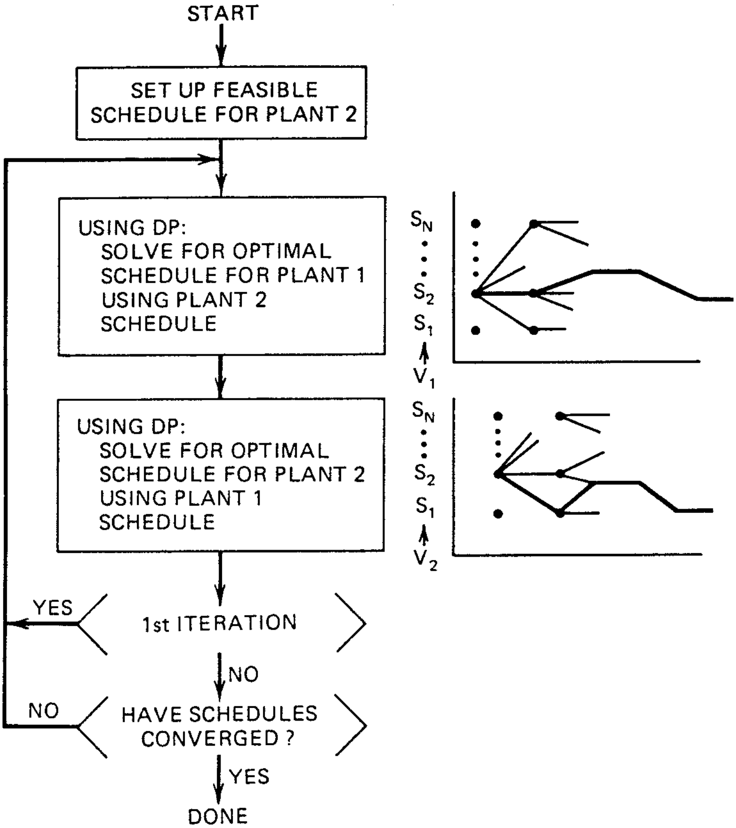


FIGURE 5.30 Successive approximation solution.

PROBLEMS

5.1 Three units are on-line all 720 h of a 30-day month. Their characteristics are as follows:

$$H_1 = 225 + 8.47P_1 + 0.0025P_1^2, \quad 50 \le P_1 \le 350$$
$$H_2 = 729 + 6.20P_2 + 0.0081P_2^2, \quad 50 \le P_2 \le 350$$
$$H_3 = 400 + 7.20P_3 + 0.0025P_3^2, \quad 50 \le P_3 \le 450$$

In these equations, the H_i are in MBtu/h and the P_i are in MW.

Fuel costs for units 2 and 3 are 0.60 \$/MBtu. Unit 1, however, is operated under a take-or-pay fuel contract where 60,000 t of coal is to be burned and/or paid for in each 30-day period. This coal costs 12 \$/t delivered and has an average heat content of 12,500 Btu/lb (1 t = 2000 lb).

www.mepcafe.com

The system monthly load-duration curve may be approximated by three steps as follows.

Load (MW)	Duration (h)	Energy (MWh)
800	50	40,000
500	550	275,000
300	120	36,000
Total	720	351,000

a. Compute the economic schedule for the month assuming all three units are on-line all the time and that the coal must be consumed. Show the MW loading for each load period, the MWh of each unit, and the value of gamma (the pseudo-fuel cost).

b. What would be the schedule if unit 1 was burning the coal at 12 \$/t with no constraint to use 60,000 t? Assume the coal may be purchased on the spot market for that price and compute all the data asked for in part a. In addition, calculate the amount of coal required for the unit.

5.2 Refer to Example 5C. The first problem solved in Example 5C left the end-**point** restrictions at 0–200,000 t for both coal piles at the end of the 3-week period. Resolve the first problem $V_1(1) = 70,000$ and $V_2(1) = 70,000$ with the added restriction that the final volume of coal at plant 2 at the end of the third week be at least 20,000 t.

5.3 Refer to Example 5C. In the second case solved with the LP algorithm (starting volumes equal to 70,000 and 50,000 for plant 1 and plant 2, respectively), we restricted the final volume of the coal pile at plant 2 to be 8000 t. What is the optimum schedule if this final volume restriction is relaxed (i.e., the final coal pile at plant 2 could go to 0)?

5.4 Using the LP problem in the text shown in Example 5C, run a linear program to find the following:

1. The coal unloading machinery at plant 2 is going to be taken out for maintenance for 1 week. During the maintenance work, no coal can be delivered to plant 2. The plant management would like to know if this should be done in week 2 or week 3.The decision will be based on the overall 3-week total cost for running both plants.
2. Could the maintenance be done in week 1? If not, why not?

Use as initial conditions those found in the beginning of the sample LP executions found in the text, that is, $V_1(1) = 70,000$ and $V_2(1) = 70,000$.

5.5 The "Cut and Shred Paper Company" of northern Minnesota has two power plants. One burns coal and the other burns natural gas supplied by the Texas Gas Company from a pipeline. The paper company has ample supplies of coal from a mine in North Dakota, and it purchases gas as take-or-pay contracts for

www.mepcafe.com

fixed periods of time. For the 8-h time period shown later, the paper company must burn 15.10^6 ft^3 of gas.
The fuel costs to the paper company are

$$\text{Coal}: \quad 0.60\ \$/\text{MBtu}$$
$$\text{Gas}: \quad 2.0\ \$/\text{ccf (where 1 ccf} = 1000\ \text{ft}^3)$$
$$\text{the gas is rated at 1100 Btu/ft}^3$$

Input–output characteristics of generators:

$$\text{Unit 1 (coal unit)}: \quad H_1(P_1) = 200 + 8.5P_1 + 0.002P_1^2\ \text{MBtu/h}\ 50 < P_1 < 500$$
$$\text{Unit 2 (gas unit)}: \quad H_2(P_2) = 300 + 6.0P_2 + 0.0025P_2^2\ \text{MBtu/h}\ 50 < P_2 < 400$$

Load (both load periods are 4 h long):

Period	Load (MW)
1	400
2	650

Assume both units are on-line for the entire 8 h.

Find the most economic operation of the paper company power plants, over the 8 h, which meets the gas consumption requirements.

5.6 An oil-fired power plant (Figure 5.31) has the following fuel consumption curve:

$$q(\text{bbl/h}) = \begin{cases} 50 + P + 0.005P^2 & \text{for} \quad 100 \le P \le 500\ \text{MW} \\ 0 & \text{for} \quad P = 0 \end{cases}$$

The plant is connected to an oil storage tank with a maximum capacity of 4000 bbl. The tank has an initial volume of oil of 3000 bbl. In addition, there is a pipeline supplying oil to the plant. The pipeline terminates in the same storage tank and must be operated by contract at 500 bbl/h. The oil-fired

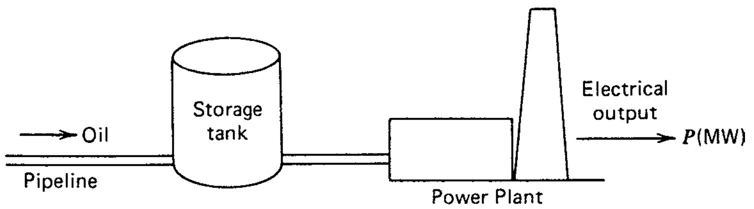


FIGURE 5.31 Oil-fired power plant with storage tank for Problem 5.6.

www.mepcafe.com

power plant supplies energy into a system, along with other units. The other units have an equivalent cost curve of

$$F_{eq} = 300 + 6P_{eq} + 0.0025P_{eq}^2$$
$$50 \le P_{eq} \le 700 \text{ MW}$$

The load to be supplied is given as follows:

Period	Load (MW)
1	400
2	900
3	700

Each time period is 2 h in length. Find the oil-fired plant's schedule using DP, such that the operating cost on the equivalent plant is minimized and the final volume in the storage tank is 2000 bbl at the end of the third period. When solving, you may use 2000, 3000, and 4000 bbl as the storage volume states for the tank. The q versus P function values you will need are included in the following table.

q (bbl/h)	P (MW)
0	0
200	100.0
250	123.6
500	216.2
750	287.3
1000	347.2
1250	400.0
1500	447.7
1800	500.0

The plant may be shut down for any of the 2-h periods with no start-up or shut-down costs.

5.7 You operate two power plants in northern Wisconsin for a paper mill. One burns coal and the other burns natural gas supplied from a pipeline coming from Canada. You have ample supplies of coal and you purchase gas as "take-or-pay" contracts for fixed periods of time. For the 24-h time period shown below, you must burn 60,000 cuft of gas.

Your fuel costs are

Coal: 2.5 \$/MBTU
Gas: the gas is rated at 1000 BTU/cuft

Input–output characteristics of generators:

Unit 1 (coal unit) $H_1(P_1) = 0.01P_1^2$ MBtu/h $\quad 50 \le P_1 \le 1200$

Unit 2 (gas unit) $q(P_2) = 0.005P_2^2$ cuft/h $\quad 50 \le P_2 \le 1000$

www.mepcafe.com

Load: (Both load periods are 12 h long)

Period	Load (MW)
1	1200
2	1500

Assume both units are on-line for the entire 24-h period

a. Write down all equations needed for solving this problem.

b. Pick any starting value you wish for gamma, then find the next value to "bracket" the answer, that is, find a value of gamma that gives a gas consumption above the 60,000 cuft and a value that gives a gas consumption below 60,000 cuft.

5.8 You are thinking of purchasing a small hydropower generator to replace the natural gas unit you use to power the paper mill in Problem 5.7. You are going to run a cost comparison between using the coal unit with the gas unit (solved in Problem 5.7) and the coal unit and the hydro-unit.

The cost of coal is (same as Problem 5.3):

Coal 2.5 \$/MBtu

Input–output characteristic of the coal plant is (same as Problem 5.3):

Unit 1 (coal unit) $H_1(P_1) = 0.01P_1^2$ MBtu / h $50 \le P_1 \le 1200$

Hydro-unit: $Q(P_H) = 50 + 5P_H$ when $P_H > 0$
and $Q(P_H) = 0$ when $P_H = 0$

where $0 \le P_H \le 700$ MW
the hydroplant reservoir holds 50,000 AF of water. The river inflow is estimated to be steady at 1500 AF/h.
The reservoir is to start at 25,000 AF at the start of the 24-h period and end at 25,000 AF.
Load: (Both load periods are 12 h long) (same as Problem 5.7)

Period	Load (MW)
1	1200
2	1500

Find the hydro-schedule that minimizes the production cost on the coal plant for the 24-h period using DP. Use the following reservoir steps in the DP: 0, 12,500, 25,000, 37,500, 50,000.
Does the hydroplant save operating cost on the coal plant?

5.9 Given the following steam-plant and hydroplant characteristics:

Steam plant:

$$\text{Incremental cost} = 2.0 + 0.002P_s \ \$/\text{MWh} \quad \text{and} \quad 100 \le P_s \le 500 \text{ MW}$$

www.mepcafe.com

Hydroplant:

$$\text{Incremental water rate} = 50 + 0.02P_H \ \text{ft}^3/\text{MW-sec} \quad 0 \le P_H \le 500 \text{ MW}$$

Load:

Time Period	P_{load} (MW)
2400–0900	350
0900–1800	700
1800–2400	350

Assume:

- The water input for $P_H = 0$ may also be assumed to be 0, that is
- $q(P_H) = 0$ for $P_H = 0$
- Neglect losses.
- The thermal plant remains on-line for the 24-h period.

Find the optimum schedule of P_s and P_H over the 24-h period that meets the restriction that the total water used is 1250 million ft^3 of water; that is,

$$q_{TOT} = 1.25 \times 10^9 \ \text{ft}^3$$

5.10 Assume that the incremental water rate in Problem 5.9 is constant at 60 ft^3/sec/MW and that the steam unit is not necessarily on all the time. Further, assume that the thermal cost is

$$F(P_s) = 250 + 2P_s + \frac{P_s^2}{1000}$$

Repeat Problem 5.9 with the same water constraint.

5.11 Hydrothermal Scheduling using Dynamic Programming

Repeat the dynamic programming example of Section 5.11 except the hydroelectric unit's water rate characteristic is now one that reflects a variable head. This characteristic also exhibits a maximum capability that is related to the net head. That is,

$$q = 0 \quad \text{for } P_H = 0$$

$$q = 260 + 10P_H\left(1.1 - \frac{\bar{V}}{100{,}000}\right) \text{acre-ft/h}$$

for

$$0 < P_H \le 2000\left(0.9 + \frac{\bar{V}}{100{,}000}\right) \text{MW}$$

www.mepcafe.com

where

$$\overline{V} = \text{average reservoir volume}$$

For this problem, assume constant rates during a period so that

$$\overline{V} = \frac{1}{2}(V_k + V_i)$$

where

$$V_k = \text{end of period volume}$$
$$V_i = \text{start of period volume}$$

The required data are

Fossil unit: On-line entire time

$$F = 770 + 5.28P_s + 0.55 \times 10^{-3} P_s^2 \ \$/\text{h}$$

for

$$200 \leq P_s \leq 1200 \ \text{MW}$$

Hydro-storage and inflow:

$$r = 1000 \text{ acre} - \text{ft/h inflow}$$
$$6000 \leq V \leq 18{,}000 \text{ acre} - \text{ft storage limits}$$
$$V = 10{,}000 \text{ acre} - \text{ft initially}$$

and

$$V = 10{,}000 \text{ acre-ft at end of period}$$

Load for 4-h periods:

J: Period	P_{load} (MW)
1	600
2	1000
3	900
4	500
5	400
6	300

Find the optimal schedule with storage volumes calculated at least to the nearest 500 acre-ft.

5.12 Pumped-Storage Plant Scheduling Problem
A thermal generation system has a composite fuel-cost characteristic as follows:

$$F = 250 + 1.5P_s + \frac{P_s^2}{200} \ \$/\text{h}$$

www.mepcafe.com

for

$$200 \le P_s \le 1200 \text{ MW}$$

In addition, it has a pumped-storage plant with the following characteristics:

1. Maximum output as a generator = 180 MW (the unit may generate between 0 and 180 MW).
2. Pumping load = 200 MW (the unit may only pump at loads of 100 or 200 MW).
3. The cycle efficiency is 70% (that is, for every 70 MWh generated, 100 MWh of pumping energy are required).
4. The reservoir storage capacity is equivalent to 1600 MWh of generation.

The system load level in chronological order is the same as that in Problem 5.11

a. Assume the reservoir is full at the start of the day and must be full at the end of the day. Schedule the pumped-storage plant to minimize the thermal system costs.

b. Repeat the solution to part a, assuming that the storage capacity of the reservoir is unknown and that it should be at the same level at the end of the day. How large should it be for minimum thermal production cost?

Note: In solving these problems you may assume that the pumped-storage plant may operate for partial time periods. That is, it does not have to stay at a constant output or pumping load for the entire 4-h load period.

5.13 Hydro Unit Commitment Problem

The "Light Up Your Life Power Company" operates one hydro-unit and four thermal-generating units. The on/off schedule of all units, as well as the MW output of the units, is to be determined for the load schedule given below.

Thermal unit data (fuel = 1.0 $/MBtu

Unit No.	Max (MW)	Min (MW)	IncrementalHeat Rate (Btu/kWh)	No-Load Energy Input (MBtu/h)	Start Up (MBtu)	Min Up Time (h)	Min Down Time (h)
1	500	70	9,950	300	800	4	4
2	250	40	10,200	210	380	4	4
3	150	30	11,000	120	110	4	8
4	150	30	11,000	120	110	4	16

Hydroplant data:

$$Q(P_H) = 1000 + 25P_H \text{ acre-ft/h}$$

where

$$0 < P_H < 200 \text{ MW}$$

min up and down time for the hydroplant is 1 h.

www.mepcafe.com

Load data (each time period is 4 h):

Time Period	P_{load} (MW)
1	600
2	800
3	700
4	1150

The starting conditions are: units 1 and 2 are running and have been up for 4 h, units 3, 4, and the hydro-unit are down and have been for 16 h. Find the schedule of the four thermal units and the hydro-unit that minimizes thermal production cost if the hydro-units starts with a full reservoir and must use 24,000 acre-ft of water over the 16-h period.

5.14 Three period hydro schedule problem

The "Lost Valley Paper Company" of northern Maine operates a very large paper plant and adjoining facilities. All of the power supplied to the paper plant must come from its own hydroplant and a group of thermal-generation facilities that we shall lump into one equivalent generating plant. The operation of the hydro-facility is tightly governed by the Maine Department of Natural Resources.

Hydroplant data:

$$Q(P_{\text{H}}) = 250 + 25P_{\text{H}} \text{ acre-ft / h}$$

and

$$0 < P_{\text{H}} < 500 \text{ MW, note } Q(P_{\text{H}}) = 0 \text{ when } P_{\text{H}} = 0$$

Equivalent steam-plant data:

$$F(P_{\text{s}}) = 600 + 5P_{\text{s}} + 0.005P_{\text{s}}^2 \text{ \$ / h}$$

and

$$100 < P_{\text{s}} < 1000 \text{ MW}$$

Load data (each period is 4 h):

Time Period	P_{load} (MW)x
1	800
2	1000
3	500

The Maine Department of Natural Resources had stated that for the 12-h period above, the hydroplant starts at a full reservoir containing 20,000 acre-ft of water and ends with a reservoir that is empty. Assume that there is no inflow to the reservoir and that both units are on-line for the entire 12 h. The Hydro unit may be shut off for all or part of a time period and does not draw any water when shut off.

Find the optimum schedule for the hydroplant using dynamic programming. Use only three volume states for this schedule: 0, 10,000, and 20,000 acre-ft.

www.mepcafe.com

6

TRANSMISSION SYSTEM EFFECTS

6.1 INTRODUCTION

The nature of the loading of the transmission components of an electric power system comprises the "power flow" solution of the network. The term "load flow" was coined before the use of computers when the load was actually moved from the demand site to the generation sites. Originally, a small physical model of the power system was built with equivalent resistors, inductors, and capacitors. Then the "solution" of the flows was achieved by measuring the current in each wire as a model was built for a segment of the system. The segments were then added using the superposition theorem of circuit theory to provide a complete solution. This was a tedious procedure. It was also quite tricky as 120 V was normally used as the source since it was available "at the wall." As demand was added, the equivalent method would be to find the impact of each new load (demand) on the generation. Then, as load is added to the system, only the impact of the new load had to be emulated. Thus, the load was "flowed" from the load location back to each generator. A more correct description is that a power flow solution is numerically found. The main output of such a solution is the power flow on all transmission and/or distribution devices. The power flow on each device is checked against the device flow capability to determine if the device will stay within operationally accepted limits.

Generally, a power flow study assumes knowledge of bus loading and generation schedule at all busses except one. The normal diagram used for wide area system awareness is shown in Figure 6.1. The major substations, the power plants,

Power Generation, Operation, and Control, Third Edition. Allen J. Wood, Bruce F. Wollenberg, and Gerald B. Sheblé.
© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

www.mepcafe.com

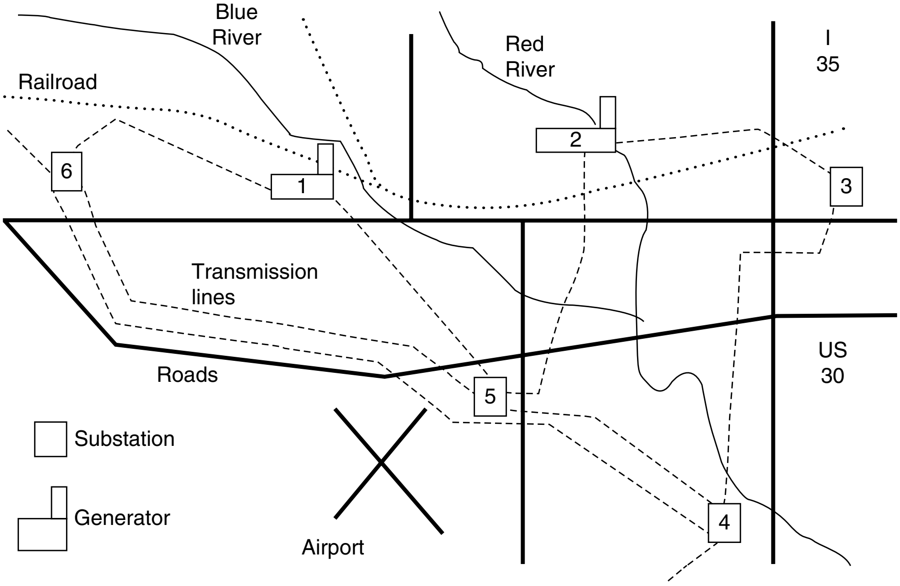


FIGURE 6.1 Small town city geographic map.

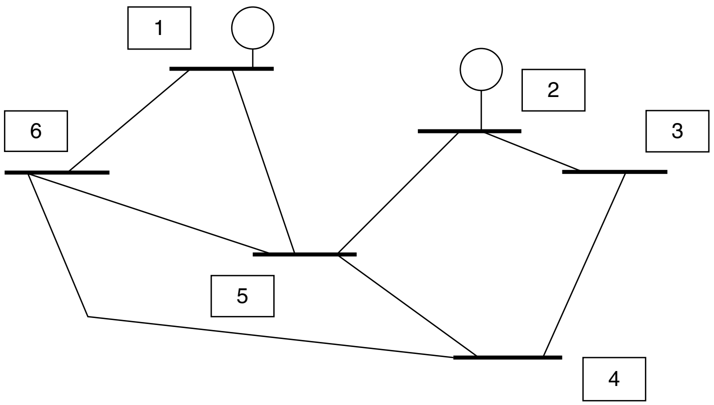


FIGURE 6.2 One line diagram.

and the transmission lines connecting the substations are often shown on a geographic map. The substations are shown on a geographic map to facilitate operator understanding of the power system state. An example one-line diagram corresponding to the geographic diagram is shown in Figure 6.2. Note that there are two generating stations, four substations feeding demand, and two interconnections to other systems. The interconnections to other systems for interchange are discussed in Chapter 11. The external systems will be modeled as equivalent generation for this chapter. The equivalent generation could be positive or negative real power as well as positive or negative reactive power. Note that most analyses have to include major sections of external systems to properly solve for the flows

www.mepcafe.com

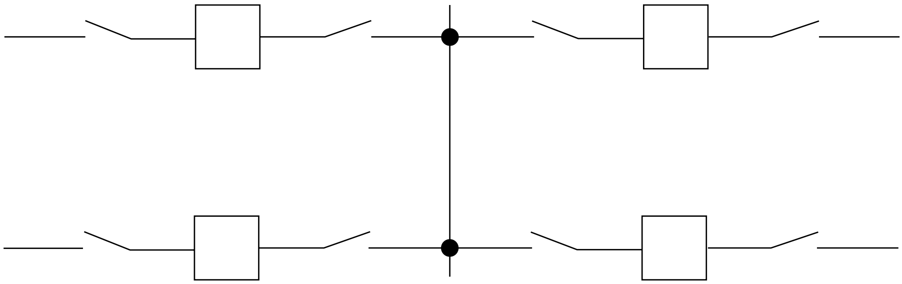


FIGURE 6.3 Single bus.

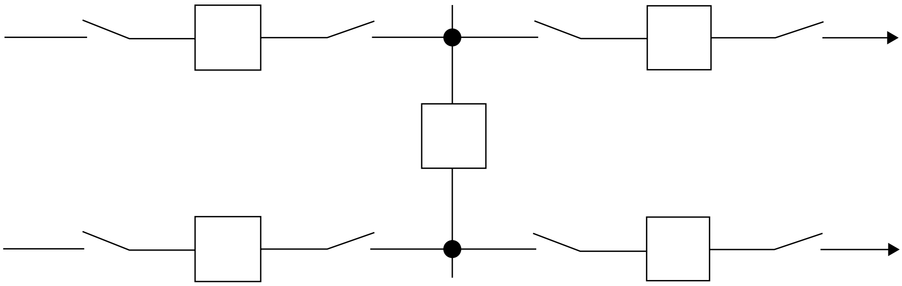


FIGURE 6.4 Sectionalized bus.

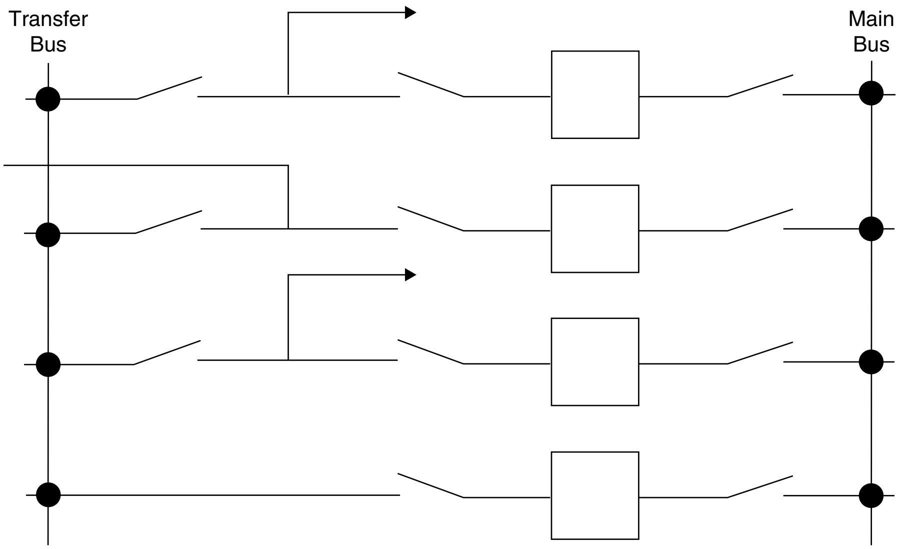


FIGURE 6.5 Main and transfer bus.

as parallel paths divide the flow as a ratio of the parallel impedances. The substation diagrams are shown in Figure 6.3, Figure 6.4, Figure 6.5, Figure 6.6, Figure 6.7, and Figure 6.8 with the detailed bus work within each substation. These figures show the disconnect switches and the breakers together. The status of the circuit

www.mepcafe.com

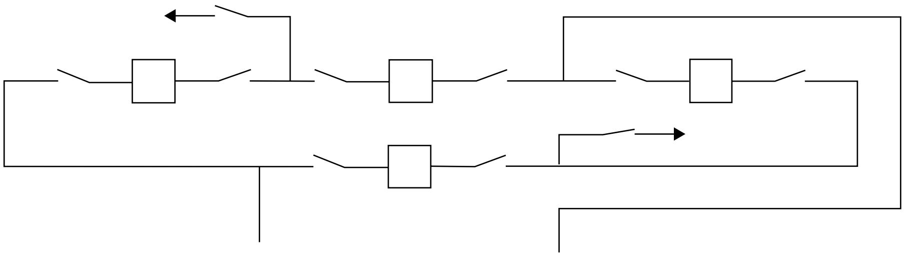


FIGURE 6.6 Ring bus.

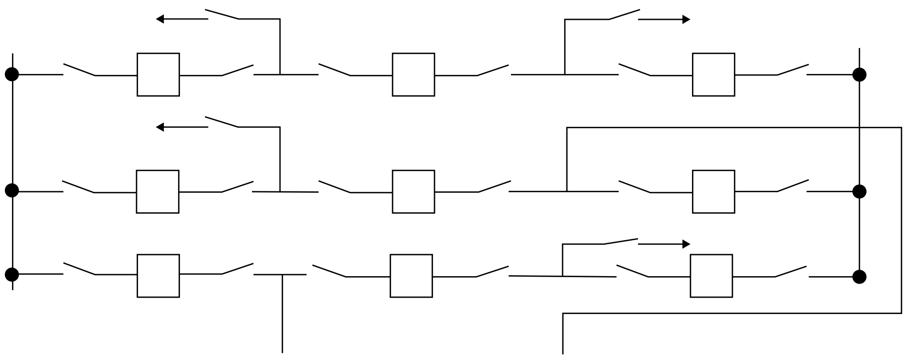


FIGURE 6.7 Breaker and a half.

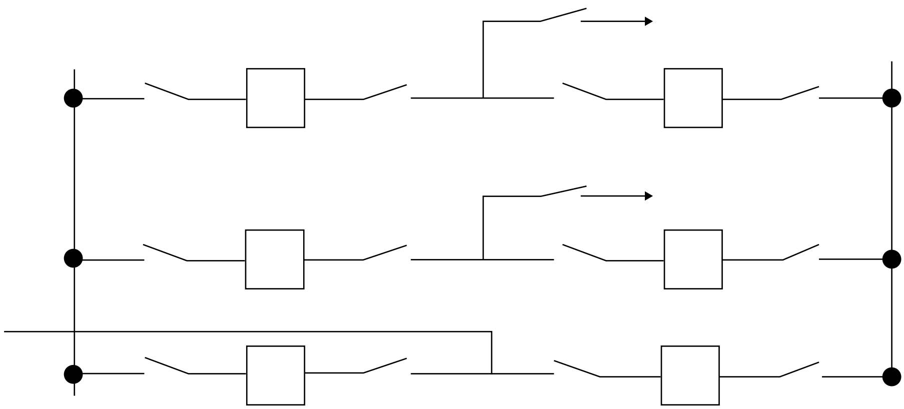


FIGURE 6.8 Double breaker double bus.

breakers and disconnects has to be processed to identify the bus and line interconnections to form the power flow model.

The term branch is used for any series device in this work. A branch could be a transmission line, a transformer, a high-voltage direct current (HVDC) link, or a FACT

www.mepcafe.com

device (FACT is the name given to high-power electronic devices used in power systems and stands for flexible alternating current transmission device).

System interconnections are modeled with equivalents if not modeled completely for direct solution. A power flow uses the knowledge of demand at each bus, the parameters for each piece of equipment (dependent current sources, dependent voltage sources, ideal transformers, resistances, reactances, and capacitances), and the power capability of each piece of equipment or device to determine a solution to the power flow equations. The power flow equations are an alternative statement of conservation of energy (COE). The sum of the power flowing into a bus (node) must be 0. Alternatively, the power flowing out of a bus has to be 0. This is commonly referred to as "Tellegen's theorem."

The transmission network includes all equipment that is interconnected to allow flow in both directions. The transmission network usually includes an equivalent model of the power system interconnections too remote to be of direct concern. Given the cost of computer capabilities, it is very common to study networks in excess of 30,000 buses. The nodes of the equivalent circuit are termed buses and are shown as dark bold lines in Figure 6.2. The branches between the buses are the transmission lines, the most common element, and are shown as light lines in Figure 6.2. Other devices, such as transformers, use overlapping circles to show transformation symbols as shown in the following.

This chapter first shows how to convert the equipment database with the status database to form a bus and branch model for power flow analysis. The models used within this work are then presented as a review. This chapter uses a single-phase per-unit equivalent assuming that the power system is balanced. Three-phase analysis is becoming more the norm as system unbalance is taken into account. Phasor measurement deployment brings forward the need to represent not only three-phase models but also the mutual coupling between transmission lines as used for system planning and fault analysis. The impact of the ground return is also noticeable when phasors are measured.

The solutions of the power flow equations are shown as an extension of circuit theory using Tellegen's theorem. Then the application of the Newton–Raphson method is shown to solve the power flow equations, control equations, etc. The more general equation form is used to model controllers, economic dispatch, etc. Approximations to the Jacobian matrix are presented to achieve faster solutions without loss of accuracy, especially the Stott's decoupled power flow.

6.2 CONVERSION OF EQUIPMENT DATA TO BUS AND BRANCH DATA

The geographic view in Figure 6.1 is the overview seen by most dispatchers in a power system's operations control center. This view gives the location and the relative distance between equipment and the interconnection of equipment and the power system state when visualization tools are added. The geographic view has to be converted to the power flow bus and branch model as shown in Figure 6.2. This is accomplished by merging the equipment static data (i.e., connection with other

www.mepcafe.com

equipment) with the status data (breaker open/closed information) acquired either by SCADA/RTU networks or phasor measurement networks (PMN).

The substation data include the majority of equipment descriptions, such as circuit breakers, disconnects, series and shunt devices, as well as transformers, flexible alternating current devices, and direct current conversion equipment. The substation data that is processed first is the disconnect and circuit breaker data with the present status of each device (open or closed). The range of substation configurations for disconnects and breakers are shown in Figure 6.3, Figure 6.4, Figure 6.5, Figure 6.6, Figure 6.7, and Figure 6.8. A circuit breaker is shown by a box. A disconnect is shown by a switch. Disconnects are often manually operated since they are for maintenance of the circuit breaker, series devices, or shunt devices. Disconnects cannot interrupt demand currents. Demand is normally shown as an arrow. Transformers and other series devices are connected between these bus configurations.

6.3 SUBSTATION BUS PROCESSING

The previous substation configurations have to be converted to a bus and line model based on the status of disconnects and circuit breakers. These device statuses are communicated to the central control facility through SCADA/RTU communication or through PMN communication.

The status determines which bus sections are connected to form a bus at this point in time. A computer code uses breaker connection specifications and real time status data to determine which closed breakers form complete buses. An example is shown in Figure 6.9. Here we illustrate that a breaker and a half scheme could result in one bus or two separate buses.

An example of the bus and branch model found is depicted in Figure 6.10.

6.4 EQUIPMENT MODELING

Generally, three-phase (3Φ), alternating-current (AC) grids are used for electric power transmission, and power flow studies rely on a single-phase per-unit equivalent circuit representation. Such a model is only valid when the system is operating in a balanced mode where the other phases have the same voltages and currents shifted by ±120°. The other phases may be obtained by simple addition or subtraction depending on the transformation used to obtain the per-phase equivalent. Since the system is in a balanced mode, only the positive sequence network is included. Per-unit equations are used throughout this section.

The transmission line model used in this chapter is the equivalent PI model consisting of two shunt capacitors and series real and inductive impedances. It is important to remember that transmission lines are lumped equivalent models of a distributed device. The impact of earth impedance is of importance as ground conductivity changes the line parameters significantly. It is important to note that short transmission lines may be treated as buses if the voltage difference across the line is negligible.

www.mepcafe.com

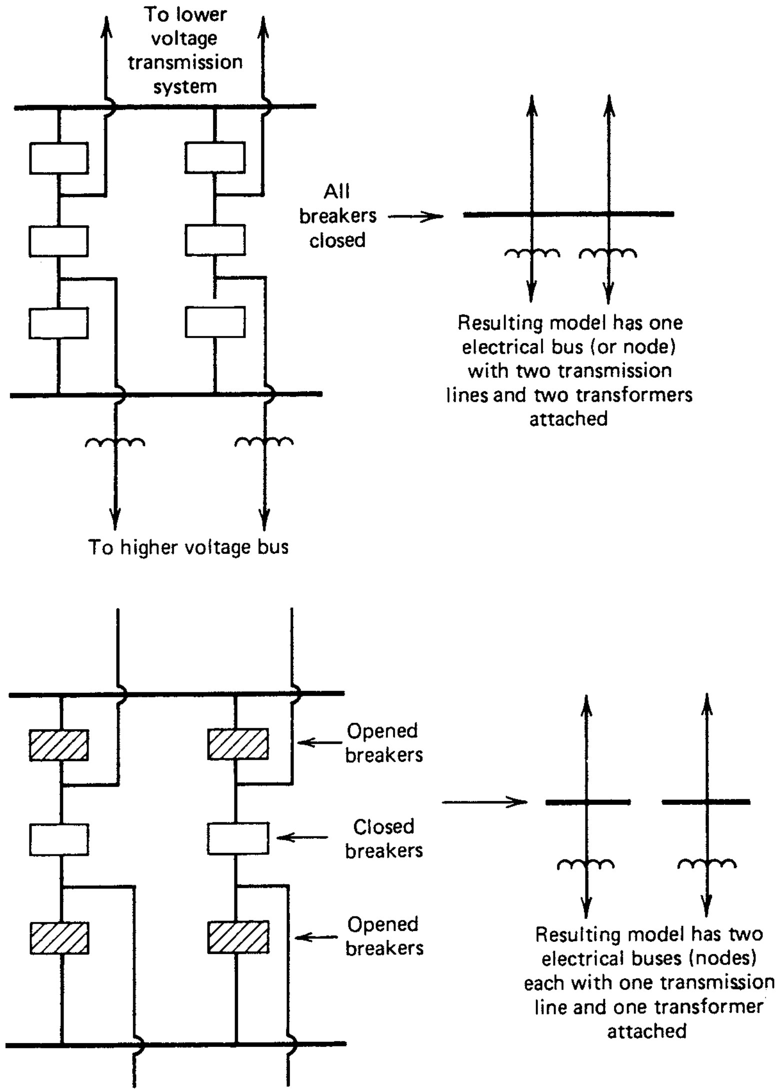


FIGURE 6.9 Breaker topology processing.

An electrical network of R, L, and C elements models a power system. It is thus very tempting to use either conventional loop or nodal analysis methods to solve for voltages and currents. Indeed, the basic circuit element analysis by phasor representation is applicable. However, this direct approach is not applicable since the demands are known as complex powers. This is a special case of a dependent current source where the independent variable is the voltage being solved. The generators should not be modeled as "voltage sources." They behave like "power sources." The generator will provide the real power independent of the voltage or current required and will provide

www.mepcafe.com

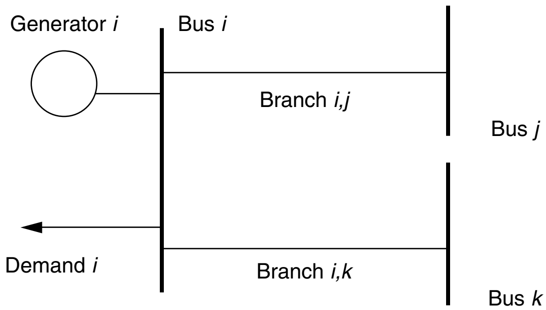


FIGURE 6.10 Example bus and branch model.

the reactive power needed to sustain the voltage desired within capability limits. The problem is to solve $2n$ nonlinear algebraic equations in $2n$ unknowns for an n bus system. This requires advanced numerical analysis techniques using optimal ordering of the equations for sparsity programming and efficient matrix manipulation.

The primary model of interest is the high-voltage AC transmission line and transformer. These systems can carry and convert large amounts of electric power. The high-voltage AC transmission line is modeled as four ideal circuit elements (capacitor, resistor, inductor, and capacitor). This model is shown in Figure 6.11. The model for a transformer is composed of three ideal circuit elements (resistor, inductor, and ideal transformer). This model is shown in Figure 6.12.

The secondary model of interest is the high-voltage electronic device that can convert and carry large amount of electric power as DC current by synchronous switching of the three-phase voltage input. Most original converters were mercury arc tube devices but all new equipment is based on solid-state thyristors.

High voltage direct current (HVDC) static converters have been commercially used for over 50 years. Various techniques for simulating HVDC systems on transient network analyzers (TNAs), analog and hybrid computers, and digital computers have been documented in the literature. TNAs have traditionally been a popular means of studying interconnected AC–DC systems. There are two common approaches: either model the total converter system or model the converter system as an equivalent power source(s) plus load(s) of the DC system to be simulated. Many advances have been achieved to simulate all aspects of HVDC and FACT devices on digital computers.

Earlier approaches solved the systems independently. The method essentially is to calculate the equivalent current sources and loads that the DC system could require of the AC system. The equivalent current sources and loads could be simulated on the TNA, since digital computer models were not available.

High voltage direct current systems have been modeled in detail. The method is to design model components that are the miniature equivalents of actual equipment. Silicon-controlled rectifiers in the normal bridge or Graetz circuit modeled the converters. Analog and hybrid computers had been used extensively to study HVDC systems due to the fast time response of the control equipment and the valves themselves.

Digital computer programs now exist for the solution of the steady-state power flow of systems containing up to 50,000 buses. The installation of HVDC systems has increased significantly in the United States over the last two decades. The operating

www.mepcafe.com

experience of the original Pacific DC Intertie, a bi-polar ±400 kV, 1440 MW system, has shown such systems to be very economic and reliable. Many similar HVDC links have been added around the world. HVDC links are often selected to provide service in underwater high-power cables (Sweden and New Zealand) and long distance heavily loaded lines (Canada, United States, West Africa, Sweden, and Russia).

The economy and reliability of HVDC systems have encouraged the consideration of FACT devices. Several FACT devices have been installed for experimental fact finding. The benefits of such systems are numerous. The costs of these systems are approaching the economic breakeven point to provide more reliable service based on the added flexibility (optionality) of this equipment. Thus, it is evident that computer power flow programs should be extended to handle all AC and DC equipment, as well as all hybrid equipment such as FACT devices.

When power conversion devices are used, such as DC links or FACT equipment, the power flow study solution methods must be modified to accommodate the rectifiers, inverters, and any series device such as the DC line. Some auxiliary equipment such as reactors, line filters, interphase transformers, etc., may also be included.

Power conversion devices are now used for many transmission and distribution control applications. Power electronic devices include AC/DC conversion equipment as used for DC links. However, more interesting devices alter the flow on the transmission or distribution network. Such devices are referred to as FACT devices. These devices are based on technology similar to the AC/DC converters. FACT devices can be categorized as series or shunt devices. Series devices are connected such that the power flow through the device is controlled. Shunt devices are connected to vary the equivalent demand at various points of the network. Series devices are used to control the real and/or reactive power through the transmission system by compensating for the natural transmission line impedances. Shunt devices are connected to ground to vary the equivalent reactive power flow at critical points of the network. Shunt devices are used to compensate customer loads that place unusual power demands on the transmission system when power factors are not unitized.

All quantities are given in per-unit. Subscript t refers to the converter ac terminals. Subscript R refers to the rectifier and I to the inverter. The per-unit system normalizes the data. The actual values are divided by the base value to find the per-unit value:

The AC system per-unit base values are: S_{base} a specified three-phase power base, V_{base} a specified Line-to-line voltage base, I_{base} the current base which is found using the complex conjugate relationship between power and voltage, and Z_{base} the impedance base found from Ohm's law.

The dc system per-unit base values are: $S_{\text{base DC}}$ is equal to, $V_{\text{base DC}}$ is equal to specified V_{base}, $I_{\text{base DC}}$ is equal to the square root of three times I_{base}, and $Z_{\text{base DC}}$ is found by Ohm's law.

6.5 DISPATCHER POWER FLOW FOR OPERATIONAL PLANNING

Dispatcher power flow studies are used for two purposes: first, the high-voltage equipment is modeled and a solution found; second, the studies are extended to determine control settings to achieve the desired operating state. Studies in an operational setting determine what the control settings should be for the given operational

www.mepcafe.com

conditions. Studies in a planning environment determine which control equipment would provide more secure and reliable operation for various future conditions.

Operational planning is given a specified system condition. The basic questions sought from a power flow study include the following:

- What are the line and transformer loads throughout the system?
- What are the voltages throughout the system?
- What are the options to unload overloaded equipment?
- What are the options to provide continued service as equipment fails unexpectedly?

Planning engineers are particularly interested in answering these questions as they evaluate proposed changes to an existing system, which are required as the demand increases as cities and industry expand or renovate:

- New generation sites
- New transmission line locations
- Which new DC links provide the best energy flows between areas?
- Which new FACT devices will provide the necessary control to maintain the system in a secure operating system?
- Where can new demand locations be added given the present system design?

6.6 CONSERVATION OF ENERGY (TELLEGEN'S THEOREM)

The basic building block is the notion of conservation of energy (COE) at a bus and for the system as a whole. The starting equation is the conservation of energy at a bus as written by the summation of the complex power:

$$\sum_{j=k}^{l} \mathbf{S}_{ij} = 0$$

where the complex power is the summation of the real and the reactive power:

$$\mathbf{S} = P + Q\sqrt{-1}$$

Alternatively, the conservation of energy can be written as two summations, one for the real power and one for the reactive power:

$$\sum_{j=k}^{l} P_{ij} = 0$$

$$\sum_{j=k}^{l} Q_{ij} = 0$$

Figure 6.10 is repeated next and shows the analysis at bus i to show conservation of energy at bus i.

www.mepcafe.com

The complex form for bus i follows:

$$\mathbf{S}_{\text{gen},i} - \mathbf{S}_{\text{demand},i} - \mathbf{S}_{ik} - \mathbf{S}_{ij} = 0$$

The real forms are found from the application of the definition for complex (apparent), real, and reactive power. Thus, COE in the real forms is as follows:

$$P_{\text{gen},i} - P_{\text{demand},i} - P_{ik} - P_{ij} = 0$$
$$Q_{\text{gen},i} - Q_{\text{demand},i} - Q_{ik} - Q_{ij} = 0$$

The power flow on any branch connecting any two buses is found as the series flow and the shunt flow as found from the π equivalent circuit of a transmission line as shown in the model development section. The conservation at each bus, except for the slack bus, is the key equation for the power flow formulation. The slack bus is also known as the swing bus. The slack bus preserves the COE for the complete system. Thus, the slack bus provides real and reactive powers for the difference between the generation, the demand, and the total system losses and reactive power storage.

6.7 EXISTING POWER FLOW TECHNIQUES

Several techniques enable the study of electric systems and interconnections on digital computers. The various techniques have a common basis: the separation of the AC system solution into two sets of linear equations. The most widely used technique divides the real and reactive equations into two independent sets of equations. The first set is the real power equations. The second is the reactive power equations. Some techniques are in fact variants of the methods used to solve sets of simultaneous linear algebraic equations. A popular variant is the Gauss–Seidel technique that is related to relaxation techniques for simultaneous equation solution. The normal procedure is to solve the power system with the embedded control systems represented. Embedded controls include automatic control of voltage magnitude or of reactive flow by tap changing under load (TCUL) transformers. Other embedded controls include real power flow control by quadrature phase shifting transformers. Real power flow control by HVDC links is also solved as an embedded system. Other FACT device controllers are included as embedded. Having solved the power system network, the extended control systems are solved and new values are calculated for the power flowing through each piece of equipment in a suboptimal manner. More recent power flow programs use optimizing algorithms to find the optimal solution. The extended control includes regional economic dispatch of generation to minimize production costs, regional shifting of generation to reduce power losses, regional shifting of generation to reduce the power flowing through equipment, etc.

www.mepcafe.com

6.8 THE NEWTON–RAPHSON METHOD USING THE AUGMENTED JACOBIAN MATRIX

6.8.1 Power Flow Statement

Consider the *i*th bus of an *n* bus system. Tellegen's theorem requires that the energy at the bus must sum to 0:

$$\mathbf{S}_{Gi} = \mathbf{S}_{Li} + \mathbf{S}_{Ti} \tag{6.1}$$

where

$\mathbf{S}_{Gi}$ = complex generated power flowing into the bus
$\mathbf{S}_{Li}$ = complex demand power flowing out of the bus
$\mathbf{S}_{Ti}$ = complex transmitted power flowing out of the bus through network equipment over all connected links to adjacent buses
(Bold letters represent complex variables.)

Since the complex power can be separated into the real and the reactive components $\left(\text{where j} = \sqrt{(-1)}\right)$:

$$\mathbf{S}_{Gi} = P_{Gi} + jQ_{Gi} \tag{6.2a}$$

$$\mathbf{S}_{Li} = P_{Li} + jQ_{Li} \tag{6.2b}$$

$$\mathbf{S}_{Ti} = P_{Ti} + jQ_{Ti} \tag{6.2c}$$

Then, it follows that the real and reactive equations may be solved independently:

$$P_{Gi} = P_{Li} + P_{Ti} \tag{6.3a}$$

$$Q_{Gi} = Q_{Li} + Q_{Ti} \tag{6.3b}$$

Note that both real and reactive powers are represented. Thus, there are six variables per bus whose relationship is dictated by two independent Equations 6.3a and 6.3b. The system size grows dramatically! An *n* bus system will yield 2*n* equations involving 6*n* variables. The power flow solution is made at specified demand conditions. Therefore, P_{Li} and Q_{Li} are known. There are four variables per bus: P_{Gi}, Q_{Gi}, P_{Ti}, and Q_{Ti}. The generation and demand terms are simple constants for most buses. However, the transmission terms represent the real and reactive flows on equipment (transmission lines, transformers, inductors, capacitors, FACTs devices, etc.) that are unique modeling problems. Note that generation and demand are external to the power flow relationships of the transmission network. The power system network is passive. It is treated as an *n* port network.

Many authors use either the impedance or the admittance matrix approach. Of these, the admittance method proves more suited. However, it is not necessary to use

www.mepcafe.com

either the admittance or the impedance matrix explicitly. We will formulate our problem directly from the power flow relationship of each piece of equipment. The bus admittance matrix is not used in the following. Instead the branch admittance matrix is used. We will use the generic term, link, for any network element (transmission line, transformer, etc.). Tellegen's theorem (conservation of energy) for each bus can be rewritten to separate the known injections from the unknown flows:

$$P_{Gi} - P_{Li} = P_{Ti} \tag{6.4a}$$

$$Q_{Gi} - Q_{Li} = Q_{Ti} \tag{6.4b}$$

Consider a bus with *m* link connections in a power system:

$$S_{Ti} = \sum_{j=1}^{m} S_{ij} \tag{6.5}$$

The complex equipment flow in each piece of equipment is composed of a real and a reactive term:

$$S_{ij} = P_{ij} + jQ_{ij} \tag{6.6}$$

The complex equipment flow is a function of the voltage at the bus and the complex conjugate of the current leaving the bus:

$$S_{ij} = V_i I_{ij}^* \tag{6.7}$$

The current is a function of the admittance between this bus and all other busses and the current to ground through the shunt impedance:

$$I_{ij} = Y_{ii} V_i + Y_{ij}(V_i - V_j) \tag{6.8}$$

We use the normal admittance notation relating the complex admittance to the real variables of conductance and susceptance: $Y = g + jb$. Y_{ii} gives the admittance representing the equipment connected only to the node and Y_{ij} gives the admittance between the buses:

$$S_{ij} = V_i Y_{ij} V_i^* + V_i V_i^* Y_{ii}^* - V_i V_j^* Y_{ij}^* \tag{6.9}$$

Note that this includes the "self" term of the bus admittance matrix:

$$P_{ij} = V_i^2 Y_{ij} \cos(-\gamma_{\mathbf{ij}}) - V_i V_j Y_{ij} \cos(\delta_{\mathbf{i}} - \delta_{\mathbf{j}} - \gamma_{\mathbf{ij}}) \tag{6.10a}$$

$$Q_{ij} = V_i^2 Y_{ij} \sin(-\gamma_{ij}) - V_i V_j Y_{ij} \sin(\delta_i - \delta_j - \gamma_{ij}) - V_i^2 Y_{ii} \tag{6.10b}$$

where V_i and V_j are the voltage magnitudes at bus *i* and bus *j*, Y_{ij} is the admittance magnitude of the link between bus *i* and bus *j*, δ_i and δ_j are the voltage angles at bus

www.mepcafe.com

i and bus j, and γ_{ij} is the admittance angle of the link between bus i and bus j. Note that the summation of all admittances to a bus still constitute the diagonal entry of the bus admittance matrix as is required of nodal analysis.

Note that we substitute the flow variables (P_{Ti}, Q_{Ti}) with the potential variables (V_i, δ_i). The complexity of the power flow problem should now be apparent. The transmitted real and reactive powers at a given bus will be a function of the voltage magnitude and angle at all other busses.

There are six variables at each bus: P_{Gi}, Q_{Gi}, P_{Li}, Q_{Li}, V_i, and δ_i. The demand is specified and is not normally altered. Such a model is based on the "obligation to serve" principle of the traditionally regulated environment. However, note that it is the net real and reactive power injection at each bus that is balanced by the transmitted flows. This leaves four variables and two equations. Thus, two more variables need to be specified. Mathematically, any two variables may be selected. However, when each bus is examined, only variables that are known should be specified. The choice is dictated by the device(s) connected to a particular bus. Several options are summarized in Table 6.1. These options define bus types to signify the variables that are known.

The first type (Vδ) is referred to as the swing or slack bus. It is normally a generator bus to satisfy the balance of demand and losses with generation. Thus, there are no constraints on the real or reactive power generated. One phasor quantity is selected as phase reference in any AC circuit. The generation required is determined by setting the phase angle of the voltage to 0 at the reference bus. It is also common to set the voltage magnitude to 1.0 per unit, since a generator normally has control of the voltage through the exciter. However, this is not necessary. The voltage magnitude may be solved if the reactive power is fixed. It is common practice to select a "tie" bus as the swing bus in some studies. A "tie" bus has no generation or demand attached. Thus, any real or reactive generation found after solution shows that the generation was not equal to the demand and losses. Then it would be necessary to redispatch the generation economically. Only one slack should be chosen per system island.

The type PQ bus would identify any bus for which P_{Gi} and Q_{Gi} are known. This includes any bus with no generation. Type PQ busses are the most common. Solve for the unknown V_i, δ_i variables at such busses.

The last bus type PV ("voltage controlled") is separated into two classifications because of control differences. The voltage can be controlled by an exciter at a generator bus, by a transformer, or by a FACT device. Different emulation calculations

TABLE 6.1 Bus Types for Power Flow Formulation

Bus Type	Code	Known	Unknowns
Slack generator	Vδ	V_i, δ_i	P_{Gi}, Q_{Gi}
Slack demand or tie	Qδ	Q_{Gi}, δ_i	P_{Gi}, V_i
Demand	PQ	P_{Gi}, Q_{Gi}	V_i, δ_i
Generator	PV	P_{Gi}, V_i	Q_{Gi}, δ_i
Controlled voltage magnitude	CV	P_{Gi}, V_i	δ_i, α

www.mepcafe.com

are used at these buses. The type PV bus is typically a bus with a generator connected to it. The two main control actions available at a generator plant enable control of P_i and V_i. Since these values are controlled, they should be specified as known. Generator operating characteristics require that the operation stays within the capability of the generator. Alternatively, other reactive power equipment may provide such control.

The limits on Q_{Gi} are unusual in that the reactive power to support a given voltage is not solved within the matrix calculations. Instead, it is calculated based on the latest update to the voltage magnitudes and angles. After this calculation, check to see if the required reactive generation is within limits. If it is not within limits, set it at the appropriate limit and release the constraint that V_i is fixed. That is, V_i and Q_{Gi} exchange roles. This changes the type of the bus from PV to PQ. During subsequent iterations, continue to check the reactive power needed to support the voltage desired. Whenever the required reactive power falls within acceptable limits, change the bus type to PV. Note that this process may occur more than once during a solution.

The second type of voltage control is a "CV" bus using a TCUL transformer. The real and reactive powers at the bus are fixed, but the setting of the tap controls the voltage. This does change the bus type at the controlled bus to a CV. The bus is modeled as normally done for the real power equation. However, the reactive equation is replaced by an equation relating the dependency between the voltage magnitude and the transformer tap ratio magnitude (α). Note that only one device should be controlling the voltage magnitude at a bus each iteration of the solution process. Such models are extended to FACT devices that provide similar control.

6.9 MATHEMATICAL OVERVIEW

The widely used method of solving AC power flow problems is the Newton–Raphson method (Newton's point form) method. The method relies on the solution of a vector-matrix nonlinear equation:

$$\mathbf{F}(\mathbf{X}) = 0 \tag{6.11}$$

where **F** is a vector-valued nonlinear function of a vector-valued argument **X**. An initial guess, $\mathbf{X}^{(O)}$, gives the vector-matrix Taylor expansion of **F**(**X**) about **X** as follows:

$$\mathbf{F}(\mathbf{X}) = 0 = \mathbf{F}(\mathbf{X}^{(O)}) + J\Delta\mathbf{X} + \text{higher order terms}$$

$$(\mathbf{J})_{ij} = \frac{\partial F_i}{\partial X_j} \tag{6.12}$$

where **J** is the system Jacobian matrix, F_i is the ith component of F, and X_j is the jth component of X. In addition, ΔX is a first-order correction to $X^{(O)}$.

www.mepcafe.com

Excluding the slack bus (node), there are usually two equations for each node of the following form:

$$\Delta P_i = \sum \frac{\partial P_{ij}}{\partial \delta_j} \quad \Delta\delta_j + \sum V_j \frac{\partial P_{ij}}{\partial V_j}\frac{\Delta V_j}{V_j}$$
$$\Delta Q_i = \sum \frac{\partial Q_{ij}}{\partial \delta_j} \quad \Delta\delta_j + \sum V_j \frac{\partial Q_{ij}}{\partial V_j}\frac{\Delta V_j}{V_j} \tag{6.13}$$

$$\Delta P_i = P_i(\text{specified}) - P_i(\text{calculated})$$
$$\Delta Q_i = Q_i(\text{specified}) - Q_i(\text{calculated})$$

where the summations are over all connections to adjacent buses, j, for the mismatch at bus i. The P_i (specified) and Q_i (specified) are the net real and net reactive power at bus i. Thus, the procedure is to calculate the partials for each bus, calculate the ΔP_i and ΔQ_i for each bus by the aforementioned equations, and test for convergence (each residual less than a prespecified tolerance). This assembles a matrix of the general form (6.15) where $(\mathbf{A})^{\mathrm{T}}$ signifies transposition and $\mathbf{A}_{\mathrm{t}}$ signifies a submatrix of matrix **A**. Note that (A_{ij}) denotes an element of matrix A.

$$\mathbf{A}x = b$$
$$x = \mathbf{A}^{-1}b$$

$$b = \left[\Delta P_2, \quad \Delta P_3,\ldots,\Delta P_n, \quad \Delta Q_2, \quad \Delta Q_3,\ldots,\Delta Q_n\right]^{\mathrm{T}}$$
$$x = \left[\Delta\delta_2, \quad \Delta\delta_3,\ldots,\Delta\delta_n, \quad \frac{\Delta V_2}{V_2}, \quad \frac{\Delta V_3}{V_3},\ldots,\frac{\Delta V_n}{V_n}\right]^{\mathrm{T}} \tag{6.14}$$

$$\mathbf{A} = \begin{bmatrix} A_{ii} & A_{ij} \\ A_{ji} & A_{jj} \end{bmatrix} \tag{6.15}$$

$$A_{ii} = \begin{bmatrix} \dfrac{\partial P_{ii}}{\partial \delta_i} & V_i \dfrac{\partial P_{ii}}{\partial V_i} \\ \dfrac{\partial Q_{ii}}{\partial \delta_i} & V_i \dfrac{\partial Q_{ii}}{\partial V_i} \end{bmatrix} \quad A_{ij} = \begin{bmatrix} \dfrac{\partial P_{ij}}{\partial \delta_j} & V_j \dfrac{\partial P_{ij}}{\partial V_j} \\ \dfrac{\partial Q_{ij}}{\partial \delta_j} & V_j \dfrac{\partial Q_{ij}}{\partial V_j} \end{bmatrix}$$

The next step is to solve the augmented matrix $[\mathbf{A} \mid b]$ by Gaussian elimination. The equations are processed one bus (one ΔP_i and ΔQ_i equation) at a time. The use of chain-linked data structures optimizes efficiency and speed. A well-known method for the rapid solution of these equations requires the back substitution of an augmented matrix. This completes the solution to find x (the voltage magnitude and angle for each bus).

www.mepcafe.com

6.10 AC SYSTEM CONTROL MODELING

The Newton–Raphson method has been expanded to directly simulate the steady-state effect of power system controls. The original work of Britton, Peterson, and Meyer provided an expanded view of the Newton–Raphson procedure as proposed by Van Ness. The Britton paper presented a rather thorough development of over 12 control models for the power flow algorithm. Typical controls that are simulated include local and remote voltage control by reactive generation, line flow control by generation, local and remote voltage control by TCUL transformers, reactive line flow control by TCUL transformers, and real line flow control by phase shifting transformers.

6.11 LOCAL VOLTAGE CONTROL

The control model for local voltage control simulates a generating station's capability to maintain a specified voltage magnitude at the system bus to which it is connected. Thus, there is only a real power equation, ΔP, with the partial derivative with respect to the voltage magnitude, $\partial P/\partial V$, deleted since the voltage magnitude is constant:

$$\Delta P_i = \frac{\partial P_{ii}}{\partial \delta_i} \Delta \delta_i + \text{partials for adjacent buses} \tag{6.16}$$

The models for the other control modes are each presented in Appendix 6A.

6.12 MODELING OF TRANSMISSION LINES AND TRANSFORMERS

6.12.1 Transmission Line Flow Equations

Transmission lines are modeled as shown in Figure 6.11. The flow as measured at each end of each transmission line is given in the following equations.

Given the voltage at each end of the transmission line:

$$\text{Voltage at bus } i\text{: } V_i = |V_i| \angle \delta_i$$
$$\text{Voltage at bus } j\text{: } V_j = |V_j| \angle \delta_j$$

Given the series and the shunt impedance of the line:

$$\text{Series impedance: } y_{ij} = |y_{ij}| \angle \gamma_{ij}$$
$$\text{Shunt impedance: } y_{ii} = |y_{ii}| \angle 90^\circ$$

Then the current from bus i to bus j measured at bus i is as follows:

$$\begin{aligned} I_{ij} &= V_i y_{ii} + (V_i - V_j) y_{ij} \\ &= V_i (y_{ii} + y_{ij}) - V_j y_{ij} \end{aligned}$$

www.mepcafe.com

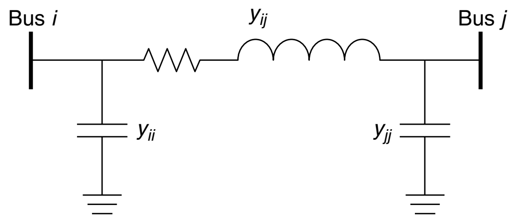


FIGURE 6.11 Transmission line model.

The power flowing from bus i to bus j measured at bus i is as follows:

$$\mathbf{S_{ij}} = \mathbf{V_i I_{ij}^*} \tag{6.17}$$

Substituting the current relationship,

$$\mathbf{S_{ij}} = |\mathbf{V_i}|^2 \mathbf{y_{ii}^*} + |\mathbf{V_i}|^2 \mathbf{y_{ij}^*} - \mathbf{V_i V_j^* y_{ij}^*} \tag{6.18}$$

This simplifies to

$$\mathbf{P_{ij}} = |\mathbf{V_i}|^2 |\mathbf{y_{ii}}| \cos(\gamma_{ij}) - |\mathbf{V_i}||\mathbf{V_j}||\mathbf{y_{ij}}| \cos(\delta_i - \delta_j - \gamma_{ij}) \tag{6.19a}$$

$$\mathbf{Q_{ij}} = -|\mathbf{V_i}|^2 |\mathbf{y_{ii}}| - |\mathbf{V_i}|^2 |\mathbf{y_{ij}}| \sin(\gamma_{ij}) - |\mathbf{V_i}||\mathbf{V_j}||\mathbf{y_{ij}}| \sin(\delta_i - \delta_j - \gamma_{ij}) \tag{6.19b}$$

These are used as the starting point for all partial derivatives.

6.12.2 Transformer Flow Equations

Transformers are modeled as shown in Figure 6.12. The flow as measured at each end of the transformer is given in the following equations.

Given the voltage at each end of the transformer: $V_m = \alpha V_j$ where α is the tap ratio and is a complex number in the case of a phase shift transformer.

Given the series impedance and tap of the transformer:

$$\text{Series impedance: } y_{ij} = |y_{ij}| \angle \gamma_{ij}$$
$$\text{Tap ratio: } \alpha_{ij} = |\alpha_{ij}| \angle \theta_t$$

Then the current from bus *i* to bus *j* measured at bus *i* is:

$$I_{ij} = (V_i - V_m) y_{ij}$$
$$I_{ij} = V_i y_{ij} - V_m y_{ij}$$
$$I_{ij} = V_i y_{ij} - \alpha_{ij} V_j y_{ij}$$

Note that α_{ij} is complex number in this model with magnitude $|\alpha_{ij}|$ and phase angle θ_t.

www.mepcafe.com

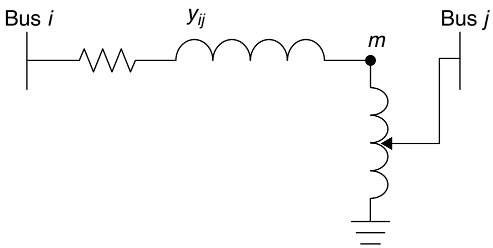


FIGURE 6.12 Transformer model.

The resulting power flow is:

$$S_{ij} = V_i I_{ij}^* = V_i V_i^* y_{ij}^* - \alpha_{ij}^* V_i V_j^* y_{ij}^* \tag{6.20}$$

Separate into two equations:

$$P_{ij} = |V_i|^2 |y_{ij}| \cos\gamma_{ij} - |\alpha_{ij}||V_i||V_j||y_{ij}| \cos(\delta_i - \delta_j - \theta_t - \gamma_{ij}) \tag{6.21}$$

$$Q_{ij} = -|V_i|^2 |y_{ij}| \sin\theta_{ij} - |\alpha_{ij}||V_i||V_j||y_{ij}| \sin(\delta_i - \delta_j - \theta_t - \gamma_{ij})$$

Correspondingly for the other bus:

$$P_{ji} = -\alpha|V_i||V_j||y_{ij}| \cos(\delta_j - \delta_i + \theta_t - \gamma_{ij}) + \alpha^2 |V_j|^2 |y_{ij}| \cos(\theta_t + \theta_{ij}) \tag{6.22}$$

$$Q_{ji} = -|\alpha_{ij}||V_i||V_j||y_{ij}| \sin(\delta_j - \delta_i + \theta_t - \gamma_{ij}) + \alpha_{ij}^2 |V_j|^2 |y_{ij}| \sin(-\theta_t - \theta_{ij})$$

Other models are also used depending on how the transformer was built, what variable tap positions are available, and how the transformer is connected into the transmission grid. Transformers are used to control voltages, real power flows, and (rarely) reactive flows.

Note that the flow equations are the normal starting point to find the partial derivatives necessary for the Jacobian matrix within industrial programs.

6.13 HVDC LINKS

The HVDC converter can be represented as a black box with a 3Φ AC line entering one side and one or two DC lines entering the other side (see Figure 6.13).

www.mepcafe.com

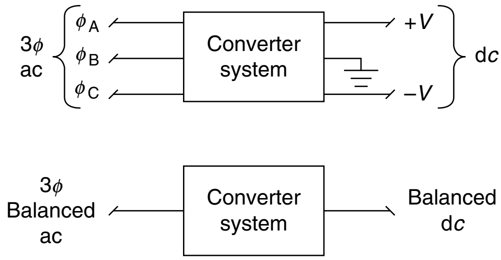


FIGURE 6.13 Single-line diagram.

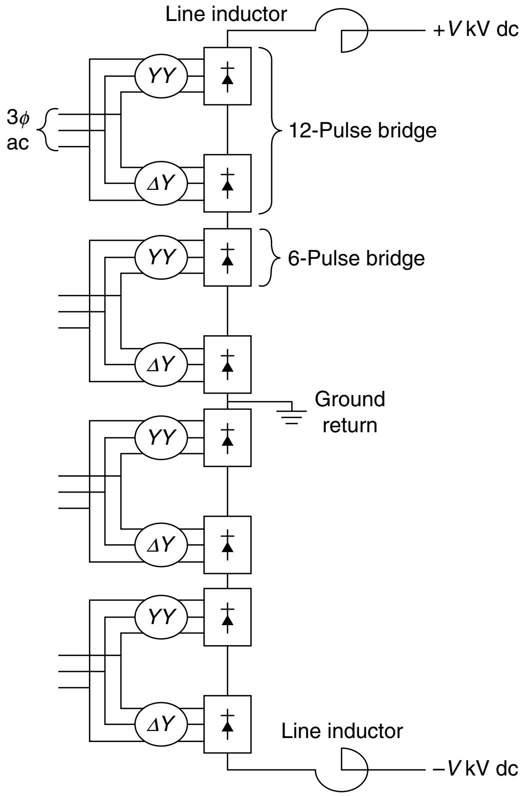


FIGURE 6.14 Converter system.

The black box represents either a 6- or a 12-pole converter that is capable of rectification or inversion. The actual converter system might be more accurately displayed as in Figure 6.14. This example shows a bipolar, eight-bridge system. Steady-state operation can reduce the previous diagram to a single line bipolar model shown in Figure 6.15. Note that this representation is symmetric with respect to an

www.mepcafe.com

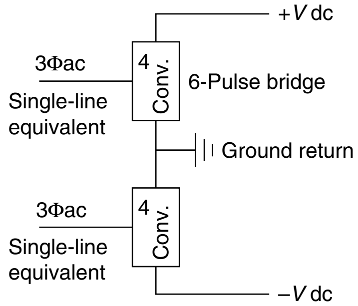


FIGURE 6.15 Single-line bipolar model.

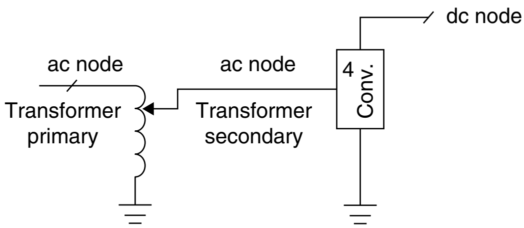


FIGURE 6.16 Monopolar model.

axis drawn through the ground return. This suggests that a monopolar model, Figure 6.16, could be adopted if the user accepted the task of reducing the system data to this model. One proposal is to use two of the monopolar models in Figure 6.16 to represent the bipolar model in Figure 6.15. This assumes that the ground return will not be a represented node in the power flow model. The equations for an AC bus i connected to a DC bus, k, are as follows:

$$\Delta P_i = \sum \frac{\partial P_{ij}}{\partial \delta_j}\Delta\delta_j + \sum V_j \frac{\partial P_{ij}}{\partial V_j}\frac{\Delta V_j}{V_j} + V_k \frac{\partial P_{ij}}{\partial V_k}\frac{\Delta V_k}{V_k} + \text{partials for adjacent buses}$$
$$\Delta Q_i = \sum \frac{\partial Q_{ij}}{\partial \delta_j}\Delta\delta_j + \sum V_j \frac{\partial Q_{ij}}{\partial V_j}\frac{\Delta V_j}{V_j} + V_k \frac{\partial Q_{ij}}{\partial V_k}\frac{\Delta V_k}{V_k} + \text{partials for adjacent buses} \quad (6.23)$$

The V_k is the average DC voltage measured with respect to ground. Compared with the general equations presented in the mathematical overview, the only change is the term corresponding to the phase angle at bus k is missing. There is no reactive power associated with any HVDC node in steady-state operation with ripple-free HVDC current.

www.mepcafe.com

The equation for a DC node, k, connected to an AC node, i, and to another DC node, j, by an HVDC transmission line is

$$\Delta P_k = V_i \frac{\partial P_{ki}}{\partial V_i}\frac{\Delta V_i}{V_i} + \frac{\partial P_{ki}}{\partial \delta_i} \quad \Delta\delta_i + V_k \frac{\partial P_{kk}}{\partial V_k}\frac{\Delta V_k}{V_k} + V_k \frac{\partial P_{kj}}{\partial V_j}\frac{\Delta V_j}{V_j} \tag{6.24}$$

The DC node is solved only for the voltage magnitude as for any DC circuit.

6.13.1 Modeling of HVDC Converters and FACT Devices

The limitations of present technology require that a converter draw substantial amounts of reactive power from the AC system to which it is connected. The reactive power is supplied by capacitors. Additionally, passive and active filters are added to remove unwanted harmonic interaction.

6.13.2 Definition of Angular Relationships in HVDC Converters

The primary concept is to find a relationship between the real and reactive flows with respect to the system voltage magnitudes and angles and any controlling parameters, such as firing angle and commutation angle. It is necessary to be able to find the first partial derivatives of these flows with respect to the voltages and the control variables. The same is true for many FACT devices since most such devices are back-to-back AC/DC/AC converters.

The complete, steady-state internal operation of a converter can largely be described by two quantities: the firing angle, α, and the commutation angle, γ. Figure 6.17 shows the relationships between the commonly defined angles for the six-pole converter equivalent circuit shown. This description uses the firing angle, α, to describe the converter mode of operation: rectification or inversion.

6.13.3 Power Equations for a Six-Pole HVDC Converter

Representative Jacobian entries for an HVDC converter are presented. The analysis is significantly simplified if the following assumptions are used. The firing voltage and the arc-drop voltage are negligible (i.e., ideal diodes). The operation of an adjacent bridge has no effect upon any other bridge. This comment is applicable when several bridges are connected in series to make up the total installation. The DC is constant and ripple-free. The power system is a balanced three-phase sinusoidal voltage of constant magnitude and frequency (infinite bus). The source impedance may be lumped with the converter transformer. The magnetizing and eddy current components of the transformer are negligible. The converter has no active power loss (the commutating resistance, R_k, is negligible).

The equations describing the steady state behavior of a monopolar DC link are summarized with the following equations:

www.mepcafe.com

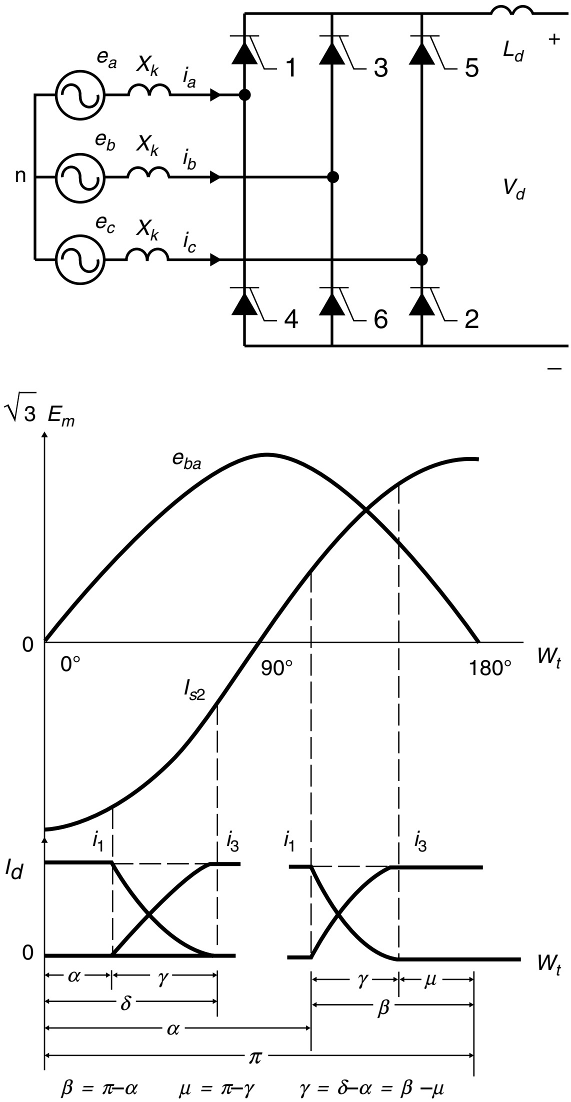


FIGURE 6.17 HVDC converter angle relationships.

$$
\begin{aligned}
V_{dR} &= \frac{3\sqrt{2}}{\pi} a_R V_{tR} \cos(\alpha_R) - \frac{3}{\pi} X_c I_d \\
V_{dI} &= \frac{3\sqrt{2}}{\pi} a_I V_{tI} \cos(\gamma_I) - \frac{3}{\pi} X_c I_d \\
V_{dR} &= V_{dI} + R_d I_d \\
P_{dR} &= V_{dR} I_d
\end{aligned}
\tag{6.25}
$$

www.mepcafe.com

$$P_{dI} = V_{dI} I_d$$
$$S_{dR} = k\frac{3\sqrt{2}}{\pi} a_R V_{tR} I_d$$
$$S_{dI} = k\frac{3\sqrt{2}}{\pi} a_I V_{tI} I_d$$
$$Q_{dR} = \sqrt{S_{dR}^2 - P_{dR}^2}$$
$$Q_{dI} = \sqrt{S_{dI}^2 - P_{dI}^2}$$

Note that k is assumed to be constant ($k = 0.995$). As with the extended variable unified method (see Section 6.19), the mismatch equations at the converter terminals AC buses have to be modified. The notation used is t for terminal bus number, R is converter that is rectifying the AC to DC, and I is the converter that is inverting DC to AC.

The DC variables satisfy the circuit equations for the DC system:

$$R\left(V_{tR}, V_{tI}, X_{dc}\right) = 0$$

Where X_{dc} is the seven DC variables. R is a set of equations given by the first three equations in (6.25) and four control specifications. The eliminated variable method solves the previous equations for X_{dc}. The real and reactive powers consumed by the converters can then be written as functions of V_{tR} and V_{tI}. It is only necessary to find a sequence of computations that the real and reactive powers and the partial derivatives with respect to the AC terminal voltages can be computed:

$$X_{DC} = f\left(V_{tR}, V_{tI}\right)$$
$$P_{dR} = P_{dR}\left(V_{tR}, V_{tI}\right)$$
$$P_L = R_d I_d^2$$
$$Q_L = \frac{3}{\pi} X_C I_d^2 \tag{6.26}$$
$$I_{ac} = \frac{3\sqrt{2}}{\pi} k I_d$$
$$k \approx 0.995$$
$$k_\alpha = \frac{k}{\cos(\alpha)}$$
$$k_\gamma = \frac{k}{\cos(\gamma)}$$

Where X_c is the commutating reactance and R_d is the dc line resistance.

www.mepcafe.com

Thus, the following partial derivatives are easily derived. The terms for the AC bus are as follows:

$$|V_{\mathrm{LL}}|\frac{\partial P_{\mathrm{AC}}}{\partial|V_{\mathrm{LL}}|} = P_{\mathrm{AC}}, \quad \frac{\partial P_{\mathrm{AC}}}{\partial \delta} = 0.0,$$

$$|V_{\mathrm{LL}}|\frac{\partial Q_{\mathrm{AC}}}{\partial|V_{\mathrm{LL}}|} = Q_{\mathrm{AC}}, \quad \frac{\partial Q_{\mathrm{AC}}}{\partial \delta} = 0.0,$$

$$V_{\mathrm{dc}}\frac{\partial P_{\mathrm{AC}}}{\partial V_{\mathrm{DC}}} = P_{\mathrm{AC}}, \quad V_{\mathrm{DC}}\frac{\partial Q_{\mathrm{AC}}}{\partial \delta} = Q_{\mathrm{AC}}$$

The terms for the DC bus are as follows:

$$V_{\mathrm{DC}}\frac{\partial P_{\mathrm{DC}}}{\partial V_{\mathrm{DC}}} = P_{\mathrm{DC}}$$

$$|V_{\mathrm{LL}}|\frac{\partial P_{\mathrm{DC}}}{\partial|V_{\mathrm{LL}}|} = P_{\mathrm{DC}}, \quad \frac{\partial P_{\mathrm{DC}}}{\partial \delta} = 0.0$$

These are the needed derivatives for the Jacobian matrix.

The resulting Jacobian matrix equation to be solved is as follows:

$$\begin{bmatrix} \Delta P_1 \\ \Delta Q_1 \\ \Delta P_2 \\ \Delta Q_2 \\ \vdots \end{bmatrix} = [J] \begin{bmatrix} \Delta\theta_1 \\ \dfrac{\Delta|V_1|}{|V_1|} \\ \Delta\theta_2 \\ \dfrac{\Delta|V_2|}{|V_2|} \\ \vdots \end{bmatrix} \tag{6.27}$$

Appendix 6B discusses the control of HVDC links.

6.14 BRIEF REVIEW OF JACOBIAN MATRIX PROCESSING

The full details of the manipulation of the Jacobian will not be given here since these details are well documented in the literature. A brief procedure is shown in Table 6.2.

The elementary form of the correction to the bus voltage vector is simply the factorization of the Jacobian matrix, augmented with the real and reactive residuals. After the factorization, the updates to the voltage magnitudes and angles are found in the column where the residuals were located.

The inclusion of DC links causes the voltage magnitude update vector to increase in dimension without increasing the dimensionality of the vector $\Delta\delta$. Similarly, the

www.mepcafe.com

TABLE 6.2 Power Flow Procedure

Step	Input	Process	Output
1	Equipment data	Merge into tables	Tabular data
2	Equipment data, initial solution or latest solution	Build Jacobian matrix	Jacobian matrix
3	Equipment data, initial solution or latest solution	Calculate residual mismatch at each bus	Net power injection at all buses, all power flows, and control statuses
4	Maximum mismatch	Compare maximum mismatch with acceptable tolerance, check iteration count	Decision to continue with next iteration
5	Jacobian matrix as well as real and reactive power mismatches	Solve AJM	Updates to voltage magnitude and angle as well as control variable updates
6		Repeat steps 2 through 6	

AJM, augmented Jacobian matrix.

ΔP vector increases in dimension but the ΔQ vector does not. Hence, the Jacobian remains square but the dimensions of the submatrices differ.
The digital computer solution is divided into the following functions:

I. Form the Jacobian matrix.
II. Augment the Jacobian with the column vector.
III. Solve the augmented Jacobian matrix (AJM) for the correction vector.
IV. Update the solution vector.

Three functions are typically used: SETUP_J, SOLVE_J, and NEWVAR.
Each routine is described in sequence.

SETUP_J

This function has two major functions:

1. Build the AJM.
2. Check each equation for solution within a specified tolerance.

SETUP_J builds the AJM bus-by-bus. All of the partial derivatives for a bus are calculated simultaneously with the flow calculations to determine the mismatch. The mismatch is actually the error for the equations. A linked list data structure is used to conserve space used.

SOLVE_J

This function performs Gaussian elimination and back substitution on the AJM using sparsity programming and linked list data techniques. An unusual feature is the limit status. This contains a code that is used to determine if there are one, two,

www.mepcafe.com

or no equations associated with each bus. The corrections are left in memory for the variable updating routine. Note that as control variables hit limits, the control may be relaxed and afterward reinstated.

NEWVAR

This function updates the independent variables (voltage magnitude and angle) by applying the corrections calculated by the solution subroutine. The corrected variables are then used to build the next iteration's Jacobian matrix and calculate the new mismatch vector.

6.15 EXAMPLE 6A: AC POWER FLOW CASE

The six-bus network shown in Figure 6.18 will be used to demonstrate several aspects of power flows here and in other chapters of this book.

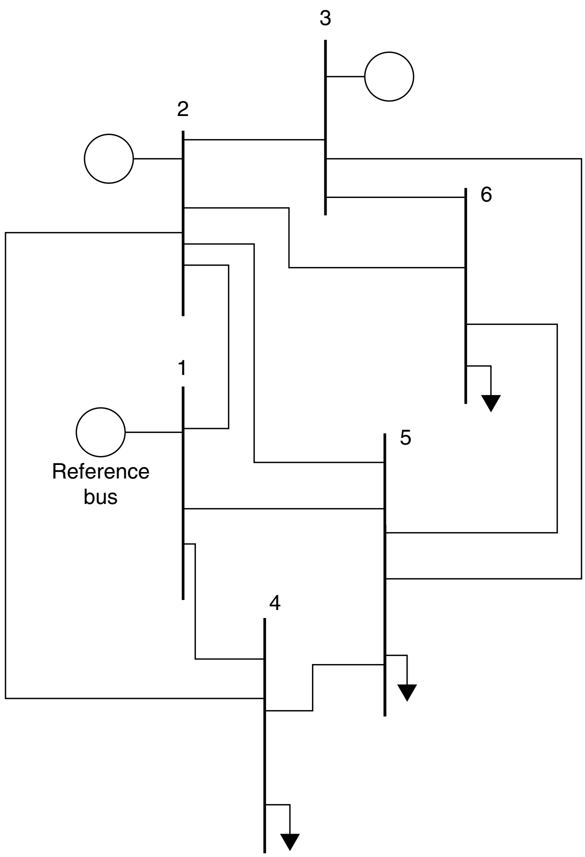


FIGURE 6.18 Six-bus sample system.

www.mepcafe.com

This example has a "base case" of 300 MW and 45 MVAR total load. The impedance values and other data for this system are shown here:

Bus Data

Bus Number	Type	P_{load}	Q_{load}	G_s	B_s	Area	V_{mag}	θ	baseKV	Zone	V_{max}	V_{min}
1	3	0	0	0	0	1	1	0	230	1	1.07	0.95
2	2	0	0	0	0	1	1	0	230	1	1.07	0.95
3	2	0	0	0	0	1	1	0	230	1	1.07	0.95
4	1	100	15	0	0	1	1	0	230	1	1.07	0.95
5	1	100	15	0	0	1	1	0	230	1	1.07	0.95
6	1	100	15	0	0	1	1	0	230	1	1.07	0.95

Type 3 = swing bus, type 2 = generator bus, type 1 = load bus

Generator Data

Bus	P_{gen}	Q_{gen}	Q_{max}	Q_{min}	V_{gen}	MVABase	Status	P_{max}	P_{min}
1	110	0	150	−100	1.07	100	1	200	50
2	50	0	150	−100	1.05	100	1	150	37.5
3	50	0	120	−100	1.05	100	1	180	45

Branch Data

Frombus	Tobus	r	x	b	rateA	rateB	rateC	Ratio	Angle	Status
1	2	0.1	0.2	0.04	100	0	0	0	0	1
1	4	0.05	0.2	0.04	100	0	0	0	0	1
1	5	0.08	0.3	0.06	100	0	0	0	0	1
2	3	0.05	0.25	0.06	60	0	0	0	0	1
2	4	0.05	0.1	0.02	60	0	0	0	0	1
2	5	0.1	0.3	0.04	60	0	0	0	0	1
2	6	0.07	0.2	0.05	60	0	0	0	0	1
3	5	0.12	0.26	0.05	60	0	0	0	0	1
3	6	0.02	0.1	0.02	60	0	0	0	0	1
4	5	0.2	0.4	0.08	60	0	0	0	0	1
5	6	0.1	0.3	0.06	60	0	0	0	0	1

Generator Cost Function Data

CostCurveType	StartUp	ShutDown	NumCoeff	c	b	a
2	0	0	3	0.00533	11.669	213.1
2	0	0	3	0.00889	10.333	200
2	0	0	3	0.00741	10.833	240

Power Flow Execution Results

Voltages are all within their limits of 0.95–1.07 pu; the reference bus is scheduled at 1.07 pu and is therefore labeled as UL for upper limit.

www.mepcafe.com

All reactive power limits are being met and all line flows are within their MW limits.

All loads are at the base case value of 100 MW and 15 MVAR.

The real and reactive power losses are small.

Newton Raphson Convergence Steps

Iter	MAXDP	MAXDPbus	MAXDQ	MAXDQbus	NumVARlim	NumVlim
1	0.825895	6	0.618486	6	0	1
2	0.097117	2	0.142304	5	0	1
3	0.001265	2	0.002313	5	0	1

POWER FLOW RESULTS

Total Pgen	=	328.58 MW	Total Qgen	=	67.74	MVAR
Total PLoad	=	300.00 MW	Total Qload	=	45.00	MVAR
Total PLosses	=	28.58 MW	Total QLosses	=	22.74	MVAR

Bus	Pmin MW	Pgen MW	Pmax MW	Qmin MVAR	Qgen MVAR	Qmax MVAR	Pload MW	Qload MVAR	Vmin pu	Vbus pu		Vbus kV	Vmax pu
1	50.0	228.6	450.0	-100.0	-32.2	150.0	0.0	0.0	0.95	1.07	UL	246.1	1.07
2	37.5	50.0	150.0	-100.0	75.7	150.0	0.0	0.0	0.95	1.05		241.5	1.07
3	45.0	50.0	180.0	-100.0	24.2	120.0	0.0	0.0	0.95	1.05		241.5	1.07
4							100.0	15.0	0.95	0.99		228.0	1.07
5							100.0	15.0	0.95	1.01		233.3	1.07
6							100.0	15.0	0.95	1.02		235.4	1.07

Bus	Vmag kV	angle deg	Pgen MW	Qgen MVAR	Pload MW	Qload MW	To Bus	Pline MW	Qline MVAR		Max Flow MW
1	246.1	0.00	228.58	-32.18	0.00	0.00					
							2	123.57	-35.59	UL	100
							5	105.01	3.42	UL	100
2	241.5	-14.46	50.00	75.71	0.00	0.00					
							1	-109.27	59.71	UL	100
							3	11.35	-5.42		60
							4	93.09	17.20	UL	60
							5	15.60	5.30		60
							6	39.22	-1.07		60
3	241.5	-15.99	50.00	24.20	0.00	0.00					
							2	-11.29	-0.89		60
							5	8.92	7.52		60
							6	52.37	17.57		60
4	228.0	-19.08			100.00	15.00					
							2	-89.01	-11.12	UL	60
							5	-10.99	-3.88		60
5	233.3	-16.57			100.00	15.00					
							1	-97.27	19.08		100
							2	-15.33	-8.74		60
							3	-8.72	-12.41		60
							4	11.23	-3.68		60
							6	10.08	-9.24		60
6	235.4	-18.58			100.00	15.00					
							2	-38.25	-1.51		60
							3	-51.81	-16.91		60
							5	-9.95	3.42		60

limits_indicator

Bus	Pgen	Qgen	Vgen
1	228.58	-32.18	1.0700 UL
2	50.00	75.71	1.0500
3	50.00	24.20	1.0500

Symbols: UL = Upper Limit Reached or exceeded
pu = per unit

6.16 THE DECOUPLED POWER FLOW

The Newton power flow is the most robust power flow algorithm used in practice. However, one drawback to its use is the fact that the terms in the Jacobian matrix must be recalculated each iteration and then the entire set of linear equations in Equation 6.27 must also be resolved each iteration.

Since thousands of complete power flows are often run for a planning or operations study, ways to speed up this process were sought. One way to speed up the calculation is to use a technique known as the "fast decoupled power flow" (it is often referred to as the "Stott decoupled power flow," in reference to its first author).

www.mepcafe.com

Starting with the terms in the Jacobian matrix, the following simplifications are made:

- Neglect and interaction between P_i and any $|V_k|$ (it was observed by power system engineers that real power was little influenced by changes in voltage magnitude—so this effect was incorporated in the algorithm). Then, all the derivatives

$$\frac{\partial P_i}{(\partial |V_k| / |V_k|)}$$

will be considered to be 0.

- Neglect any interaction between Q_i and θ_k (see the note mentioned earlier—a similar observation was made on the insensitivity of reactive power to changes in phase angle). Then, all the derivatives

$$\frac{\partial Q_i}{\partial \theta_k}$$

are also considered to be 0.

- Let $cos(\theta_i - \theta_j) \cong 1$, which is a good approximation since $(\theta_i - \theta_j)$ is usually small.
- Assume that

$$G_{ik} \sin(\theta_i - \theta_k) \ll B_{ik}$$

- Assume that

$$Q_i \ll B_{ii} |V_i|^2$$

This leaves the derivatives as

$$\frac{\partial P_i}{\partial \theta_k} = -|V_i||V_k|B_{ik} \tag{6.28}$$

$$\frac{\partial Q_i}{(\partial |E_k| / |E_k|)} = -|V_i||V_k|B_{ik} \tag{6.29}$$

If we now write the power flow adjustment equations as

$$\Delta P_i = \left(\frac{\partial P_i}{\partial \theta_k}\right)\Delta\theta_k \tag{6.30}$$

$$\Delta Q_i = \left[\frac{\partial Q_i}{(\partial |V_k| / |V_k|)}\right]\frac{\Delta|V_k|}{|V_k|} \tag{6.31}$$

then substituting Equation 6.28 into Equation 6.30, and Equation 6.29 into Equation 6.31, we obtain

$$\Delta P_i = -|V_i||V_k|B_{ik}\Delta\theta_k \tag{6.32}$$

www.mepcafe.com

$$\Delta Q_i = -|V_i||V_k|B_{ik}\frac{\Delta|V_k|}{|V_k|} \quad (6.33)$$

Further simplification can then be made:
Divide Equations 6.32 and 6.33 by $|V_i|$.
Assume $|V_k| \cong 1$ in Equation 6.33.

This results in

$$\frac{\Delta P_i}{|V_i|} = -B_{ik}\Delta\theta_k \quad (6.34)$$

$$\frac{\Delta Q_i}{|V_i|} = -B_{ik}\Delta|V_k| \quad (6.35)$$

We now build Equations 6.34 and 6.35 into two matrix equations:

$$\begin{bmatrix} \frac{\Delta P_1}{|V_1|} \\ \frac{\Delta P_2}{|V_2|} \\ \vdots \end{bmatrix} = \begin{bmatrix} -B_{11} & -B_{12} & \cdots \\ -B_{21} & -B_{22} & \cdots \\ \vdots & & \end{bmatrix} \begin{bmatrix} \Delta\theta_1 \\ \Delta\theta_2 \\ \vdots \end{bmatrix} \quad (6.36)$$

$$\begin{bmatrix} \frac{\Delta Q_1}{|V_1|} \\ \frac{\Delta Q_2}{|V_2|} \\ \vdots \end{bmatrix} = \begin{bmatrix} -B_{11} & -B_{12} & \cdots \\ -B_{21} & -B_{22} & \cdots \\ \vdots & & \end{bmatrix} \begin{bmatrix} \Delta|V_1| \\ \Delta|V_2| \\ \vdots \end{bmatrix} \quad (6.37)$$

Note that both Equations 6.36 and 6.37 use the same matrix. By further simplification, however, we shall make them different.

Simplifying the $\Delta P - \Delta\theta$ relationship of Equation 6.34:
Assume $r_{ik} \ll x_{ik}$; this changes $-B_{ik}$ to $-1/x_{ik}$.
Eliminate all shunt reactances to ground.
Eliminate all shunts to ground that arise from autotransformers.

Simplifying the $\Delta Q - \Delta|E|$ relationship of Equation 6.35:

Omit all effects from phase shift transformers.

www.mepcafe.com

The resulting equations are as follows:

$$\begin{bmatrix} \dfrac{\Delta P_1}{|V_1|} \\ \dfrac{\Delta P_2}{|V_2|} \\ \vdots \end{bmatrix} = [B'] \begin{bmatrix} \Delta\theta_1 \\ \Delta\theta_2 \\ \vdots \end{bmatrix} \tag{6.38}$$

$$\begin{bmatrix} \dfrac{\Delta Q_1}{|V_1|} \\ \dfrac{\Delta Q_2}{|V_2|} \\ \vdots \end{bmatrix} = [B''] \begin{bmatrix} \Delta |V_1| \\ \Delta |V_2| \\ \vdots \end{bmatrix} \tag{6.39}$$

where the terms in the matrices are

$$B_{ik} = -\frac{1}{x_{ik}}, \text{ assuming a branch from } i \text{ to } k \text{ (0 otherwise)}$$

$$B_{ii} = \sum_{k=1}^{N_{\text{bus}}} \frac{1}{x_{ik}}$$

$$B_{ik} = -B_{ik} = -\frac{x_{ik}}{r_{ik}^2 + x_{ik}^2}$$

$$B_{ii} = \sum_{k=1}^{N_{\text{bus}}} -B_{ik}$$

The decoupled power flow has several advantages and disadvantages over the Newton power flow. (Note: Since the introduction and widespread use of the decoupled power flow, the Newton power flow is often referred to as the “full Newton” power flow.)

Advantages:

B' and B'' are constant; therefore, they can be calculated once and, except for changes to B'' resulting from generation volt-ampere reactive (VAR) limiting, they are not updated.

Since B' and B'' are each about one-quarter of the number of terms in [*J*] (the full Newton power flow Jacobian matrix), there is much less arithmetic to solve Equations 6.38 and 6.39.

Disadvantages:

The decoupled power flow algorithm may fail to converge when some of the underlying assumptions (such as $r_{ik} \ll x_{ik}$) do not hold. In such cases, one must switch to using the full Newton power flow.

www.mepcafe.com

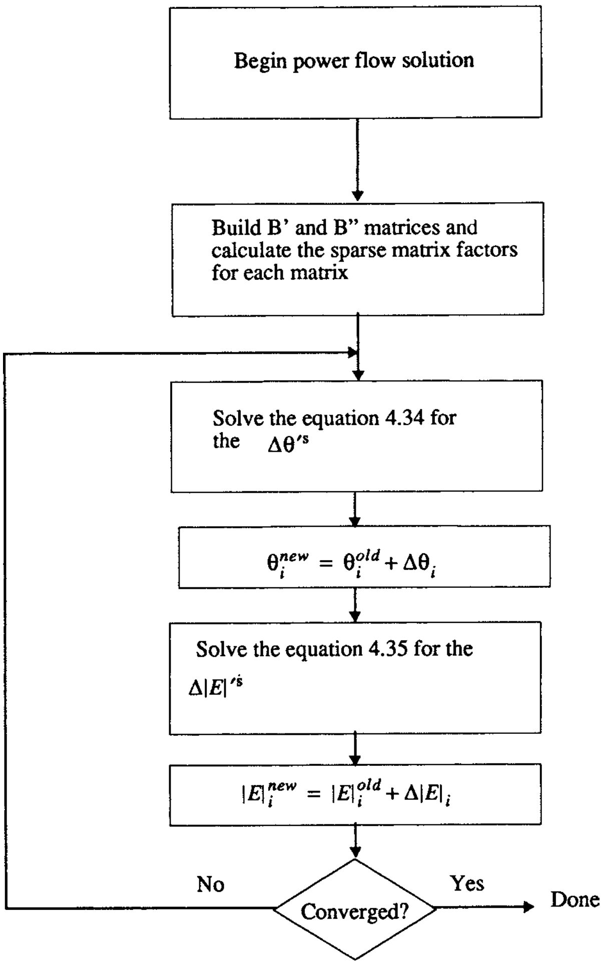


FIGURE 6.19 Decoupled power flow algorithm.

Note that Equation 6.38 is often referred to as the $P-\theta$ equation and Equation 6.39 as the $Q-V$ equation.

A flowchart of the decoupled power flow algorithm is shown in Figure 6.19.

6.17 THE GAUSS–SEIDEL METHOD

The Gauss–Seidel method was the first AC power flow method to be developed for solution on digital computers. This method is characteristically slow in solving due to its slow convergence, and often difficulty is experienced with unusual network conditions such as negative reactance branches.

www.mepcafe.com

The equation of net power injection at a bus is usually written as

$$P_k + jQ_k = V_k I_k^* = V_k \left(\sum_{j=1}^{N_{\text{bus}}} Y_{kj} V_j \right)^*$$

The voltages at each bus can be solved for by using the Gauss–Seidel method. The equation in this case is

$$V_k^{(\alpha)} = \frac{1}{Y_{kk}} \frac{(P_k - jQ_k)}{V_k^{(\alpha-1)^*}} - \frac{1}{Y_{kk}} \left[\sum_{j<k} Y_{kj} V_j^{(\alpha)} + \sum_{j>k} Y_{kj} V_j^{(\alpha-1)} \right]$$

where $V_k^{(\alpha)}$ is the new or updated voltage for bus k at iteration α; if the bus is a voltage-controlled bus, the updated voltage is readjusted to maintain its magnitude at the scheduled value.

Here each bus is updated one at a time. A pass through all the buses constitutes an iteration; then the process restarts for another iteration by updating the buses in the same order as the first iteration. The process usually takes hundreds of iterations to converge.

A comparison of the convergence of the Gauss–Seidel, decoupled, and Newton power flow algorithms is shown in Figure 6.20.

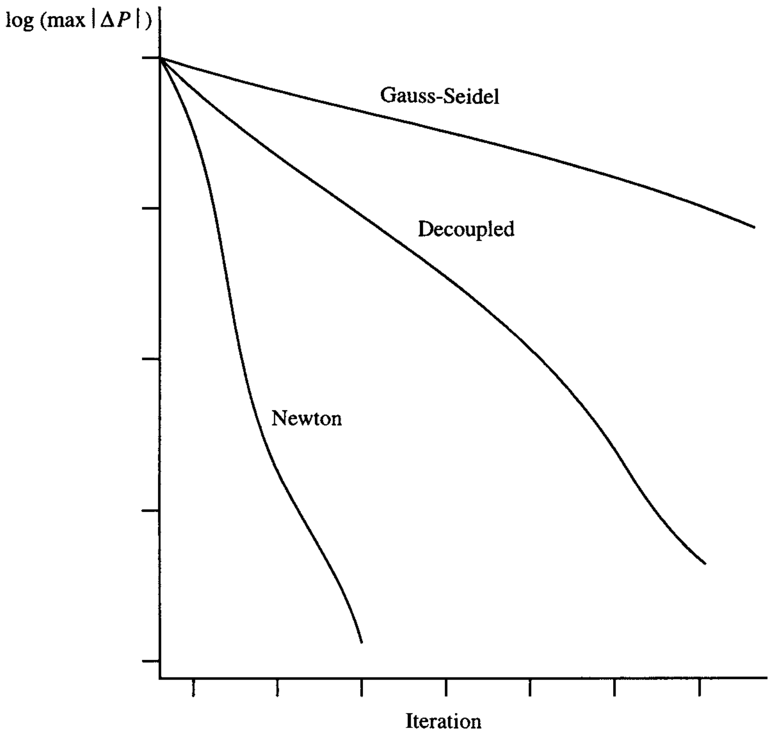


FIGURE 6.20 Comparison of three power flow algorithm convergence characteristics.

www.mepcafe.com

6.18 THE "DC" OR LINEAR POWER FLOW

A further simplification of the power flow algorithm involves simply dropping the $Q-V$ equation (Equation 6.39) altogether. This results in a completely linear, noniterative power flow algorithm. To carry this out, we simply assume that all $|V_i| = 1.0$ per unit. Then Equation 6.38 becomes

$$\begin{bmatrix} \Delta P_1 \\ \Delta P_2 \\ \vdots \end{bmatrix} = [B'] \begin{bmatrix} \Delta\theta_1 \\ \Delta\theta_2 \\ \vdots \end{bmatrix} \tag{6.40}$$

where the terms in B' are as described previously. The DC power flow is only good for calculating MW flows on transmission lines and transformers. It gives no indication of what happens to voltage magnitudes or MVAR or MVA flows. The power flowing on each line using the DC power flow is then

$$P_{ik} = \frac{1}{x_{ik}}(\theta_i - \theta_k) \tag{6.41}$$

and

$$P_i = \sum_{\substack{k=\text{buses} \\ \text{connected to } i}} P_k$$

6.18.1 DC Power Flow Calculation

The megawatt flows on the network in Figure 6.21 will be solved using the DC power flow. The B' matrix equation is

$$\begin{bmatrix} 7.5 & -5.0 \\ -5.0 & 9.0 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}$$
$$\text{and we assume } \theta_3 = 0$$

Note that all megawatt quantities and network quantities are expressed in pu (per unit on 100 MVA base). All phase angles will then be in radians.

The solution to the preceding matrix equation is as follows:

$$\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 0.2118 & 0.1177 \\ 0.1177 & 0.1765 \end{bmatrix} \begin{bmatrix} 0.65 \\ -1.00 \end{bmatrix} = \begin{bmatrix} 0.02 \\ -0.1 \end{bmatrix}$$

The resulting flows are shown in Figure 6.22 and calculated using Equation 6.39. Note that all flows in Figure 6.22 were converted to actual megawatt values.

www.mepcafe.com

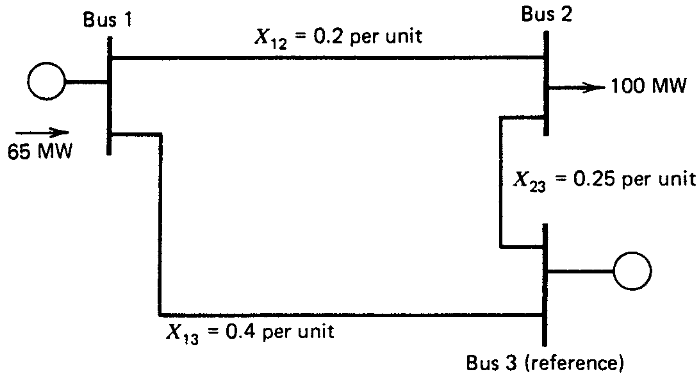


FIGURE 6.21 Three-bus network.

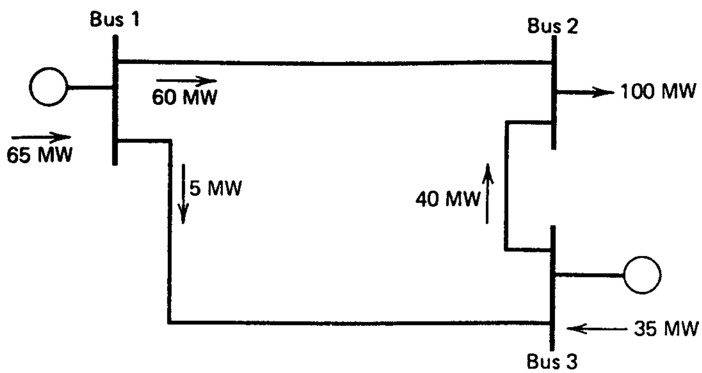


FIGURE 6.22 Three-bus network showing flows calculated by DC power flow.

6.18.2 EXAMPLE 6B: DC Power Flow Example on the Six-Bus Sample System

The network of Example 6A was solved using the DC power flow with resulting power flows as shown in Figure 6.23. The DC power flow is useful for rapid calculations of real power flows, and, as will be shown later, it is very useful in security analysis studies.

6.19 UNIFIED ELIMINATED VARIABLE HVDC METHOD

The solution of HVDC links, especially multiterminal, requires fewer assumptions than those with the extended variable unified method. The Unified Eliminated Variable method simplifies the extended variable unified method as the states are not increased but the interaction of the AC and DC system is solved directly through the Jacobian. This method alters the original Newton Raphson algorithm with far fewer changes and can be implemented with a Fast Decoupled Power algorithm. This method does not require that the linked AC systems are synchronous. The basic idea is to treat the real and reactive powers

www.mepcafe.com

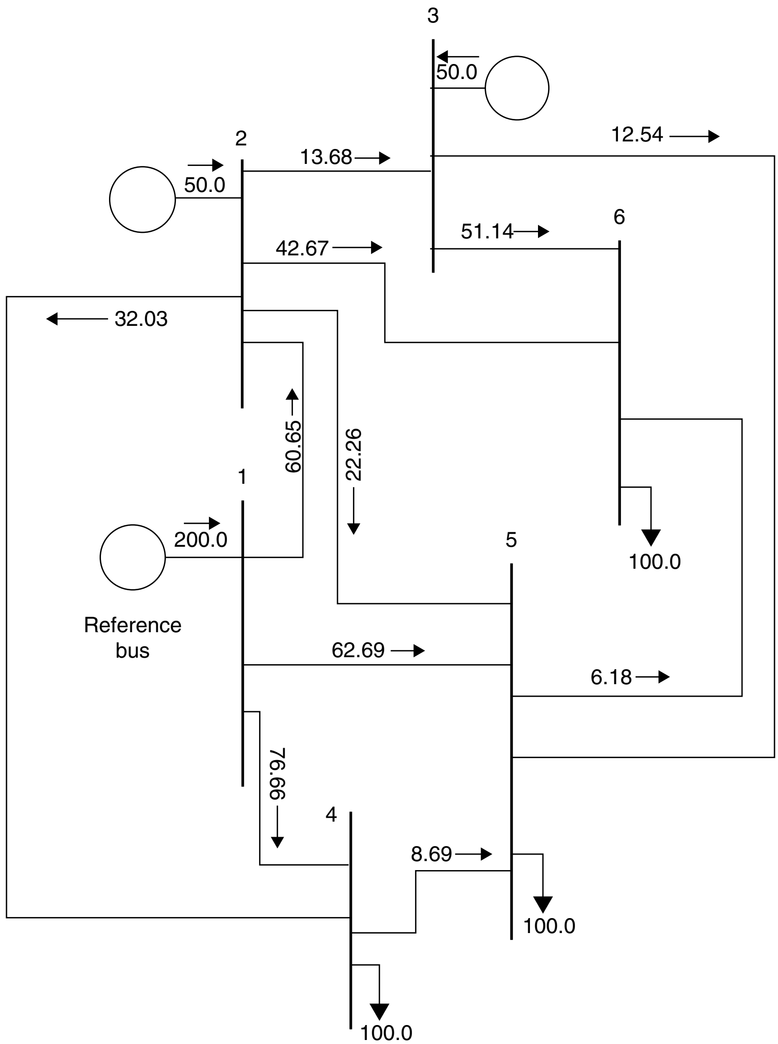


FIGURE 6.23 Six-bus network base case DC power flow for Example 6A.

as voltage dependent demands. The DC equations are solved analytically or numerically. The DC variables are eliminated from the power flow equations. The method is unified since the Jacobian sensitivities are updated with the effect of the DC link.

6.19.1 Changes to Jacobian Matrix Reduced

Then it is only necessary to alter the N and L submatrices of the Newton Raphson algorithm:

$$\begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} = \begin{bmatrix} H & N' \\ J & L' \end{bmatrix} \begin{bmatrix} \Delta \delta \\ \dfrac{\Delta V}{V} \end{bmatrix}$$

www.mepcafe.com

The partial derivatives can be found by the following equations. Explicit functions are not required:

$$N'(tR,tR) = V_{tR}\frac{\partial P_{tR}^{ac}}{\partial V_{tR}} + V_{tR}\frac{\partial P_{dR}(V_{tR},V_{tI})}{\partial V_{tR}}$$

$$N'(tR,tI) = V_{tI}\frac{\partial P_{tR}^{ac}}{\partial V_{tI}} + V_{tI}\frac{\partial P_{dR}(V_{tR},V_{tI})}{\partial V_{tI}}$$

$$N'(tI,tR) = V_{tR}\frac{\partial P_{tI}^{ac}}{\partial V_{tR}} - V_{tR}\frac{\partial P_{dI}(V_{tR},V_{tI})}{\partial V_{tR}}$$

$$N'(tI,tI) = V_{tI}\frac{\partial P_{tI}^{ac}}{\partial V_{tI}} - V_{tI}\frac{\partial P_{dI}(V_{tR},V_{tI})}{\partial V_{tI}}$$

L' is modified analogously. Only four powers mismatch equations and up to eight elements of the Jacobian have to be modified. No new variables are added to the solution vector. The DC variables are not kept constant. The Jacobian in (6.19) has been found to be more numerically stable (well-conditioned) than the Jacobian of the extended variable method.

6.19.2 Control Modes

Since there are seven variables and three independent equations for a DC link, four specifications have to be made to define a unique solution. The control modes in the following table are used to illustrate the analytical elimination method.

Control mode A is the based case, the well-known current margin control corresponds to one terminal controlling the voltage and the other one the current, or equivalently, the power. The control angles and the DC voltage are specified. The transformer tap positions are varied to meet these specifications. The other modes in Table 6.3 are obtained from mode A, if variables hit their limits during the power flow computations or if the time scale does not allow the taps to be moved. The HVDC scheme determines which modes are obtained when limits are encountered. The rectifier tap angle determines the firing angle and the inverter tap angle determines the direct voltage for modes B–D. The rectifier tap ratio determines the direct voltage for modes E–G. Subscript I refers to constant current control. The tap ratios are assumed to be continuous variables.

6.19.3 Analytical Elimination

The analytical elimination is detailed for Control Mode A. It is sufficient to find P_d and S_d at each converter, since Q_d then can be computed with equations (6.25). The partial derivatives for several modes are shown in the following.

www.mepcafe.com

Table 6.3 Control Mode Specified Variables

Control Mode	Specified Variables			
A	αR	γ_I	V_{dI}	P_{dI}
B	aR	γ_I	V_{dI}	P_{dI}
C	αR	γ_I	a_I	P_{dI}
D	αR	γ_I	a_I	P_{dI}
E	αR	γ_I	a_R	P_{dI}
F	αR	a_I	V_{dI}	P_{dI}
G	αR	a_I	a_R	P_{dI}
AI	αR	γ_I	V_{dI}	Id
BI	aR	γ_I	V_{dI}	Id
CI	αR	γ_I	a_I	Id
DI	aR	γ_I	a_I	Id
EI	αR	γ_I	a_R	Id
FI	αR	a_I	V_{dI}	Id
GI	αR	a_I	a_R	Id

Since both the voltage and power at the inverter are specified, the direct current can be computed and P_{dR} can then be found by combining equations in (6.25):

$$P_{dR} = P_{dI} + R_d I_d^2$$

S_{dR} can be found by combining terms in (6.25) to obtain the following:

$$S_{dR} = k_\alpha \left(P_{dI} + P_L + Q_L \right)$$

Analogously, S_{dI} is found to be the following:

$$S_{dI} = k_\gamma \left(P_{dI} + Q_L \right)$$

Thus, all real and reactive powers consumed by the converters can be precomputed. Including the DC link in the power flow is trivial for this control mode. The same is true for many specifications.

Control Mode B occurs if the tap changer at the rectifier hits a limit in control mode A under current control in the rectifier. Since P_{dI} and V_{dI} are specified, I_d, V_{dR}, P_{dR}, and S_{dI} are computed as for mode A. Since a_R is specified, S_{dR} is computed with the S_{dR} expression in (6.25) equations instead of the earlier equation:

$$V_{tR} \frac{\partial S_{dR}}{\partial V_{tR}} = S_{dR}$$

$$V_{tR} \frac{\partial Q_{dR}}{\partial V_{tR}} = \frac{S_{dR}^2}{Q_{dR}}$$

The formulas for mode B_I are essentially identical. The only difference is that P_{dI}, rather than I_d, is computed with (6.25). When two of the variables are specified, the other three can be computed from (6.25) in general.

www.mepcafe.com

The partial derivatives for control modes A and A_I follow:

$$\frac{\partial P_{dR}}{\partial V_{tR}} = V_{tR}\frac{\partial Q_{dR}}{\partial V_{tR}} = V_{tI}\frac{\partial P_{dR}}{\partial V_{tI}} = V_{tI}\frac{\partial Q_{dR}}{\partial V_{tI}} = \frac{\partial P_{dI}}{\partial V_{tI}} = \frac{\partial Q_{dI}}{\partial V_{tR}} = V_{tI}\frac{\partial P_{dI}}{\partial V_{tI}} = V_{tI}\frac{\partial Q_{dI}}{\partial V_{tI}} = 0$$

The partial derivatives for the control modes B and B_I are:

$$\frac{\partial P_{dR}}{\partial V_{tR}} = V_{tI}\frac{\partial P_{dR}}{\partial V_{tI}} = V_{tI}\frac{\partial Q_{dR}}{\partial V_{tI}} = \frac{\partial P_{dI}}{\partial V_{tI}} = \frac{\partial Q_{dI}}{\partial V_{tR}} = V_{tI}\frac{\partial P_{dI}}{\partial V_{tI}} = V_{tI}\frac{\partial Q_{dI}}{\partial V_{tI}} = 0$$

$$V_{tR}\frac{\partial Q_{dR}}{\partial V_{tR}} = \frac{S_{dR}^2}{Q_{dR}}$$

The partial derivatives for mode C are:

$$\frac{\partial P_{dR}}{\partial V_{tR}} = V_{tR}\frac{\partial Q_{dR}}{\partial V_{tR}} = V_{tI}\frac{\partial P_{dR}}{\partial V_{tI}} = V_{tI}\frac{\partial Q_{dR}}{\partial V_{tI}} = \frac{\partial P_{dI}}{\partial V_{tI}} = \frac{\partial Q_{dI}}{\partial V_{tR}} = V_{tI}\frac{\partial P_{dI}}{\partial V_{tI}} = V_{tI}\frac{\partial Q_{dI}}{\partial V_{tI}} = 0$$

The partial derivatives for mode C_I are:

$$\frac{\partial P_{dR}}{\partial V_{tR}} = V_{tR}\frac{\partial Q_{dR}}{\partial V_{tR}} = V_{tI}\frac{\partial P_{dR}}{\partial V_{tI}} = V_{tI}\frac{\partial Q_{dR}}{\partial V_{tI}} = \frac{\partial P_{dI}}{\partial V_{tI}} = \frac{\partial Q_{dI}}{\partial V_{tR}} = V_{tI}\frac{\partial P_{dI}}{\partial V_{tI}} = V_{tI}\frac{\partial Q_{dI}}{\partial V_{tI}} = 0$$

$$V_{tI}\frac{\partial P_{dR}}{\partial V_{tI}} = 2P_{dI}\partial I_I$$

$$V_{tI}\frac{\partial Q_{dR}}{\partial V_{tI}} = \frac{\partial I_I}{\partial Q_{tR}}\left[k_a S_{dI}\left(P_L + Q_L\right) - P_{dR}P_L\right]$$

$$V_{tI}\frac{\partial Q_{dI}}{\partial V_{tI}} = \frac{S_{dI}^2}{Q_{dI}}\left(1 + \partial I_I\right)$$

The partial derivatives for the other control modes can be derived analogously. Two other examples of partial derivative calculation follow:

Numerical Elimination

The analytical elimination has a drawback that the formulas have to be rederived for other DC system configurations or if other specifications are used. The numerical elimination procedure uses finite differences to find the partial derivatives. Let $P_{dR}(V_{tR}, V_{tI})$ denote a function that solves the DC equations and computes P_{dR} given the terminal voltages. Then the following approximation can be used:

$$\frac{\partial P_{dR}\left(V_{tR}, V_{tI}\right)}{\partial V_{tI}} = \frac{P_{dR}\left(V_{tR}, V_{tI} + h\right) - P_{dR}\left(V_{tR}, V_{tI}\right)}{h}$$

www.mepcafe.com

The choice of h is a trade-off between the truncation error and the computational precision. An initial estimate of $h \approx \sqrt{toll}$. If $toll = 10^{-6}$, then $h = 10^{-3}$. This method can be used with any voltage dependent device.

6.19.4 Control Mode Switching

If a dc variable, for example a tap changer or a control angle, hits a limit during the power flow computation, a switch to another control mode is necessary. In the Newton Raphson power flow, this is a minor problem if the eliminated variable is used, since the components of the solution vector are the same.

Hitting a limit normally makes a refactorization of the reactive power portion of the approximated Jacobian necessary for the fast decoupled power flow. The fact that a maximum of four elements of the reactive power part of the approximated Jacobian can change only makes a partial refactorization necessary. If area-decoupling can be used, the switch is trivial if the corresponding mismatch equation is placed last; the appropriate value of partial derivative only has to be added and subtracted to the lower right element of L in the LU decomposition.

If area-decoupling cannot be performed, we can keep the simple refactorization by making suitable approximations, just as for removing a line in the security analysis chapter.

There are some similarities in the handling of switches in the new method and the extended variable based method for fast decoupled power flow, but we want to point out the following differences. A rather large partial refactorization, including all dc equations has to be performed every iteration. In the eliminated variable method, partial refactorization is necessary only when a variable has hit a limit. If area-decoupling or the approximated fast decoupled ac/dc power flow is used, the refactorization is trivial. In the eliminated variable method, all variables in the solution vector and all mismatch equations remain the same before and after a switch. Either some variables or some mismatch equations has to be replaced when switching between control modes in the other methods.

6.19.5 Bipolar and 12-Pulse Converters

The modeling of 12-pulse converters is straightforward. Denote the number of 6-pulse bridges by n_B. Then, if V_d is replaced by V_d/n_B in (6.25) and the right hand sides of (6.25) are multiplied by n_B, all previous results hold.

Bipolar converters could be considered as two independent DC links, except for the coupling via the ground return resistance:

$$V_{dR1} = V_{dIr} + R_d I_{dl} + R_g \left(I_{d1} - I_{da} \right)$$

The partial derivatives differs to those in the following in a nontrivial way only for the not very common case of unbalanced operation in constant power control, with nonzero ground return current (this does not include the case when one pole is out of operation, since the remaining pole can be treated as a monopole). The modification for this case consists of taking into account the extra terms introduced by the ground resistance.

If the ac systems are nonsynchronous, the power flow for the system at the voltage setting terminal should be solved first.

www.mepcafe.com

6.20 TRANSMISSION LOSSES

6.20.1 A Two-Generator System Example

We are given the power system in Figure 6.24. The losses on the transmission line are proportional to the square of the power flow. The generating units are identical, and the production cost is modeled using a quadratic equation. If both units were loaded to 250 MW, we would fall short of the 500 MW load value by 12.5 MW lost on the transmission line, as shown in Figure 6.25.

Where should the extra 12.5 MW be generated? Solve the Lagrange equation that was given in Chapter 3.

$$\mathcal{L} = F_1(P_1) + F_2(P_2) + \lambda(500 + P_{\text{loss}} - P_1 - P_2) \tag{6.42}$$

where

$$F_1(P_1) = 400 + 2P_1 + 0.002P_1^2$$
$$F_2(P_2) = 400 + 2P_2 + 0.002P_2^2$$

$$P_{\text{loss}} = 0.0002P_1^2$$

then

$$\frac{\partial \mathcal{L}}{\partial P_1} = \frac{\mathrm{d}F_1(P_1)}{\mathrm{d}P_1} - \lambda\left(1 - \frac{\partial P_{\text{loss}}}{\partial P_1}\right) = 0$$

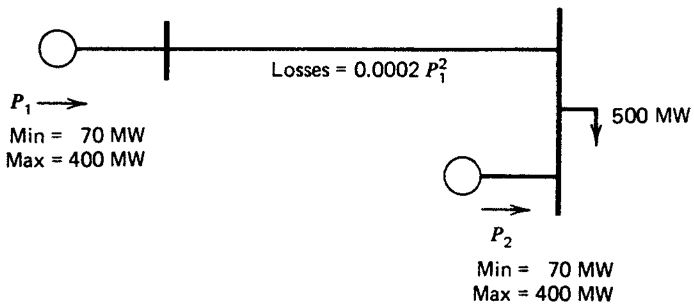


FIGURE 6.24 Two-generator system.

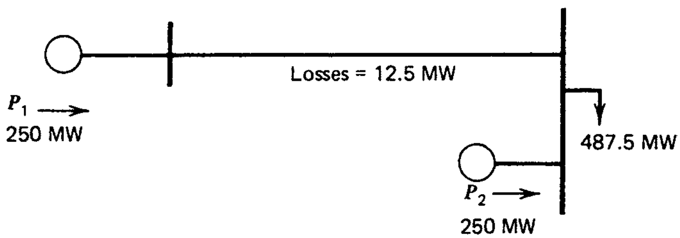


FIGURE 6.25 Two-generator system with both generators at 250 MW output.

www.mepcafe.com

$$\frac{\partial \mathcal{L}}{\partial P_2} = \frac{dF_2(P_2)}{dP_2} - \lambda\left(1 - \frac{\partial P_{loss}}{\partial P_2}\right) = 0$$
$$P_1 + P_2 - 500 - P_{loss} = 0 \tag{6.43}$$

Substituting into Equation 6.43, we obtain

$$7.0 + 0.004P_1 - \lambda(1 - 0.0004P_1) = 0$$
$$7.0 + 0.004P_2 - \lambda = 0$$
$$P_1 + P_2 - 500 - 0.0002P_1^2 = 0$$

$$\textbf{Solution}: \quad P_1 = 178.882$$
$$P_2 = 327.496$$
$$\textbf{Production cost}: \quad F_1(P_1) + F_2(P_2) = 4623.15\,\$/\text{h}$$
$$\textbf{Losses}: \quad 6.378 \text{ MW}$$

Suppose we had decided simply to ignore the economic influence of losses and ran unit 1 up until it supplied all the losses. It would need to be run at 263.932 MW, as shown in Figure 6.26. In this case, the total production cost would be

$$F_1(263.932) + F_2(250) = 4661.84\,\$/\text{h}$$

Note that the optimum dispatch tends toward supplying the losses from the unit close to the load, and it also resulted in a lower value of losses. Also note that best economics are not necessarily attained at minimum losses. The minimum loss solution for this case would simply run unit 1 down and unit 2 up as far as possible. The result is unit 2 on high limit.

$$P_1 = 102.084 \text{ MW}$$
$$P_2 = 400.00 \text{ MW (high limit)}$$

The minimum loss production cost would be

$$F_1(102.084) + F_2(400) = 4655.43\ \$/\text{h}$$
$$\text{Min losses} = 2.084 \text{ MW}$$

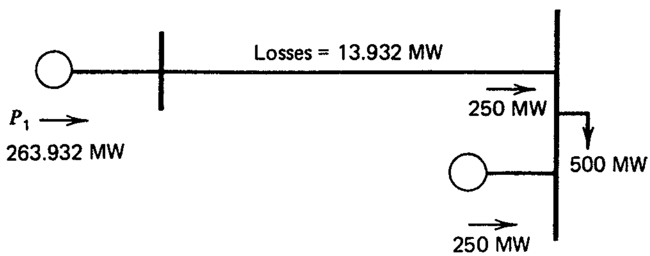


FIGURE 6.26 Two-generator system with generator 1 supplying all losses.

www.mepcafe.com

6.20.2 Coordination Equations, Incremental Losses, and Penalty Factors

The classic Lagrange multiplier solution to the economic dispatch problem was given in Chapter 3. This is repeated here and expanded.

$$\textbf{Minimize}: \quad \mathcal{L} = F_{\mathrm{T}} + \lambda\phi$$

$$\textbf{Where}: \quad F_{\mathrm{T}} = \sum_{i=1}^{N_{\mathrm{bus}}} F_i(P_i)$$

$$\phi = P_{\mathrm{load}} + P_{\mathrm{loss}}(P_1, P_2, \ldots, P_N) - \sum_{i=1}^{N_{\mathrm{bus}}} P_i$$

$$\textbf{Solution}: \quad \frac{\partial \mathcal{L}}{\partial P_i} = 0 \quad \text{for all} \;\; P_{i\min} \le P_i \le P_{i\max}$$

Then

$$\frac{\partial \mathcal{L}}{\partial P_1} = \frac{\mathrm{d}F_i}{\mathrm{d}P_i} - \lambda\left(1 - \frac{\partial P_{\mathrm{loss}}}{\partial P_i}\right) = 0$$

The equations are rearranged

$$\left(\frac{1}{1 - \dfrac{\partial P_{\mathrm{loss}}}{\partial P_i}}\right)\frac{\mathrm{d}F_i(P_i)}{\mathrm{d}P_i} = \lambda \tag{6.44}$$

where

$$\frac{\partial P_{\mathrm{loss}}}{\partial P_i}$$

is called the *incremental loss* for bus i, and

$$\mathrm{Pf}_i = \left(\frac{1}{1 - \dfrac{\partial P_{\mathrm{loss}}}{\partial P_i}}\right)$$

is called the *penalty factor* for bus i. Note that if the losses increase for an increase in power from bus i, the incremental loss is positive and the penalty factor is greater than unity.

When we did not take account of transmission losses, the economic dispatch problem was solved by making the incremental cost at eacsh unit the same. We can still use this concept by observing that the penalty factor, Pf_i, will have the following effect. For $\mathrm{Pf}_i > 1$ (positive increase in P_i results in an increase in losses),

www.mepcafe.com

$$\mathrm{Pf}_i \frac{\mathrm{d}F_i(P_i)}{\mathrm{d}P_i}$$

acts as if

$$\frac{\mathrm{d}F_i(P_i)}{\mathrm{d}P_i}$$

had been slightly increased (moved up). For $\mathrm{Pf}_i < 1$ (positive increase in P_i results in decrease in losses),

$$\mathrm{Pf}_i \frac{\mathrm{d}F_i(P_i)}{\mathrm{d}P_i}$$

acts as if

$$\frac{\mathrm{d}F_i(P_i)}{\mathrm{d}P_i}$$

had been slightly decreased (moved down). The resulting set of equations look like

$$\mathrm{Pf}_i \frac{\mathrm{d}F_i(P_i)}{\mathrm{d}P_i} = \lambda \quad \text{for all } P_{i\min} \le P_i \le P_{i\max} \tag{6.45}$$

and are called *coordination equations*. The P_i values that result when penalty factors are used will be somewhat different from the dispatch that ignores the losses (depending on the Pf_i and $\mathrm{d}F_i(P_i)/\mathrm{d}P_i$ values). This is illustrated in Figure 6.27.

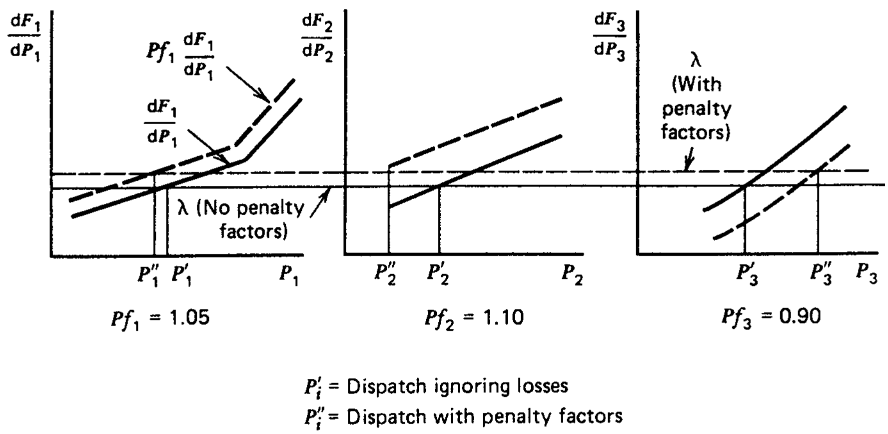


FIGURE 6.27 Economic dispatch, with and without penalty factors.

www.mepcafe.com

6.21 DISCUSSION OF REFERENCE BUS PENALTY FACTORS

An alternative approach to economic dispatch is to use a reference bus that always moves when an increment in generation is made on a different generator bus. Suppose we change the generation on bus i by ΔP_i:

$$P_i^{\text{new}} = P_i^{\text{old}} + \Delta P_i$$

Furthermore, we will assume that *load stays constanteverywhere* and that to compensate for the increase in ΔP_i, the reference bus just drops off by ΔP_{ref}:

$$P_{\text{ref}}^{\text{new}} = P_{\text{ref}}^{\text{old}} + \Delta P_{\text{ref}}$$

If nothing else changed, ΔP_{ref} would be the negative of ΔP_i; however, the flows on the system can change as a result of the two generation adjustments. The change in flow is apt to cause a change in losses so that ΔP_{ref} is not necessarily equal to ΔP_i. That is:

$$\Delta P_{\text{ref}} = -\Delta P_i + \Delta P_{\text{loss}}$$

Next, we can define β_i as the ratio of the negative change in the reference-bus power to the change ΔP_i:

$$\beta_i = \frac{-\Delta P_{\text{ref}}}{\Delta P_i} = \frac{\left(\Delta P_i - \Delta P_{\text{loss}}\right)}{\Delta P_i}$$

or

$$\beta_i = 1 - \frac{\partial P_{\text{loss}}}{\partial P_i}$$

Note that it is often useful to express the reference bus penalty factor for the reference bus itself as:

$$\beta_{\text{ref}} = 1.0$$

We can define economic dispatch as follows.

All generators are in economic dispatch when a shift of ΔP MW from any generator to the reference bus results in no change in net production cost; where ΔP is arbitrarily small.

That is, if

$$\text{Total production cost} = \sum F_i\left(P_i\right)$$

then the change in production cost with a shift ΔP_i from plant i is:

$$\Delta\text{Production cost} = \frac{\mathrm{d}F_i\left(P_i\right)}{\mathrm{d}P_i}\Delta P_i + \frac{\mathrm{d}F_{\text{ref}}\left(P_{\text{ref}}\right)}{\mathrm{d}P_{\text{ref}}}\Delta P_{\text{ref}}$$

but

$$\Delta P_{\text{ref}} = -\beta_i \Delta P_i$$

www.mepcafe.com

then

$$\Delta\text{Production cost} = \frac{dF_i(P_i)}{dP_i}\Delta P_i - \beta_i \frac{dF_{\text{ref}}(P_{\text{ref}})}{dP_{\text{ref}}}\Delta P_i$$

To satisfy the economic conditions:

$$\Delta\text{Production cost} = 0$$

or

$$\frac{dF_i(P_i)}{dP_i} = \beta_i \frac{dF_{\text{ref}}(P_{\text{ref}})}{dP_{\text{ref}}}$$

which could be written as:

$$\frac{1}{\beta_i}\frac{dF_i(P_i)}{dP_i} = \frac{dF_{\text{ref}}(P_{\text{ref}})}{dP_{\text{ref}}}$$

To obtain an economic dispatch solution, pick a value of generation on the reference bus and then set all other generators according to the previous equation, and check for total demand and readjust reference generation as needed until a solution is reached.

Further, we note the following relationship:

$$\Delta F_T = \left[\frac{dF_i}{dP_i} - \beta_i \frac{dF_{\text{ref}}}{dP_{\text{ref}}}\right]\Delta P_i$$

Then

$$\frac{\Delta F_T}{\Delta P_i} = \left[\frac{dF_i}{dP_i} - \beta_i \frac{dF_{\text{ref}}}{dP_{\text{ref}}}\right]$$

the ratio $\Delta F_T/\Delta P_i$ will be zero when the system is at optimum economic dispatch, thus:

$$0 = \left[\frac{dF_i}{dP_i} - \beta_i \frac{dF_{\text{ref}}}{dP_{\text{ref}}}\right] \quad \text{for all buses } i, i \neq \text{ref}$$

or

$$\frac{dF_i}{dP_i} = \beta_i \frac{dF_{\text{ref}}}{dP_{\text{ref}}} = \left(1 - \frac{dP_{\text{loss}}}{dP_i}\right)\frac{dF_{\text{ref}}}{dP_{\text{ref}}}$$

In Chapter 8, Section 8.11 this result is shown to be the locational marginal price, LMP, for bus *i* as a function of the reference bus LMP and incremental loss for bus *i*. The next section shows how to calculate the β_i factors directly from the AC Newton power flow.

6.22 BUS PENALTY FACTORS DIRECT FROM THE AC POWER FLOW

The penalty factors may be derived using the Newton–Raphson power flow. What we wish to know is the ratio of change in power on the reference bus when a change ΔP_i is made.

www.mepcafe.com

Where P_{ref} is a function of the voltage magnitude and phase angle on the network, when a change in ΔP_i is made, all phase angles and voltages in the network will change. Then

$$\begin{aligned}\Delta P_{ref} &= \sum_i \frac{\partial P_{ref}}{\partial \theta_i}\Delta\theta_i + \sum_i \frac{\partial P_{ref}}{\partial |V_i|}\Delta |V_i| \\ &= \sum_i \frac{\partial P_{ref}}{\partial \theta_i}\frac{\partial \theta_i}{\partial P_i}\Delta P_i + \sum_i \frac{\partial P_{ref}}{\partial |V_i|}\frac{\partial |V_i|}{\partial P_i}\Delta P_i\end{aligned} \tag{6.46}$$

To carry out the matrix manipulations, we will also need the following:

$$\begin{aligned}\Delta P_{ref} &= \sum_i \frac{\partial P_{ref}}{\partial \theta_i}\Delta\theta_i + \sum_i \frac{\partial P_{ref}}{\partial |V_i|}\Delta |V_i| \\ &= \sum_i \frac{\partial P_{ref}}{\partial \theta_i}\frac{\partial \theta_i}{\partial Q_i}\Delta Q_i + \sum_i \frac{\partial P_{ref}}{\partial |V_i|}\frac{\partial |V_i|}{\partial Q_i}\Delta Q_i\end{aligned} \tag{6.47}$$

The terms $\partial P_{ref}/\partial \theta_i$ and $\partial P_{ref}/|V_i|$ *are derived by differentiating the standard power flow equations for the reference bus. The terms* $\partial \theta_i/\partial P_i$ and $\partial |V_i|/\partial P_i$ are from the inverse Jacobian matrix (see Equation 6.15). We can write Equations 6.47 and 6.48 for every bus *i* in the network. The resulting equation is

$$\left[\frac{\partial P_{ref}}{\partial P_1}\ \frac{\partial P_{ref}}{\partial Q_1}\ \frac{\partial P_{ref}}{\partial P_2}\ \frac{\partial P_{ref}}{\partial Q_2}\cdots\frac{\partial P_{ref}}{\partial P_N}\ \frac{\partial P_{ref}}{\partial Q_N}\right] = \left[\frac{\partial P_{ref}}{\partial \theta_1}\ \frac{\partial P_{ref}}{\partial |V_1|}\ \frac{\partial P_{ref}}{\partial \theta_2}\ \frac{\partial P_{ref}}{\partial |V_2|}\cdots\frac{\partial P_{ref}}{\partial \theta_N}\ \frac{\partial P_{ref}}{\partial |V_N|}\right]\left[J^{-1}\right] \tag{6.48}$$

By transposing, we get

$$\begin{bmatrix}\frac{\partial P_{ref}}{\partial P_1}\\ \frac{\partial P_{ref}}{\partial Q_1}\\ \frac{\partial P_{ref}}{\partial P_2}\\ \frac{\partial P_{ref}}{\partial Q_2}\\ \vdots \\ \frac{\partial P_{ref}}{\partial P_N}\\ \frac{\partial P_{ref}}{\partial Q_N}\end{bmatrix} = [J^{T-1}]\begin{bmatrix}\frac{\partial P_{ref}}{\partial \theta_1}\\ \frac{\partial P_{ref}}{\partial |V_1|}\\ \frac{\partial P_{ref}}{\partial \theta_2}\\ \frac{\partial P_{ref}}{\partial |V_2|}\\ \vdots \\ \frac{\partial P_{ref}}{\partial \theta_N}\\ \frac{\partial P_{ref}}{\partial |V_N|}\end{bmatrix} \tag{6.49}$$

In practice, instead of calculating J^{T-1} explicitly, we use Gaussian elimination on J^T in the same way we operate on J in the Newton power flow solution.

www.mepcafe.com

PROBLEMS

6.1 Given the following one line diagram, find the equivalent admittance for line charging capacitance connected between each line to ground if the solution is as listed below:

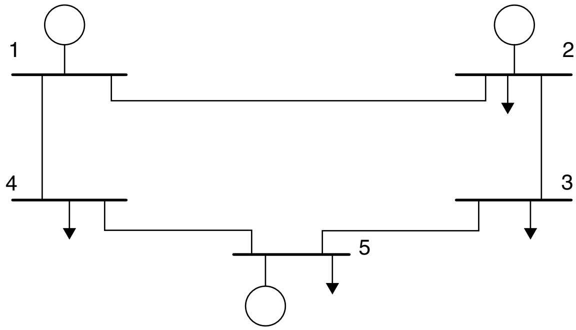


FIGURE 6.28 Network for Problem 6.1.

Bus	V magnitude (per unit)	Angle (degrees)
1	1	0
2	0.96	–5.91
3	1.1	–13.662
4	1.04	–8.718
5	1.05	–6.36

LINE DATA(pu)	BUS DATA(pu)
$Y_{12} = 2 - j5$	$S1inj = 2.1328 - j0.9607$
$Y_{14} = 1 - j10$	$S2inj = 2.0 - j1.7111$
$Y_{23} = 1 - j20$	$S3inj = -4.0 - j0.50$
$Y_{35} = 3 - j10$	$S4inj = -2.0 - j0.40$
$Y_{45} = 3 - j10$	$S5inj = 2.0 - j1.35$

6.2 The pi-equivalent per-unit model of the transformer is given as follows (Figure 6.29):

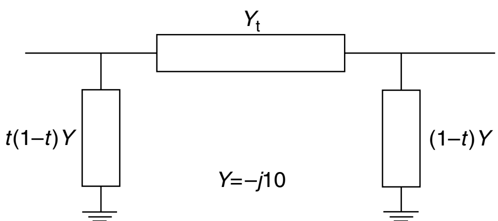


FIGURE 6.29 Pi equivalent for transformer in Problem 6.2.

www.mepcafe.com

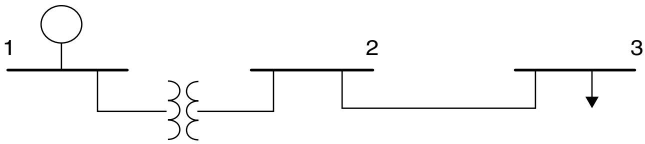


FIGURE 6.30 Network for Problem 6.2.

The turn ratio is selected such that the voltage at bus three (3) is 1.0 pu magnitude. Develop a Newton–Raphson procedure to find the tap ratio, t, and show the solution. You should assume that Z_{23} is a known impedance and your answer will be in terms of Z_{23} (Figure 6.30). (You may assume also that Z_{23} has no charging capacitance to ground.)

$$\begin{aligned}|V_1| &= 1.0\\ \delta_1 &= 0.0\\ P_2 &= 0\\ Q_2 &= 0\\ P_{D3} &= 2.0\\ Q_{D3} &= 0.5\end{aligned}$$

Data for network in Problem 6.2

6.3 Using the fast decoupled power flow for the network that follows, show the equations to find the capacitance at bus 5 such that the voltage at that bus is exactly 1.0 pu. Assume that the capacitance can be modeled as a PV bus for this solution (Figure 6.31). Why is this true?

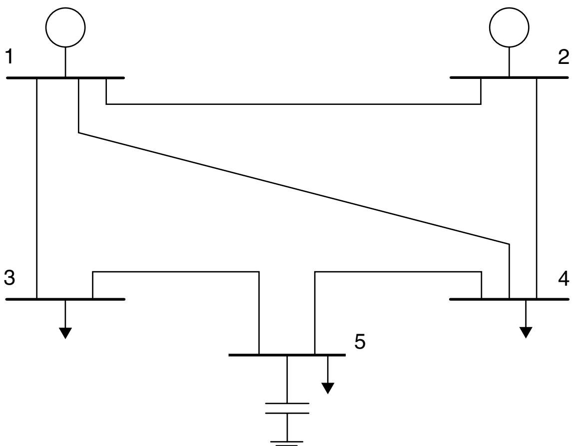


FIGURE 6.31 Network for Problem 6.3.

www.mepcafe.com

LINE DATA	BUS DATA
$Y_{12} = 1 - j10$	$\lvert V_1 \rvert = 0.98$
$Y_{13} = 1 - j10$	$\delta_1 = 0.0$
$Y_{14} = 1 - j10$	$\lvert V_2 \rvert = 0.97$
$Y_{24} = 2 - j10$	$P_{G2} = 1.0$
$Y_{35} = 1 - j5$	$P_{D3} = 0.5$
$Y_{45} = 1 - j8$	$Q_{D3} = 0.0$
$Y_{S12} = 0$	$P_{D4} = 2.0$
$Y_{S13} = 0$	$Q_{D4} = 0.5$
$Y_{S14} = 0$	$P_{D5} = 0.5$
$Y_{S24} = 0$	$Q_{D5} = 0.1$
$Y_{S35} = 0$	
$Y_{S45} = j0.5$	
$Y_{S54} = j0.5$	
$Y_{S5} = j0.5$ Capacitor to ground	

Data for network in Problem 6.3

6.4 Find the admittance of the transformer where the remaining parameters are in the following table (see Figure 6.32):

Turns Ratio (a)	$10e^{-j\pi/6}$
Z_s	$+j\ 20\ \Omega$
Z_m	$+j\ 500\ \Omega$

6.5 You are to resolve Problem 3.17 in Chapter 3 but now you are to assume that there are transmission losses in the system and the incremental losses for the generators are

$$\frac{dP_{loss}}{dP_1} = -0.05263$$

and

$$\frac{dP_{loss}}{dP_2} = 0.04762$$

Find the optimum schedule for a total power delivery of 650 MW, that is, 650 equals the load plus the losses.

6.6 Two generators are connected to a load of 1000 MW (Figure 6.33).

Given the two generating unit operating cost functions (in \$/MWh):

$$\text{Gen 1:} \quad F_1(P_1) = 2200 + 25P_1 + 0.025 \times P_1^2 \quad \text{where } 220 \le P_1 \le 600 \text{ MW}$$

$$\text{Gen 2:} \quad F_2(P_2) = 1500 + 11P_2 + 0.02 \times P_2^2 \quad \text{where } 350 \le P_2 \le 800 \text{ MW}$$

$$\text{MW losses} = 0.0002 \times P_1^2 + 0.0001 \times P_2^2 \quad \text{where } P_1 \text{ and } P_2 \text{ are in MW}$$

www.mepcafe.com

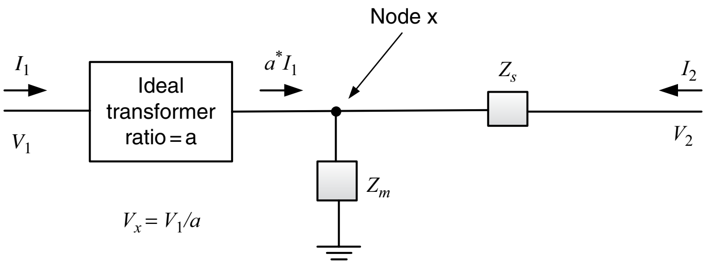


FIGURE 6.32 Network for Problem 6.4.

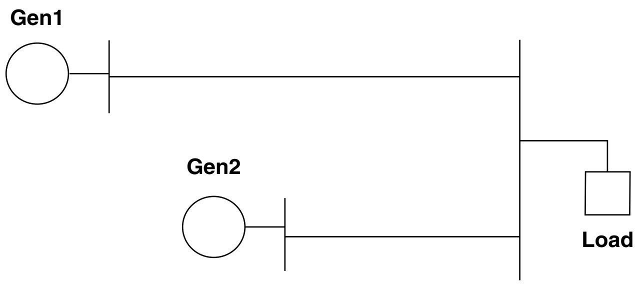


FIGURE 6.33 Power system diagram for Problem 6.6.

Solve for the economic dispatch to supply the 1000 MW load with line losses taken into account. Start from $P_1 = 500$ MW and $P_2 = 500$ MW, and give the losses win MW and the incremental losses and the penalty factors for Gen 1 and Gen 2.

Solve for P_1 and P_2 then recalculate the losses, incremental losses, penalty factors and then resolve for P_1 and P_2, etc., until the changes in P_1 and P_2 are less than 1 MW.

6.7 You are given the four-bus problem as follows (Figure 6.34):

The data for this problem are as follows:

$$x12 = 0.2;$$
$$x14 = 0.4;$$
$$x13 = 0.25;$$
$$x24 = 0.5;$$
$$x34 = 0.1;$$

a. Form the 4 × 4 **Bx** matrix for this network. Use of Matlab is strongly recommended.

The first row will be

$$\mathbf{Bx} = \left[\left(1/x12 + 1/x13 + 1/x14\right) - 1/x12 - 1/x13 - 1/x14 \right]$$

b. You cannot invert the **Bx** matrix, so you should replace all elements in the fourth column by zeros, and replace all elements in the fourth row by zeros. Now replace the term **Bx**(4,4) by 1.0 and the matrix can be inverted. The result,

www.mepcafe.com

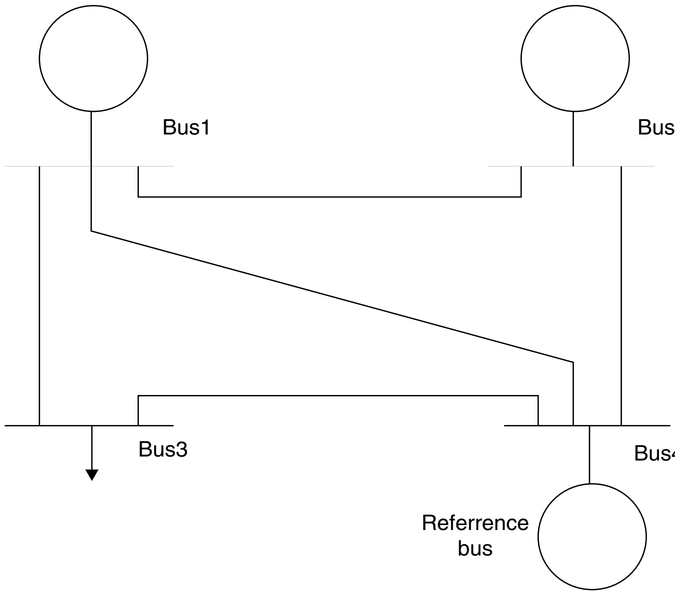


FIGURE 6.34 Four-bus network for Problem 6.7.

which we shall call the **X** matrix, will be a 3 × 3 matrix corresponding to the first three rows and columns and the fourth row and column unchanged.

c. The generation and load on the system are

$$P_{gen} \text{ on bus } 1 = 100.00 \text{ MW}$$
$$P_{gen} \text{ on bus } 2 = 150.00 \text{ MW}$$
$$P_{load} \text{ on bus } 3 = 350.00 \text{ MW}$$

Assuming bus 4 as the reference, let theta for bus 4 be 0.0 radians.

Assume S_{base} is 100 MVA.

d. You are going to solve a DC or linear power flow for this system. First convert all P_{gen} and P_{load} values to per unit, then solve for the bus phase angles on buses 1, 2, and 3. Then solve for the line flows on all transmission lines.

www.mepcafe.com

7

POWER SYSTEM SECURITY

7.1 INTRODUCTION

Up until now we have been mainly concerned with minimizing the cost of operating a power system. An overriding factor in the operation of a power system is the desire to maintain system security. System security involves practices designed to keep the system operating when components fail. For example, a generating unit may have to be taken offline because of auxiliary equipment failure. By maintaining proper amounts of spinning reserve, the remaining units on the system can make up the deficit without too low a frequency drop or need to shed any load. Similarly, a transmission line may be damaged by a storm and taken out by automatic relaying. If, in committing and dispatching generation, proper regard for transmission flows is maintained, the remaining transmission lines can take the increased loading and still remain within limit.

All equipment in a power system is designed such that it can be disconnected from the network. The reasons for these disconnections are generally divided into two categories: scheduled outages and forced outages.

Scheduled outages are typically done to perform maintenance or replacement of the equipment, and, as its name implies, the time of disconnect is scheduled by operators to minimize the impact on the reliability of the system.

Forced outages are those that happen at random and may be due to internal component failures or outside influences such as lightning, wind storms, ice buildup, etc.

Because the specific times at which forced outages occur are unpredictable, the system must be operated at all times in such a way that the system will not be left in a dangerous condition should any credible outage event occur. Since power system

Power Generation, Operation, and Control, Third Edition. Allen J. Wood, Bruce F. Wollenberg, and Gerald B. Sheblé.
© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

www.mepcafe.com

equipment is designed to be operated within certain limits, most pieces of equipment are protected by automatic devices that can cause equipment to be switched out of the system if these limits are violated. If a forced outage occurs on a system that leaves it operating with limits violated on other components, the event may be followed by a series of further actions that switch other equipment out of service. If this process of cascading failures continues, the entire system or large parts of it may completely collapse. This is usually referred to as a *system blackout*.

An example of the type of event sequence that can cause a blackout might start with a single line being opened due to an insulation failure; the remaining transmission circuits in the system will take up the flow that was flowing on the now-opened line. If one of the remaining lines is now too heavily loaded, it may open due to relay action, thereby causing even more load on the remaining lines. This type of process is often termed a *cascading outage*. Most power systems are operated such that any single initial failure event will not leave other components heavily overloaded, specifically to avoid cascading failures.

In the United States, this is referred to as the NERC ($n - 1$) rule.[1] NERC is the North American Electric Reliability Corporation that sets the reliability standards for all electric systems. $n - 1$ refers to a system with n components, and $n - 1$ is its state with one component out. The NERC ($n - 1$) rule states that no single outage will result in other components experiencing flow or voltage limit violations.

Most large power systems install equipment to allow operations personnel to monitor and operate the system in a reliable manner. This chapter will deal with the techniques and equipment used in these systems. We will lump these under the commonly used title *system security*.

System security can be broken down into three major functions that are carried out in an operations control center:

1. System monitoring
2. Contingency analysis
3. Security-constrained optimal power flow

System monitoring provides the operators of the power system with pertinent up-to-date information on the conditions on the power system. Generally speaking, it is the most important function of the three. From the time that utilities went beyond systems of one unit supplying a group of loads, effective operation of the system required that critical quantities be measured and the values of the measurements be transmitted to a central location. Such systems of measurement and data transmission, called *energy management systems* (EMS), have evolved to schemes that can monitor voltages, currents, power flows, and the status of circuit breakers and switches in every substation in a power system transmission network. In addition, other critical information such as frequency, generator unit outputs, and transformer tap positions can also be telemetered. With so much information telemetered simultaneously, no human operator could hope to check all of it in a reasonable time frame. For this reason, digital computers are installed in operations control centers to gather the telemetered data, process them, and place them in a database from which operators can display information on large display monitors.

[1] See NERC *Reliability Concepts* v 1.0.2

www.mepcafe.com

More importantly, the computer can check incoming information against prestored limits and alarm the operators in the event of an overload or out-of-limit voltage.

State estimation is often used in such systems to combine telemetered system data with system models to produce the best estimate (in a statistical sense) of the current power system conditions or "state." We will discuss some of the highlights of these techniques in Chapter 9.

Such systems are usually combined with supervisory control systems that allow operators to control circuit breakers and disconnect switches and transformer taps remotely. Together, these systems are often referred to as *SCADA systems*, standing for supervisory control and data acquisition system. The SCADA system allows a few operators to monitor the generation and high-voltage transmission systems and to take action to correct overloads or out-of-limit voltages.

The power system as seen by power system operators, whether at the highest level such as Independent System Operators in the United States or individual SCADA system operators at a small electric company, all have to deal with the power system in what has been characterized as one of four modes:

- Normal
- Alert
- Emergency
- Restoration

These variations have been used for many years, see, for example, the New York ISO operating manual.[2]

- Normal usually means that there are no alarms being presented and contingency analysis is not reporting any contingencies that would cause overloads or voltage violations.
- Alert means that either an alarm has been presented to the operator or the contingency analysis programs have presented the possibility of a contingency problem. Of course, other problems could also present the operator with a situation that would mean that he or she must take action as well.
- Emergency would indicate serious alarm messages that the operators must act on immediately and threaten to cause major shutdowns of power system equipment or even parts of the system.
- Restoration comes if the system does in fact lose equipment or part of the system or even most of it is shut down or blacked out. In restoration, equipment must be investigated to see if it can be brought back on line and then switched back into the system. Loads that were dropped are brought back on line, sometimes in small blocks. Restoration can take many hours especially if large generators are involved.

The second major security function is contingency analysis. The results of this type of analysis allow systems to be operated defensively. Many of the problems that occur on a power system can cause serious trouble within such a quick time period

[2] New York Independent System Operator, *Transmission and Dispatching Operations Manual, Manual 12*, October 2012

www.mepcafe.com

that the operator cannot take action fast enough once the process is started. This is often the case with cascading failures. Because of this aspect of systems operation, modern operations computers are equipped with contingency analysis programs that model possible system troubles before they arise. These programs are based on a model of the power system and are used to study outage events and alarm the operators to any potential overloads or out-of-limit voltages. For example, the simplest form of contingency analysis can be put together with a standard power flow program such as that described in Chapter 6, together with procedures to set up the power flow data for each outage to be studied by the power flow program. Several variations of this type of contingency analysis scheme involve fast solution methods, automatic contingency event selection, and automatic initializing of the contingency power flows using actual system data and state estimation procedures.

The third major security function is security-constrained optimal power flow. In this function, a contingency analysis is combined with an optimal power flow that seeks to make changes to the optimal dispatch of generation, as well as other adjustments, so that when a security analysis is run, no contingencies result in violations. To show how this can be done, we shall divide the power system into four operating objectives.

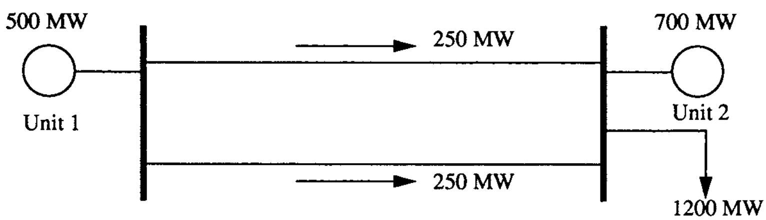
Normal state dispatch: This is the state that the power system is in prior to any contingency. It is optimal with respect to economic operation, but it may not be secure.

Postcontingency: This is the objective after a contingency has occurred. We shall assume here that this condition has a security violation (line or transformer beyond its flow limit or a bus voltage outside the limit).

Secure dispatch: This is the objective with no contingency outages is to correct the operating parameters to account for security violations.

Secure postcontingency: The objective is to re-mediate the contingency as applied to the base-operating condition with corrections.

We shall illustrate the aforementioned states with examples. Suppose the trivial power system consisting of two generators, a load, and a double circuit line is to be operated with both generators supplying the load as shown in the following (ignore losses):

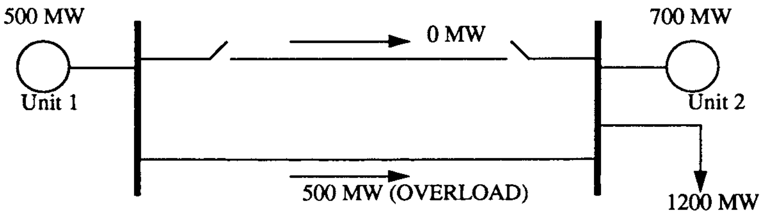


NORMAL DISPATCH

We assume that the system as shown is in economic dispatch; that is, the 500 MW from unit 1 and the 700 MW from unit 2 are the optimum dispatch. Further, we assert that each circuit of the double circuit line can carry a maximum of 400 MW, so that there is no loading problem in the base-operating condition.

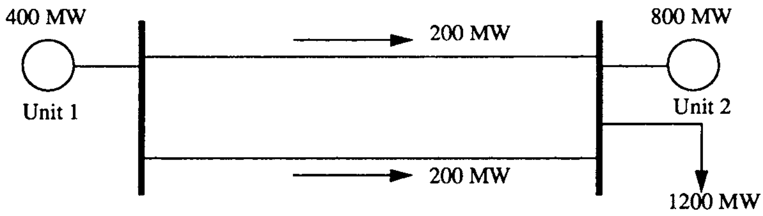
www.mepcafe.com

Now, we shall postulate that one of the two circuits making up the transmission line suffered a forced outage and opened. This results in the flows shown in the following:



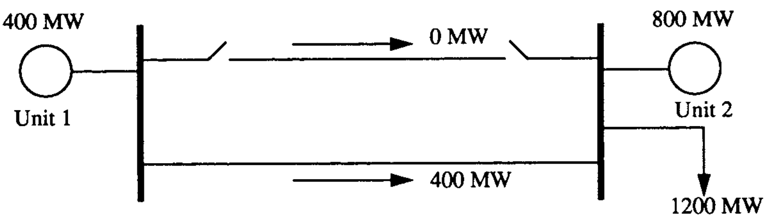
POST CONTINGENCY STATE

Now there is an overload on the remaining circuit. We shall assume for this example that we do not want this condition to arise and that we will correct the condition by lowering the generation on unit 1 to 400 MW. The secure dispatch is



SECURE DISPATCH

Now, if the same contingency analysis is done, the postcontingency condition is shown below.



SECURE POST CONTINGENCY STATE

By adjusting the generation on unit 1 and unit 2, we have prevented the postcontingency operating state from having an overload. This is the essence of what is called "security corrections." Programs that can make control adjustments to the base or precontingency operation to prevent violations in the postcontingency conditions are called "security-constrained optimal power flows" or SCOPF. These programs can take account of many contingencies and calculate adjustments to generator MW, generator voltages, transformer taps, interchange, etc. We shall show how the SCOPF is formed in Chapter 8.

Together, the functions of system monitoring, contingency analysis, and corrective action analysis comprise a very complex set of tools that can aid in the secure operation of a power system. This chapter concentrates on contingency analysis.

www.mepcafe.com

7.2 FACTORS AFFECTING POWER SYSTEM SECURITY

As a consequence of many widespread blackouts in interconnected power systems, the priorities for operation of modern power systems have evolved to the following:

Operate the system in such a way that power is delivered reliably.

Within the constraints placed on the system operation by reliability considerations, the system will be operated most economically.

The greater part of this book is devoted to developing methods to operate a power system to gain maximum economy. But what factors affect its operation from a reliability standpoint? We will assume that the engineering groups who have designed the power system's transmission and generation systems have done so with reliability in mind. This means that adequate generation has been installed to meet the load and that adequate transmission has been installed to deliver the generated power to the load. If the operation of the system went on without sudden failures or without experiencing unanticipated operating states, we would probably have no reliability problems. However, any piece of equipment in the system can fail, either due to internal causes or due to external causes such as lightning strikes, objects hitting transmission towers, or human errors in setting relays.

It is impossible to build a power system with so much redundancy (i.e., extra transmission lines, reserve generation, etc.) that failures never cause load to be dropped on a system. Rather, systems are designed so that the probability of dropping load is acceptably small. Thus, most power systems are designed to have sufficient redundancy to withstand all major failure events, but this does not guarantee that the system will be 100% reliable.

Within the design and economic limitations, it is the job of the operators to try to maximize the reliability of the system they have at any given time. Usually, a power system is never operated with all equipment "in" (i.e., connected) since failures occur or maintenance may require taking equipment out of service. Thus, the operators play a considerable role in seeing that the system is reliable.

In this chapter, we will not be concerned with all the events that can cause trouble on a power system. Instead, we will concentrate on the possible consequences and remedial actions required by two major types of failure events: transmission-line outages and generation-unit failures.

Transmission-line failures cause changes in the flows and voltages on the transmission equipment remaining connected to the system.

Therefore, the analysis of transmission failures requires methods to predict these flows and voltages so as to be sure they are within their respective limits. Generation failures can also cause flows and voltages to change in the transmission system, with the addition of dynamic problems involving system frequency and generator output.

7.3 CONTINGENCY ANALYSIS: DETECTION OF NETWORK PROBLEMS

7.3.1 Generation Outages

When a generator suffers a forced outage, it causes changes in other generators as well as changes in the transmission system.

www.mepcafe.com

Effect on Other Generations. When a generator fails, its power output is lost, and the result is an imbalance between total load plus losses and total generation. This imbalance results in a drop in frequency, which must be restored. To restore frequency back to its nominal value (50Hz or 60Hz), other generators must make up the loss of power from the outaged generator. The proportion of the lost power made up by each generator is strictly determined by its governor droop characteristic (see Chapter 10). Alternatively, we can model the external systems as needed for accuracy by swapping data through the Common Information Model (CIM) with an equivalent system at the edges. A rough approximation is to assume that each generator makes up the lost power in proportion to its own capacity. However, if a generator is at its maximum output at the time when it is called on to help restore frequency, then it cannot go up further and it cannot participate in frequency restoration. In this case, other generators will have to keep increasing until the total deficit is made up. If there was not sufficient generation that could be raised because too many were at or near their maximum, then the system would need to shed load to restore frequency. To guard against this, all generators are checked against their maximum limits. The sum of generation that can be applied to make up a loss must be greater than the largest generator output. This sum is called spinning reserve and must be sufficiently large to meet reliability criteria.

If the system being modeled is part of a large interconnected network, the lost generation will be picked up by a large number of generating units outside the system's immediate control area. When this happens, the pickup in generation is seen as an increase in flow over the tie lines to the neighboring systems. To model this, we can build a network model of our own system plus an equivalent network of our neighbor's system and place the swing bus or reference bus in the equivalent system. Alternatively, we can model the external systems as needed for accuracy by swapping data through the Common Information Model (CIM) with an equivalent system at the edges. A generator outage is then modeled so that all lost generation is picked up on the swing bus, which then appears as an increase on the tie flows, thus approximately modeling the generation loss when interconnected. If, however, the system of interest is not interconnected, then the loss of generation must be shown as a pickup in output on the other generation units within the system. An approximate method of doing this is shown in Section 7.4.1.1.

Effects on Transmission. When generation is lost, much of the made up power will come from tie lines, and this can mean line flow limit or bus voltage limit violations.

In summary, the system must monitor two things to be sure generator outages do not cause problems when one is lost: check spinning reserve at all times to be sure it is adequate and model generator outages and their effect on transmission flows and voltages.

7.3.2 Transmission Outages

When a transmission line or transformer fails and is disconnected, the flow on that line goes to zero and all flows nearby will be affected. The result can be a line flow limit or bus voltage limit violation. There is no way to know which line or transformer outage is going to cause the worst violations. The operators therefore usually want to check as

www.mepcafe.com

many of them as possible, as often as possible. Thus, the operators may seek to model and calculate the outage effects from an outage of every line and transformer in the system.

Double Outages. An even more difficult analysis is to check all pairs of possible simultaneous outages, which is denoted (n – 2). Thus, all pairs of generators, and all pairs of transmission lines as well as pairs of single generator outages plus a possible single transmission-line outage at the same time would have to be analyzed. This (n – 2) analysis is much more difficult because of the extremely large number of cases to model. The usual practice is to only study a few of the (n – 2) cases that are known by experience to be the most serious cases.

Reactive Losses. The reactive losses in the transmission system have a big effect on the voltages at the buses. There are two components to the reactive losses: the MVAR consumed by the line and transformer inductive reactance calculated as

$$\sum_{\text{all lines } \ell} I_\ell^2 x_\ell$$

Note that the reactive power consumed by the transmission lines is proportional to the square of the line current. As lines become heavily loaded, this term goes up and more reactive power must be supplied from some other resource.

The next term is the reactive power injected back into the power system by the capacitive charging of the transmission line, which we model as a capacitance at each end of the line. This term, because it is feeding reactive power into the network, is a negative loss. Its calculation is

$$-\sum_{\text{all lines } \ell} \left(V^2_{\substack{\text{from end} \\ \text{of line } \ell}} B_{\text{cap}_\ell} + V^2_{\substack{\text{to end} \\ \text{of line } \ell}} B_{\text{cap}_\ell} \right)$$

Last of all, we need to account for any fixed capacitors injecting reactive power into buses:

$$-\sum_{\text{all buses } i} V_i^2 B_{\substack{\text{fixed cap} \\ \text{at bus } i}}$$

Then the reactive losses are calculated using the sum of these three terms:

$$\text{Reactive power losses} = \sum_{\text{all lines } \ell} I_\ell^2 x_\ell - \sum_{\text{all lines } \ell} \left(V^2_{\substack{\text{from end} \\ \text{of line } \ell}} B_{\text{cap}_\ell} + V^2_{\substack{\text{to end} \\ \text{of line } \ell}} B_{\text{cap}_\ell} \right) - \sum_{\text{all buses } i} V_i^2 B_{\substack{\text{fixed cap} \\ \text{at bus } i}}$$

In reviewing the power flow cases in this section, we will make reference to this term as well as the real power losses that are defined as

$$\text{Real power losses} = \sum_{\text{all lines } \ell} I_\ell^2 r_\ell$$

www.mepcafe.com

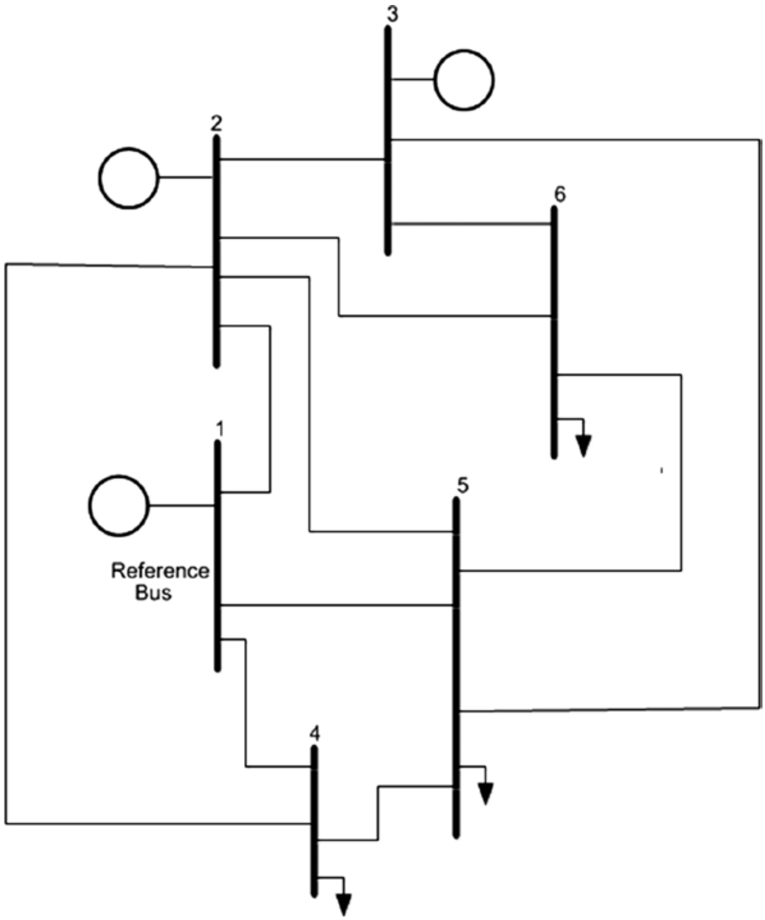


FIGURE 7.1 Six-bus network (see Chapter 6).

Sample Power Flows. Several examples are shown in Appendix 7A and are all based on the six-bus sample system introduced in Chapter 6 (Figure 7.1).

Students are encouraged to study the sample cases in Appendix 7A. From the base case, the appendix shows what happens when loads are high or a line or generator is lost—leading to low voltages and some overloads. Of particular interest are the cases shown in the appendix, when there are high loads and a line is lost leading to excessive line currents and high reactive losses resulting in generators at VAR limits with low generator bus voltages and extremely low load bus voltages.

Operations Needs to Know What Outages Will Cause Problems. Operations personnel must know which line or generation outages will cause flows or voltages to fall outside limits. To predict the effects of outages, contingency analysis techniques are used. Contingency analysis procedures model single failure events (i.e., one-line outage or one-generator outage) or multiple equipment failure events (i.e., two transmission

www.mepcafe.com

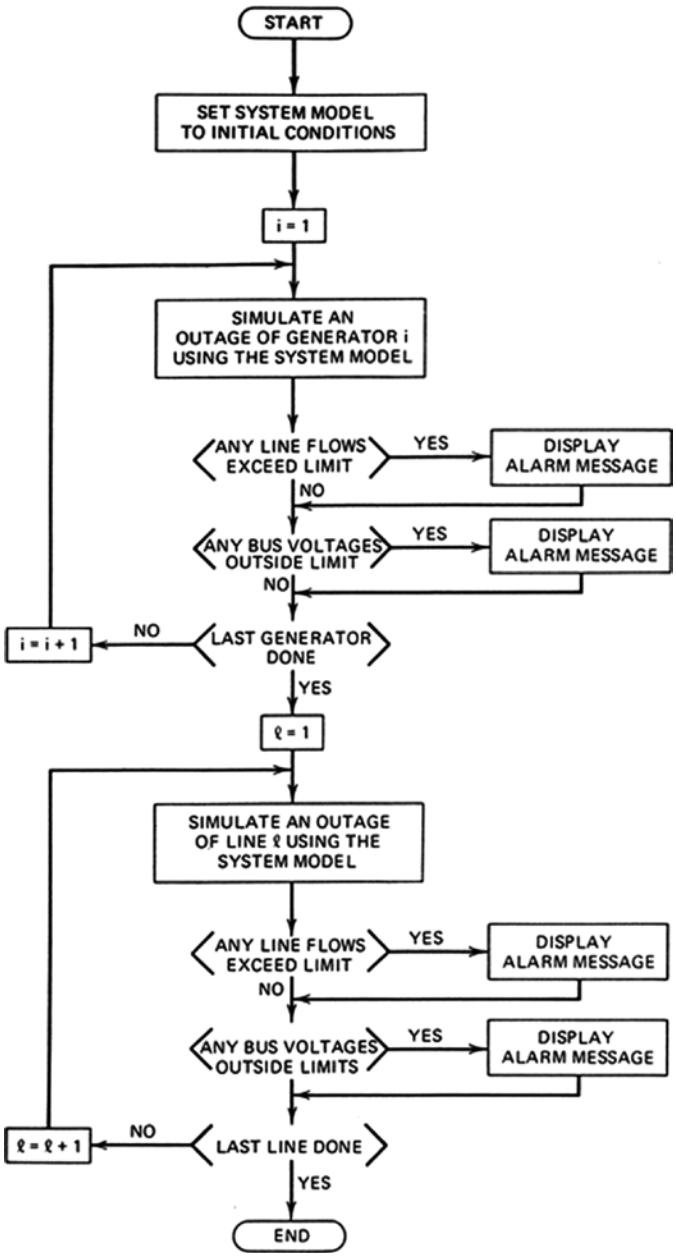


FIGURE 7.2 Full AC power flow contingency analysis procedure.

lines, one transmission line plus one generator, etc.), one after another in sequence until "all credible outages" have been studied. For each outage tested, the contingency analysis procedure checks all lines and voltages in the network against their respective limits. The simplest form of such a contingency analysis technique is shown in Figure 7.2.

www.mepcafe.com

The most difficult methodological problem to cope with in contingency analysis is the speed of solution of the model used. The most difficult logical problem is the selection of "all credible outages." If each outage case studied were to solve in 1 s and several thousand outages were of concern, it would take close to 1 h before all cases could be reported. This would be useful if the system conditions did not change over that period of time. However, power systems are constantly undergoing changes and the operators usually need to know if the present operation of the system is safe, without waiting too long for the answer. Contingency analysis execution times of less than 1 min for several thousand outage cases are typical of computer and analytical technology as of 2014.

One way to gain speed of solution in a contingency analysis procedure is to use an approximate model of the power system. For many systems, the use of DC load flow models provides adequate capability. In such systems, the voltage magnitudes may not be of great concern and the DC load flow provides sufficient accuracy with respect to the megawatt flows. For other systems, voltage is a concern and full AC load flow analysis is required.

7.4 AN OVERVIEW OF SECURITY ANALYSIS

A security analysis study that is run in an operations center must be executed very quickly in order to be of any use to operators. There are three basic ways to accomplish this:

1. Study the power system with approximate but very fast algorithms.
2. Select only the important cases for detailed analysis.
3. Use a computer system made up of multiple processors or vector processors to gain speed.

The first method has been in use for many years and goes under various names such as "D factor methods," "linear sensitivity methods," "DC power flow methods," etc. This approach is useful if one only desires an approximate analysis of the effect of each outage. This book presents these methods under the name linear sensitivity factors and uses the same derivation as was presented in Chapter 6 under the DC power flow methods. It has all the limitations attributed to the DC power flow; that is, only branch MW flows are calculated and these are only within about 5% accuracy. There is no knowledge of MVAR flows or bus voltage magnitudes. Linear sensitivity factors are presented in Section 7.4.1.

If it is necessary to know a power system's MVA flows and bus voltage magnitudes after a contingency outage, then some form of complete AC power flow must be used. This presents a great deal of difficulty when thousands of cases must be checked. It is simply impossible, even on the fastest processors, to execute thousands of complete AC power flows quickly enough. Fortunately, this need not be done as most of the cases result in power flow results that do not have flow or voltage limit violations. We need a technique to eliminate all or most of the nonviolation cases and only run complete power flows on the "critical" cases. These techniques go under the names of "contingency selection" or "contingency screening" and are introduced in Section 7.6.2.

www.mepcafe.com

Last of all, it must be mentioned that there are ways of running thousands of contingency power flows if special computing facilities are used. These facilities involve the use of many processors running separate cases in parallel or vector processors that achieve parallel operation by "unwinding" the looping instruction sets in the computer code used.

7.4.1 Linear Sensitivity Factors

The problem of studying thousands of possible outages becomes very difficult to solve if it is desired to present the results quickly. One of the easiest ways to provide a quick calculation of possible overloads is to use *linear sensitivity factors*. These factors show the approximate change in line flows for changes in generation on the network configuration and are derived from the DC load flow presented in Chapter 6. These factors can be derived in a variety of ways and basically come down to two types:

1. Power Transfer Distribution Factors (PTDFs)
2. Line Outage Distribution Factors (LODFs)

Here, we shall describe how these factors are used. The derivation of sensitivity factors is given in Appendix 7B.

7.4.1.1 Power Transfer Distribution Factors. The PTDF factors are designated and have the following definition:

$$\mathrm{PTDF}_{i,j,\ell} = \frac{\Delta f_\ell}{\Delta P}$$

where

ℓ = line index

i = bus where power is injected

j = bus where power is taken out

Δf_ℓ = change in megawatt power flow on line ℓ when a power transfer of ΔP is made between i and j

ΔP = power transferred from bus i to bus j

The PTDF factor then represents the sensitivity of the flow on line ℓ to a shift of power from i to j. Suppose one wanted to study the outage of a large generating unit and it was assumed that all the generation lost would be made up by the reference generation (we will deal with the case where the generation is picked up by many machines shortly). If the generator in question was generating P_i^0 MW and it was lost, we would represent ΔP as

$$\Delta P = -P_i^0$$

www.mepcafe.com

and the new power flow on each line in the network could be calculated using a precalculated set of "PTDF" factors as follows:

$$\hat{f}_\ell = f_\ell^0 + \mathrm{PTDF}_{\mathrm{i,ref},\ell}\Delta P$$

for $\ell = 1 \ldots L$
where

$\hat{f}_\ell$ = flow on line ℓ after the generator on bus i fails
f_ℓ^0 = flow before the failure

Note that in this case we substitute "ref" for "j" to indicate that the shift is from bus i to the reference bus.

The "outage flow," $\hat{f}_\ell$, on each line can be compared to its limit and those exceeding their limit flagged for alarming. This would tell the operations personnel that the loss of the generator on bus i would result in an overload on line ℓ.

The PTDF factors are linear estimates of the change in flow on a line with a shift in power from one bus to another. Therefore, the effects of simultaneous changes on several generating buses can be calculated using superposition. Suppose, for example, that the loss of the generator on bus i were compensated by governor action on machines throughout the interconnected system. One frequently used method assumes that the remaining generators pick up in proportion to their maximum MW rating. Thus, the proportion of generation pickup from unit $j (j \neq i)$ would be

$$\gamma_{ij} = \frac{P_j^{\max}}{\sum\limits_{\substack{k \\ k \neq i}} P_k^{\max}}$$

where

$P_k^{\max}$ = maximum MW rating for generator k
γ_{ij} = proportionality factor for pickup on generating unit j when unit i fails

Then, to test for the flow on line ℓ, under the assumption that all the generators in the interconnection participate in making up the loss, use the following:

$$\hat{f}_\ell = f_\ell^0 + \mathrm{PTDF}_{i,\mathrm{ref},\ell}\Delta P_i - \sum_{j \neq i}\left[\mathrm{PTDF}_{\mathrm{ref},j,\ell}\gamma_{ji}\Delta P_i\right]$$

Note that this assumes that no unit will actually hit its maximum. If this is apt to be the case, a more detailed generation pickup algorithm that took account of generation limits would be required.

www.mepcafe.com

7.4.1.2 Line Outage Distribution Factors. The LODF factors are used in a similar manner, only they apply to the testing for overloads when transmission circuits are lost. By definition, the line outage distribution factor has the following meaning:

$$\mathrm{LODF}_{\ell,k} = \frac{\Delta f_\ell}{f_k^0}$$

where

$\mathrm{LODF}_{\ell,k}$ = line outage distribution factor when monitoring line ℓ after an outage on line k

Δf_ℓ = change in MW flow on line ℓ

f_k^0 = original flow on line k before it was outaged (opened)

If one knows the power on line ℓ and line k, the flow on line ℓ with line k out can be determined using "LODF" factors:

$$\hat{f}_\ell = f_\ell^0 + \mathrm{LODF}_{\ell,k} f_k^0$$

where

f_ℓ^0, f_k^0 preoutage flows on lines ℓ and k, respectively

$\hat{f}_\ell$ = flow on line ℓ with line k out

By precalculating the LODFs, a very fast procedure can be set up to test all lines in the network for overload for the outage of a particular line. Furthermore, this procedure can be repeated for the outage of each line in turn, with overloads reported to the operations personnel in the form of alarm messages.

Using the generator and line outage procedures described earlier, one can program a digital computer to execute a contingency analysis study of the power system as shown in Figure 7.3. Note that a line flow can be positive or negative so that, as shown in Figure 7.3, we must check $\hat{f}_\ell$ against $-f_\ell^{\max}$ as well as $f_\ell^{\max}$. This Figure makes several assumptions. First, it assumes that the generator output for each of the generators in the system is available and that the line flow for each transmission line in the network is also available. Second, it assumes that the sensitivity factors have been calculated and stored and that they are correct. The assumption that the generation and line flow MWs are available can be satisfied with telemetry systems or with state estimation techniques. The assumption that the sensitivity factors are correct is valid as long as the transmission network has not undergone any significant switching operations that would change its structure. For this reason, control systems that use sensitivity factors must have provision for updating the factors when the network is switched. A third assumption is that all generation pickup will be made on the reference bus. If this is not the case, substitute the equation at the end of Section 7.4.1.1 in the generator outage loop.

www.mepcafe.com

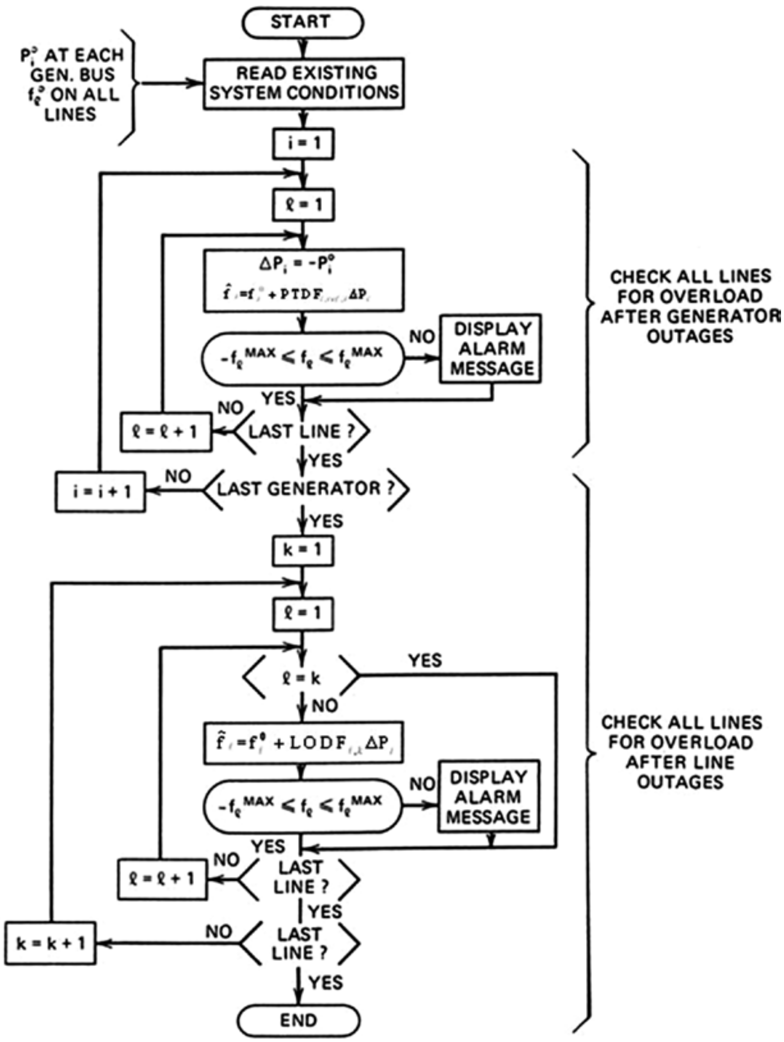


FIGURE 7.3 Contingency analysis using sensitivity factors.

Example 7A: The [*X*] matrix for our six-bus sample network is shown in Figure 7.4, together with the LODF factors. The PTDF factors are shown in Figure 7.5.

The LODF matrix is stored such that each row and column corresponds to one line in the network, with rows corresponding to monitored line and columns corresponding to the outaged lines. The LODF for a particular outage and monitored line is obtained by finding the monitored line ℓ down the rows and then finding the outaged line k along that row in the appropriate column. For instance, the line outage distribution factor that gives the fraction of flow picked up on line 3–5 for an outage on line 3–6 is found in the eighth row and ninth column and has the value 0.4747. Case 2 shown

www.mepcafe.com

X MATRIX

Bus	BUS 1	2	3	4	5	6
1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.0000	0.0941	0.0805	0.0630	0.0643	0.0813
3	0.0000	0.0805	0.1659	0.0590	0.0908	0.1290
4	0.0000	0.0630	0.0590	0.1009	0.0542	0.0592
5	0.0000	0.0643	0.0908	0.0542	0.1222	0.0893
6	0.0000	0.0813	0.1290	0.0592	0.0893	0.1633

LINE OUTAGE DISTRIBUTION FACTOR (LODF) MATRIX

Monitored line	Outage of line From bus - To bus 1 to 2	1 to 4	1 to 5	2 to 3	2 to 4	2 to 5	2 to 6	3 to 5	3 to 6	4 to 5	5 to 6
1 to 2	0.0000	0.6353	0.5427	-0.1127	-0.5031	-0.2103	-0.1221	-0.1369	0.0135	0.0096	0.1316
1 to 4	0.5948	0.0000	0.4573	-0.0331	0.6121	-0.0618	-0.0359	-0.0403	0.0040	-0.3269	0.0387
1 to 5	0.4052	0.3647	0.0000	0.1458	-0.1090	0.2721	0.1580	0.1772	-0.0174	0.3174	-0.1703
2 to 3	-0.1029	-0.0323	0.1783	0.0000	0.1242	0.2262	0.4662	-0.3995	-0.5253	0.1706	0.1320
2 to 4	-0.5884	0.7647	-0.1708	0.1591	0.0000	0.2969	0.1724	0.1933	-0.0190	-0.6731	-0.1858
2 to 5	-0.1875	-0.0589	0.3250	0.2209	0.2264	0.0000	0.2394	0.2685	-0.0264	0.3110	-0.2580
2 to 6	-0.1213	-0.0381	0.2102	0.5073	0.1464	0.2667	0.0000	-0.1992	0.5842	0.2011	0.4433
3 to 5	-0.1175	-0.0369	0.2036	-0.3755	0.1418	0.2583	-0.1720	0.0000	0.4747	0.1948	-0.4246
3 to 6	0.0146	0.0046	-0.0253	-0.6245	-0.0176	-0.0321	0.6382	0.6005	0.0000	-0.0242	0.5567
4 to 5	0.0065	-0.2353	0.2865	0.1259	-0.3879	0.2350	0.1365	0.1530	-0.0150	0.0000	-0.1471
5 to 6	0.1067	0.0335	-0.1849	0.1172	-0.1288	-0.2346	0.3618	-0.4013	0.4158	-0.1769	0.0000

FIGURE 7.4 X matrix and LODF matrix for the six-bus system.

www.mepcafe.com

POWER TRANSFER DISTRIBUTION FACTOR (PTDF) MATRIX

Transfer power
From bus - To bus

Monitored line	1 to 2	1 to 3	1 to 4	1 to 5	1 to 6	2 to 3	2 to 4
1 to 2	0.4706	0.4026	0.3149	0.3217	0.4064	-0.0681	-0.1557
1 to 4	0.3149	0.2949	0.5044	0.2711	0.2960	-0.0200	0.1895
1 to 5	0.2145	0.3026	0.1807	0.4072	0.2976	0.0881	-0.0338
2 to 3	-0.0544	0.3416	-0.0160	0.1057	0.1907	0.3960	0.0384
2 to 4	-0.3115	-0.2154	0.3790	-0.1013	-0.2208	0.0961	0.6904
2 to 5	-0.0993	0.0342	-0.0292	0.1927	0.0266	0.1335	0.0701
2 to 6	-0.0642	0.2422	-0.0189	0.1246	0.4100	0.3064	0.0453
3 to 5	-0.0622	-0.2890	-0.0183	0.1207	-0.1526	-0.2268	0.0439
3 to 6	0.0077	-0.3695	0.0023	-0.0150	0.3433	-0.3772	-0.0055
4 to 5	0.0034	0.0795	-0.1166	0.1698	0.0752	0.0761	-0.1201
5 to 6	0.0565	0.1273	0.0166	-0.1096	0.2467	0.0708	-0.0399

PTDF MATRIX (continued)

Transfer power
From bus - To bus

Monitored line	2 to 5	2 to 6	3 to 4	3 to 5	3 to 6	4 to 5	4 to 6	5 to 6
1 to 2	-0.1489	-0.0642	-0.0877	-0.0808	0.0039	0.0068	0.0915	0.0847
1 to 4	-0.0438	-0.0189	0.2095	-0.0238	0.0011	-0.2333	-0.2084	0.0249
1 to 5	0.1927	0.0831	-0.1218	0.1046	-0.0050	0.2264	0.1168	-0.1096
2 to 3	0.1601	0.2451	-0.3576	-0.2359	-0.1509	0.1217	0.2067	0.0850
2 to 4	0.2102	0.0906	0.5944	0.1141	-0.0055	-0.4802	-0.5998	-0.1196
2 to 5	0.2919	0.1259	-0.0634	0.1585	-0.0076	0.2219	0.0558	-0.1661
2 to 6	0.1888	0.4742	-0.2611	-0.1176	0.1678	0.1435	0.4289	0.2854
3 to 5	0.1829	-0.0905	0.2707	0.4097	0.1363	0.1390	-0.1343	-0.2733
3 to 6	-0.0227	0.3356	0.3718	0.3545	0.7128	-0.0173	0.3410	0.3583
4 to 5	0.1664	0.0717	-0.1961	0.0903	-0.0043	0.2865	0.1918	-0.0947
5 to 6	-0.1661	0.1902	-0.1107	-0.2369	0.1194	-0.1262	0.2301	0.3563

FIGURE 7.5 PTDF matrix for six-bus system.

www.mepcafe.com

in Appendix 7A gives the results for an outage of the generator on bus 3 with all pickup of lost generation coming on the reference bus generator at bus 1. To calculate the flow on line 1–4 after the outage of the generator on bus 3, we need the base-case flows that are given in Case 1 in Appendix 7A:

Base-case flow on line 1–4 = 82.8 MW

Base-case generation on bus 3 = 50 MW

$\mathrm{PTDF}_{3,1,\mathrm{line}\ 1-4} = -\mathrm{PTDF}_{1,3,\mathrm{line}\ 1-4} = -0.2949$ (note that the table in Figure 7.5 has $\mathrm{PTDF}_{1,3,\mathrm{line}\ 1-4} = 0.2949$, so we use −0.2949)

Then the flow on line 1–4 after generator outage is

$= \text{base case flow}_{1-4} + \mathrm{PTDF}_{3,1,\mathrm{line}1-4}\Delta P_{\mathrm{gen}3}$

= 82.8 + (−0.2949) (−60 MW) = 100.49 MW, which is the flow on line 1–4 shown in Case 2 of Appendix 7A.

To show how the line outage and generation shift factors are used, calculate the flows for an outage of line 1–4 as given in Case 3 of Appendix 7A. If we wish to calculate the power flowing on lines 1–2 with line 1–4 opened, we would need the following:

Base-case flow on line 1–4 = 82.8 MW (the line to be taken out)

Base-case flow on line 1–2 = 62.8 MW (the monitored line)

Line outage distribution factor: $\mathrm{LODF}_{1-2,1-4} = 0.6353$

Then the flow on 1–2 after the outage of line 1–4 is

$$= \text{base flow}_{1-2} + \mathrm{LODF}_{1-2,1-4} \times \text{base flow}_{1-4}$$

$$= 62.8 + (0.6353) \times (82.8) = 115.4\,\mathrm{MW}$$

The result is not equal to the AC power flow result but still would give the operator an indication of an overload.

In both outages, the generator outage and the line outage, the flows calculated by the sensitivity methods are reasonably close to the values calculated by the full AC load flows as shown in Appendix 7A Case 3.

7.5 MONITORING POWER TRANSACTIONS USING "FLOWGATES"

A "flowgate" is a collection of transmission lines that are treated together for the purposes of monitoring transactions between electric companies. The term came about due to a desire to present transmission system limitations to persons selling and buying power (Figure 7.6).[3]

[3] The original development of the flowgate concept came from Paul Barber (now with NERC), Ben Li of the Independent Electricity System Operator (IESO) in Ontario, Canada, and Bob Cummings also of NERC. "We noted that network flow limitations seldom depended upon single lines but rather on groups of lines that could be modeled by partial cutsets. We first called them tollgates because we had envisioned a revenue-sharing scheme to get transmission owners paid for the use of their investments but it created quite an uproar so I suggested calling them flowgates to avoid the revenue arguments." (Communication from Paul Barber)

www.mepcafe.com

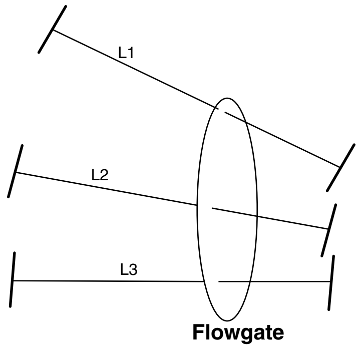


FIGURE 7.6 Flowgate example.

Flowgates are simply linear sums of line flows, and therefore we can create PTDF and LODF factors for flowgates and then use these new factors in the same way as we do for individual lines. For example, using the flowgate pictured in Figure 7.5,

$$f_{\text{flowgate}} = f_{\text{L1}} + f_{\text{L2}} + f_{\text{L3}}$$

Then the total flow through the flowgate after a transaction from bus i to bus j can be calculated as follows:

$$\begin{aligned}
\hat{f}_{\text{flowgate}} &= \sum_{\substack{\text{all lines } \ell \\ \text{in flowgate}}} \hat{f}_\ell = \sum_{\substack{\text{all lines } \ell \\ \text{in flowgate}}} \left(f_\ell^0 + \text{PTDF}_{ij\ell} \cdot \Delta P_{ij} \right) \\
&= \sum_{\substack{\text{all lines } \ell \\ \text{in flowgate}}} f_\ell^0 + \left(\sum_{\substack{\text{all lines } \ell \\ \text{in flowgate}}} \text{PTDF}_{ij\ell} \right) \Delta P_{ij} \\
&= f_{\text{flowgate}}^0 + \text{PTDF}_{ij\ \text{flowgate}} \Delta P_{ij}
\end{aligned}$$

where

f_{flowgate}^0 is the initial flow on the flowgate before the transaction is placed on the system
$\hat{f}_{\text{flowgate}}$ is the flow on the flowgate after the transaction is made
$\text{PTDF}_{ij\ \text{flowgate}}$ is the PTDF that tells us what fraction of the transfer of power from bus i to bus j will flow on the flowgate.

Similarly, to determine the effect that a line outage on branch k has on the flowgate, we use the LODF for each constituent line in the flowgate:

$$\begin{aligned}
\hat{f}_{\text{flowgate}} &= \sum_{\substack{\text{all lines } \ell \\ \text{in flowgate}}} \hat{f}_\ell = \sum_{\substack{\text{all lines } \ell \\ \text{in flowgate}}} \left(f_\ell^0 + \text{LODF}_{\ell,k} \cdot f_k^0 \right) \\
&= \sum_{\substack{\text{all lines } \ell \\ \text{in flowgate}}} f_\ell^0 + \left(\sum_{\substack{\text{all lines } \ell \\ \text{in flowgate}}} \text{LODF}_{\ell,k} \right) f_k^0 \\
&= f_{\text{flowgate}}^0 + \text{LODF}_{\text{flowgate},k} f_k^0
\end{aligned}$$

www.mepcafe.com

where

f^0_{flowgate} is the initial flow on the flowgate before the line outage on line k

$\hat{f}_{\text{flowgate}}$ is the flow on the flowgate after the line outage on line k

$\text{LODF}_{ij\text{ flowgate}}$ is the LODF that tells us what fraction of the original power flowing on line k before it was taken out ends up flowing on the flowgate

Flowgates can have limits that are simply the sum of the flow limits on each line making up the flowgate—or they may be limits set by reliability or stability studies to limit the flow across an important system interface. Given the flowgate limits and the flowgate PTDF and LODF factors, we can perform security analysis checks for line outage cases in the same manner as is done for single transmission lines. Flowgate PTDF and LODF factors can be combined to form "composite" OTDF factors as shown in the appendix to this chapter so that flowgate PTDF's under line outage conditions can be calculated.

7.6 VOLTAGE COLLAPSE

When the reactive power required by the transmission system becomes inadequate, we say that the power system goes through a "voltage collapse." Voltage collapse can be best explained by the trivial example shown in Figure 7.7.

If we solve the power flow equations for the load bus, we get

$$P(V,\theta) = -V \cdot \sin(\theta) \cdot \frac{E}{X}$$

and

$$Q(V,\theta) = -V \cdot \frac{(-\cos(\theta) \cdot E + V)}{X}$$

Lastly, if we give the relationship between Q and P as $Q = \tan(\phi) \cdot P$, then we get the plot for various values of $\tan(\phi)$ as shown in Figure 7.8.

Several things can be deduced from Figure 7.8. When the load draws reactive power ($\tan(\theta) \geq 0$), the voltage drops off faster as P increases. Similarly, if the load produces reactive power ($\tan(\theta) \leq 0$), the voltage actually rises and stays above 1.0 pu for much of the range. This is the phenomena of reactive compensation wherein one can supply more real power, P, if reactive power is supplied at the load.

During a serious emergency on the transmission system or when a generator is lost, the reactive power being consumed by the transmission system will cause the voltage to drop. More importantly, the curve, often called the "nose curve" because of its shape, contracts and the result is as shown in Figure 7.9.

www.mepcafe.com

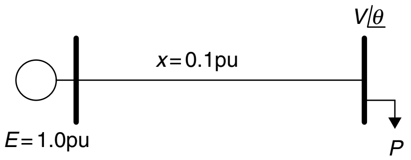


FIGURE 7.7 Example power system to show voltage collapse.

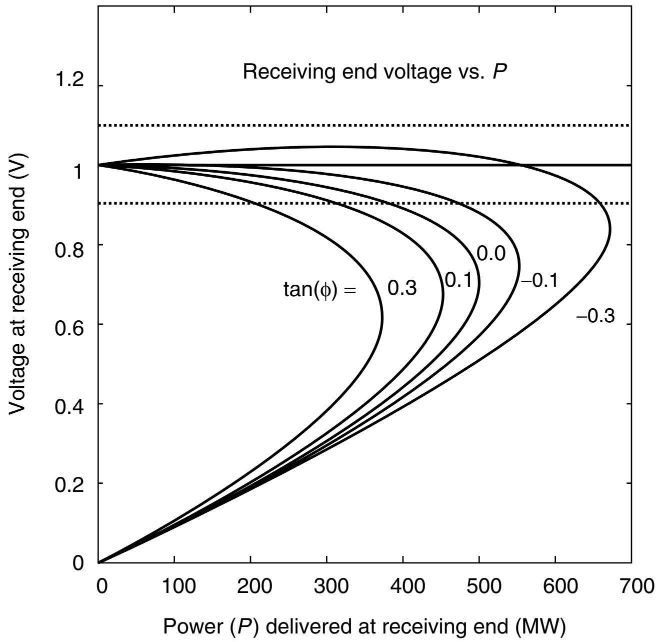


FIGURE 7.8 Voltage at receiving bus versus power delivered.

In this case, we shall assume that the system was supplying about 330 MW and gave adequate reactive support to result in a predisturbance voltage that was within limits. The disturbance results in loss of transmission, and the resulting new voltage characteristic no longer even intersects the vertical line at 330 MW and we then have a situation where the power flow cannot be solved. Under such conditions, the voltage will collapse and the whole power system will go down. Thus, there is a need to study AC power systems with AC power flow methods to be sure this kind of condition is detected.

Note, when the power flow cannot be solved for a given contingency, case engineers often resort to the use of "continuation power flow" algorithms that are able to get to a solution when conventional power flow algorithms, such as the Newton algorithm, fail. The failure of the Newton algorithm is due to the fact that, as the power system approaches the right most edge of the nose curve, the Jacobian matrix used in the Newton algorithm becomes singular (has no inverse). Generally, failure

www.mepcafe.com

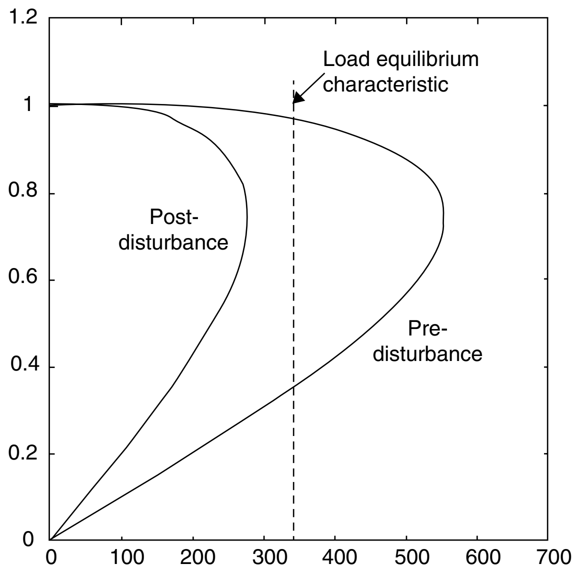


FIGURE 7.9 Contraction of the voltage characteristic during a transmission outage.

of the power flow is a sign that the power system is not secure and should be alarmed to operators. The next sections detail some of the methods now in use to allow the calculation of an AC power flow for contingency analysis.

7.6.1 AC Power Flow Methods

The calculations made by network sensitivity methods are faster than those made by AC power flow methods and therefore find wide use in operations control systems. However, there are many power systems where voltage magnitudes are the critical factor in assessing contingencies. In addition, there are some systems where VAR flows predominate on some circuits, such as underground cables, and an analysis of only the MW flows will not be adequate to indicate overloads. When such situations present themselves, the network sensitivity methods may not be adequate and the operations control system will have to incorporate a full AC power flow for contingency analysis.

When an AC power flow is to be used to study each contingency case, the speed of solution and the number of cases to be studied are critical. To repeat what was said before, if the contingency alarms come too late for operators to act, they are worthless. Most operations control centers that use an AC power flow program for contingency analysis use either a Newton–Raphson or the decoupled power flow. These solution algorithms are used because of their speed of solution and the fact that they are reasonably reliable in convergence when solving difficult cases. The decoupled power flow has the further advantage that a matrix alteration formula can be incorporated into it to simulate the outage of transmission lines without reinverting the system Jacobian matrix at each iteration.

www.mepcafe.com

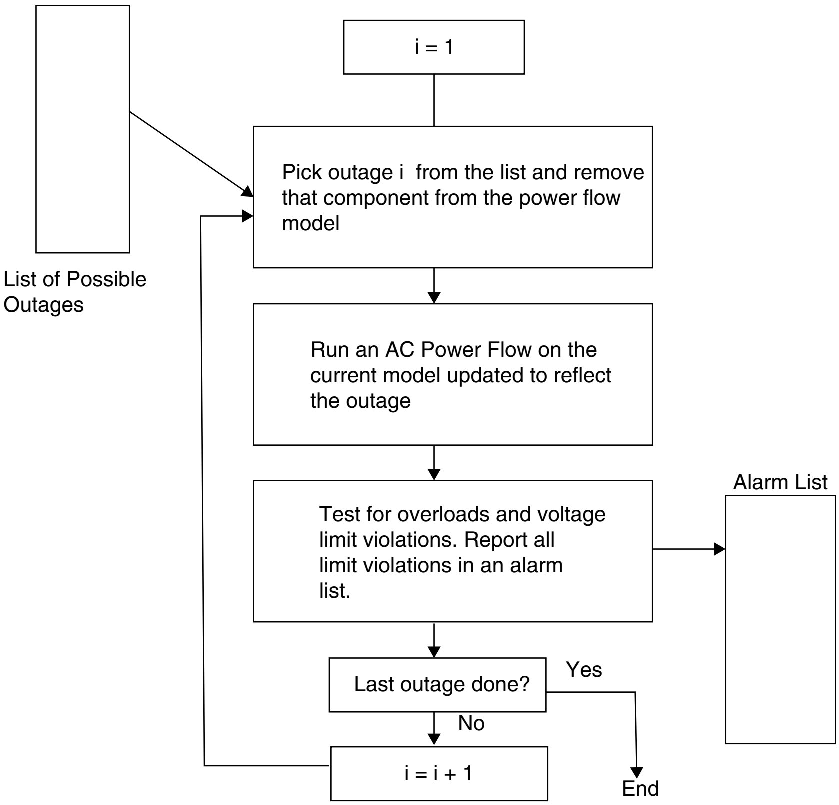


FIGURE 7.10 AC power flow security analysis.

The simplest AC security analysis procedure consists of running an AC power flow analysis for each possible generator, transmission line, and transformer outage as shown in Figure 7.10. This procedure will determine the overloads and voltage limit violations accurately (at least within the accuracy of the power flow program, the accuracy of the model data, and the accuracy with which we have obtained the initial conditions for the power flow). It does suffer a major drawback, however, and that concerns the time such a program takes to execute. If the list of outages has several thousand entries, then the total time to test for all of the outages can be too long (Figure 7.10).

We are thus confronted with a dilemma. Fast, but inaccurate, methods involving the PTDF and LODF factors can be used to give rapid analysis of the system, but they cannot give information about MVAR flows and voltages. Slower, full AC power flow methods give full accuracy but take too long.

Fortunately, there is a way out of this dilemma. Because of the way the power system is designed and operated, very few of the outages will actually cause trouble. That is, most of the time spent running AC power flows will go for solutions of the power flow model that discover that there are no problems. Only a few of the power flow solutions will, in fact, conclude that an overload or voltage violation exists.

www.mepcafe.com

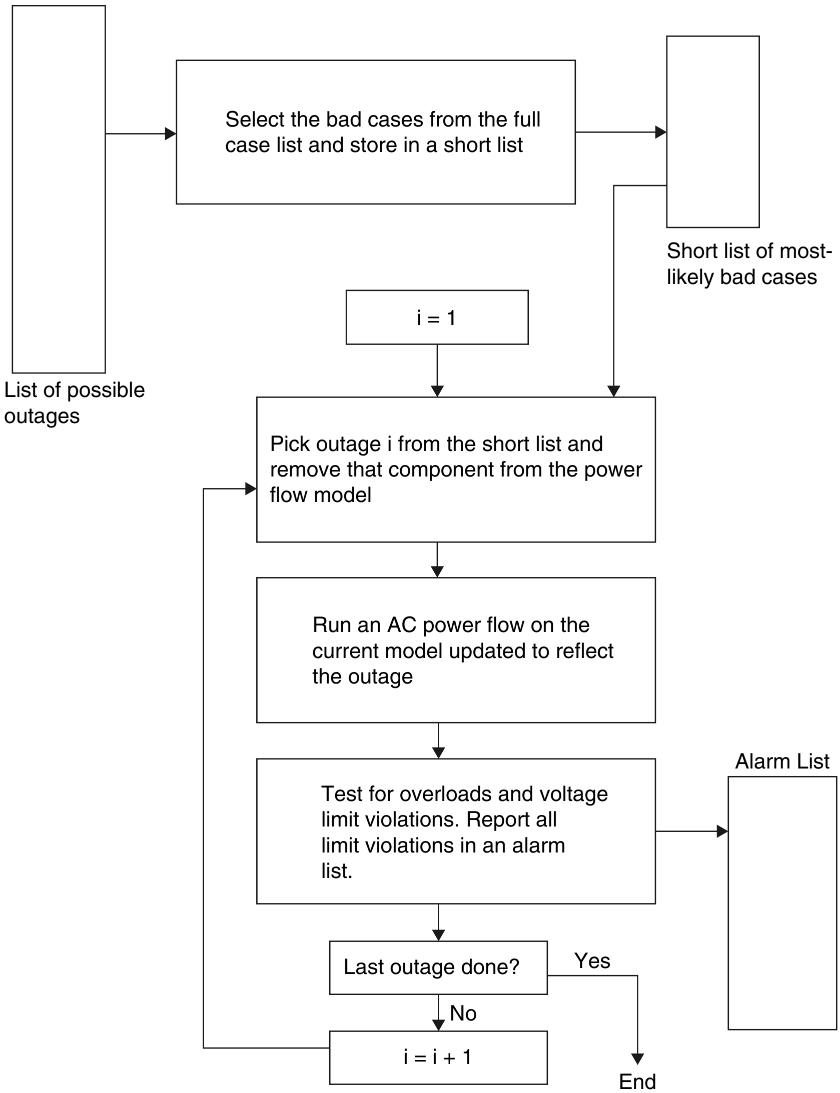


FIGURE 7.11 AC power flow security analysis with contingency case selection.

The solution to this dilemma is to find a way to select contingencies in such a way that only those that are likely to result in an overload or voltage limit violation will actually be studied in detail and the other cases will go unanalyzed.

A flowchart for a process like this appears in Figure 7.11.

Selecting the highest impact contingency cases from the full outage case list is not an exact procedure and has been the subject of intense research for many years. Two sources of error can arise.

www.mepcafe.com

1. **Placing too many cases on the shortlist**: This is essentially the "conservative" approach and simply leads to longer run times for the security analysis procedure to execute.
2. **Skipping cases**: Here, a case that would have shown a problem is not placed on the shortlist and results in possibly having that outage take place and cause trouble without the operators being warned.

7.6.2 Contingency Selection

We would like to get some measure as to how much a particular outage might affect the power system. The idea of a performance index (PI) for line outage on line i seems to fulfill this need. The definition for the flow performance index, PI_flow_i, is as follows:

For each contingency outage i:

$$\text{PI\_flow}_i = \sum_{\substack{\text{all branches} \\ \ell, \ell \neq i}} \left(\frac{P_{\text{flow } \ell, i}}{P_\ell^{\max}} \right)^{2n_{\text{PI flow}}}$$

for $\quad i = 1 \ldots N_{\text{lines}}$

and $\quad P_{\text{flow } \ell, i}$ is the flow on line ℓ with line i out

If $n_{\text{PI flow}}$ is a large number, the PI_flow_i will be a small number if all flows are within limit, and it will be large if one or more lines are overloaded. The problem then is how to use this performance index.

Various techniques have been tried to obtain the value of PI_flow_i when a branch is taken out. These calculations can be made exactly if $n_{\text{PI flow}} = 1$; that is, a table of PI_flow_i values, one for each line in the network, can be calculated quite quickly. The selection procedure then involves ordering the PI_flow_i table from largest value to least. The lines corresponding to the top of the list are then the candidates for the shortlist. One procedure simply ordered the PI_flow_i table and then picked the top N_c entries from this list and placed them on the shortlist.

However, when $n_{\text{PI flow}} = 1$, the PI_flow_i does not snap from near zero to near infinity as the branch exceeds its limit. Instead, it rises as a quadratic function. A line that is just below its limit contributes to PI_flow_i almost as much as one that is just over its limit. The result is a PI_flow_i that may be large when many lines are loaded just below their limit. Thus, the PI_flow_i's ability to distinguish or detect bad cases is limited when $n_{\text{PI flow}} = 1$. Ordering the PI_flow_i values when $n_{\text{PI flow}} = 1$ usually results in a list that is not at all representative of one with the truly bad cases at the top. Trying to develop an algorithm that can quickly calculate PI_flow_i when $n_{\text{PI flow}} = 2$ or larger has proven extremely difficult.

One way to perform an outage case selection is to perform what has been called the *1P1Q method*. Here, a decoupled power flow is used. As shown in Figure 7.12, the solution procedure is interrupted after one iteration (one $P-\theta$ calculation and one Q–V calculation; thus, the name 1P1Q).

With this procedure, the PI_flow_i can use as large an n value as desired, say $n_{\text{PI flow}} = 5$. There appears to be sufficient information in the solution at the end of the

www.mepcafe.com

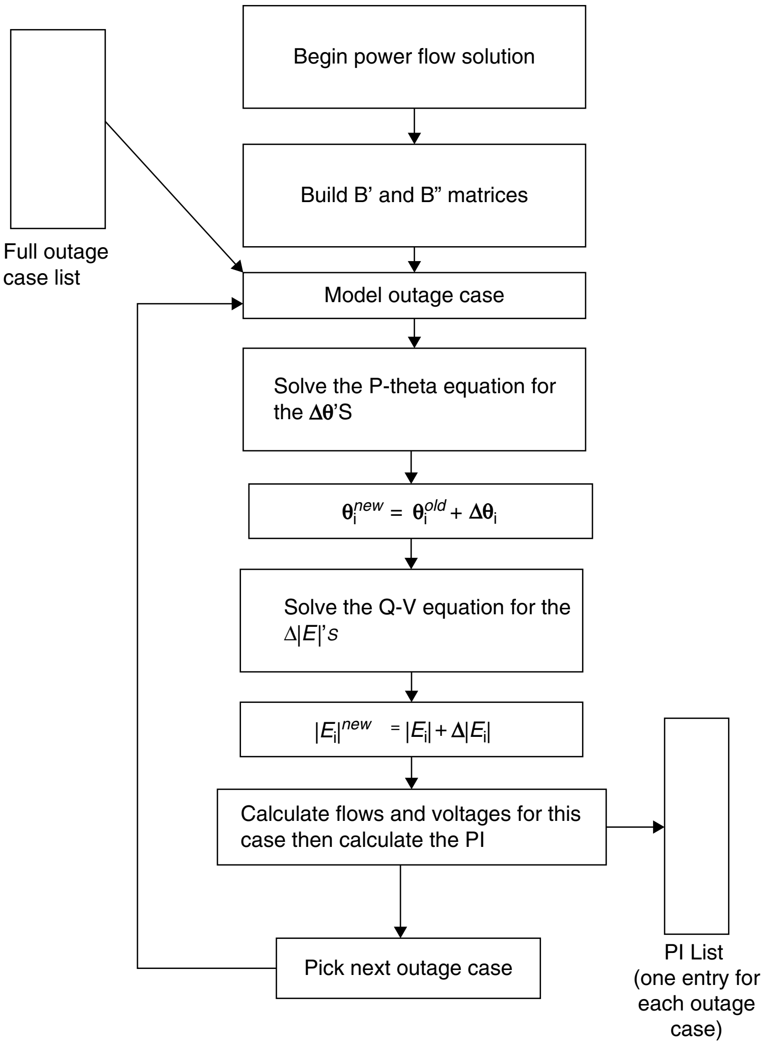


FIGURE 7.12 The 1P1Q contingency selection procedure.

first iteration of the decoupled power flow to give a reasonable PI_flow_i. Another advantage to this procedure is the fact that the voltages can also be included in the PI. Thus, a different PI can be used, such as

$$PI\_volt_i = \sum_{\substack{\text{all buses} \\ j}} \left(\frac{V_j^{\min}}{V_{j,i}} \right)^{n_{\text{PIV}}} + \sum_{\substack{\text{all buses} \\ j}} \left(\frac{V_{j,i}}{V_j^{\max}} \right)^{n_{\text{PIV}}}$$

where

$V_{j,i}$ = voltage on bus j with line i out

www.mepcafe.com

TABLE 7.1a Solving the Full AC Power Flow for Each Contingency

Full AC Power Flow Contingency Results						PI_flow	PI_V
Contingency	1	from	1	to	2	45.342	4.75
Contingency	2	from	1	to	4	289.253	30.024
Contingency	3	from	1	to	5	47.753	4.8
Contingency	4	from	2	to	3	19.312	4.697
Contingency	5	from	2	to	4	32.746	18.936
Contingency	6	from	2	to	5	20.6	4.717
Contingency	7	from	2	to	6	25.227	4.679
Contingency	8	from	3	to	5	20.217	4.718
Contingency	9	from	3	to	6	22.436	4.93
Contingency	10	from	4	to	5	20.948	4.794
Contingency	11	from	5	to	6	19.517	4.702

TABLE 7.1b Solving Each Contingency with 1P1Q Only

1P1Q Contingency Results						PI_flow	PI_V
Contingency	1	from	1	to	2	34.147	7.076
Contingency	2	from	1	to	4	180.975	7.574
Contingency	3	from	1	to	5	25.874	7.079
Contingency	4	from	2	to	3	17.467	7.06
Contingency	5	from	2	to	4	21.708	8.34
Contingency	6	from	2	to	5	18.915	7.061
Contingency	7	from	2	to	6	23.101	7.049
Contingency	8	from	3	to	5	16.832	7.062
Contingency	9	from	3	to	6	19.847	7.122
Contingency	10	from	4	to	5	18.192	7.094
Contingency	11	from	5	to	6	17.279	7.058

Here, the idea is to add to the PI when voltages are below the low limit or above the high limit.

To complete the security analysis, the PI list is sorted so that the largest PI appears at the top. The security analysis can then start by executing full power flows with the case that is at the top of the list, then solve the case that is second, and so on down the list. This continues until either a fixed number of cases are solved or until a predetermined number of cases are solved that do not have any alarms.

In the following tables, we see how this would work for the six-bus system given in Appendix 7A. Table 7.1a and 7.1b shows the comparison of the results for a full AC power flow compared to the 1P1Q method for each contingency.

Now these results can be compared by sorting the contingency cases using first the PI_flow values only (Table 7.2a).

Note that sorting on the PI_flow is not the same as sorting on the PI_V for voltages. Apparently, contingency 5 dropping line 2–4 causes a voltage problem but not that much of a flow overload problem (Table 7.2b).

Here, we see that the worst case is indeed found using the 1P1Q method, but that the ordering is somewhat different for cases immediately below the top case.

www.mepcafe.com

TABLE 7.2a Sorting Contingency Cases on PI_flow for the Full AC Solutions

Sorted on PI_flow

Full AC Power Flow Contingency Results						PI_flow	PI_V
Contingency	2	from	1	to	4	289.253	30.024
Contingency	3	from	1	to	5	47.753	4.8
Contingency	1	from	1	to	2	45.342	4.75
Contingency	5	from	2	to	4	32.746	18.936
Contingency	7	from	2	to	6	25.227	4.679
Contingency	9	from	3	to	6	22.436	4.93
Contingency	10	from	4	to	5	20.948	4.794
Contingency	6	from	2	to	5	20.6	4.717
Contingency	8	from	3	to	5	20.217	4.718
Contingency	11	from	5	to	6	19.517	4.702
Contingency	4	from	2	to	3	19.312	4.697

TABLE 7.2b Sorting Contingency Cases on PI_flow for the 1P1Q Solutions

Sorted on PI_flow

1P1Q Contingency Results						PI_flow	PI_V
Contingency	2	from	1	To	4	180.975	7.574
Contingency	1	from	1	To	2	34.147	7.076
Contingency	3	from	1	To	5	25.874	7.079
Contingency	7	from	2	To	6	23.101	7.049
Contingency	5	from	2	To	4	21.708	8.34
Contingency	9	from	3	To	6	19.847	7.122
Contingency	6	from	2	To	5	18.915	7.061
Contingency	10	from	4	To	5	18.192	7.094
Contingency	4	from	2	To	3	17.467	7.06
Contingency	11	from	5	To	6	17.279	7.058
Contingency	8	from	3	To	5	16.832	7.062

Note also that the 1P1Q method has very little variation in the PI_V column but that its values in the PI_flow column are quite similar.

Finally, we sort on the sum of PI_flow and PI_V as shown in Table 7.3a and 7.3b

7.6.3 Concentric Relaxation

Another idea to enter the field of security analysis in power systems is that an outage only has a limited geographical effect. The loss of a transmission line does not cause much effect a thousand miles away; in fact, we might hope that it does not cause much trouble beyond 20 miles from the outage, although if the line were a heavily loaded, high-voltage line, its loss will most likely be felt more than 20 miles away.

To realize any benefit from the limited geographical effect of an outage, the power system must be divided into two parts: the affected part and the part that is unaffected. To make this division, the buses at the end of the outaged line are marked as layer zero. The buses that are one transmission line or transformer from layer zero are then

www.mepcafe.com

TABLE 7.3a Sorting Contingency Cases on PI_flow +PI_V for the Full AC Solutions

Sorted on Sum of PI_V and PI_flow

Full AC Power Flow Contingency Results						PI_flow	PI_V	Sum
Contingency	2	from	1	to	4	289.253	30.024	319.277
Contingency	3	from	1	to	5	47.753	4.8	52.553
Contingency	5	from	2	to	4	32.746	18.936	51.682
Contingency	1	from	1	to	2	45.342	4.75	50.092
Contingency	7	from	2	to	6	25.227	4.679	29.906
Contingency	9	from	3	to	6	22.436	4.93	27.366
Contingency	10	from	4	to	5	20.948	4.794	25.742
Contingency	6	from	2	to	5	20.6	4.717	25.317
Contingency	8	from	3	to	5	20.217	4.718	24.935
Contingency	11	from	5	to	6	19.517	4.702	24.219
Contingency	4	from	2	to	3	19.312	4.697	24.009

TABLE 7.3b Sorting Contingency Cases on PI_flow + PI_V for the 1P1Q Solutions

Sorted on Sum of PI_V and PI_flow

1P1Q Contingency Results						PI_flow	PI_V	Sum
Contingency	2	from	1	to	4	180.975	7.574	188.549
Contingency	1	from	1	to	2	34.147	7.076	41.223
Contingency	3	from	1	to	5	25.874	7.079	32.953
Contingency	7	from	2	to	6	23.101	7.049	30.15
Contingency	5	from	2	to	4	21.708	8.34	30.048
Contingency	9	from	3	to	6	19.847	7.122	26.969
Contingency	6	from	2	to	5	18.915	7.061	25.976
Contingency	10	from	4	to	5	18.192	7.094	25.286
Contingency	4	from	2	to	3	17.467	7.06	24.527
Contingency	11	from	5	to	6	17.279	7.058	24.337
Contingency	8	from	3	to	5	16.832	7.062	23.894

labeled layer one. This same process can be carried out, layer by layer, until all the buses in the entire network are included. Some arbitrary number of layers is chosen and all buses included in that layer and lower-numbered layers are solved as a power flow with the outage in place. The buses in the higher-numbered layers are kept as constant voltage and phase angle (i.e., as reference buses).

This procedure can be used in two ways: either the solution of the layers included becomes the final solution of that case and all overloads and voltage violations are determined from this power flow or the solution simply is used to form a PI for that outage. Figure 7.13 illustrates this layering procedure. The concentric relaxation procedure was originally proposed by Zaborsky. The trouble with the concentric relaxation technique is that it requires more layers for circuits whose influence is felt further from the outage. Instead of layers, several researchers have used the term neighbors. The buses that are immediately linked to the bus of the outaged equipment are called first neighbors (Heydt).

www.mepcafe.com

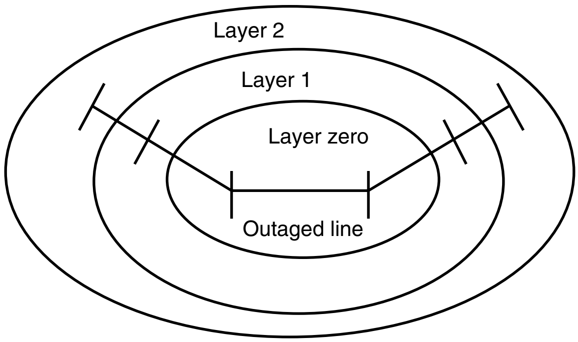


FIGURE 7.13 Layering of outage effects.

7.6.4 Bounding

A paper by Brandwajn solves at least one of the problems in using the concentric relaxation method. Namely, it uses an adjustable region around the outage to solve for the outage case overloads. In reference 1, this is applied only to the linear (DC) power flow; it has subsequently been extended for AC network analysis. To perform the analysis in the bounding technique, we define three subsystems of the power system as follows:

N1 = the subsystem immediately surrounding the outaged line
N2 = the external subsystem that we shall not solve in detail
N3 = the set of boundary buses that separate N1 and N2

The subsystems appear as shown in Figure 7.14. The bounding method is based on the fact that we can make certain assumptions about the phase angle spread across the lines in N2, given the injections in N1 and the maximum phase angle appearing across any two buses in N3.

7.6.5 Adaptive Localization

The concentric relaxation and the bounding methods described above are both based on the fact that a power system line outage will not affect all bus voltages, bus phase angles, and line flows equally, but will in fact produce greater changes in the neighborhood of the outage than at a distance from the outaged line. This concept led to an entire power flow method that took advantage of this.

The adaptive localization method (see reference 2) is best used to solve for contingency outage cases starting at a solved base case. If a power flow is solved and then a line is removed but the bus voltage magnitudes and phase angles are kept the same as the base case, the power flow algorithm will only see significant mismatch[4]

[4] Mismatch is the error between the scheduled P and Q at the bus and the P and Q calculated from the bus voltage and current, where the current is a function of the bus voltages and phase angles at the buses connected to that bus.

www.mepcafe.com

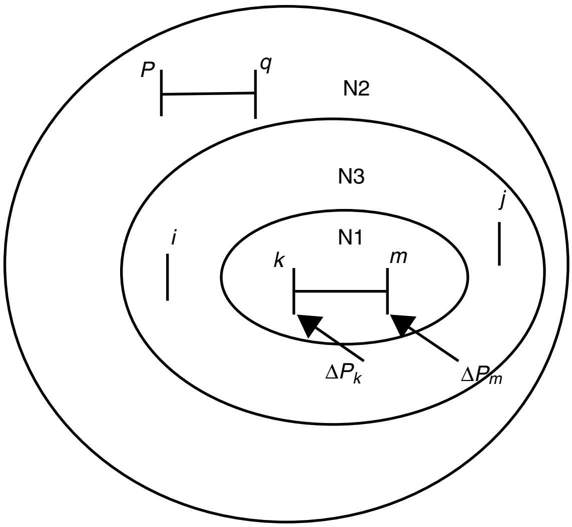


FIGURE 7.14 Layers used in bounding analysis.

error at the two buses at the ends of the line that is outaged. In developing the adaptive localization method, we need to avoid what happens in a conventional power flow where we solve from a default "flat start"[5] and then all the buses become part of the solution and all bus voltage magnitudes and phase angles would be adjusted each iteration. The adaptive localization method developers simply reasoned that in a contingency case only the buses at the end of the outaged line have a significant error, so they built an algorithm that only calculated changes to the voltage magnitude and phase angle at the buses with the significant mismatch error and left all the other buses at their initial solved value. If after the first iteration, the calculated changes to the buses at the ends of the outaged line caused more changes to appear at other buses nearby, then the algorithm adjusted those as well. The method can be explained, then, as a power flow that only adjusts buses that have a significant mismatch; in fact, it only calculates the mismatch for those buses until the algorithm detects that mismatch errors are spreading. Typically, the algorithm starts with errors at the outaged line, then more errors appear at other buses, and, as the algorithm converges to the new solution, the number of buses with mismatch errors needing adjustment declines until only a few are left and then no mismatches greater than the convergence tolerance remain at the very end.

An example showing this is a power flow on a case with 2730 buses shown in reference 3. The first table shows the iterations for the standard power flow solution using a flat start (Table 7.4a and 7.4b).

[5] Flat start for a power flow means all buses are set to 1.0 per unit at 0°, with generators set to their scheduled voltage at 0°.

www.mepcafe.com

Table 7.4a Iterations from a Flat Start on 2730 Bus Case

Iteration Number	Number of Buses Above 10^{-5} ΔP	Number of Buses Above 10^{-5} ΔQ
1	561	1282
2	1672	153
3	1308	108
4	458	74
5	7	0
6	0	0

Table 7.4b Iterations from the Solved 2730 Bus Case Using the Adaptive Localization Algorithm

Iteration Number	Number of Buses Above 10^{-5} ΔP	Number of Buses Above 10^{-5} ΔQ
1	1	0
2	194	6
3	22	0
4	0	0

Next, we see the tremendous savings using the adaptive localization algorithm. Here the same outage case is solved but starting at the base-solved values and only adjusting buses needing correction.

Not only can the adaptive localization algorithm solve contingency outage cases much faster, it can be used in the same manner as the 1P1Q method to do a partial solution on each case to order them and then solve them to completion starting at the worst case.

One can examine how an outage spreads by using the Gauss-Seidel power flow algorithm in a recursive manner. Start the solution at the buses connected to the outaged equipment, then solve the buses in the first layer, then the second layer, until the changes are within the solution tolerance desired. This will show the extent of the outage impact. Iterate by starting at the buses connected at the outaged equipment and repeating the previous process. The number of iterations is small if the impact is small, large if the impact is severe.

APPENDIX 7A

AC Power Flow Sample Cases

This appendix presents seven different AC power flows run on the six-bus test system shown in Figure 7.15.

www.mepcafe.com

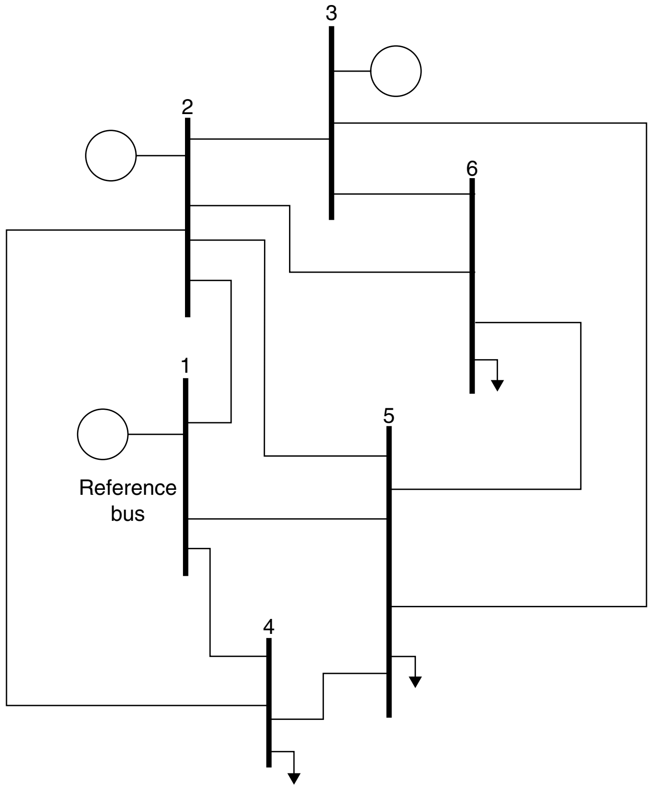


FIGURE 7.15 Six-bus test system

Case 1 Base Case

Voltages are all within their limits of 0.95–1.07 pu; the reference bus is scheduled at 1.07 pu and is therefore labeled at UL for upper limit.
All reactive power limits are being met and all line flows are within their MW limits.
All loads are at the base-case value of 100 MW and 15 MVAR
The real power losses are small and the reactive power losses are negative (meaning more reactive power is coming into the system from capacitive MVARs than is lost by the transmission lines).

www.mepcafe.com

Note: UL = quantity at or above upper limit
LL = quantity at or below lower limit

Total Pgen	=	312.96	MW	Total Qgen	=	30.01	MVAR
Total PLoad	=	300.00	MW	Total Qload	=	45.00	MVAR
Total PLosses	=	12.96	MW	Total QLosses	=	-14.99	MVAR

Bus	Pmin MW	Pgen MW	Pmax MW	Qmin MVAR	Qgen MVAR	Qmax MVAR	Pload MW	Qload MVAR	Vmin pu	Vbus pu		Vbus kV	Vmax pu
1	50.0	213.0	450.0	-100.0	-10.8	150.0	0.0	0.0	0.95	1.07	UL	246.1	1.07
2	37.5	50.0	150.0	-100.0	21.8	150.0	0.0	0.0	0.95	1.05		241.5	1.07
3	45.0	50.0	180.0	-100.0	19.0	120.0	0.0	0.0	0.95	1.05		241.5	1.07
4							100.0	15.0	0.95	1.03		236.3	1.07
5							100.0	15.0	0.95	1.02		235.1	1.07
6							100.0	15.0	0.95	1.02		235.7	1.07

Bus	Vmag kV	angle deg	Pgen MW	Qgen MVAR	Pload MW	Qload MW	To Bus	Pline MW	Qline MVAR	Max Flow MW
1	246.1	0.00	212.96	-10.76	0.00	0.00				
							2	62.18	-18.28	100
							4	82.80	5.87	100
							5	67.98	1.65	100
2	241.5	-7.18	50.00	21.76	0.00	0.00				
							1	-58.58	20.99	100
							3	14.76	-5.99	60
							4	28.86	8.66	60
							5	21.94	0.85	60
							6	43.01	-2.74	60
3	241.5	-9.16	50.00	19.02	0.00	0.00				
							2	-14.66	-0.11	60
							5	12.43	2.88	60
							6	52.23	16.25	60
4	236.3	-8.45			100.00	15.00				
							1	-79.77	1.82	100
							2	-28.44	-9.98	60
							5	8.21	-6.85	60
5	235.1	-10.53			100.00	15.00				
							1	-64.73	3.95	100
							2	-21.50	-3.81	60
							3	-12.23	-7.81	60
							4	-8.07	-1.27	60
							6	6.53	-6.07	60
6	235.7	-11.76			100.00	15.00				
							2	-41.84	0.72	60
							3	-51.68	-15.65	60
							5	-6.48	-0.07	60

www.mepcafe.com

Case 2 Generator Outage

The generator on bus 3 is lost with base-case (Case 1) loads. The deficit in generation is made up by an increase in the output of the reference generator on bus 1.

Bus voltages on 4, 5, and 6 drop slightly and one slight overload appears on line 1–4.

Total Pgen	=	320.41	MW	Total Qgen	=	53.40	MVAR
Total PLoad	=	300.00	MW	Total QLoad	=	45.00	MVAR
Total PLosses	=	20.41	MW	Total QLosses	=	8.40	MVAR

Bus	Pmin MW	Pgen MW	Pmax MW	Qmin MVAR	Qgen MVAR	Qmax MVAR	Pload MW	Qload MVAR	Vmin pu	Vbus pu	Vbus kV	Vmax pu
1	50.0	270.4	450.0	-100.0	-14.6	150.0	0.0	0.0	0.95	1.07 UL	246.1	1.07
2	37.5	50.0	150.0	-100.0	68.0	150.0	0.0	0.0	0.95	1.05	241.5	1.07
3							0.0	0.0	0.95	1.02	233.6	1.07
4							100.0	15.0	0.95	1.03	235.9	1.07
5							100.0	15.0	0.95	1.01	231.7	1.07
6							100.0	15.0	0.95	1.00	230.2	1.07

Bus	Vmag kV	angle deg	Pgen MW	Qgen MVAR	Pload MW	Qload MW	To Bus	Pline MW	Qline MVAR		Max Flow MW
1	246.1	0.00	270.41	-14.61	0.00	0.00					
							2	84.85	-25.66		100
							4	100.41	5.22	UL	100
							5	85.15	5.83		100
2	241.5	-9.89	50.00	68.01	0.00	0.00					
							1	-78.08	34.69		100
							3	31.61	5.77		60
							4	16.97	15.95		60
							5	23.53	5.48		60
							6	55.98	6.12		60
3	233.6	-13.90			0.00	0.00					
							2	-31.12	-9.72		60
							5	-2.11	1.70		60
							6	33.23	8.01		60
4	235.9	-10.34			100.00	15.00					
							1	-95.98	8.10		100
							2	-16.70	-17.58		60
							5	12.69	-5.51		60
5	231.7	-13.30			100.00	15.00					
							1	-80.02	6.92		100
							2	-22.97	-8.05		60
							3	2.14	-6.76		60
							4	-12.38	-2.14		60
							6	13.24	-4.97		60
6	230.2	-15.67			100.00	15.00					
							2	-53.94	-5.55		60
							3	-33.00	-8.90		60
							5	-13.06	-0.55		60

www.mepcafe.com

Case 3 Line Outage Case

Line outage on line 1–4, base-case conditions on load buses.
Voltage at load buses 4, 5, and 6 fall but not below low limits.
Lines 1–2, 1–5, and 2–4 are now above their max MW limits.
Reactive losses reverse from −14 to +22 MVAR; real power losses double reflecting increased line currents.

```
Total Pgen     =   328.58  MW    Total Qgen     =   67.74  MVAR
Total PLoad    =   300.00  MW    Total Qload    =   45.00  MVAR
Total PLosses  =    28.58  MW    Total QLosses  =   22.74  MVAR
```

Bus	Pmin MW	Pgen MW	Pmax MW	Qmin MVAR	Qgen MVAR	Qmax MVAR	Pload MW	Qload MVAR	Vmin pu	Vbus pu		Vbus kV	Vmax pu
1	50.0	228.6	450.0	-100.0	-32.2	150.0	0.0	0.0	0.95	1.07	UL	246.1	1.07
2	37.5	50.0	150.0	-100.0	75.7	150.0	0.0	0.0	0.95	1.05		241.5	1.07
3	45.0	50.0	180.0	-100.0	24.2	120.0	0.0	0.0	0.95	1.05		241.5	1.07
4							100.0	15.0	0.95	0.99		228.0	1.07
5							100.0	15.0	0.95	1.01		233.3	1.07
6							100.0	15.0	0.95	1.02		235.4	1.07

Bus	Vmag kV	angle deg	Pgen MW	Qgen MVAR	Pload MW	Qload MW	To Bus	Pline MW	Qline MVAR		Max Flow MW
1	246.1	0.00	228.58	-32.18	0.00	0.00					
							2	123.57	-35.59	UL	100
							5	105.01	3.42	UL	100
2	241.5	-14.46	50.00	75.71	0.00	0.00					
							1	-109.27	59.71	UL	100
							3	11.35	-5.42		60
							4	93.09	17.20	UL	60
							5	15.60	5.30		60
							6	39.22	-1.07		60
3	241.5	-15.99	50.00	24.20	0.00	0.00					
							2	-11.29	-0.89		60
							5	8.92	7.52		60
							6	52.37	17.57		60
4	228.0	-19.08			100.00	15.00					
							2	-89.01	-11.12	UL	60
							5	-10.99	-3.88		60
5	233.3	-16.57			100.00	15.00					
							1	-97.27	19.08		100
							2	-15.33	-8.74		60
							3	-8.72	-12.41		60
							4	11.23	-3.68		60
							6	10.08	-9.24		60
6	235.4	-18.58			100.00	15.00					
							2	-38.25	-1.51		60
							3	-51.81	-16.91		60
							5	-9.95	3.42		60

www.mepcafe.com

Case 4 Double Load on Bus 4

All lines and generators in.
Bus 4 load goes to 200 MW and 30 MVAR.
Reactive power losses go positive and large, and bus voltages at bus 4, 5, and 6 are low.
Line flow overloads appear on lines 1–4 and 2–4.

Total Pgen	=	428.79	MW	Total Qgen	=	93.04	MVAR
Total PLoad	=	400.00	MW	Total QLoad	=	60.00	MVAR
Total PLosses	=	28.79	MW	Total QLosses	=	33.04	MVAR

Bus	Pmin MW	Pgen MW	Pmax MW	Qmin MVAR	Qgen MVAR	Qmax MVAR	Pload MW	Qload MVAR	Vmin pu	Vbus pu		Vbus kV	Vmax pu
1	50.0	328.8	450.0	-100.0	-2.9	150.0	0.0	0.0	0.95	1.07	UL	246.1	1.07
2	37.5	50.0	150.0	-100.0	72.3	150.0	0.0	0.0	0.95	1.05		241.5	1.07
3	45.0	50.0	180.0	-100.0	23.6	120.0	0.0	0.0	0.95	1.05		241.5	1.07
4							200.0	30.0	0.95	0.99		227.2	1.07
5							100.0	15.0	0.95	1.02		233.5	1.07
6							100.0	15.0	0.95	1.02		235.4	1.07

Bus	Vmag kV	angle deg	Pgen MW	Qgen MVAR	Pload MW	Qload MW	To Bus	Pline MW	Qline MVAR		Max Flow MW
1	246.1	0.00	328.79	-2.86	0.00	0.00					
							2	99.40	-29.78		100
							4	138.95	23.82	UL	100
							5	90.44	3.09		100
2	241.5	-11.62	50.00	72.34	0.00	0.00					
							1	-90.11	43.86		100
							3	12.73	-5.66		60
							4	68.54	31.29	UL	60
							5	18.07	4.37		60
							6	40.77	-1.52		60
3	241.5	-13.34	50.00	23.57	0.00	0.00					
							2	-12.66	-0.58		60
							5	10.39	6.62		60
							6	52.27	17.52		60
4	227.2	-14.51			200.00	30.00					
							1	-130.22	6.86	UL	100
							2	-65.93	-28.15	UL	60
							5	-3.85	-8.70		60
5	233.5	-14.18			100.00	15.00					
							1	-84.69	11.93		100
							2	-17.74	-7.63		60
							3	-10.17	-11.49		60
							4	3.92	0.84		60
							6	8.68	-8.64		60
6	235.4	-15.92			100.00	15.00					
							2	-39.71	-0.84		60
							3	-51.71	-16.88		60
							5	-8.58	2.72		60

www.mepcafe.com

Case 5 Max Load on Bus 4

Load on bus 4 at 220 MW and 33 MVAR
Extremely high reactive losses, voltage at bus 4 is at 0.98 pu

Total Pgen	=	453.37	MW	Total Qgen	=	109.62	MVAR
Total PLoad	=	420.00	MW	Total Qload	=	63.00	MVAR
Total PLosses	=	33.37	MW	Total QLosses	=	46.62	MVAR

Bus	Pmin MW	Pgen MW	Pmax MW	Qmin MVAR	Qgen MVAR	Qmax MVAR	Pload MW	Qload MVAR	Vmin pu	Vbus pu	Vbus kV	Vmax pu
1	50.0	353.4	450.0	-100.0	0.5	150.0	0.0	0.0	0.95	1.07 UL	246.1	1.07
2	37.5	50.0	150.0	-100.0	84.4	150.0	0.0	0.0	0.95	1.05	241.5	1.07
3	45.0	50.0	180.0	-100.0	24.7	120.0	0.0	0.0	0.95	1.05	241.5	1.07
4							220.0	33.0	0.95	0.98	225.1	1.07
5							100.0	15.0	0.95	1.01	233.0	1.07
6							100.0	15.0	0.95	1.02	235.3	1.07

Bus	Vmag kV	angle deg	Pgen MW	Qgen MVAR	Pload MW	Qload MW	To Bus	Pline MW	Qline MVAR		Max Flow MW
1	246.1	0.00	353.37	0.49	0.00	0.00					
							2	107.61	-31.90	UL	100
							4	150.48	28.71	UL	100
							5	95.27	3.68		100
2	241.5	-12.59	50.00	84.42	0.00	0.00					
							1	-96.73	49.16		100
							3	12.31	-5.58		60
							4	76.89	36.84	UL	60
							5	17.25	5.22		60
							6	40.29	-1.23		60
3	241.5	-14.24	50.00	24.71	0.00	0.00					
							2	-12.23	-0.68		60
							5	9.96	7.53		60
							6	52.28	17.85		60
4	225.1	-15.82			220.00	33.00					
							1	-140.17	8.32	UL	100
							2	-73.55	-32.23	UL	60
							5	-6.28	-9.09		60
5	233.0	-14.97			100.00	15.00					
							1	-88.89	13.72		100
							2	-16.93	-8.52		60
							3	-9.73	-12.37		60
							4	6.42	1.43		60
							6	9.14	-9.26		60
6	235.3	-16.83			100.00	15.00					
							2	-39.26	-1.20		60
							3	-51.72	-17.20		60
							5	-9.02	3.39		60

www.mepcafe.com

Case 6 Combination of High Load and Line Outage

Line 1–4 out and bus 4 load at 200 MW and 30 MVAR
Lines 1–2, 1–5 and 2–4 overload
Voltage at bus 4 is 0.87 pu, well below its low limit
Generator 2 is at its max VAR limit, so its voltage is now below scheduled
Extremely high reactive power losses and real power losses.

Total Pgen	=	489.49	MW	Total Qgen	=	229.51	MVAR
Total PLoad	=	400.00	MW	Total QLoad	=	60.00	MVAR
Total PLosses	=	89.49	MW	Total QLosses	=	169.51	MVAR

Bus	Pmin MW	Pgen MW	Pmax MW	Qmin MVAR	Qgen MVAR	Qmax MVAR	Pload MW	Qload MVAR	Vmin pu	Vbus pu	Vbus kV	Vmax pu
1	50.0	389.5	450.0	-100.0	-1.5	150.0	0.0	0.0	0.95	1.07 UL	246.1	1.07
2	37.5	50.0	150.0	-100.0	150.0 UL	150.0	0.0	0.0	0.95	1.01	231.9	1.07
3	45.0	50.0	180.0	-100.0	81.0	120.0	0.0	0.0	0.95	1.05	241.5	1.07
4							200.0	30.0	0.95	0.87 LL	200.1	1.07
5							100.0	15.0	0.95	0.97	223.3	1.07
6							100.0	15.0	0.95	1.00	230.8	1.07

Bus	Vmag kV	angle deg	Pgen MW	Qgen MVAR	Pload MW	Qload MW	To Bus	Pline MW	Qline MVAR		Max Flow MW
1	246.1	0.00	389.49	-1.50	0.00	0.00					
							2	226.22	-26.81	UL	100
							5	163.27	25.31	UL	100
2	231.9	-26.24	50.00	150.00	0.00	0.00					
							1	-181.00	112.94	UL	100
							3	8.60	-21.41		60
							4	182.74	60.47	UL	60
							5	6.16	8.50		60
							6	33.51	-10.50		60
3	241.5	-27.90	50.00	81.01	0.00	0.00					
							2	-8.40	16.06		60
							5	5.44	26.77		60
							6	52.96	38.18		60
4	200.1	-36.22			200.00	30.00					
							2	-164.45	-25.68	UL	60
							5	-35.55	-4.32		60
5	223.3	-26.70			100.00	15.00					
							1	-144.06	40.44	UL	100
							2	-6.01	-11.98		60
							3	-4.46	-29.76		60
							4	38.89	4.21		60
							6	15.64	-17.91		60
6	230.8	-30.35			100.00	15.00					
							2	-32.69	7.78		60
							3	-52.17	-36.35		60
							5	-15.14	13.57		60

www.mepcafe.com

Case 7 Max Load on Bus 4 of 220 MW and 33 MVAR, and Line 1–4 Out

Bus 4 voltage is 0.8 pu, and bus 5 is down to 0.95 pu.
Generators 2 and 3 are both at max VARs and their bus voltages are below scheduled.
Note: Any more load added to bus 4 results in a case where the power flow will not solve.

Total Pgen	=	537.44	MW	Total Qgen	=	301.56	MVAR
Total PLoad	=	420.00	MW	Total Qload	=	63.00	MVAR
Total PLosses	=	117.44	MW	Total QLosses	=	238.56	MVAR

Bus	Pmin MW	Pgen MW	Pmax MW	Qmin MVAR	Qgen MVAR		Qmax MVAR	Pload MW	Qload MVAR	Vmin pu	Vbus pu		Vbus kV	Vmax pu
1	50.0	437.4	450.0	-100.0	31.6		150.0	0.0	0.0	0.95	1.07	UL	246.1	1.07
2	37.5	50.0	150.0	-100.0	150.0	UL	150.0	0.0	0.0	0.95	0.97		223.4	1.07
3	45.0	50.0	180.0	-100.0	120.0	UL	120.0	0.0	0.0	0.95	1.05		240.6	1.07
4								220.0	33.0	0.95	0.80	LL	185.0	1.07
5								100.0	15.0	0.95	0.94	LL	216.6	1.07
6								100.0	15.0	0.95	0.99		226.6	1.07

Bus	Vmag kV	angle deg	Pgen MW	Qgen MVAR	Pload MW	Qload MW	To Bus	Pline MW	Qline MVAR		Max Flow MW
1	246.1	0.00	437.44	31.56	0.00	0.00					
							2	256.97	-8.21	UL	100
							5	180.47	39.77	UL	100
2	223.4	-30.02	50.00	150.00	0.00	0.00					
							1	-199.26	119.45	UL	100
							3	9.21	-33.49		60
							4	204.29	76.52	UL	60
							5	3.73	6.41		60
							6	32.03	-18.88		60
3	240.6	-32.18	50.00	120.00	0.00	0.00					
							2	-8.67	30.10		60
							5	5.26	37.07		60
							6	53.41	52.83		60
4	185.0	-42.25			220.00	33.00					
							2	-178.99	-27.49	UL	60
							5	-41.01	-5.51		60
5	216.6	-30.20			100.00	15.00					
							1	-156.41	44.36	UL	100
							2	-3.64	-9.81		60
							3	-3.49	-38.19		60
							4	46.23	9.82		60
							6	17.30	-21.18		60
6	226.6	-34.55			100.00	15.00					
							2	-31.07	16.85		60
							3	-52.35	-49.64		60
							5	-16.58	17.78		60

NOTE AT LOAD BUS 4 SET TO 250 Mw AND 37.5 MVAR POWER FLOW WILL NOT CONVERGE

www.mepcafe.com

APPENDIX 7B
Calculation of Network Sensitivity Factors

7B.1 CALCULATION OF PTDF FACTORS

The PTDF assumes a transfer of power from a sending bus (bus s) to a receiving bus (bus r). That is, an amount of power is injected into bus s and is taken from bus r as shown in the Figure 7.16. The PTDF gives the fraction of that transferred power that ends up flowing on line ℓ.

By definition,

$$\text{PTDF}_{s,r,\ell} = \frac{\Delta f_\ell}{\Delta P_{s\text{ to }r}}$$

or

$$\Delta f_\ell = \text{PTDF}_{s,r,\ell}\Delta P_{s\text{ to }r}$$

or

$$\tilde{f}_\ell = f_\ell^0 + \text{PTDF}_{s,r,\ell}\Delta P_{s\text{ to }r}$$

If the transfer is in the opposite direction from bus r to bus s, then

$$\text{PTDF}_{r,s,\ell} = -\text{PTDF}_{s,r,\ell}$$

Numerical range: $-1 \le \text{PTDF}_{s,r,\ell} \le +1$

If we model the power system using the linear power flow, then the effect on bus phase angles for a transfer of one MW from bus s to bus r can be modeled using the linear power flow equation:

$$\underline{\Delta\theta} = [X]\underline{\Delta P_{s\text{ to }r}}$$

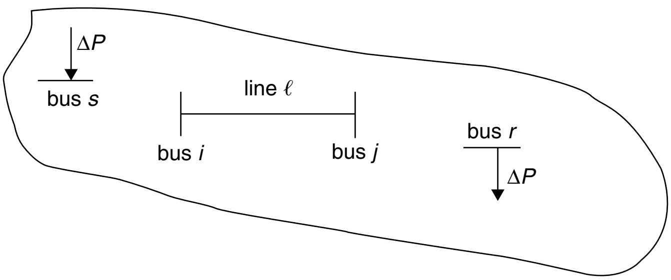


FIGURE 7.16 Effect of flow on line ℓ of a transfer of power from bus s to bus r.

www.mepcafe.com

where

$$\underline{\Delta P_{s\text{ to }r}} = \begin{bmatrix} 0 \\ \vdots \\ +1 \\ -1 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \\ \\ s \\ r \\ \\ \\ \end{matrix}$$

Then the phase angle changes are

$$\begin{bmatrix} \Delta\theta_1 \\ \Delta\theta_2 \\ \vdots \\ \\ \Delta\theta_{n\text{bus}} \end{bmatrix} = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1,n\text{bus}} \\ X_{21} & X_{22} & & \\ \vdots & & \ddots & \\ & & & \\ X_{n\text{bus},1} & & & X_{n\text{bus},n\text{bus}} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ +1 \\ -1 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \\ \\ s \\ r \\ \\ \\ \end{matrix}$$

Then the phase angle change on buses i and j are

$$\Delta\theta_i = X_{is} - X_{ir}$$

$$\Delta\theta_j = X_{js} - X_{jr}$$

The change in flow on the line ℓ is

$$\Delta f_\ell = \frac{1}{x_\ell}\left(\Delta\theta_i - \Delta\theta_j\right)$$

then

$$\Delta f_\ell = \frac{1}{x_\ell}((X_{is} - X_{ir}) - (X_{js} - X_{jr}))$$

Therefore, the PTDF is

$$\text{PTDF}_{s,r,\ell} = \frac{1}{x_\ell}((X_{is} - X_{ir}) - (X_{js} - X_{jr}))$$

www.mepcafe.com

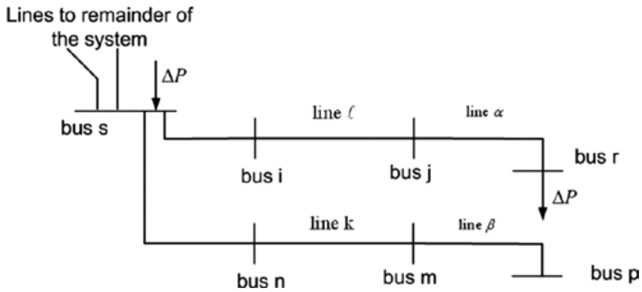


FIGURE 7.17 Effect of flow on line ℓ of a transfer of power from bus *s* to bus *r* with additional lines shown.

$$\text{if } s \text{ is the reference bus, } X_{is} = 0 \text{ and } X_{js} = 0$$
$$\text{if } r \text{ is the reference bus, } X_{ir} = 0 \text{ and } X_{jr} = 0$$
$$\text{if } i \text{ is the reference bus, } X_{is} = 0 \text{ and } X_{ir} = 0$$
$$\text{if } j \text{ is the reference bus, } X_{js} = 0 \text{ and } X_{jr} = 0$$

Note that the PTDF depends only on the network parameters and is not affected by the loading or voltages on the network. Furthermore, the PTDF does not depend on the location of the reference bus in the network.

Some special cases in calculating the PTDFs for a network are illustrated in this simple Figure (Figure 7.17):

Case 1 $\text{PTDF}_{s,r,l} = 1$

When the PTDF is exactly 1, it is an indication that all of the transferred power from *s* to *r* must flow through line ℓ; it is also important to note that if line ℓ is opened, buses *j* and *r* will be separated, or islanded from buses *s*, *i*, *n*, *m*, and *p*, so that PTDF = 1 indicates a line that if opened will cause islanding. This will play a key role when we discuss LODF factors in the next section.

Case 2 $\text{PTDF}_{s,r,k} = 0$

If the PTDF is exactly 0, then there is no loop in the network that allows power to flow from *s* to *r* with some of that power transfer passing through line *k*. Here line *k* will have a $\text{PTDF}_{s,r,k}$ of 0 since none of the transfer of power from *s* to *r* goes through line *k*.

Case 3 A Radial Line

Both line α and line β in the previous Figure are radial lines. That is, a radial line has at least one bus that has no connections to other lines. In this case, $\text{PTDF}_{j,r,\alpha} = 1$ and all other PTDFs for transfer between *j* and *r* for lines other than α will be 0. Similarly, $\text{PTDF}_{m,p,\beta} = 1$ and all other PTDFs for transfer between *m* and *p* for lines other than β will be 0.

www.mepcafe.com

7B.2 CALCULATION OF LODF FACTORS

In this section, we shall derive the LODF factors that were defined previously as

$$\text{LODF}_{\ell,k} = \frac{\Delta f_\ell}{f_k^0}$$

where

$\text{LODF}_{\ell,k}$ = line outage distribution factor when monitoring line ℓ after an outage on line k
Δf_ℓ = change in MW flow on line ℓ
f_k^0 = original flow on line k before it was outaged (opened)

If one knows the power on line ℓ and line k, the flow on line ℓ with line k out can be determined using "LODF" factors.

$$\hat{f}_\ell = f_\ell^0 + \text{LODF}_{\ell,k} f_k^0$$

where
f_ℓ^0, f_k^0 preoutage flows on lines ℓ and k, respectively
$\hat{f}_\ell$ = flow on line ℓ with line k out

It is useful to see this in a diagram as shown in Figure 7.18.

First of all we will describe a method to simulate the opening of line k by adding an injection into bus n and bus m. This is shown in Figure 7.19, where line k is connected to lines in the remainder of the system at bus n through a breaker, and similarly line k is connected to lines in the remainder of the system at bus m through a breaker. When the breakers are opened, middle of the Figure, the line flow is 0 on line k.

We note that the original power flowing on line k from bus n to bus m is P_{nm} and that when the injections ΔP_n and ΔP_m are added to bus n and bus m, respectively, the resulting flow on line k is $\tilde{P}_{nm}$, as in the illustration at the bottom of the Figure.

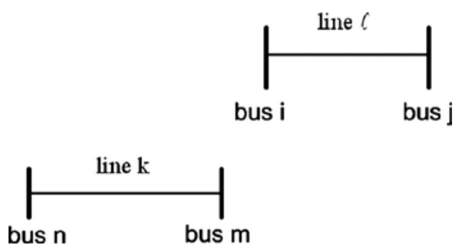


FIGURE 7.18 Effect of flow on line ℓ with line k out.

www.mepcafe.com

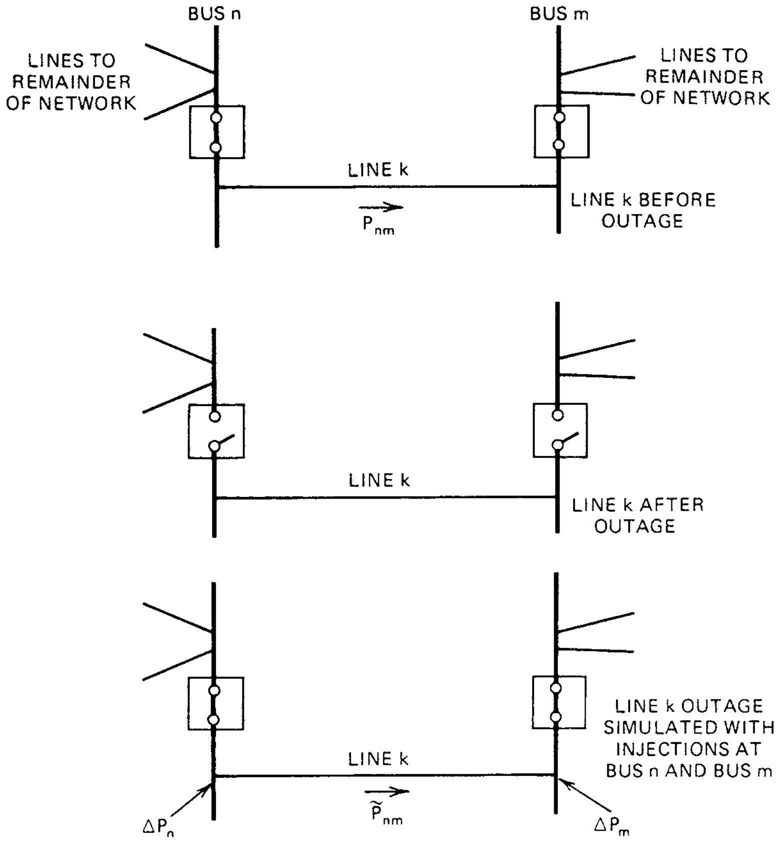


FIGURE 7.19 Effect of flow on line ℓ with line *k* out, modeled as injections at buses *n* and *m*.

The opening of line *k* can then be simulated if

$$\Delta P_n = \tilde{P}_{nm}$$

and

$$\Delta P_m = -\tilde{P}_{nm}$$

This means that all of the injected power into bus *n* flows in line *k* and out of bus *m* so that there is no flow through the breaker connecting bus *n* to the remainder of the system and no flow through the breaker connecting bus *m* to the system. Zero flow in the two breakers is the same as if they were opened.

$\tilde{P}_{nm}$ can be calculated easily if we note that the flow on line *k* for an injection into bus *n* and out of bus *m* is simply

$$\tilde{P}_{nm} = P_{nm} + \text{PTDF}_{n,m,k}\Delta P_n$$

www.mepcafe.com

We are using the PTDF to calculate how much of the injection ΔP_n ends up flowing on line k, but by definition $\Delta P_n = \tilde{P}_{nm}$, then

$$\tilde{P}_{nm} = P_{nm} + \text{PTDF}_{n,m,k}\tilde{P}_{nm}$$

or

$$\tilde{P}_{nm}\left(1 - \text{PTDF}_{n,m,k}\right) = P_{nm}$$

or

$$\tilde{P}_{nm} = \left(\frac{1}{1 - \text{PTDF}_{n,m,k}}\right)P_{nm}$$

The change in flow on line ℓ from i to j is

$$\Delta f_\ell = \text{PTDF}_{n,m,\ell}\tilde{P}_{nm} = \text{PTDF}_{n,m,\ell}\left(\frac{1}{1 - \text{PTDF}_{n,m,k}}\right)P_{nm}$$

Thus, the LODF giving the change in flow on line ℓ is simply

$$\text{LODF}_{\ell,k} = \text{PTDF}_{n,m,\ell}\left(\frac{1}{1 - \text{PTDF}_{n,m,k}}\right)$$

So that

$$\Delta f_\ell = \text{LODF}_{\ell,k}P_{nm}$$

Thus, we simply multiply the preoutage flow on line k, P_{nm}, times $\text{LODF}_{\ell,k}$ to get the change in flow on line ℓ, and then the new flow on line ℓ, $\tilde{f}_\ell$, with an outage on line k is as follows:

$$\tilde{f}_\ell = f_\ell^0 + \text{LODF}_{\ell,k}P_{nm}$$

or

$$\tilde{f}_\ell = f_\ell^0 + \text{LODF}_{\ell,k}f_k^0$$

7B.2.1 Special Cases

Case 1

Line k is a line that leaves the system islanded if it is opened. In such a case, $\text{PTDF}_{n,m,k} = 1$ and the expression for the LODF results in 1 over 0. In reality, it is not possible using a linear power flow to tell what effect of islanding the system will have on any given line; therefore, the best means to handle this is to set LODF = 0.

www.mepcafe.com

Case 2

Line k is a radial line. In such a case, the PTDF for line k is also 0, but opening it has a very predictable effect on the remainder of the system. Specifically, any flow that was flowing on the radial line now becomes an injection into the bus still connected to the system. For example, if the outaged line were line α in Figure 7.17 then any flow from bus j to r can be modeled as an injection back into bus j. This flow is actually a transfer from bus j to the reference bus. The effect of this transfer is simply the PTDF. We previously derived the formula for the PTDF repeated here:

$$\text{PTDF}_{s,r,\ell} = \frac{1}{x_\ell}((X_{is} - X_{ir}) - (X_{js} - X_{jr}))$$

$$\begin{aligned}
&\text{if } s \text{ is the reference bus, } X_{is} = 0 \text{ and } X_{js} = 0\\
&\text{if } r \text{ is the reference bus, } X_{ir} = 0 \text{ and } X_{jr} = 0\\
&\text{if } i \text{ is the reference bus, } X_{is} = 0 \text{ and } X_{ir} = 0\\
&\text{if } j \text{ is the reference bus, } X_{js} = 0 \text{ and } X_{jr} = 0
\end{aligned}$$

In Figure 7.20, we redraw Figure 7.16 with radial line α opened and we indicate an injection into bus j.

Given the definition for the PTDF factor, the flow on line ℓ when an injection is made into bus j, results in a transfer from bus j to the reference bus:

$$\text{PTDF}_{j,\text{ref},\ell} = \frac{1}{x_\ell}(X_{is} - X_{js})$$

In the first and second editions of this text, this PTDF from a bus to the reference was called an "a" factor and the matrix of such factors was the "a matrix."

Now the flow on line ℓ when a radial line α is opened is simply

$$\tilde{f}_\ell = f_\ell^0 + \text{PTDF}_{m,\text{ref},\alpha} f_\alpha^0$$

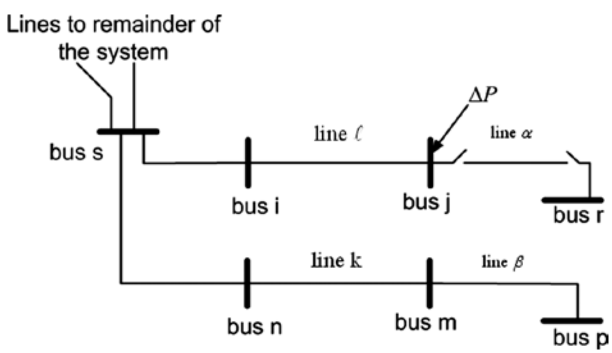


FIGURE 7.20 Effect of flow on line ℓ with line α opened.

www.mepcafe.com

7B.3 COMPENSATED PTDF FACTORS

Often it is useful to have a means to calculate the PTDF for a transfer from bus s to bus r when line k is opened at the same time. We start with the calculation of flow on line ℓ with line k opened; this was done earlier as follows:

$$\hat{f}_\ell = f_\ell^0 + \text{LODF}_{\ell,k} f_k^0$$

where

f_ℓ^0, f_k^0 preoutage flows on lines ℓ and k, respectively
$\hat{f}_\ell$ = flow on line ℓ with line k out

Now if we calculate the new flow resulting on both lines ℓ and k due to a transfer of power from bus s to bus r using the PTDF factors, we get

$$\tilde{f}_\ell = f_\ell^0 + \text{PTDF}_{s,r,\ell}\ \Delta P_{s \text{ to } r}$$

and

$$\tilde{f}_k = f_k^0 + \text{PTDF}_{s,r,k} \Delta P_{s \text{ to } r}$$

Then by superposition, we can add the effect of the transfer from bus s to bus r to the effect of the line outage on line k and we get

$$\begin{aligned}\hat{f}_\ell &= f_\ell^0 + \text{LODF}_{\ell,k} f_k^0 \\ &= \left(f_\ell^0 + \text{PTDF}_{s,r,\ell} \Delta P_{s \text{ to } r}\right) + \text{LODF}_{\ell,k}\left(f_k^0 + \text{PTDF}_{s,r,k} \Delta P_{s \text{ to } r}\right) \\ &= \left(f_\ell^0 + \text{LODF}_{\ell,k} f_k^0\right) + \left(\text{PTDF}_{s,r,\ell} + \text{LODF}_{\ell,k}\text{PTDF}_{s,r,k}\right)\Delta P_{s \text{ to } r}\end{aligned}$$

Now we can see that the PTDF for the transfer s to r with line k out is

$$\text{PTDF}_{s,r,\ell} + \text{LODF}_{\ell,k}\text{PTDF}_{s,r,k}$$

These, “compensated” PTDFs play a key role in building “contingency constraints” or “outage constraints” into an optimal power flow using the linear power flow model.

PROBLEMS

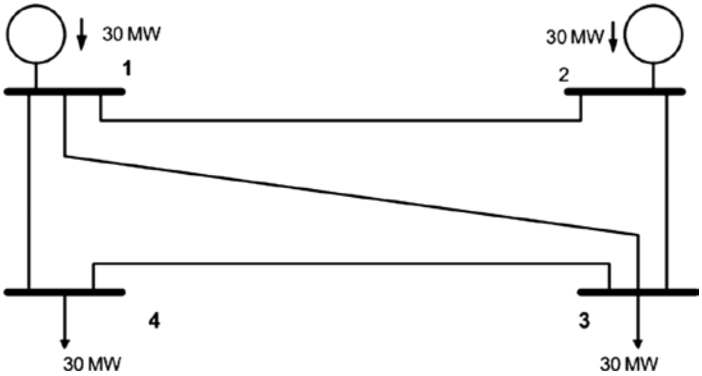
7.1 This problem uses the network and data found in Problem 6.7. You will have to solve all parts of 6.7 before you can solve this problem.

a. Build the PTDF matrix for transactions from each generator to the load on bus 3, that is, from generator on bus 1 to the load on bus 3, from generator on bus 2 to the load on bus 3, and from the generator on bus 4 to the load on bus 3.

b. Calculate the LODF factors for an outage of line 1–3.

7.2 Consider the following power system network. Both generators are at 30 MW output and both loads are consuming 30 MW. Bus 1 is the reference bus:

www.mepcafe.com



The data for this network is as follows:

From Bus	To Bus	x	MW Limit
1	2	0.20	35.0
1	3	0.30	35.0
1	4	0.30	25.0
2	3	0.30	40.0
3	4	0.40	25.0

The initial power flows are calculated with a DC power flow and result in the following base flows. The line flow limits are also shown with the percent loading:

BASE TRANSMISSION LOADING

Path	From	To	Low	Flow	High	Percent Loading
1	1	2	–35.0	–8.87	35.0	25.4
2	1	3	–35.0	15.21	35.0	43.5
3	1	4	–25.0	23.66	25.0	94.6
4	2	3	–40.0	21.13	40.0	52.8
5	3	4	–25.0	6.34	25.0	25.4

The PTDF factors and the LODF factors for this system are as follows:

POWER TRANSFER DISTRIBUTION FACTOR (PTDF) MATRIX

Monitored Line	Transaction From(Sell) – To(Buy)			
	1 to 3	1 to 4	2 to 3	2 to 4
1 to 2	0.2958	0.1268	–0.4225	–0.5915
1 to 3	0.4930	0.2113	0.2958	0.0141
1 to 4	0.2113	0.6620	0.1268	0.5775
2 to 3	0.2958	0.1268	0.5775	0.4085
3 to 4	–0.2113	0.3380	–0.1268	0.4225

www.mepcafe.com

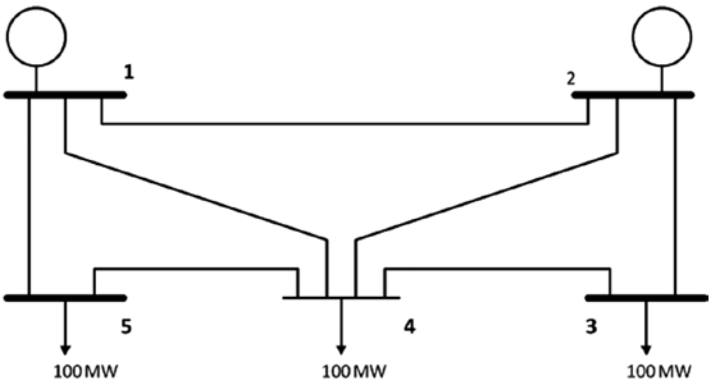
LINE OUTAGE DISTRIBUTION FACTOR (LODF) MATRIX
Monitored Line Outage of one circuit From – To

	1 to 2	1 to 3	1 to 4	2 to 3	3 to 4
1 to 2	0.0000	0.5833	0.3750	−1.0000	−0.3750
1 to 3	0.7000	0.0000	0.6250	0.7000	−0.6250
1 to 4	0.3000	0.4167	0.0000	0.3000	1.0000
2 to 3	−1.0000	0.5833	0.3750	0.0000	−0.3750
3 to 4	−0.3000	−0.4167	1.0000	−0.3000	0.0000

a. In this problem we are only concerned with outages on lines 1–2, 1–3, and 1–4. Do any of these outages, taken one outage at a time, result in overloads? If so, how much and what lines are overloaded?

b. The generator at bus 1 is going to reduce its output and at the same time the load at bus 4 is going to reduce its load until there are no overloads due to the lines listed in part (a). How much should the load on bus 4 and the generation on bus 1 be reduced to eliminate all overloads?

7.3 Consider the following power system network (both generators are at 150 MW and all loads are 100 MW). Consider bus 1 as the reference bus:



The initial flows calculated with a DC power flow and the LODF and PTDF factors for this system are as shown in the following:

a. Given the initial transmission system loading, and considering outages of only lines 12, 1–4, and 1–5, calculate any overloads due to line outages. Only check the contingency flows on lines 1–2, 1–4, and 1–5. Identify which line is overloaded and the overload amount.

b. Assume all lines are in. The generator at bus 1 is initially selling 100 MW (out of its 150 MW) to a customer on bus 4. The flow on line 1–5 is 88.05 MW and the operators want it reduced to only 75 MW by adjusting the 100 MW transaction from bus 1 to bus 4. How much should this transaction be adjusted? What is the new transaction MW amount?

c. As a result of the adjustment taken in part (b) to reduce the 1–5 flow, are there any flows that increase? If so, which flows are they and what are their flows after the adjustment?

www.mepcafe.com

Data for Problem 1:

Initial power flows:

From	To	Low	Flow	High	Percent Loading
1	2	-100.0	-6.19	100.0	6.2
1	5	-100.0	88.05	100.0	88.1
1	4	-100.0	68.14	100.0	68.1
2	3	-100.0	89.60	100.0	89.6
2	4	-100.0	54.20	100.0	54.2
3	4	-100.0	-10.40	100.0	10.4
4	5	-100.0	11.95	100.0	11.9

LINE OUTAGE DISTRIBUTION FACTOR (LODF) MATRIX

Monitored Line	Outage of one circuit From -To						
	1 to 2	1 to 5	1 to 4	2 to 3	2 to 4	3 to 4	4 to 5
1 to 2	0.0000	0.3913	0.6316	-0.4889	-0.6567	-0.4889	-0.3913
1 to 5	0.2727	0.0000	0.3684	0.1333	0.1791	0.1333	1.0000
1 to 4	0.7273	0.6087	0.0000	0.3556	0.4776	0.3556	-0.6087
2 to 3	-0.3333	0.1304	0.2105	0.0000	0.3433	-1.0000	-0.1304
2 to 4	-0.6667	0.2609	0.4211	0.5111	0.0000	0.5111	-0.2609
3 to 4	-0.3333	0.1304	0.2105	-1.0000	0.3433	0.0000	-0.1304
4 to 5	-0.2727	1.0000	-0.3684	-0.1333	-0.1791	-0.1333	0.0000

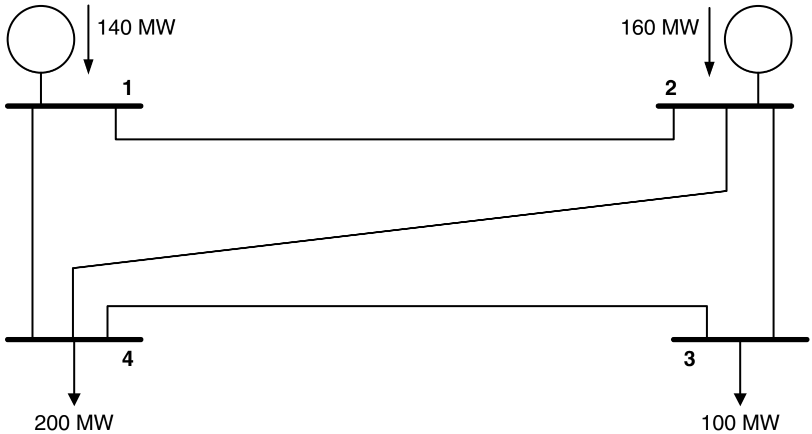
POWER TRANSFER DISTRIBUTION FACTOR (PTDF) MATRIX

Monitored Line	Transaction From(Sell) - To(Buy)					
	1 to 3	1 to 4	1 to 5	2 to 3	2 to 4	2 to 5
1 to 2	0.5619	0.3186	0.1195	-0.1460	-0.3894	-0.5885
1 to 5	0.1195	0.1858	0.6947	0.0398	0.1062	0.6150
1 to 4	0.3186	0.4956	0.1858	0.1062	0.2832	-0.0265
2 to 3	0.6040	0.1062	0.0398	0.7013	0.2035	0.1372
2 to 4	-0.0420	0.2124	0.0796	0.1527	0.4071	0.2743
3 to 4	-0.3960	0.1062	0.0398	-0.2987	0.2035	0.1372
4 to 5	-0.1195	-0.1858	0.3053	-0.0398	-0.1062	0.3850

www.mepcafe.com

Data for Problem 7.3

7.4 Consider the following four bus system with generation and load as shown. Bus 1 is the reference bus.



The data for this problem is shown here:

BASE TRANSMISSION LOADING

From	To	Low	Flow	High	Percent Loading
1	2	-200.0	28.73	200.0	14.4
2	4	-200.0	92.11	200.0	46.1
1	4	-300.0	111.27	300.0	37.1
2	3	-250.0	96.62	250.0	38.6
3	4	-300.0	-3.38	300.0	1.1

POWER TRANSFER DISTRIBUTION FACTOR (PTDF) MATRIX

Monitored Line	Transaction From(Sell) - To(Buy)			
	1 to 3	1 to 4	2 to 3	2 to 4
1 to 2	0.5915	0.4225	-0.1268	-0.2958
2 to 4	0.0141	0.2958	0.2113	0.4930
1 to 4	0.4085	0.5775	0.1268	0.2958
2 to 3	0.5775	0.1268	0.6620	0.2113
3 to 4	-0.4225	0.1268	-0.3380	0.2113

LINE OUTAGE DISTRIBUTION FACTOR (LODF) MATRIX

Monitored Line	Outage of one circuit From - To				
	1 to 2	2 to 4	1 to 4	2 to 3	3 to 4
1 to 2	0.0000	-0.5833	1.0000	-0.3750	-0.3750
2 to 4	-0.7000	0.0000	0.7000	0.6250	0.6250
1 to 4	1.0000	0.5833	0.0000	0.3750	0.3750
2 to 3	-0.3000	0.4167	0.3000	0.0000	-1.0000
3 to 4	-0.3000	0.4167	0.3000	-1.0000	0.0000

www.mepcafe.com

Line limits are as follows:

Line 1-2	200 MW
Line 2-4	200 MW
Line 1-4	300 MW
Line 2-3	250 MW
Line 3-4	300 MW

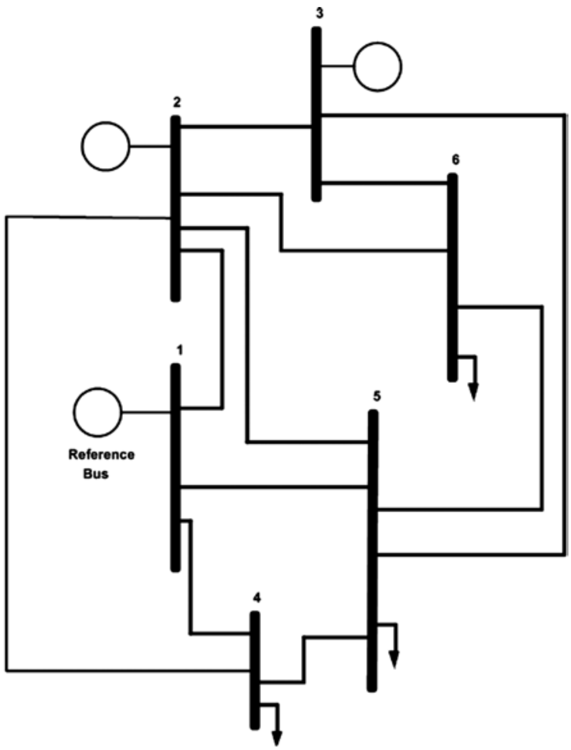
The power system operations department is concerned only with overloads on line 2–4.

a. Given the base transmission system loading and considering outages of only lines 1–2 and 1–4, calculate any overloads on 2–4 due to line outages. Give the outage and the overload amount.

b. A large industrial customer wants to build a factory located at bus 4, what is the largest additional (additional to the base flows shown) amount of power that could be generated at bus 2 and sent to bus 4. Ignore outages and only worry about the loading of line 2–4.

c. What is the largest additional (additional to the base flows shown) amount of power that could be transmitted from bus 2 to bus 4 and still be safe against an outage of line 1–2, causing an overload on line 2–4?

7.5 Consider the following six-bus power system network introduced in Chapter 6. Initially, all loads at 40 MW generators are each at 40 MW output. Bus 1 is the reference bus:



www.mepcafe.com

The initial power flows are calculated with a DC power flow and result in the following base flows. The line flow limits are also shown along with the percent loading.

BASE TRANSMISSION LOADING

Path	From	To	Low	Flow	High	Percent Loading
1	1	2	-40.0	9.24	40.0	23.1
2	1	4	-45.0	17.01	45.0	37.8
3	1	5	-45.0	13.75	45.0	30.5
4	2	3	-45.0	-0.02	45.0	0.0
5	2	4	-45.0	24.79	45.0	55.1
6	2	5	-45.0	10.67	45.0	23.7
7	2	6	-45.0	13.81	45.0	30.7
8	3	5	-45.0	12.32	45.0	27.4
9	3	6	-45.0	27.66	45.0	61.5
10	4	5	-45.0	1.80	45.0	4.0
11	5	6	-45.0	-1.46	45.0	3.2

The generator on bus 2 now agrees to sell an extra 50 MW to the load on bus 6. This requires adding 50 MW to the generation on bus 2 and adding an extra 50 MW to the load on bus 6. With this sale added to the original generation and load in operation, the power system must pass the $(n - 1)$ security test for line outages on lines 6–2, 6–3, and 6–5.

The LODF factors and the PTDF factors for this system are given in Figure 7.4 and Figure 7.5.

a. In the base case, there are no contingency outages that cause overloads. (You do not have to calculate this.)

b. Do any of the line outages on lines 6–2, 6–3, and 6–5, taken one outage at a time, result in overloads when the additional 50 MW is being transferred? If so, how much and what lines are overloaded? You only need to test for overloads on the same set of lines 6–2, 6–3, and 6–5.

c. If there are overloads found, the generator at bus 2 will reduce its output and at the same time the load at bus 6 is going to reduce its load until there are no overloads due to the line outages found in part b. How much should the generation on bus 2 and the load on bus 6 be reduced to eliminate all overloads?

REFERENCES

1. Brandwajn, V., "Efficient Bounding Method for Linear Contingency Analysis," *IEEE Transactions on Power Systems*, Vol. 3, No. 1, February 1988, pp. 38–43.
2. Ejebe, G. C., Paliza, R. F., Tinney, W. F., "An Adaptive Localization Method for Real-Time Security Analysis," *IEEE Transactions onPower Systems*, Vol. 7, No. 2, 1992, pp. 777–783.
3. Alsaffar, M. A., *Voltage Collapse and Power Flow Algorithms*, PhD thesis, University of Minnesota, November 2005.

www.mepcafe.com

8

OPTIMAL POWER FLOW

8.1 INTRODUCTION

Economic dispatch (ED) is the calculation introduced in Chapter 3 that finds the lowest-cost generation dispatch for a set of generators that is constrained within the individual generator limits and results in a total generation that equals the total load plus losses. The ED calculation ignores the specific details of the network that the generators are connected to and lumps all the network effects into the losses and total load demand. As a result, the ED calculation ignores the effect that the dispatch of generation has on the loading of transmission branches or the effect it has on bus voltages. The dispatch of generation does, in fact, have an important effect on transmission flows, and under some circumstances these affects need to be taken into account.

The optimal power flow (OPF), as the name implies, couples the ED calculation with a power flow calculation so that the ED and the power flow are solved simultaneously. The total losses of the power system are simply part of the power flow calculation and are reflected in the loading of the generation on the reference bus—thus, there is no need to specifically calculate the losses since they are inherently a part of the power flow. More importantly, the ED can be constrained to meet transmission system limits such as MW or MVA flow limits on lines or transformers or voltage limits on buses. The result is the generation dispatch representing the minimum $/h total generation cost and that also solves the power flow at that optimum.

Power Generation, Operation, and Control, Third Edition. Allen J. Wood, Bruce F. Wollenberg, and Gerald B. Sheblé.
© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

www.mepcafe.com

8.2 THE ECONOMIC DISPATCH FORMULATION

In the ED calculation, we solve the following problem:

- Objective function = total generation cost in \$/h.
- Each generator is within its minimum and maximum limits.
- Sum of all generator outputs = total load plus losses.

Objective function:

$$\min \sum_{j=1}^{N_{\text{gen}}} F_j(P_{\text{gen}_j})$$

where j is an index over the generators only.

Generator limit inequality constraints:

$$P_{\text{gen}_j}^{\min} \le P_{\text{gen}_j} \le P_{\text{gen}_j}^{\max}, \quad \text{for} \quad j = 1 \ldots N_{\text{gen}}$$

Generation, load, loss balance equality constraint:

$$P_{\text{total load}} + P_{\text{total loss}} - \sum_{j=1}^{N_{\text{gen}}} P_{\text{gen}_j} = 0$$

where N_{gen} is the number of generators.

Using the nomenclature of the appendix to Chapter 3, we can write the ED problem as

$$\min f(\underline{P_{\text{gen}}, u})$$

where

$$\underline{P_{\text{gen}}} = \begin{bmatrix} P_{\text{gen}_1} \\ \vdots \\ P_{\text{gen}_{NG}} \end{bmatrix} \text{ and } f(\underline{P_{\text{gen}}, u}) = \sum_{j=1}^{N_{\text{gen}}} F_j(P_{\text{gen}_j})$$

(the $\underline{u}$ vector appearing in the $f(\underline{P_{\text{gen}}, u})$ now contains the parameters of the generator cost functions)

and is subject to

$$\omega\left(\underline{P_{\text{gen}}}, P_{\text{total load}}, P_{\text{total loss}}\right) = 0 \quad \text{where} \quad \omega = P_{\text{total load}} + P_{\text{total loss}} - \sum_{j=1}^{N_{\text{gen}}} P_{\text{gen}_j}$$

$$g\left(\underline{P_{\text{gen}}}, P_{\text{gen}_j}^{\min}, P_{\text{gen}_j}^{\max}\right) \le 0 \quad \text{where} \quad g = P_{\text{gen}_j}^{\min} \le P_{\text{gen}_j} \le P_{\text{gen}_j}^{\max}$$

www.mepcafe.com

The terms shown earlier that are fixed numbers, such as $P_{\text{total-load}}, P_{\text{total-loss}}, P^{\min}_{\text{gen}_j}$, and $P^{\max}_{\text{gen}_j}$, are often lumped into a vector of “parameters” $\underline{u}$.

Then the ED calculation can be given in a compact notation as

$$\min f(\underline{P_{\text{gen}}}, \underline{u})$$

subject to

$$\omega(\underline{P_{\text{gen}}}, \underline{u}) = 0$$

$$g(\underline{P_{\text{gen}}}, \underline{u}) \le 0$$

8.3 THE OPTIMAL POWER FLOW CALCULATION COMBINING ECONOMIC DISPATCH AND THE POWER FLOW

In the OPF calculation, we solve the following more complex problem:

- Objective function = total generation cost in \$/h (same as ED)
- Generator limit inequality constraints (same as ED)
- For each bus,
  - Generation minus load at the bus = the net power flow at the bus. This is done using the Y matrix power flow equation for the bus (Equation 4.18, 2nd ed.). This introduces the voltage magnitude and phase angle of each bus, except for the reference bus, as unknowns into the problem. This also adds $2(N_{\text{bus}} - 1)$ new equality constraints into the problem representing the $N_{\text{bus}} - 1$ complex Y Matrix equations for all buses except the reference bus.

The notation is changed slightly, where the variable P_{gen_i} is the generation on bus i. From now on in this chapter, we shall always refer to the generators as being at a specific bus and not by generator number.

The OPF equations now look like this:

Objective function:

$$\min \sum_{i=1}^{N_{\text{bus}}} F_i(P_{\text{gen}_i})$$

where i is an index over the buses and $F_i(P_{\text{gen}_i})$ is the cost function of the generator at bus i.

Generator limit inequality constraints:

$$P^{\min}_{\text{gen}_i} \le P_{\text{gen}_i} \le P^{\max}_{\text{gen}_i}, \quad \text{for} \quad i = 1 \ldots N_{\text{bus}}$$

Generator var limit inequality constraints:

$$Q^{\min}_{\text{gen}_i} \le Q_{\text{gen}_i} \le Q^{\max}_{\text{gen}_i}, \quad \text{for} \quad i = 1 \ldots N_{\text{bus}}$$

www.mepcafe.com

Power flow Y matrix equation for each bus i:

$$P_{\text{net}_i} + jQ_{\text{net}_i} = V_i \left(\sum_{k=1}^{N_{\text{bus}}} Y_{ik} V_k \right)^*$$

where P_{net_i} and Q_{net_i} are the net real and reactive power at bus i, V_i is the complex voltage at bus i, and Y_{ik} is the Y matrix i,k term. For the OPF the net real and reactive power are replaced with the generation minus the load so that we have

$$(P_{\text{gen}_i} - P_{\text{load}_i}) + j(Q_{\text{gen}_i} - Q_{\text{load}_i}) = V_i \left(\sum_{k=1}^{N_{\text{bus}}} Y_{ik} V_k \right)^*$$

which is broken down into separate real and reactive power equality constraint equations for each bus i as follows:

$$(P_{\text{gen}_i} - P_{\text{load}_i}) - \text{Real}\left\{ V_i \left(\sum_{k=1}^{N_{\text{bus}}} Y_{ik} V_{ik} \right)^* \right\} = 0$$

$$(Q_{\text{gen}_i} - Q_{\text{load}_i}) - \text{Imag}\left\{ V_i \left(\sum_{k=1}^{N_{\text{bus}}} Y_{ik} V_k \right)^* \right\} = 0$$

As indicated earlier, the fact that we are solving the power flow itself results in a set of generator outputs that equals the load plus losses as required. We do not have to explicitly calculate the losses or the generator incremental losses as this is all incorporated in the AC power flow equations. As in a power flow, we must force the reference-bus voltage magnitude and phase angle to be fixed quantities.

We can again use the compact notation:

$$\min f(\underline{P_{\text{gen}}}, \underline{u})$$

subject to

$$\omega(\underline{P_{\text{gen}}}, \underline{u}) = 0$$

$$g(\underline{P_{\text{gen}}}, \underline{u}) \le 0$$

where the vector of parameters $\underline{u}$ now contains the generator cost function parameters, the generator real and reactive power limits, and all the fixed parameters for the power system's transmission system. This includes the Y matrix, the bus fixed voltages at generators that are regulating, the reference-bus fixed voltage magnitude and phase angle.

The equation $\omega(\underline{P_{\text{gen}}}, \underline{u}) = 0$ is now actually $2N_{\text{bus}}$ equality constraints representing the power flow Y matrix equations, and $g(\underline{P_{\text{gen}}}, \underline{u}) \le 0$ is a set of inequality constraints representing every generator real power limits and reactive power limits.

www.mepcafe.com

The great advantage to this formulation is the fact that in addition to a very accurate calculation of losses inside the OPF, other transmission system constraints can be added to the set inequality constraints. For example,

- For each transmission line or transformer,
  - MW flow or MVA flow is less than or equal to the branch's MW flow or MVA flow limit:

$$\text{MW flow}_{ij} = \text{Real}\left\{ V_i \left[(V_i - V_j) y_{ij} + V_i^2 y_{\text{charging}_{ij}} \right]^* \right\} \le \text{MW flow}_{ij}^{\max}$$

or

$$\text{MVA flow}_{ij} = \text{abs}\left\{ V_i \left[(V_i - V_j) y_{ij} + V_i^2 y_{\text{charging}_{ij}} \right]^* \right\} \le \text{MVA flow}_{ij}^{\max}$$

The result is a dispatch of generation that does not overload the transmission line. In addition, each bus voltage magnitude in the power system can be forced to be within limits:

- For each bus in the network,
  - The voltage magnitude at the bus is within a low and high voltage limit:

$$V_i^{\min} \le V_i \le V_i^{\max}, \quad \text{for} \quad i = 1 \ldots N_{\text{bus}}$$

Finally, we can again use the compact notation:

$$\min f(\underline{P_{\text{gen}}}, \underline{u})$$

subject to

$$\omega(\underline{P_{\text{gen}}}, \underline{u}) = 0$$

$$g(\underline{P_{\text{gen}}}, \underline{u}) \le 0$$

where the equation $g(\underline{P_{\text{gen}}}, \underline{u}) \le 0$ now contains the generator real and reactive power limits plus the line flow limit equations and the bus voltage limits.

Before we go further in showing how the OPF with the AC power flow equations is solved, we will first show how to formulate the OPF using a DC power flow and how it is solved.

8.4 OPTIMAL POWER FLOW USING THE DC POWER FLOW

We can also formulate the OPF with the transmission system constraints represented using a DC power flow. This is used in many types of studies, and as with the DC power flow itself, it gives a good approximation to the AC power flow and is much faster and easier to set up and solve. We shall refer to this calculation as the DCOPF algorithm.

www.mepcafe.com

Once again we start with the objective function:

$$\min \sum_{i=1}^{N_{\text{bus}}} F_i(P_{\text{gen}_i})$$

where i is an index over the buses and $F_i(P_{\text{gen}_i})$ is the cost function of the generator at bus i.

Generator limit inequality constraints:

$$P_{\text{gen}_i}^{\min} \leq P_{\text{gen}_i} \leq P_{\text{gen}_i}^{\max}, \quad \text{for} \quad i = 1 \ldots N_{\text{bus}}$$

This formulation, however, uses the DC power flow formulation as follows:

$$[B_x]\underline{\theta} = \underline{P_{\text{gen}}} - \underline{P_{\text{load}}}$$

where $\underline{\theta} = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_{N_{\text{bus}}} \end{bmatrix}$ is in radians and $[B_x]$ is in per unit

We wish to keep the P_{gen_i} and P_{load_i} values in MW, and we will assume that the θ_i values are in radians. However, the B_x matrix values are in per unit so that $[B_x]\underline{\theta}$ gives the bus power in per unit. To convert the bus power from per unit to MW with an MVA system base of 100 MVA, we shall need to multiply the $[B_x]\underline{\theta}$ equation by 100, or more simply the values in the $[B_x]$ matrix will all be multiplied by 100 so that the resulting bus powers are in MW. Then our DC power flow becomes

$$100[B_x]\underline{\theta} = \underline{P_{\text{gen}}} - \underline{P_{\text{load}}}$$

Our compact form is still applicable as long as we define the parameters of the problem to use the DC power flow.

We will now proceed to formulate the DCOPF problem using the Lagrange multiplier approach of Chapter 3. We can write the basic formulation of the problem as

$$\mathcal{L} = \sum_{i=1}^{N_{\text{bus}}} F_i\left(P_{\text{gen}_i}\right) + \underline{\lambda}^{\text{T}}\left(100[B_x]\underline{\theta} - (\underline{P_{\text{gen}}} - \underline{P_{\text{load}}})\right) + \lambda_{N_{\text{bus}}+1}\left(\theta_{\text{refbus}} - 0\right)$$

This formulation includes a term at the right-hand end that is there as an equality constraint to force the reference-bus phase angle to 0 (without this constraint the OPF will not solve properly). The generator inequality constraints can then be added so that we obtain $\mathcal{L}$:

$$\mathcal{L} = \sum_{i=1}^{N_{\text{bus}}} F_i(P_{\text{gen}_i}) + \underline{\lambda}^{\text{T}}\left(100[B_x]\underline{\theta} - (\underline{P_{\text{gen}}} - \underline{P_{\text{load}}})\right)$$
$$+\lambda_{N_{\text{bus}}+1}(\theta_{\text{refbus}} - 0) + \underline{\mu}^{\text{T}}[g(\underline{P_{\text{gen}}}, P_{\text{gen}_j}^{\min}, P_{\text{gen}_j}^{\max})]$$

www.mepcafe.com

8.5 EXAMPLE 8A: SOLUTION OF THE DC POWER FLOW OPF

Using the three-generator ED problem of Example 3A, we have

$$F_1(P_1) = 561 + 7.92P_1 + 0.001562P_1^2$$
$$F_2(P_2) = 310 + 7.85P_2 + 0.00194P_2^2$$
$$F_3(P_3) = 78 + 7.97P_3 + 0.00482P_3^2$$

with generator min and max power limits of

Generator	P^{min}	P^{max}
1	150	600
2	100	400
3	50	200

Example 3A ignores losses and requires a delivery of 850 MW resulting in a dispatch of

$$P_1 = 393.2 \text{ MW}$$
$$P_2 = 334.6 \text{ MW}$$
$$P_3 = 122.2 \text{ MW}$$

The Lagrange multiplier is

$$\lambda = 9.148 \text{ \$/MWh}$$

We shall now pair the generators of Example 3A with a three-bus network (Figure 8.1) whose data is as follows:

Line	Line Reactance (Per Unit)
1–2	x12 = 0.1
1–3	x13 = 0.125
2–3	x23 = 0.2

The B_x matrix is formed using the line reactances as follows:

$$\begin{pmatrix} B_{x11} & B_{x12} & B_{x13} \\ B_{x21} & B_{x22} & B_{x23} \\ B_{x31} & B_{x32} & B_{x33} \end{pmatrix} = \begin{pmatrix} (1/x_{12} + 1/x_{13}) & -1/x_{12} & -1/x_{13} \\ -1/x_{12} & (1/x_{12} + 1/x_{23}) & -1/x_{23} \\ -1/x_{13} & -1/x_{23} & (1/x_{13} + 1/x_{23}) \end{pmatrix}$$
$$= \begin{pmatrix} 18 & -10 & -8 \\ -10 & 15 & -5 \\ -8 & -5 & 13 \end{pmatrix}$$

www.mepcafe.com

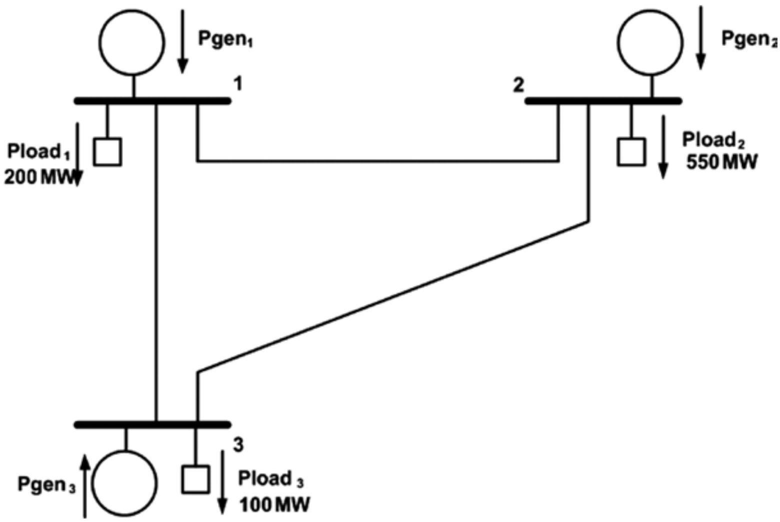


FIGURE 8.1 Three-bus test system.

To keep the power injections in MW, we multiply $[B_x]$ by 100, and the result is

$$\begin{pmatrix} 1800 & -1000 & -800 \\ -1000 & 1500 & -500 \\ -800 & -500 & 1300 \end{pmatrix}$$

We shall do the calculations using the Lagrangian while ignoring the generation limits. This allows us to have only equality constraints, and the solutions of the Lagrangian are simply a solution of a set of linear equations.
The expression for the Lagrangian with the power flow equations written out becomes

$$\begin{aligned}
\mathcal{L} = &\sum_{i=1}^{N_{\text{bus}}} (a_i + b_i P_{gen_i} + c_i P_{gen_i}{}^2) \\
&+\lambda_1(100B_{x11}\theta_1 + 100B_{x12}\theta_2 + 100B_{x13}\theta_3 - P_{\text{gen}_1} + P_{\text{load}_1}) \\
&+\lambda_2(100B_{x21}\theta_1 + 100B_{x22}\theta_2 + 100B_{x23}\theta_3 - P_{\text{gen}_2} + P_{\text{load}_2}) \\
&+\lambda_3(100B_{x31}\theta_1 + 100B_{x32}\theta_2 + 100B_{x33}\theta_3 - P_{\text{gen}_3} + P_{\text{load}_3}) \\
&+\lambda_4(\theta_1 - 0)
\end{aligned}$$

where we have now assumed the reference bus to be bus 1.

www.mepcafe.com

We now must solve the Lagrangian by taking the derivatives of L with respect to each independent variable in the problem. The independent variables are

$$P_{\text{gen}_1}, P_{\text{gen}_2}, P_{\text{gen}_3}, \theta_1, \theta_2, \theta_3, \lambda_1, \lambda_2, \lambda_3 \text{ and } \lambda_4$$

Then we need derivatives of $\mathcal{L}$ with respect to each of these variables as shown next:

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}P_{\text{gen}_1}} = b_1 + 2c_1 P_{\text{gen}_1} - \lambda_1 = 0$$

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}P_{\text{gen}_2}} = b_2 + 2c_2 P_{\text{gen}_2} - \lambda_2 = 0$$

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}P_{\text{gen}_3}} = b_3 + 2c_3 P_{\text{gen}_3} - \lambda_3 = 0$$

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\theta_1} = 100B_{x11}\lambda_1 + 100B_{x21}\lambda_2 + 100B_{x31}\lambda_3 + \lambda_4 = 0$$

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\theta_2} = 100B_{x12}\lambda_1 + 100B_{x22}\lambda_2 + 100B_{x32}\lambda_3 = 0$$

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\theta_3} = 100B_{x13}\lambda_1 + 100B_{x23}\lambda_2 + 100B_{x33}\lambda_3 = 0$$

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\lambda_1} = 100B_{x11}\theta_1 + 100B_{x12}\theta_2 + 100B_{x13}\theta_3 - P_{\text{gen}_1} + P_{\text{load}_1} = 0$$

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\lambda_2} = 100B_{x21}\theta_1 + 100B_{x22}\theta_2 + 100B_{x23}\theta_3 - P_{\text{gen}_2} + P_{\text{load}_2} = 0$$

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\lambda_3} = 100B_{x31}\theta_1 + 100B_{x32}\theta_2 + 100B_{x33}\theta_3 - P_{\text{gen}_3} + P_{\text{load}_3} = 0$$

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\lambda_4} = \theta_1 - 0 = 0$$

The resulting matrix equation is shown in the following text. The rows of the matrix are the derivatives and the columns are associated with the independent variables.

www.mepcafe.com

$$
\begin{pmatrix}
2c_1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 2c_2 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & 2c_3 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 100\,B_{x11} & 100\,B_{x21} & 100\,B_{x31} & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 100\,B_{x12} & 100\,B_{x22} & 100\,B_{x32} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 100\,B_{x13} & 100\,B_{x23} & 100\,B_{x33} & 0 \\
-1 & 0 & 0 & 100\,B_{x11} & 100\,B_{x12} & 100\,B_{x13} & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 100\,B_{x21} & 100\,B_{x22} & 100\,B_{x23} & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 100\,B_{x31} & 100\,B_{x32} & 100\,B_{x33} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
Pgen_1 \\ Pgen_2 \\ Pgen_3 \\ \theta_1 \\ \theta_2 \\ \theta_3 \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4
\end{pmatrix}
=
\begin{pmatrix}
-b_1 \\ -b_2 \\ -b_3 \\ 0 \\ 0 \\ 0 \\ -Pload_1 \\ -Pload_2 \\ -Pload_3 \\ 0
\end{pmatrix}
$$

The Numerical values for this matrix are (where the variable associated with each variable is shown above the matrix):

Pg1	Pg2	Pg3	theta1	theta2	theta3	lambda1	lambda2	lambda3	lambda4
0.003124	0	0	0	0	0	−1	0	0	0
0	0.003880	0	0	0	0	0	−1	0	0
0	0	0.009640	0	0	0	0	0	−1	0
0	0	0	0	0	0	1800	−1000	−800	1
0	0	0	0	0	0	−1000	1500	−500	0
0	0	0	0	0	0	−800	−500	1300	0
−1	0	0	1800	−1000	−800	0	0	0	0
0	−1	0	−1000	1500	−500	0	0	0	0
0	0	−1	−800	−500	1300	0	0	0	0
0	0	0	1	0	0	0	0	0	0

www.mepcafe.com

The right-hand side is

```
-7.92
-7.85
-7.97
 0
 0
 0
-200
-550
-100
 0
```

The solution to this set of equations is

$$P_{\text{gen}_1} = 393.1698$$

$$P_{\text{gen}_2} = 334.6083$$

$$P_{\text{gen}_3} = 122.2264$$

which matches the ED solution found in Chapter 3. The lambda values are

$$\lambda_1 = \lambda_2 = \lambda_3 = 9.1483$$

which is the same lambda as found in Chapter 3. Finally, the bus phase angles and power flows are

$$\theta_1 = 0$$

$$\theta_2 = -0.1582 \text{ rad}$$

$$\theta_3 = -0.0437 \text{ rad}$$

And the line power flows are

$$P_{\text{flow}_{12}} = 158.1776 \text{ MW}$$

$$P_{\text{flow}_{13}} = 34.9922 \text{ MW}$$

$$P_{\text{flow}_{23}} = -57.2186 \text{ MW}$$

The network flows look as follows on the one-line diagram (Figure 8.2):

We now can see that the OPF model has given us both the ED and the solution of the network equations that reflect the generators at the ED.

www.mepcafe.com

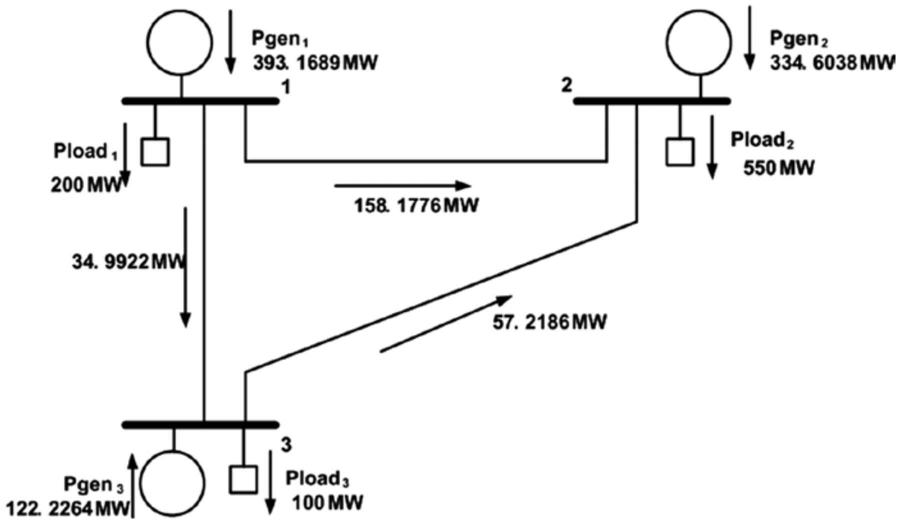


FIGURE 8.2 Resulting dispatch and power flows with no flow constraints.

8.6 EXAMPLE 8B: DCOPF WITH TRANSMISSION LINE LIMIT IMPOSED

We now return to our three-bus system and add a constraint to the system representing the flow limit on line 1–2. The limit will be set to 150 MW. Since the flow on line 1–2 was at 158.1776 MW when we dispatched it earlier, imposing the 150 MW limit will result in the line being held at the limit. Thus, we can simply add another equality constraint to the Lagrangian. Note that normally we shall use an inequality constraint for line flow limits, but in this introduction we will use the equality constraint. See Section 8.7 for more details on how to formulate the DCOPF. The added constraint is

$$\frac{100}{x_{12}}(\theta_1 - \theta_2) = 150$$

This constraint is added to the Lagrangian with its own Lagrange multiplier, and the result is

$$\begin{aligned}
\mathcal{L} = & \sum_{i=1}^{N_{\text{bus}}} (a_i + b_i P_{\text{gen}_i} + c_i P_{\text{gen}_i}{}^2) \\
& + \lambda_1 (100 B_{x11}\theta_1 + 100 B_{x12}\theta_2 + 100 B_{x13}\theta_3 - P_{\text{gen}_1} + P_{\text{load}_1}) \\
& + \lambda_2 (100 B_{x21}\theta_1 + 100 B_{x22}\theta_2 + 100 B_{x23}\theta_3 - P_{\text{gen}_2} + P_{\text{load}_2}) \\
& + \lambda_3 (100 B_{x31}\theta_1 + 100 B_{x32}\theta_2 + 100 B_{x33}\theta_3 - P_{\text{gen}_3} + P_{\text{load}_3}) \\
& + \lambda_4 (\theta_1 - 0) \\
& + \lambda_5 \left(\frac{100}{x_{12}}\theta_1 - \frac{100}{x_{12}}\theta_2 - 150 \right)
\end{aligned}$$

www.mepcafe.com

The derivatives of the Lagrangian also reflect the new constraint and new Lagrange multiplier:

$$\frac{d\mathcal{L}}{dP_{\text{gen}_1}} = b_1 + 2c_1 P_{\text{gen}_1} - \lambda_1 = 0$$

$$\frac{d\mathcal{L}}{dP_{\text{gen}_2}} = b_2 + 2c_2 P_{\text{gen}_2} - \lambda_2 = 0$$

$$\frac{d\mathcal{L}}{dP_{\text{gen}_3}} = b_3 + 2c_3 P_{\text{gen}_3} - \lambda_3 = 0$$

$$\frac{d\mathcal{L}}{d\theta_1} = 100B_{x11}\lambda_1 + 100B_{x21}\lambda_2 + 100B_{x31}\lambda_3 + \lambda_4 + \lambda_5 \frac{100}{x_{12}} = 0$$

$$\frac{d\mathcal{L}}{d\theta_2} = 100B_{x12}\lambda_1 + 100B_{x22}\lambda_2 + 100B_{x32}\lambda_3 - \lambda_5 \frac{100}{x_{12}} = 0$$

$$\frac{d\mathcal{L}}{d\theta_3} = 100B_{x13}\lambda_1 + 100B_{x23}\lambda_2 + 100B_{x33}\lambda_3 = 0$$

$$\frac{d\mathcal{L}}{d\lambda_1} = 100B_{x11}\theta_1 + 100B_{x12}\theta_2 + 100B_{x13}\theta_3 - P_{\text{gen}_1} + P_{\text{load}_1} = 0$$

$$\frac{d\mathcal{L}}{d\lambda_2} = 100B_{x21}\theta_1 + 100B_{x22}\theta_2 + 100B_{x23}\theta_3 - P_{\text{gen}_2} + P_{\text{load}_2} = 0$$

$$\frac{d\mathcal{L}}{d\lambda_3} = 100B_{x31}\theta_1 + 100B_{x32}\theta_2 + 100B_{x33}\theta_3 - P_{\text{gen}_3} + P_{\text{load}_3} = 0$$

$$\frac{d\mathcal{L}}{d\lambda_4} = \theta_1 - 0 = 0$$

$$\frac{d\mathcal{L}}{d\lambda_5} = \frac{100}{x_{12}}\theta_1 - \frac{100}{x_{12}}\theta_2 - 150 = 0$$

Note that new terms appear in the derivatives with respect to θ_1 and θ_2 as well as the new derivative with respect to λ_5.

The set of linear equations is shown as follows.

www.mepcafe.com

$$
\begin{pmatrix}
2c_1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 2c_2 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 2c_3 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 100B_{x11} & 100B_{x21} & 100B_{x31} & 1 & 100/\mathrm{x}_{12} \\
0 & 0 & 0 & 0 & 0 & 0 & 100B_{x12} & 100B_{x22} & 100B_{x32} & 0 & -100/\mathrm{x}_{12} \\
0 & 0 & 0 & 0 & 0 & 0 & 100B_{x13} & 100B_{x23} & 100B_{x33} & 0 & 0 \\
-1 & 0 & 0 & 100B_{x11} & 100B_{x12} & 100B_{x13} & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 100B_{x21} & 100B_{x22} & 100B_{x23} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 100B_{x31} & 100B_{x32} & 100B_{x33} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 100/\mathrm{x}_{12} & -100/\mathrm{x}_{12} & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
Pgen_1 \\ Pgen_2 \\ Pgen_3 \\ \theta_1 \\ \theta_2 \\ \theta_3 \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5
\end{pmatrix}
=
\begin{pmatrix}
-b_1 \\ -b_2 \\ -b_3 \\ 0 \\ 0 \\ 0 \\ -Pload_1 \\ -Pload_2 \\ -Pload_3 \\ 0 \\ 150
\end{pmatrix}
$$

The Numerical values for this matrix are:

Pg1	Pg2	Pg3	theta1	theta2	theta3	lambda1	lambda2	lambda3	lambda4	Lambda5
0.003124	0	0	0	0	0	−1	0	0	0	0
0	0.003880	0	0	0	0	0	−1	0	0	0
0	0	0.009640	0	0	0	0	0	−1	0	0
0	0	0	0	0	0	1800	−1000	−800	1	1000
0	0	0	0	0	0	−1000	1500	−500	0	-1000
0	0	0	0	0	0	−800	−500	1300	0	0
−1	0	0	1800	−1000	−800	0	0	0	0	0
0	−1	0	−1000	1500	−500	0	0	0	0	0
0	0	−1	−800	−500	1300	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0
0	0	0	1000	−1000	0	0	0	0	0	0

www.mepcafe.com

The right-hand side vector is

```
-7.92
-7.85
-7.97
0
0
0
-200
-550
-100
0
150
```

The solution to this set of equations is

$$P_{\text{gen}_1} = 382.7244$$

$$P_{\text{gen}_2} = 345.4528$$

$$P_{\text{gen}_3} = 121.8228$$

which matches the ED solution found in Chapter 3. The lambda values are

$$\lambda_1 = 9.1156$$

$$\lambda_2 = 9.1904$$

$$\lambda_3 = 9.1444$$

Because the line flow on line 1–2 is at its limit, the lambda values are different as explained in Chapter 3, Section 3.10. The Locational Marginal Price (LMP) values for the buses are equal to these lambda values.

Finally, the bus phase angles and power flows are

$$\theta_1 = 0$$

$$\theta_2 = -0.1500 \text{ rad}$$

$$\theta_3 = -0.0409 \text{ rad}$$

And the line power flows are

$$P_{\text{flow}_{12}} = 150.0 \text{ MW}$$

$$P_{\text{flow}_{13}} = 32.7244 \text{ MW}$$

$$P_{\text{flow}_{23}} = -54.5472 \text{ MW}$$

www.mepcafe.com

8.7 FORMAL SOLUTION OF THE DCOPF

The DCOPF solution shown in the previous section ignored generator inequality constraints, and it imposed a line flow limit as an equality constraint. When we solved the ED problem in Chapter 3, we were able to formulate it with inequality constraints for the generator limits (see Appendix 3A) and then apply the Karush–Kuhn–Tucker (KKT) conditions to test the optimality. To obtain the solution, we solve with no generator limits applied, then test the result with the KKT conditions, apply the appropriate generator limits, and then resolve. This procedure, which we called *lambda search*, leads to the correct solution. Thus, the solution for the ED with generator limits is made as a series of solutions with generator limits ignored or generator output forced to the limits for those generators violating the KKT conditions.

The KKT conditions must be met when solving the DCOPF; however, when we have generator limits and line flow limits, the KKT conditions do not give a clear indication of which variable(s) should be placed on limit. Thus, the KKT conditions are a necessary condition to test the optimum, but they are not sufficient to find that optimum and we need a more powerful method. The two most popular methods of solving the DCOPF are linear programming (LP) and quadratic programming (QP).

The student is advised to seek references on the theory of LP and QP.

8.8 ADDING LINE FLOW CONSTRAINTS TO THE LINEAR PROGRAMMING SOLUTION

In Chapter 3, LP was introduced as a means of solving the ED problem. LP can also be used to solve OPFs if one uses the linear or DC power flow to model the transmission system.

If we add a flow constraint to the ED problem, we can easily solve it using LP. Picking up from the LP ED solution in Chapter 3, Section 3.3, the problem now becomes as follows:

$$\text{Minimize} \sum_{i=1}^{N_{\text{bus}}} (F_i \left(P_{\text{gen } i}^{\min}\right) + s_{i1} P_{\text{gen}_{i1}} + s_{i2} P_{\text{gen}_{i2}} + s_{i3} P_{\text{gen}_{i3}})$$

$$0 \le P_{\text{gen}_{ik}} \le P_{\text{gen}_{ik}}^{\max} \quad \text{for} \quad k = 1,\ 2,\ 3$$

and finally

$$P_i = P_i^{\min} + P_{\text{gen}_{i1}} + P_{\text{gen}_{i2}} + P_{\text{gen}_{i3}}$$

subject to

$$100\,[B_x]\,\underline{\theta} = \underline{P_{\text{gen}}} - \underline{P_{\text{load}}}$$

$$P_{\text{gen}_i}^{\min} \le P_{\text{gen}_i} \le P_{\text{gen}_i}^{\max}, \quad \text{for} \quad i = 1 \ldots N_{\text{bus}}$$

www.mepcafe.com

and line flow constraints of the form

$$\frac{100}{x_{ij}}\left(\theta_i - \theta_j\right) \leq P_{ij}^{\max}$$

For our three-bus system with a line flow limit on the line from bus 1 to bus 2, we would add a constraint like this:

$$\frac{100}{x_{12}}\left(\theta_1 - \theta_2\right) \leq 150$$

This constraint guarantees that the flow from bus 1 to bus 2 is less than or equal to 150 MW.

Actual transmission lines are not usually limited to flow in one direction but are usually limited to the same limit on flow in either direction. This can be accomplished by adding another constraint to the one shown earlier that limits the flow from bus 2 to bus 1 to the same 150 MW, as is given in the following where we simply place a minus sign in front of the constraint equation:

$$-\frac{100}{x_{12}}(\theta_1 - \theta_2) \leq 150$$

A better way to limit flow on a transmission line to the same MW limit in either direction is to use a slack variable added to the constraint with limits on the slack variable as given in the following:

$$\frac{100}{x_{12}}(\theta_1 - \theta_2) + S_{12} = 150$$

where $0 \leq S_{12} \leq 2 * P_{\text{flow } 12}^{\max}$ or for our example $0 \leq S_{12} \leq 300$.

When the flow on the 1–2 line approaches +150, the slack variable S_{12} goes to its lower limits of 0; when the flow on the 1–2 line approaches −150 (or 150 MW in the 2–1 direction), the slack variable approaches its upper limit of twice the limit. Therefore, by adding the slack variable with upper and lower limits, we can add only one constraint to the LP and still achieve limiting its flow in both directions.

The aforementioned also shows the advantage of having an LP algorithm with an "upper bounding" feature where the user can specify an upper and lower bound for each variable in the LP.

Note that the LP given previously has the entire power flow embedded in it with an LP variable for each bus's phase angle added to the problem.

The solution to the LP with 10 segments and including the line flow limit is

$$P_{\text{gen}_1} = 380.77$$

$$P_{\text{gen}_2} = 344.23$$

$$P_{\text{gen}_3} = 125.00$$

www.mepcafe.com

and

$$P_{\text{flow}_{12}} = 150.0 \text{ MW}$$

The LP can be built with all line flow constraints built into it that makes the LP simplex matrix quite large; an alternative is to solve the LP without any line flow limits and then solve for the optimum and calculate all flows that result and only put line limits into the LP for those over limit. This process has been called "iterative constraint search" (ICS) and works quite well and gets to the same solution as the LP with all line flow limits built in.

8.8.1 Solving the DCOPF Using Quadratic Programming

Quadratic Programming allows a full quadratic cost function for each variable together with linear equality and inequality constraints as well as upper bound limits on the variables. Thus, the cost function for our problem becomes

$$\sum_{i=1}^{N_{\text{bus}}} F_i\left(P_{\text{gen}_i}\right) = \sum_{i=1}^{N_{\text{bus}}} \left(a_i + b_i P_{\text{gen}_i} + c_i P_{\text{gen}_i}^2\right)$$

or in matrix form:

$$\sum_{i=1}^{N_{\text{bus}}} F_i\left(P_{\text{gen}_i}\right) = \begin{bmatrix} a_1 & a_2 & \cdots \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} b_1 & b_2 & \cdots \end{bmatrix} \begin{bmatrix} P_{\text{gen}_1} \\ P_{\text{gen}_2} \\ \vdots \end{bmatrix}$$
$$+ \begin{bmatrix} P_{\text{gen}_1} & P_{\text{gen}_2} & \cdots \end{bmatrix} \begin{pmatrix} c_1 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & 0 & \ddots \end{pmatrix} \begin{bmatrix} P_{\text{gen}_1} \\ P_{\text{gen}_2} \\ \vdots \end{bmatrix}$$

The remainder of the constraints are identical to the LP; in this case we will write the equation as

$$\underline{P}_{\text{gen}} - \underline{P}_{\text{load}} - 100\left[B_x\right]\underline{\theta} = \underline{0}$$

$$P_{\text{gen}\ i}^{\min} \le P_{\text{gen}_i} \le P_{\text{gen}\ i}^{\max}, \quad \text{for} \quad i = 1 \ldots N_{\text{bus}}$$

and

$$\frac{100}{x_{12}}\left(\theta_1 - \theta_2\right) + S_{12} = 150 \text{ where } 0 \le S_{12} \le 2P_{\text{flow }12}^{\max} \text{ for a tie line.}$$

The results are identical to those found Section 8.3 and shown in Figure 8.3.

www.mepcafe.com

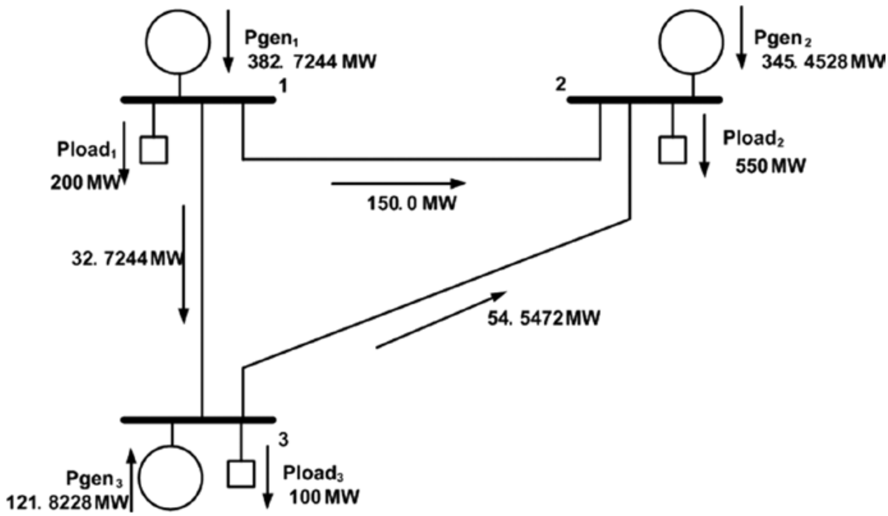


FIGURE 8.3 Resulting dispatch and power flows with flow constraint on line 1–2.

8.9 SOLUTION OF THE ACOPF

To solve the ACOPF, we go back to the formulation shown in the beginning of this chapter where the bus power equations require a real and an imaginary part to take care of the real power and reactive power.

Objective Function: The objective function is identical for the ACOPF.

$$\min \sum_{i=1}^{N_{\text{bus}}} F_i\left(P_{\text{gen}_i}\right)$$

Network Model: We must replace the DC power flow equations with a complete set of AC power flow equations. For each bus i except the reference bus, we have these two equations:

$$\left(P_{\text{gen}_i} - P_{\text{load}_i}\right) - \text{Real}\left\{V_i\left(\sum_{k=1}^{N_{\text{bus}}} Y_{ik} V_k\right)^*\right\} = 0$$

$$\left(Q_{\text{gen}_i} - Q_{\text{load}_i}\right) - \text{Imag}\left\{V_i\left(\sum_{k=1}^{N_{\text{bus}}} Y_{ik} V_k\right)^*\right\} = 0$$

Inequality Constraints:

- Generator limit inequality constraints:

$$P_{\text{gen}_i}^{\min} \le P_{\text{gen}_i} \le P_{\text{gen}_i}^{\max}, \quad \text{for} \quad i = 1 \ldots N_{\text{bus}}$$

www.mepcafe.com

- Generator reactive power limit inequality constraints:

$$Q_{\mathrm{gen}_i}^{\min} \le Q_{\mathrm{gen}_i} \le Q_{\mathrm{gen}_i}^{\max}, \quad \text{for} \quad i = 1 \ldots N_{\mathrm{bus}}$$

- Line MW flow or MVA flow inequality constraints are less than or equal to the branch's MW flow or MVA flow limit:

$$\text{MW flow}_{ij} = \text{Real}\left\{V_i\left[\left(V_i - V_j\right)y_{ij} + V_i^2 y_{\mathrm{charging}_{ij}}\right]^*\right\} \le \text{MW flow}_{ij}^{\max}$$

Or

$$\text{MVA flow}_{ij} = \text{abs}\left\{V_i\left[\left(V_i - V_j\right)y_{ij} + V_i^2 y_{\mathrm{charging}_{ij}}\right]^*\right\} \le \text{MVA flow}_{ij}^{\max}$$

- the voltage magnitude inequality constraints

$$V_i^{\min} \le V_i \le V_i^{\max}, \quad \text{for} \quad i = 1 \ldots N_{\mathrm{B}}$$

8.10 ALGORITHMS FOR SOLUTION OF THE ACOPF

The ACOPF has approximately twice as many variables as the DC power flow OPF (DCOPF); in addition, the network equations are not linear and this makes the solution of the ACOPF much more difficult.

Comparison of Network Equations

If we compare the DC power flow network equations to the AC power flow network equations using Newton's method, we find the following.

DC Power Flow

$$\left[B_x\right]\underline{\theta} = \underline{P_{\mathrm{gen}}} - \underline{P_{\mathrm{load}}}$$

The solution is obtained by solving a set of linear equations, a straightforward calculation.

AC Power Flow

Using Newton's method, we solve the following equation:

$$\begin{bmatrix} \dfrac{dP_1}{d\theta_1} & \dfrac{dP_1}{dV_1} & \dfrac{dP_1}{d\theta_2} & \cdots \\ \dfrac{dQ_1}{d\theta_1} & \dfrac{dQ_1}{dV_1} & \dfrac{dQ_1}{d\theta_2} & \cdots \\ \dfrac{dP_2}{d\theta_1} & \dfrac{dP_2}{dV_1} & \dfrac{dP_2}{d\theta_2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \Delta\theta_1 \\ \Delta V_1 \\ \Delta\theta_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} P_{\mathrm{scheduled}_1} - P_1\left(\underline{V},\underline{\theta}\right) \\ Q_{\mathrm{scheduled}_1} - Q_1\left(\underline{V},\underline{\theta}\right) \\ P_{\mathrm{scheduled}_2} - P_2\left(\underline{V},\underline{\theta}\right) \\ \vdots \end{bmatrix}$$

www.mepcafe.com

Here the Jacobian matrix is twice the dimension of the $[B_x]$ matrix, and the Jacobian matrix terms depend on the value of $\underline{V}$ and $\underline{\theta}$, so we must solve the equation once, then update the values of $\underline{V}$ and $\underline{\theta}$, and then start over again to recalculate the Jacobian matrix. This iterative method will converge, but the overall process is much more time consuming than the DC power flow solution.

In addition, the constraints are linear for the DCOPF and nonlinear for the ACOPF.

Algorithms

Without doubt, the OPF problem has been a challenge to power system analysis research for a long time. The greatest difficulty in obtaining a successful algorithm is to use one that handles constraints easily. In fact, the constraint handling is seen by many as more important than how the nonlinear equations are handled. For example, we can use a pure Newton's method for solving the OPF.

$$\text{Minimize } f(\underline{x})$$

$$\text{subject to } g(\underline{x}) = 0$$

$$\text{and } h(\underline{x}) \le 0$$

where

$f(\underline{x})$ is the objective function $\sum_{i=1}^{N_{\text{bus}}} F_i(P_{\text{gen}_i})$

$g(\underline{x}) = 0$ is the set of network equality constraint equations

$h(\underline{x}) \le 0$ is the set of all inequality constraints

The first step is to write a Lagrange equation for this problem. Here we have added a slack variable to each inequality constraints. The variables λ and μ are vectors of Lagrange variables for the equality and inequality constraints, respectively.

$$L = f(x) + \lambda^{\mathrm{T}} g(x) + \gamma^{\mathrm{T}}(h(x) + s)$$

The classical solution to obtain the optimum requires that we set the gradient of the Lagrange equation with respect to x, λ, μ and s:

$$\Delta_x L = \Delta_x f(x) + \Delta_x g(x)\lambda + \Delta_x h(x)\gamma = 0$$

$$\Delta_\lambda L = g(x) = 0$$

$$\Delta_\mu L = h(x) + s = 0$$

$$\Delta_s L = \gamma = 0$$

Driving a vector of derivatives—the gradient—to 0 is easily done with Newton's method. The problem with using Newton's method on the Lagrange equation gradient, however, is the inequality constraints—when are they active (i.e., $s = 0$) and when are they inactive (i.e., not on limit so that $s \ge 0$). Two primary algorithms have been used for ACOPF; both handle constraints well.

www.mepcafe.com

Iterative LP

The first is simply to take the nonlinear equations and linearize them so that you have a linear objective function and linear constraint equations and then solve them with LP. Because the objective and equations are nonlinear, we would need to restrict the amount of change in the basic variables of the problem. Thus, the LP would solve; we would relinearize all equations, move the window of limits on the variables, and resolve the LP. This process works well because the LP is so good at handling inequality constraints. The entire optimization will now be written in terms of

$$\Delta P_{\text{gen}}, \Delta Q_{\text{gen}}, \Delta V, \text{and } \Delta\theta$$

and we will assume that before the LP is set up and run, a power flow solution is obtained with solution values:

$$P^0_{\text{gen}}, Q^0_{\text{gen}}, V^0, \text{and } \theta^0$$

To start, we linearize the objective function about the starting point:

$$\min \sum_{j=1}^{N_{\text{gen}}} \left[F_j(P^0_{\text{gen}\,j}) + \frac{\mathrm{d}F_j(P_{\text{gen}_j})}{\mathrm{d}P_{\text{gen}_j}} \Delta P_{\text{gen}_j} \right]$$

where $F_j(P^0_{\text{gen}j})$ is considered to be constant and

$$\frac{\mathrm{d}F_j(P_{\text{gen}_j})}{\mathrm{d}P_{\text{gen}_j}} \text{ is evaluated at } P^0_{\text{gen}\,j}.$$

Thus, the nonlinear objective is now simply a linear function

$$\min \sum_{j=1}^{N_{\text{gen}}} \left[\frac{\mathrm{d}F_j(P_{\text{gen}_j})}{\mathrm{d}P_{\text{gen}_j}} \Delta P_{\text{gen}_j} \right]$$

Next, the power flow constraints are linearized in a manner identical to the Newton power flow, except that all rows and columns of the Jacobian are included:

$$\begin{bmatrix} \frac{\mathrm{d}P_1}{\mathrm{d}\theta_1} & \frac{\mathrm{d}P_1}{\mathrm{d}V_1} & \frac{\mathrm{d}P_1}{\mathrm{d}\theta_2} & \cdots \\ \frac{\mathrm{d}Q_1}{\mathrm{d}\theta_1} & \frac{\mathrm{d}Q_1}{\mathrm{d}V_1} & \frac{\mathrm{d}Q_1}{\mathrm{d}\theta_2} & \cdots \\ \frac{\mathrm{d}P_2}{\mathrm{d}\theta_1} & \frac{\mathrm{d}P_2}{\mathrm{d}V_1} & \frac{\mathrm{d}P_2}{\mathrm{d}\theta_2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \Delta\theta_1 \\ \Delta V_1 \\ \Delta\theta_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} \Delta P_{\text{gen}_j} \\ \Delta Q_{\text{gen}_j} \\ \vdots \end{bmatrix}$$

www.mepcafe.com

These constitute a set of linear equality constraints. To make this algorithm work properly, we force the swing bus voltage magnitude and phase angle to be constant, that is,

$$\Delta V_{\text{swing}} = 0 \quad \text{and} \quad \Delta\theta_{\text{swing}} = 0$$

In addition, all load bus P and Q are forced to be constant:

$$\Delta P_{\text{load}} = 0 \quad \text{and} \quad \Delta Q_{\text{load}} = 0 \text{ for each load bus.}$$

Last of all, we place upper and lower limits on the generator real and reactive power:

$$\left.\begin{array}{l} P_{\text{gen}_j}^{\min} - P_{\text{gen}_j} \le \Delta P_{\text{gen}_j} \le P_{\text{gen}_j}^{\max} - P_{\text{gen}_j} \\ Q_{\text{gen}_j}^{\min} - Q_{\text{gen}_j} \le \Delta Q_{\text{gen}_j} \le Q_{\text{gen}_j}^{\max} - Q_{\text{gen}_j} \end{array}\right. \text{ for all generators } j$$

and limits on the bus voltage magnitudes:

$$|V_i|^{\min} - |V_i| \le \Delta|V_i| \le |V_i|^{\max} - |V_i| \text{ for all buses } i.$$

For each LP variable ΔP_{gen_j}, ΔQ_{gen_j}, and $\Delta|V_i|$, the limits earlier are the maximum that should be used; however, it is best to limit the range of each variable to much smaller values so that the linearization is more accurate. This smaller set of limits can be referred to as a window within which the variables are allowed to move on any LP execution. At the end of that execution, the limits of the window are moved but always stay within the limits in the equations shown earlier. Thus, the LP solves one small region about a starting point, then relinearizes about the solution and solves another LP within a small region about the solution. Methodologies to adjust the "window size" refer to it as a "trust region."

The steps for the incremental LP solution to the ACOPF are as follows:

1. Solve a base power flow.
2. Linearize the objective function.
3. Linearize the constraints.
4. Set variable limits.
5. Solve the LP.
6. If there is significant change in any of the variables (ΔP, ΔQ, or $\Delta|V|$), then go back to step 1, else done.

Example 8C: Comparison of DCOPF and ACOPF

We shall compare the results of an OPF using the QP algorithm and a DC power flow with the ACOPF using an iterative LP OPF. The problem to be optimized is the six-bus test system from Chapter 6 shown next (Figure 8.4).

The generator cost function is a quadratic function

$$F_i(P_i) = a_i + b_i P_i + c_i P_i^2 \text{ with } P_i^{\min} \le P_i \le P_i^{\max}$$

www.mepcafe.com

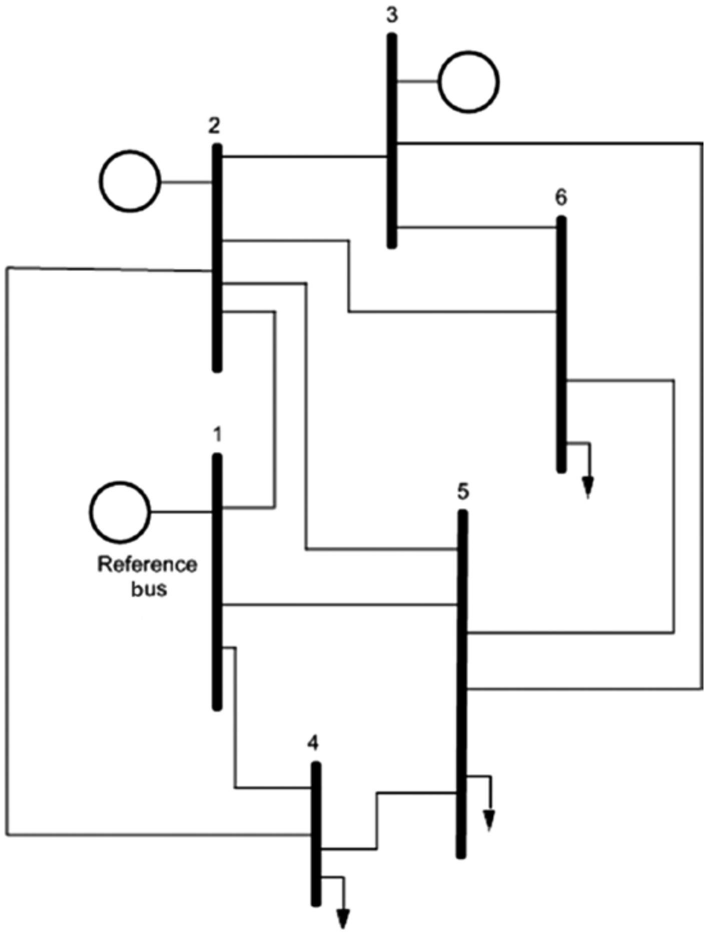


FIGURE 8.4 Six-bus power flow system.

The data for this system as given in Chapter 6 is

Generator Bus	a	b	c	$P^{\min}$ (MW)	$P^{\max}$ (MW)	$Q^{\max}$ (MVAR)	$Q^{\min}$ (MVAR)	V_g (pu Volts)
1	213.1	11.669	0.00533	50.0	200.0	150.0	−100.0	1.07
2	200.0	10.333	0.00889	37.5	150.0	150.0	−100.0	1.05
3	240.0	10.833	0.00741	45.0	180.0	120.0	−100.0	1.05

The bus loads are

Load Bus	P_{load} (MW)	Q_{load} (MVAR)
4	100.0	15.0
5	100.0	15.0
6	100.0	15.0

www.mepcafe.com

The branch impedance data is

From Bus	To Bus	R (pu)	X (pu)	B_{cap} (pu)	Flow Limit
1	2	0.10	0.20	0.04	100.0
1	4	0.05	0.20	0.04	100.0
1	5	0.08	0.30	0.06	100.0
2	3	0.05	0.25	0.06	60.0
2	4	0.05	0.10	0.02	60.0
2	5	0.10	0.30	0.04	60.0
2	6	0.07	0.20	0.05	60.0
3	5	0.12	0.26	0.05	60.0
3	6	0.02	0.10	0.02	60.0
4	5	0.20	0.40	0.08	60.0
5	6	0.10	0.30	0.06	60.0

All buses have the following voltage limits:

$$0.95 \text{ pu} \le |V_i| \le 1.07 \text{ pu}$$

The OPF will be solved using the DC power flow (linear power flow) model as well as the full AC power flow model.

The first set of results are for the DCOPF and ACOPF running so as to ignore line limits.

DCOPF Results

Bus	P^{min} (MW)	P_{gen} (MW)	P^{max} (MW)	P_{load} (MW)	Lambda ($/MWh)
1	50.00	72.64	200	0	12.443
2	37.50	118.69	150	0	12.443
3	45.00	108.66	180	0	12.443
4		0		100	12.443
5		0		100	12.443
6		0		100	12.443
Total generation cost =			4145.2	$/h	

ACOPF Results

Total	P_{gen}	=	307	Total	Q_{gen}	=	8.56
Total	P_{load}	=	300	Total	Q_{load}	=	45
Total	MW	Losses =	7	Total	MVAR	Losses	−36.44
Total generation cost =			4232	$/MWh			

www.mepcafe.com

Bus	P^{min} (MW)	P_{gen} (MW)	P^{max} (MW)	Q^{min} (MVAR)	Q_{gen} (MVAR)	Q^{max} (MVAR)	P_{load} (MW)	Q_{load} (MVAR)	Lambda P ($/MWh)	Lambda Q ($/MVARh)
1	50.0	78.5	200	−100	2.3	150	0	0	12.506	0.000
2	37.5	118.8	150	−100	1.0	150	0	0	12.446	0.000
3	45.0	109.6	180	−100	5.2	120	0	0	12.459	0.000
4							100	15	13.114	0.084
5							100	15	13.183	0.058
6							100	15	12.892	0.060

V^{min} (pu Volts)	V_{bus} (pu Volts)	V^{max} (pu Volts)
0.95	1.07UL	1.07
0.95	1.07UL	1.07
0.95	1.07UL	1.07
0.95	1.04	1.07
0.95	1.04	1.07
0.95	1.05	1.07

where UL is an indicator that a voltage is at its upper limit.

The solutions differ in the total generation, 300 MW for the DCOPF and 307 MW for the ACOPF indicating that the ACOPF has to supply the 7 MW of losses. Further, the bus incremental costs or bus lambda values are all 12.443 for the DCOPF and vary from 12.446 to 13.183 for the ACOPF. More on this later. Last, we observe that the ACOPF has run all generator controlled voltages to their upper limit of 1.07 per unit.

Now the line limits case will be run for each.

DCOPF Results

Bus	P^{min} (MW)	P_{gen} (MW)	P^{max} (MW)	P_{load} (MW)	Lambda ($/MWh)
1	50	102.34	200	0	12.760
2	37.5	122.18	150	0	12.505
3	45	75.48	180	0	11.952
4		0		100	13.089
5		0		100	12.648
6		0		100	13.149
Total generation cost =			4158.2	$/h	

Line Limits Results

Line Flows at Limit				
From Bus	To Bus	P_{flow} (MW)	$P_{flowmax}$ (MW)	Lambda
2	4	60	60	0.860
3	6	60	60	1.690

www.mepcafe.com

ACOPF Results

Total	P_{gen}		306.99	Total	Q_{gen}		8.56
Total	P_{load}		300	Total	Q_{load}		45
Total	MW	Losses	6.99	Total	MVAR	Losses	−36.44
Total generation cost =			4232.4	\$/h			

Bus	P^{min} (MW)	P_{gen} (MW)	P^{max} (MW)	Q^{min} (MVAR)	Q_{gen} (MVAR)	Q^{max} (MVAR)	P_{load} (MW)	Q_{load} (MVAR)	Lambda P (\$/MWh)	Lambda Q (\$/MVARh)
1	50.0	78.5	200	−100	2.3	150	0	0	12.506	0.000
2	37.5	118.8	150	−100	1.0	150	0	0	12.446	0.000
3	45.0	109.6	180	−100	5.2	120	0	0	12.459	0.000
4							100	15	13.114	0.084
5							100	15	13.183	0.058
6							100	15	12.892	0.060

V_{min} (pu Volts)	V_{bus} (pu Volts)	V_{max} (pu Volts)
0.95	1.07UL	1.07
0.95	1.07UL	1.07
0.95	1.06	1.07
0.95	1.04	1.07
0.95	1.03	1.07
0.95	1.04	1.07

Line Limits Results

Line Flows at Limit				
From Bus	To Bus	P_{flow} (MW)	$P_{flowmax}$ (MW)	Lambda
2	4	59.98	60	1.036
3	6	59.97	60	2.752

8.11 RELATIONSHIP BETWEEN LMP, INCREMENTAL LOSSES, AND LINE FLOW CONSTRAINTS

First, it has to be noted that one must take into account whether a bus generation increase or a bus load increase is being dealt with in the following sections. If one uses the incremental loss

$$\frac{\partial P_{loss}}{\partial P_i}$$

www.mepcafe.com

it matters whether it is a generation that is increasing its output by +1 MW into the bus or a load that is increasing its load by +1 MW that takes 1 MW out of the bus. Thus,

$$\Delta P_{\text{loss}} = \frac{\partial P_{\text{loss}}}{\partial P_i} \Delta P_i \text{ if bus } i \text{ is a generator bus}$$

$$\Delta P_{\text{loss}} = -\frac{\partial P_{\text{loss}}}{\partial P_i} \Delta P_i \text{ if bus } i \text{ is a load bus}$$

Similarly the sensitivity of line flow on line ℓ to a 1 MW change at bus i, with the reference bus adjusting to keep the system load and generation balanced, is

$$a_{\ell i} = \frac{\partial P_{\text{flow}_\ell}}{\partial P_i}$$

Then as the preceding,

$$\Delta \text{flow}_\ell = a_{\ell i} \Delta P_i \text{ if bus } i \text{ is a generator bus}$$

$$\Delta \text{flow}_\ell = -a_{\ell i} \Delta P_i \text{ if bus } i \text{ is a load bus}$$

In the following sections, we will use the following bus designations:

i = an index for any bus
g = an index for any generator
ref = the reference bus
k = a load bus
j = a generator bus
ℓ = an index for a line in the system

8.11.1 Locational Marginal Price at a Bus with No Lines Being Held at Limit

The definition of LMP is the marginal increase in cost to the system to supply one additional MW of load at bus k as shown in Figure 8.5 that follows. We shall assume here that the entire system is at its minimum cost as solved by an OPF and that when the additional 1 MW is added at bus k, the additional power is supplied by the reference bus, which is a generator.

Then

$$\Delta P_{\text{ref}} = \Delta P_k - \frac{\partial P_{\text{loss}}}{\partial P_k} \Delta P_k$$

www.mepcafe.com

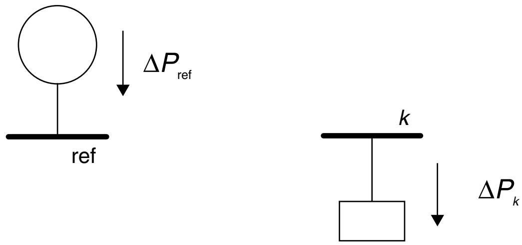


FIGURE 8.5 LMP at load bus with no line flows at limit.

where the minus sign in front of the incremental loss is due to bus k being a load bus. The total cost of supplying power is

$$T_{cost} = \sum_{\substack{\text{all generating} \\ \text{units } g}} C_g(P_g)$$

but only the reference-bus generation is going to change when 1 MW is added to the load; so then

$$\Delta T_{cost} = \frac{\partial C_{ref}(P_{ref})}{\partial P_{ref}} \Delta P_{ref}$$

Then

$$\Delta T_{cost} = \frac{\partial C_{ref}(P_{ref})}{\partial P_{ref}} \left(1 - \frac{\partial P_{loss}}{\partial P_k}\right) \Delta P_k$$

By definition, LMP_k is

$$\text{LMP}_k = \frac{\Delta T_{cost}}{\Delta P_k} = \frac{\partial C_{ref}(P_{ref})}{\partial P_{ref}} \left(1 - \frac{\partial P_{loss}}{\partial P_k}\right) = \text{LMP}_{ref} - \frac{\partial P_{loss}}{\partial P_k} \text{LMP}_{ref}$$

or

$$\text{LMP}_k = \text{LMP}_{ref} - \frac{\partial P_{loss}}{\partial P_k} \text{LMP}_{ref}$$

Thus, without any line flows at their limit, the LMP at any load bus has two components:

$$\text{LMP}_k = \text{system marginal price at the reference bus} + \text{system loss component}$$

8.11.2 Locational Marginal Price with a Line Held at Its Limit

By definition the LMP at load bus k in this case will be the cost to the power system to add 1 MW to the load at bus k without changing the flow on line ℓ. To add 1 MW to bus k without changing the flow on line ℓ, we must introduce another generator at

www.mepcafe.com

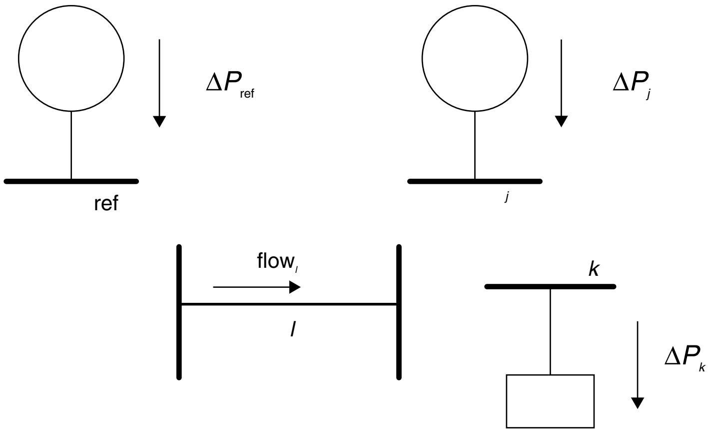


FIGURE 8.6 LMP at load bus with a line flow held at limit.

bus *j*. When the 1 MW is added to bus *k*, the reference bus takes up the additional load plus losses and the generator at bus *j* will change to keep line flow at line ℓ at its limit. That is, there will be no change in flow on line ℓ. We will model the system shown next that has a reference bus, a load bus, a generator bus, and a transmission line as in Figure 8.6 that follows:

We first must go back to the Lagrange function to get the equation for the Lagrange multiplier for the line flow constraint:

$$\mathcal{L} = C_{\text{ref}}(P_{\text{ref}}) + C_j(P_j) + \lambda(P_{\text{load}} + P_{\text{loss}} - P_{\text{ref}} - P_j) + \mu(\text{flow}_\ell - \text{flow}_\ell^+)$$

where flow_ℓ^+ is the maximum flow on line ℓ and this limit is being held so that by the KKT criteria, μ is positive. To model the power flow on line ℓ, we write the following:

$$\text{flow}_\ell = \text{flow}_\ell^0 + \Delta\text{flow}_\ell$$

where

$$\Delta\text{flow}_\ell = a_{\ell j}\Delta P_j - a_{\ell k}\Delta P_k$$

or finally

$$\text{flow}_\ell = \text{flow}_\ell^o + a_{\ell j}(P_j - P_j^0) - a_{\ell k}(P_k - P_k^0)$$

The conditions at the optimum are

$$\frac{\partial \mathcal{L}}{\partial \text{flow}_\ell} = -\mu$$

$$\frac{\partial \mathcal{L}}{\partial P_{\text{ref}}} = \frac{\partial C_{\text{ref}}(P_{\text{ref}})}{\partial P_{\text{ref}}} - \lambda = 0 \text{ so that } \lambda = \frac{\partial C_{\text{ref}}(P_{\text{ref}})}{\partial P_{\text{ref}}}$$

www.mepcafe.com

$$\frac{\partial \mathcal{L}}{\partial P_j} = \frac{\partial C_j(P_j)}{\partial P_j} + \lambda\left(\frac{\partial P_{\text{loss}}}{\partial P_j} - 1\right) + \mu\left(\frac{\partial \text{flow}_\ell}{\partial P_j}\right) = 0$$

or

$$\frac{\partial \mathcal{L}}{\partial P_j} = \frac{\partial C_j(P_j)}{\partial P_j} + \lambda\left(\frac{\partial P_{\text{loss}}}{\partial P_j} - 1\right) + \mu(a_{\ell j}) = 0$$

then

$$\mu = \left(\frac{\partial C_{\text{ref}}(P_{\text{ref}})}{\partial P_{\text{ref}}}\left(1 - \frac{\partial P_{\text{loss}}}{\partial P_j}\right) - \frac{\partial C_j(P_j)}{\partial P_j}\right)\frac{1}{a_{\ell j}}$$

The flow on line ℓ is to be held constant, so $\Delta \text{flow}_\ell = 0$ then $\Delta P_j = \frac{a_{\ell k}}{a_{\ell j}}\Delta P_k$.

We now go back to the equation relating the change in system cost to all the changes in generation in order to supply 1 MW extra to bus k while line ℓ is held to its limit:

$$\Delta P_{\text{ref}} = \Delta P_k - \frac{\partial P_{\text{loss}}}{\partial P_k}\Delta P_k - \Delta P_j + \frac{\partial P_{\text{loss}}}{\partial P_j}\Delta P_j$$

This equation takes into account the change in the load at bus k and its effect on losses as well as the change in generation on bus j and its effect on losses. By rearranging this we obtain

$$\Delta P_{\text{ref}} = \left(1 - \frac{\partial P_{\text{loss}}}{\partial P_k}\right)\Delta P_k + \left(-1 + \frac{\partial P_{\text{loss}}}{\partial P_j}\right)\Delta P_j$$

Next, we substitute $\Delta P_j = \frac{a_{\ell k}}{a_{\ell j}}\Delta P_k$ since the line flow on line ℓ is held constant, and this gives us

$$\Delta P_{\text{ref}} = \left(1 - \frac{\partial P_{\text{loss}}}{\partial P_k}\right)\Delta P_k + \left(-1 + \frac{\partial P_{\text{loss}}}{\partial P_j}\right)\left(\frac{a_{\ell k}}{a_{\ell j}}\right)\Delta P_k$$

or

$$\Delta P_{\text{ref}} = \left(\left(1 - \frac{\partial P_{\text{loss}}}{\partial P_k}\right) + \left(-1 + \frac{\partial P_{\text{loss}}}{\partial P_j}\right)\left(\frac{a_{\ell k}}{a_{\ell j}}\right)\right)\Delta P_k$$

Finally, we go back to the relationship in the change in total generation cost:

$$\Delta T_{\text{cost}} = \frac{\partial C_{\text{ref}}(P_{\text{ref}})}{\partial P_{\text{ref}}}\Delta P_{\text{ref}} + \frac{\partial C_j(P_j)}{\partial P_j}\Delta P_j$$

www.mepcafe.com

We then make substitutions for ΔP_{ref} and ΔP_j to obtain

$$\Delta T_{\text{cost}} = \frac{\partial C_{\text{ref}}(P_{\text{ref}})}{\partial P_{\text{ref}}}\left(\left(1-\frac{\partial P_{\text{loss}}}{\partial P_k}\right)+\left(-1+\frac{\partial P_{\text{loss}}}{\partial P_j}\right)\left(\frac{a_{\ell k}}{a_{\ell j}}\right)\right)\Delta P_k + \frac{\partial C_j(P_j)}{\partial P_j}\left(\frac{a_{\ell k}}{a_{\ell j}}\right)\Delta P_k$$

This can now be arranged into a final form

$$\text{LMP}_k = \frac{\Delta T_{\text{cost}}}{\Delta P_k} = \frac{\partial C_{\text{ref}}(P_{\text{ref}})}{\partial P_{\text{ref}}} - \frac{\partial P_{\text{loss}}}{\partial P_k}\frac{\partial C_{\text{ref}}(P_{\text{ref}})}{\partial P_{\text{ref}}} - \left(\frac{\partial C_{\text{ref}}(P_{\text{ref}})}{\partial P_{\text{ref}}}\left(1-\frac{\partial P_{\text{loss}}}{\partial P_j}\right)\left(\frac{1}{a_{\ell j}}\right) - \frac{\partial C_j(P_j)}{\partial P_j}\left(\frac{1}{a_{\ell j}}\right)\right)a_{\ell k}$$

but the last term in the preceding equation in the large brackets is μ; then

$$\text{LMP}_k = \frac{\Delta T_{\text{cost}}}{\Delta P_k} = \frac{\partial C_{\text{ref}}(P_{\text{ref}})}{\partial P_{\text{ref}}} - \frac{\partial P_{\text{loss}}}{\partial P_k}\frac{\partial C_{\text{ref}}(P_{\text{ref}})}{\partial P_{\text{ref}}} - \mu a_{\ell k}$$

or

$$\text{LMP}_k = \frac{\partial C_{\text{ref}}(P_{\text{ref}})}{\partial P_{\text{ref}}} - \frac{\partial P_{\text{loss}}}{\partial P_k}\frac{\partial C_{\text{ref}}(P_{\text{ref}})}{\partial P_{\text{ref}}} - \mu a_{\ell k} = \text{LMP}_{\text{ref}} - \frac{\partial P_{\text{loss}}}{\partial P_k}\text{LMP}_{\text{ref}} - \mu a_{\ell k}$$

If there are multiple lines being held at their limit, this becomes

$$\text{LMP}_k = \text{LMP}_{\text{ref}} - \frac{\partial P_{\text{loss}}}{\partial P_k}\text{LMP}_{\text{ref}} - \sum_{\ell=1}^{\text{NLL}} \mu_\ell a_{\ell k}$$

where NLL = number of lines at limit and μ_ℓ is the Lagrange multiplier for line ℓ.

We see now that the LMP at bus k with lines held at their limit has three components:

$$LMP_k = \text{system marginal price at the reference bus} + \text{system loss component} + \text{congestion (line flow limit) component}$$

The relationship between the bus incremental cost at the reference bus and the bus LMP values for the six bus case with no line flow limits are shown in the following table. The LMP Calc column is calculated using $\text{LMP}_k = \text{LMP}_{\text{ref}} - \frac{\partial P_{\text{loss}}}{\partial P_k}\text{LMP}_{\text{ref}}$ and is compared to the LMP from the OPF algorithm itself.

www.mepcafe.com

				From OPF
Bus	Lambda Ref	dP_{loss}/dP_i	LMP Calc	Lambda P ($/MWh)
1				12.506
2	12.506	0.0052	12.4409688	12.446
3	12.506	0.0044	12.4509736	12.459
4	12.506	−0.0484	13.1112904	13.114
5	12.506	−0.0538	13.1788228	13.183
6	12.506	−0.0303	12.8849318	12.892

Similar results can be generated for the case with line limits enforced (here lines 2–4 and 3–6 are at limit). The LMP Calc column is calculated using

$$\mathrm{LMP}_k = \mathrm{LMP}_{\mathrm{ref}} - \frac{\partial P_{\mathrm{loss}}}{\partial P_k}\mathrm{LMP}_{\mathrm{ref}} - \sum_{i=1}^{\mathrm{NLL}} \mu_i a_{\ell ik}$$

								From OPF
Bus	Lambda Ref	dP_{loss}/dP_i	Mu(2–4)	$a_{\ell i} = d$flow (2–4)/dP_i	Mu(3–6)	$a_{\ell i} = d$flow (3–6)/dP_i	LMP Calc	Lambda P ($/MWh)
1								12.796
2	12.796	−0.0124	1.036	0.3115	2.752	−0.0077	12.65337	12.628
3	12.796	−0.0302	1.036	0.2154	2.752	0.3695	11.94279	11.910
4	12.796	−0.0592	1.036	−0.379	2.752	−0.0023	13.95245	13.950
5	12.796	−0.0726	1.036	0.1013	2.752	0.015	13.57892	13.562
6	12.796	−0.0605	1.036	0.2208	2.752	−0.3433	14.28619	14.293

8.12 SECURITY-CONSTRAINED OPF

In Chapter 7, we introduced the concept of security analysis and the idea that a power system could be constrained to operate in a secure manner. Programs that can make control adjustments to the base or precontingency operation to prevent violations in the postcontingency conditions are called "security-constrained OPFs," or SCOPF.

We have seen previously that an OPF is distinguished from an ED by the fact that it constantly updates a power flow of the transmission system as it progresses toward the minimum of the objective function. One advantage of having the power flow updated is the fact that constraints can be added to the OPF that reflect the limits that must be respected in the transmission system. Thus, the OPF allows us to reach an optimum with limits on network components recognized.

An extension to this procedure is to add constraints that model the limits on components during contingency conditions. That is, these new "security constraints" or "contingency constraints" allow the OPF to meet precontingency limits as well as postcontingency limits. There is a price to pay, however, and that is the fact that as we iterate the OPF with an AC power flow, we must also run power flows for all the contingency cases being observed. This is illustrated in Figure 8.7.

www.mepcafe.com

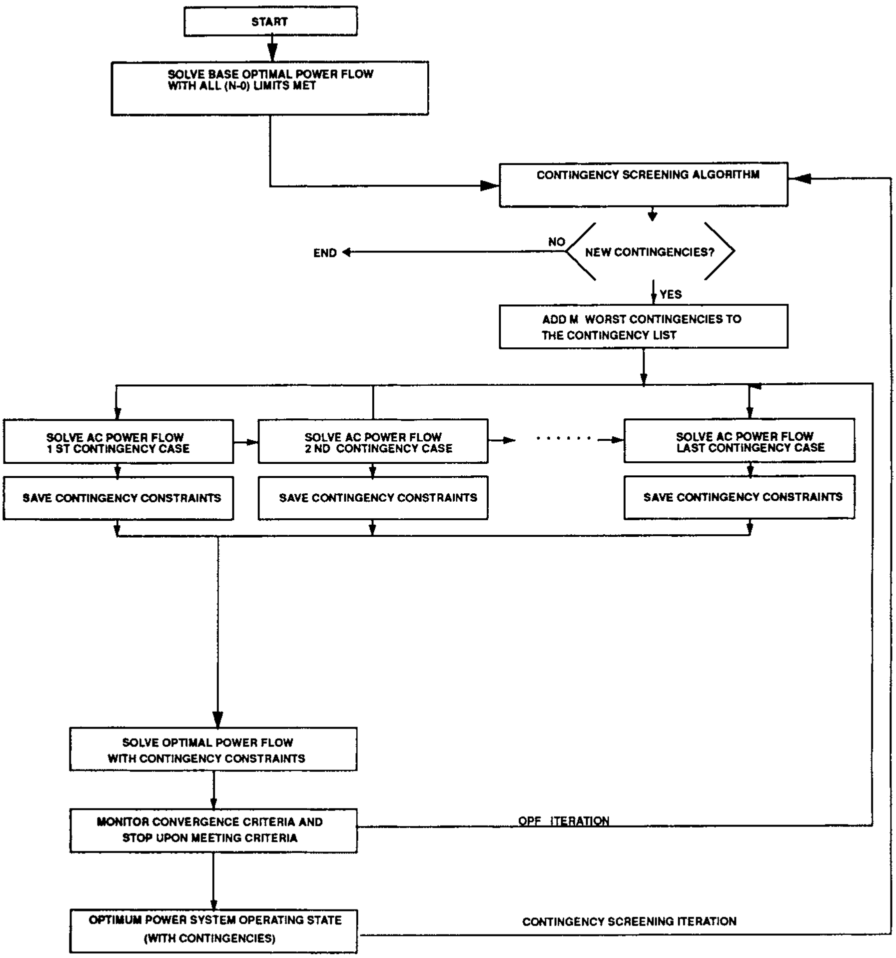


FIGURE 8.7 Security-constrained optimal power flow.

The SCOPF shown in Figure 8.7 starts by solving an OPF with $(n-0)$ constraints only. Only when it has solved for the optimal, constrained conditions is the contingency analysis executed. In Figure 8.7, the contingency analysis starts by screening the power system and identifying the potential worst-contingency cases. As was pointed out in Chapter 7, not all of these cases are going to result in a postcontingency violation and it is important to limit the number of full power flows that are executed. This is especially important in the SCOPF, where each contingency power flow may result in new contingency constraints being added to the OPF. We assume here that only the M worst cases screened by the screening algorithm are added. It is possible to make $M=1$, in which case only the worst potential contingency is added.

Next, all the $(n-1)$ contingency cases that are under consideration must be solved by running a power flow with that contingency reflected in alterations to the power flow model. When the power flow results in a security violation, the power system model is

www.mepcafe.com

used to create a contingency constraint. In fact, what is done is to run a network sensitivity calculation (see Section 7.4.1) on the model with the contingency outage and save the resulting constraint sensitivities. When all contingency power flows are complete, all the contingency constraints are added to the OPF model and it is solved.

Note, in Figure 8.7, there are two main loops to be executed. The loop labeled "OPF Iteration" requires the OPF and each of the contingency power flows to be re-executed until the OPF has solved with all contingency constraints met. Next, the outer loop labeled "Contingency Screening Iteration" is tried. If the contingency screening algorithm does not pick up any new contingencies, the SCOPF can end; if new contingencies are found, it must add them to the list and continue.

Why is all this necessary? The optimum operation conditions for a power system will often result in violation of system security. An example occurs when a large amount of interchange power is available at a favorable price. In this instance, the selling power system can be modeled in the OPF with its price of production set accordingly, and the OPF will then raise the interchange up to the point where transmission system components are limiting. Now, when the contingency analysis is run, there may be many cases that result in contingency violations, and the OPF, with contingency constraints added, will have to back off the interchange power in order to meet the contingency limits.

It should also be noted that when some contingency constraints are added to the OPF, it will redispatch generation and adjust voltages and transformers to meet these constraints. The process of adjustments may result in many new contingency violations when the screening algorithm and the power flows are run. The need to iterate between the OPF and the contingency screening represents an effort to find the "most constraining" contingencies.

8.12.1 Security Constrained OPF Using the DC Power Flow and Quadratic Programming

As shown in Section 8.8.1, the cost function will be represented as

$$\sum_{i=1}^{N_{\text{gen}}} F_i(P_{\text{gen}_i}) = \sum_{i=1}^{N_{\text{gen}}} (a_i + b_i P_{\text{gen}_i} + c_i P_{\text{gen}_i}^2)$$

and the power flow as a set of linear equations:

$$\underline{P_{\text{gen}}} - \underline{P_{\text{load}}} - 100[B_x]\underline{\theta} = \underline{0}$$

$$P_{\text{gen}_i}^{\min} \le P_{\text{gen}_i} \le P_{\text{gen}_i}^{\max}, \quad \text{for} \quad i = 1 \ldots N_{\text{gen}}$$

Instead of making the load fixed constants, we will enter them as variables. Later we will actually allow the loads to vary in response to the so-called load benefit functions; for now, assume that the P_{load} variables are entered as variables into the QP and then simply held fixed by making the upper and lower limits equal to the value desired.

www.mepcafe.com

The variables in our *QP* now consist of P_{gen}'s, and P_{load}'s. Next, we add a variable called theta for each bus, including the reference bus. We shall add more variables later. Thus, the variables vector now looks like this:

$P_{gen1}\ P_{gen2} \ldots P_{load1}\ P_{load2} \ldots \text{theta}_1\ \text{theta}_2 \ldots$

8.12.2 DC Power Flow

The program develops two matrices:

$B_x = \text{sum}(1/x_{ij})$ for all lines connected to bus i

$B_x = -1/x_{ij}$ for each line ij connected to bus i

And we solve the power flow using: $P_{net} = [B_x] \times \underline{\text{theta}}$

where

$\underline{P_{net}}$ is the vector of net power at each bus ($P_{gen} - P_{load}$)

$\underline{\text{theta}}$ is the vector of phase angles

The *QP* constraint matrix is shown in Figure 8.8.

The number of variables in the *QP* constraint matrix is now $N_{gen} + N_{loads} + N_{bus}$ and the number of rows is N_{bus}. The *QP* with just this set of equality constraints will solve for the minimum of the cost function and solve the DC power flow at the same time.

The MATLAB *QP* returns the Lagrange multiplier for each equality and each inequality constraint as well as for each upper and lower bound constraint (see Example 8D). The Lagrange multipliers for the set of constraints mentioned earlier—that is, for each bus in the DC power flow—are the LMP. In this case, since there are no line constraints all, LMPs are equal.

8.12.3 Line Flow Limits

To add line flow limits, we develop a "Flow" matrix [Flow] where

$$\text{Flow}_{\text{row}\,n} = \begin{bmatrix} & 1/x_{ij} & -1/x_{ij} & \\ & \text{col}\,i & \text{col}\,j & \end{bmatrix}$$

where line n connects bus i and bus j.

Then flows on all lines (1 to n lines) are calculated using the Flow matrix:

$$\text{Line flow} = [\text{Flow}] \times \text{theta}$$

where Lineflow is vector of all line flows with positive value being flow from bus i toward bus j.

Each line has a maximum flow called flowmax, so we have the inequality

$$-\text{flow max}_n \leq \text{lineflow}_n \leq \text{flow max}_n$$

This is actually a pair of inequality constraints, but can be represented by a single constraint if we have the ability to express variable limits on variables (which we do

www.mepcafe.com

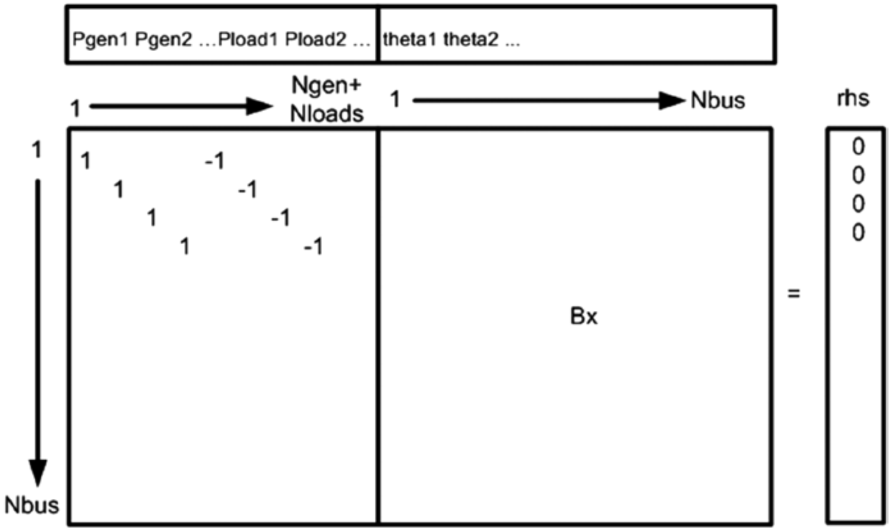


FIGURE 8.8 QP constraint matrix with DC power flow only.

in the MATLAB QP). So this pair of constraints becomes (now expressed as an equality for line "n" with a slack variable S_n)

$$\text{lineflow}_n - S_n = \text{flow}max_n$$

where upper and lower bounds on the variable S_n are

$$0 \le S_n \le 2 * \text{flowmax}_n$$

If we wish to add line flow limits, we add the line flow constraints to the *QP* constraint matrix as shown in Figure 8.9:

Now the number of *LP* variables is $N_{\text{gen}} + N_{\text{loads}} + N_{\text{bus}} + N_{\text{lines}}$, and the number of rows is $N_{\text{bus}} + N_{\text{lines}}$.

This *QP* will solve for the minimum cost while solving the DC power flow and meet all line flow constraints. The LMPs will now reflect differences if there are binding flow limits. In addition, the Lagrange multipliers for the flow constraints are referred to as the congestion costs.

8.12.4 Contingency Limits

We calculate a table of LODF as shown in Chapter 7 such that

$$\text{lineflow}_n(\text{after outage } of \text{ line } m) = \text{lineflow}_n \text{ (before outage)}$$
$$+ \text{LODF}(n,m) \times \text{lineflow}_m(\text{before outage})$$

We do not try to generate a matrix that represents the effect of all possible (n − 1) contingencies on all remaining lines. This would be of the order of N_{lines} squared. The best

www.mepcafe.com

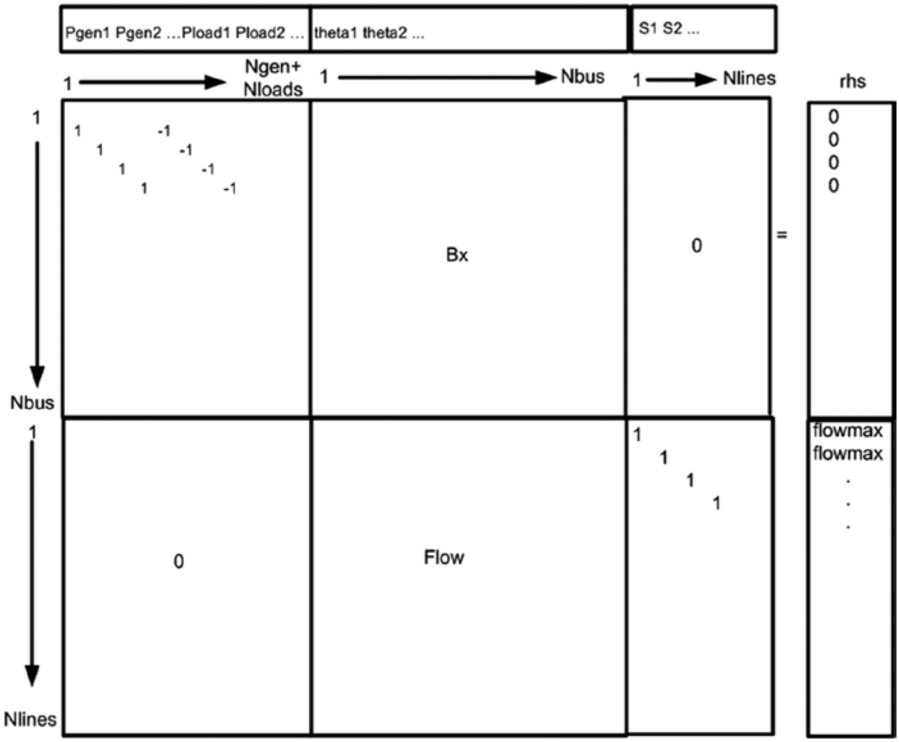


FIGURE 8.9 QP constraint matrix with DC power flow and line limit constraints for all lines.

procedure here is to run the base LP with line limits, then test all contingencies and build constraints for only those that represent violations as shown in Figure 8.10. We put a constraint for the most violated contingency into the *QP* constraint matrix, one at a time, and resolve. Typically the contingency constraints end up binding and the flow constraints are slack (nonbinding). When the *QP* is done, we simply rerun the contingency analysis and see if any new line outages produce violations and put the worst into the *QP* matrix and resolve again. This process is called the iterative constraint search (ICS). ICS usually takes only a few extra iterations to find the most binding set of contingencies and then produces a solution with no additional contingency violations.

To put the contingencies in the *QP*, we need the sensitivities to the contingencies. These are simply

$$\text{Flow}_{\text{row }n} + \text{LODF}(n,m) * \text{Flow}_{\text{row }m}$$

The limits on the contingency flows can be different from those on the flow constraints, the *QP*; for example, flowmax is multiplied a multiplier like 1.1, meaning that all contingency flows are allowed to be 10% higher. The altered contingency loading limits are simply listed as flowmax in Figure 8.10. The number of active contingency constraints in the *QP* is N_{cont}.

www.mepcafe.com

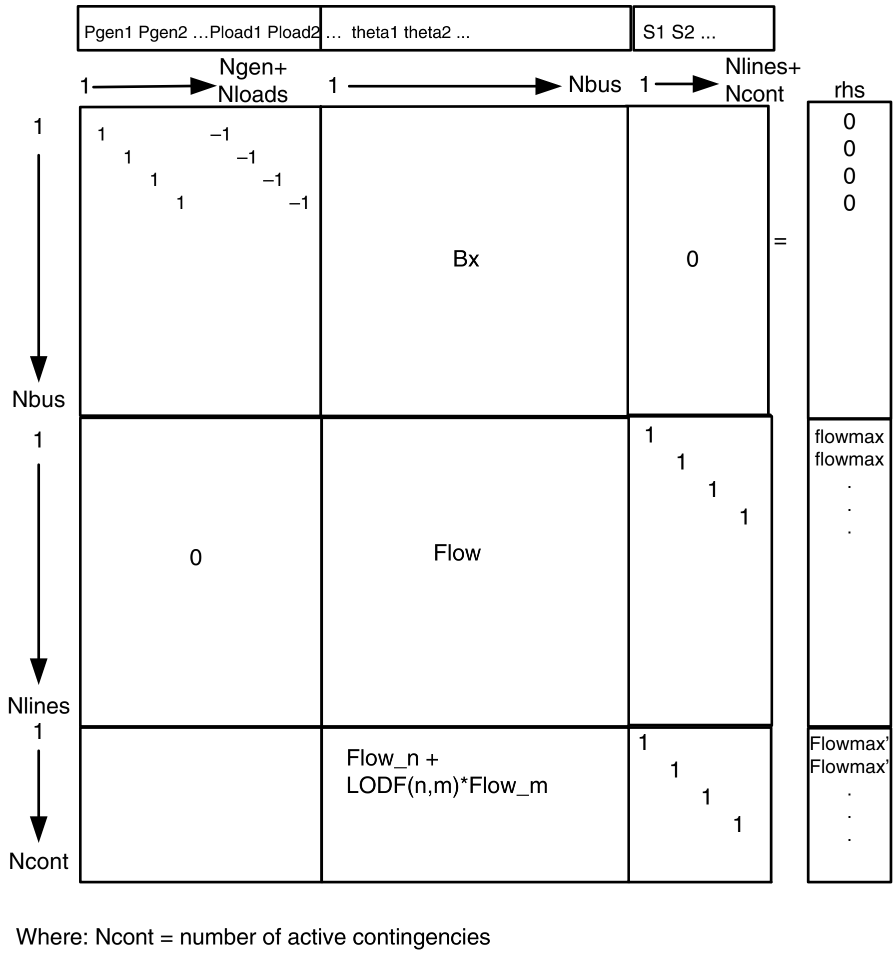


FIGURE 8.10 QP constraint matrix with DC power flow, line flow limits, and contingency limits.

The *QP* with flow constraints and contingency constraints is now as follows.

The *QP* returns bus LMPs, as well as congestion costs for binding flow limits and congestion costs for binding contingency constraints. Note that there are now $N_{\text{lines}} + N_{\text{contingency}}$ slack variables.

Example 8D: This example shows the execution of the SCOPF on a 12-bus system made by joining two copies of the original 6-bus system we started using in Chapter 6 (see Figure 6.14). In the 12-bus system, we have doubled the line capacity of some of the lines and added the three interconnections shown. In addition, we have changed the bus loading and the generation cost functions. The diagram of the 12-bus system follows (Figure 8.11).

The data for the 12-bus system is given in Appendix 8B. The first step is to solve the OPF for the network, loads, and generators shown in Appendix 8B. The generators in

www.mepcafe.com

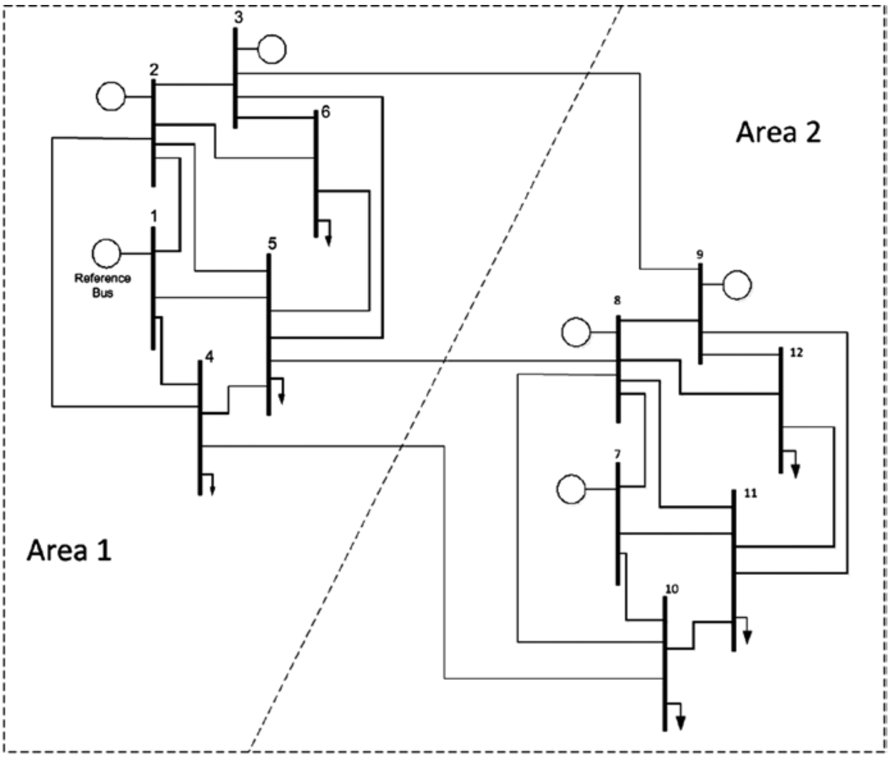


FIGURE 8.11 12-bus example system.

area 1 are much more expensive, and the OPF places all three generators (busses 1, 2, and 3) at their minimum. The bus loads, generations, and lambdas are

Bus Number	Generation at Bus	Load at Bus	Bus Lambda
1	50.00	0.00	12.662
2	37.50	0.00	12.662
3	45.00	0.00	12.662
4	0.00	110.00	12.662
5	0.00	110.00	12.662
6	0.00	110.00	12.662
7	93.13	0.00	12.662
8	130.98	0.00	12.662
9	123.40	0.00	12.662
10	0.00	50.00	12.662
11	0.00	50.00	12.662
12	0.00	50.00	12.662

www.mepcafe.com

The lambda at each bus is identical since no line flows are at limit. The power flowing over the three tie lines is

From Bus	To Bus	Low Limit	Flow	High Limit	Percent Loading
3	9	-80.0	-64.6	80.0	80.7
5	8	-80.0	-70.3	80.0	87.8
4	10	-80.0	-62.7	80.0	78.3

Area 1 has a load of 330 MW but is only generating 132.5 MW and is importing 197.6 MW over the tie lines.

When this same system with the same loads and generator cost functions is now subjected to contingency analysis, the following is observed. Note that we are allowing the contingency flows to go to 125% of the normal line loading limits. The worst $(n - 1)$ outage is

```
Drop line 7 10 Overload on 8 10 Flow=93.22 Limit=75.00
```

In this case loss of line 7–10 results in an overload on line 8–10; the contingency flow is 93.22 MW, well above the limit of 75 MW. This contingency is modeled using the compensated flow factors shown in Figure 8.10. The resulting OPF is then subjected to another complete set of $(n - 1)$ outage tests and results in another contingency overload:

```
Drop line 3 6 Overload on 2 6 Flow=75.06 Limit=75.00
```

The overload is very small but it is incorporated into the OPF. The final execution of the OPF followed by a last complete contingency analysis finds no new $(n - 1)$ contingencies. Both contingencies result in loading exactly at their respective limits. The new OPF shows that the bus loading is

Bus Number	Generation at Bus	Load at Bus	Bus Lambda
1	50.00	0.00	17.926
2	83.82	0.00	17.735
3	45.00	0.00	17.195
4	0.00	110.00	18.470
5	0.00	110.00	17.395
6	0.00	110.00	18.309
7	92.96	0.00	12.660
8	43.08	0.00	11.099
9	165.14	0.00	13.280
10	0.00	50.00	28.681
11	0.00	50.00	15.007
12	0.00	50.00	12.999

The bus lambdas are now all different because of the active contingency constraints. More of the area 1 generation is being supplied by its own generation on buses 1, 2, and 3, and the interchange between areas is

www.mepcafe.com

From Bus	To Bus	Low Limit	Flow	High Limit	Percent Loading
3	9	-80.0	-53.0	80.0	66.2
5	8	-80.0	-52.3	80.0	65.4
4	10	-80.0	-45.9	80.0	57.3

Area 1 is now importing only 151.2 MW. The two contingency constraints that are in the OPF are

Contingency	Contflow	Contlimit	Rating	Overload%	Lambda
drop 3 to 6 overload on 2 to 6	75.0	75.0	60.0	100.0	1.58
drop 7 to 10 overload on 8 to 10	75.0	75.0	60.0	100.0	22.27

where

Contflow is the resulting flow if the line is dropped.
Contlimit is the "contingency limit" = 1.25 * the flow limit.
Rating is the normal line rating.
Overload% is the normal line rating.
Lambda is the contingency constraint lambda.

Thus, the final solution took three OPF executions, one for the base solution with no contingency constraints, one after incorporating the first contingency constraint, and one after incorporating the second contingency constraint.

APPENDIX 8A
Interior Point Method

The interior point method converts the inequality constraints to equality constraints using the same technique of adding a slack variable to each constraint. Then a "penalty function" is added to the objective function so that the equations become as follows:

Minimize $f_\mu = f(\underline{x}) - \mu^k \sum_{i=1}^{N_{\text{ineq}}} \ln(s_i)$, where N_{ineq} is the number if inequality constraints.

Subject to $g(\underline{x}) = 0$

and $h(\underline{x}) + s = 0$

where $\mu^k \geq 0$ is the *barrier parameter* that is forced to decrease toward 0 as the algorithm iterates to a solution (k is the iteration counter). The Lagrangian now looks like this:

$$L_\mu = f(x) - \mu^k \sum_{i=1}^{N_{\text{ineq}}} \ln(s_i) + \lambda^{\mathrm{T}} g(x) + \gamma^{\mathrm{T}}(h(x) + s)$$

And the gradient of this Lagrangian becomes

$$\nabla_x L_\mu = \nabla_x f(x) + \nabla_x g(x)\lambda + \nabla_x h(x)\gamma = 0$$

www.mepcafe.com

$$\nabla_\lambda L_\mu = g(x) = 0$$

$$\nabla_\mu L_\mu = h(x) + s = 0$$

$$\nabla_s L_\mu = -\mu^k S^{-1} e + \gamma = 0$$

$$\text{where } S = \begin{bmatrix} s_1 & & & \\ & s_2 & & \\ & & s_3 & \\ & & & \ddots \end{bmatrix} \text{ and } e = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \end{bmatrix}$$

The algorithm at iteration k requires the solution of the following set of equations:

$$\begin{bmatrix} \nabla_x^2 L_\mu & \nabla_x g(x) & \nabla_x h(x) & 0 \\ \nabla_x g(x)^{\mathrm{T}} & 0 & 0 & 0 \\ \nabla_x h(x)^{\mathrm{T}} & 0 & 0 & I \\ 0 & 0 & I & \nabla_s^2 L_\mu \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta \gamma \\ \Delta s \end{bmatrix} = \begin{bmatrix} \nabla_x L_\mu \\ \nabla_\lambda L_\mu \\ \nabla_\gamma L_\mu \\ \nabla_s L_\mu \end{bmatrix}$$

Solution of the Optimization Problem of Appendix 3A Using Interior Point Algorithm[1]

The optimization problem of Appendix 3A with one equality constraint and one inequality constraint is as given in the following:

$$\min f(x_1, x_2) = 0.25x_1^2 + x_2^2$$

$$g(x_1, x_2) = 5 - x_1 - x_2 = 0$$

$$h(x_1, x_2) = x_1 + 0.2x_2 - 3 \leq 0$$

To make this a more interesting problem, we shall drop the equality constraint and reverse the direction of the inequality constraint; we then solve the problem:

$$\min f(x_1, x_2) = 0.25x_1^2 + x_2^2$$

$$h(x_1, x_2) = -x_1 - 0.2x_2 + 3 \leq 0$$

First, we add a slack variable to the inequality constraint and the log barrier function to the objective:

$$\min f(x_1, x_2) = 0.25x_1^2 + x_2^2 - \mu^k \ln(s)$$

$$h(x_1, x_2) = -x_1 - 0.2x_2 + 3 + s = 0$$

[1] This solution was done for a class homework problem by Volker Landenberger.

www.mepcafe.com

The Lagrangian function is

$$L(x_1, x_2, \gamma, s) = 0.25x_1^2 + x_2^2 - \mu^k \ln(s) + \gamma(-x_1 - 0.2x_2 + 3 + s)$$

And the minimum is found at the point $(x_1^*, x_2^*, \gamma^*, s^*)$ where

$$\nabla_{x_1} L_\mu = 0.5x_1^* - \gamma^* = 0$$

$$\nabla_{x_2} L_\mu = 2x_2^* - 0.2\gamma^* = 0$$

$$\nabla_\lambda L_\mu = -x_1^* - 0.2x_2^* + 3 + s^* = 0$$

$$\nabla_s L_\mu = -\frac{\mu^k}{s^*} + \gamma^* = 0$$

Starting at $x_1 = 10$, $x_2 = 10$, and $\mu = 20$, this solution of these equations with a reduction of μ by one half each iteration converges to the solution:

$$x_1 = 2.9704$$

$$x_2 = 0.1485$$

$$\gamma = 1.4852$$

$$\mu = 0.0002$$

$$s = 0.0001$$

APPENDIX 8B
Data for the 12-Bus System

Bus Data

Bus Number	Type	Pd	Qd	Gs	Bs	Area	Vm	Va	BaseKV	Zone	Vmax	Vmin
1	3	0	0	0	0	1	1	0	0	1	1.07	0.95
2	2	0	0	0	0	1	1	0	0	1	1.07	0.95
3	2	0	0	0	0	1	1	0	0	1	1.07	0.95
4	1	110	15	0	0	1	1	0	0	1	1.07	0.95
5	1	110	15	0	0	1	1	0	0	1	1.07	0.95
6	1	110	15	0	0	1	1	0	0	1	1.07	0.95
7	2	0	0	0	0	2	1	0	0	1	1.07	0.95
8	2	0	0	0	0	2	1	0	0	1	1.07	0.95
9	2	0	0	0	0	2	1	0	0	1	1.07	0.95
10	1	50	15	0	0	2	1	0	0	1	1.07	0.95
11	1	50	15	0	0	2	1	0	0	1	1.07	0.95
12	1	50	15	0	0	2	1	0	0	1	1.07	0.95

www.mepcafe.com

Generator Data

Bus	Pg	Qg	Qmax	Qmin	Vg	MVABase	Status	Pmax	Pmin
1	110	0	150	-100	1.07	100	1	200	50
2	50	0	150	-100	1.05	100	1	150	37.5
3	50	0	120	-100	1.05	100	1	180	45
7	110	0	150	-100	1.07	100	1	200	50
8	50	0	150	-100	1.05	100	1	150	37.5
9	50	0	120	-100	1.05	100	1	180	45

Generator Cost Functions

Gen Bus	a	b	c
1	319.65	17.5035	0.007995
2	300	15.4995	0.013335
3	360	16.2495	0.011115
7	213.1	11.669	0.00533
8	200	10.333	0.00889
9	240	10.833	0.00741

where $f(P_{gen}) = a + b * P_{gen} + c * P_{gen}^2$.

Lines and Tie Line Data

fbus	tbus	r	x	b	Rate A	Rate B	Rate C	Ratio	Angle	Status
1	2	0.1	0.2	0.04	100	0	0	0	0	1
1	4	0.05	0.2	0.04	100	0	0	0	0	1
1	5	0.08	0.3	0.06	100	0	0	0	0	1
2	3	0.05	0.25	0.06	60	0	0	0	0	1
2	4	0.05	0.1	0.02	120	0	0	0	0	1
2	5	0.1	0.3	0.04	60	0	0	0	0	1
2	6	0.07	0.2	0.05	60	0	0	0	0	1
3	5	0.12	0.26	0.05	60	0	0	0	0	1
3	6	0.02	0.1	0.02	120	0	0	0	0	1
4	5	0.2	0.4	0.08	60	0	0	0	0	1
5	6	0.1	0.3	0.06	60	0	0	0	0	1
7	8	0.1	0.2	0.04	100	0	0	0	0	1
7	10	0.05	0.2	0.04	100	0	0	0	0	1
7	11	0.08	0.3	0.06	100	0	0	0	0	1
8	9	0.05	0.25	0.06	60	0	0	0	0	1
8	10	0.05	0.1	0.02	60	0	0	0	0	1
8	11	0.1	0.3	0.04	60	0	0	0	0	1
8	12	0.07	0.2	0.05	60	0	0	0	0	1
9	11	0.12	0.26	0.05	60	0	0	0	0	1
9	12	0.02	0.1	0.02	120	0	0	0	0	1
10	11	0.2	0.4	0.08	60	0	0	0	0	1
11	12	0.1	0.3	0.06	60	0	0	0	0	1
3	9	0.4	0.8	0.16	80	0	0	0	0	1
5	8	0.4	0.8	0.16	80	0	0	0	0	1
4	10	0.4	0.8	0.16	80	0	0	0	0	1

www.mepcafe.com

APPENDIX 8C
Line Flow Sensitivity Factors

In Chapter 7, we developed the concepts of PTDF and LODF factors. The PTDF factors are used to measure the linear effect of a transaction between bus *s* and bus *r*. However, if we are going to use LP, it may be useful to have a different factor that measures the effect of a change flow on line ℓ to a change in the power at bus *i*. Here we shall assume that this change in power at bus *i* is compensated by a change in the reference bus. We will use the notation used in the first and second editions of this text and call the *line flow sensitivity factors* $a_{\ell i}$ as given in the following text.[2]

The line flow sensitivity factors are designated $a_{\ell i}$ and have the following definition:

$$a_{\ell i} = \frac{\Delta f_\ell}{\Delta P_i}$$

where

$$\begin{aligned} \ell &= \text{line index} \\ i &= \text{bus index} \\ \Delta f_\ell &= \text{change in megawatt power flow on line } \ell \text{ when a change in} \\ &\quad\ \text{generation}, \Delta P_i, \text{ occurs at bus } i \\ \Delta P_i &= \text{change in generation at bus } i \end{aligned}$$

It is assumed in this definition that the change in generation, ΔP_i, is exactly compensated by an opposite change in generation at the reference bus and that all other generators remain fixed. The $a_{\ell i}$ factor then represents the sensitivity of the flow on line ℓ to a change in generation at bus *i*. The new power flow on each line in the network could be calculated using a set of "*a*" factors as follows:

$$\hat{f}_\ell = f_\ell^0 + a_{\ell i}\Delta P_i \quad \text{for} \quad l = 1 \ldots L$$

where

$$\begin{aligned} \hat{f}_\ell &= \text{flow on line } \ell \text{ after the generator on bus } i \text{ changes} \\ f_\ell^0 &= \text{the original flow on line } \ell \end{aligned}$$

The line flow sensitivity factors are linear estimates of the change in flow with a change in power at a bus. Therefore, the effects of simultaneous changes on several generating buses can be calculated using superposition.

Calculation of Line Flow Sensitivity Factors

To derive the line flow sensitivity factors, we note that the derivation in Appendix 7B gave us the equation for the change in flow for a transfer of power from bus s to bus r as

[2] People who have used the first and second editions of this text will recognize this as the "generation shift factor."

www.mepcafe.com

$$\Delta f_\ell = \frac{1}{x_\ell}(\Delta\theta_i - \Delta\theta_j)$$

Then

$$\Delta f_\ell = \frac{1}{x_\ell}\Big((X_{is} - X_{ir}) - (X_{js} - X_{jr})\Big)$$

In the case of the line flow sensitivity factors being developed here, the vector of bus power injection perturbations represents a 1-pu power increase that is made at bus i and is compensated by a 1-pu decrease in power at the reference bus. The $\Delta\theta$ values are thus equal to the derivative of the bus angles with respect to a change in power injection at bus i. Then, the required sensitivity factors are

$$a_{\ell i} = \frac{\mathrm{d}f_\ell}{\mathrm{d}P_i} = \frac{\mathrm{d}}{\mathrm{d}P_i}\left[\frac{1}{x_\ell}(\theta_n - \theta_m)\right]$$
$$= \frac{1}{x_\ell}\left(\frac{\mathrm{d}\theta_n}{\mathrm{d}P_i} - \frac{\mathrm{d}\theta_m}{\mathrm{d}P_i}\right) = \frac{1}{x_\ell}(X_{ni} - X_{mi})$$

where

$$X_{ni} = \frac{\mathrm{d}\theta_n}{\mathrm{d}P_i} = n\text{th element from the } \Delta\theta \text{ vector}$$

$$X_{mi} = \frac{\mathrm{d}\theta_m}{\mathrm{d}P_i} = m\text{th element from the } \Delta\theta \text{ vector}$$

$$x_\ell = \text{line reactance for line } \ell$$

We can now write the formula for the change in flow on line ℓ as shown earlier:

$$\hat{f}_\ell = f_\ell^0 + a_{\ell i}\Delta P_i \text{ for } l = 1 \ldots L$$

and if we desire to include the flow itself in an LP with generators P_i for $i = 1$ to N_{gen}, then

$$\hat{f}_\ell = f_\ell^0 + \sum_{i=1}^{N_{\text{gen}}} a_{\ell i}\Delta P_i$$

Or this can be expressed in terms of the generation itself on bus i:

$$\hat{f}_\ell = f_\ell^0 + \sum_{i=1}^{N_{\text{gen}}} a_{\ell i}(P_i - P_i^0)$$

And a line limit constraint would be written as

$$\sum_{i=1}^{N_{\text{gen}}} a_{\ell i}P_i \le f_\ell^{\max} + \sum_{i=1}^{N_{\text{gen}}} a_{\ell i}P_i^0 - f_\ell^0$$

www.mepcafe.com

The student can use this formulation for line limits in an *LP* without having to build the entire X matrix, and all the bus phase angles, theta, into the *LP* as we did in Examples 8A and 8B and in Section 8.12. See Problem 8.2.

APPENDIX 8D
Linear Sensitivity Analysis of the AC Power Flow

Linear sensitivity coefficients give an indication of the change in one system quantity (e.g., MW flow, MVA flow, bus voltage) as another quantity is varied (e.g., generator MW output, transformer tap position). These linear relationships are essential for the application of LP. Note that as the adjustable variable is changed, we assume that the power system reacts so as to keep all of the power flow equations solved. As such, linear sensitivity coefficients can be expressed as partial derivatives. For example,

$$\frac{\partial \text{MVA flow}_{ij}}{\partial \text{MW gen}_k}$$

shows the sensitivity of the flow (MVA) on line $(i-j)$ with respect to the power generated at bus k.

Some sensitivity coefficients may change rapidly as the adjustment is made and the power flow conditions are updated. This is because some system quantities vary in a nonlinear relationship with the adjustment and resolution of the power flow equations. This is especially true for quantities that have to do with voltage and MVAR flows. Sensitivities such as the variation of MW flow with respect to a change in generator MW output are rather linear across a wide range of adjustments and lead to the usefulness of the DC power flow equations and the "PTDF" and "LODF" factors introduced in 7.

For this reason, the value represented by a sensitivity coefficient is only good for small adjustments and the sensitivities must be recalculated often.

Sensitivity Coefficients of an AC Network Model

The following procedure is used to linearize the AC transmission system model for a power system. To start, we shall define two general equations giving the power injection at a bus, that is, the net power flowing into a transmission system from the bus. This function represents the power flowing into transmission lines and shunts at the bus:

$$P_i(|V|,\theta) = \text{Re}\left[\left(\sum_j V_i\left[(V_i - t_{ij}V_j)\,y_{ij}\right]^*\right) + V_i\left(V_i \sum_i y_{\text{shunt}_i}\right)^*\right]$$

$$Q_i(|V|,\theta) = \text{Im}\left[\left(\sum_j V_i\left[\left(V_i - t_{ij}V_j\right)y_{ij}\right]^*\right) + V_i\left(V_i \sum_i y_{\text{shunt}_i}\right)^*\right]$$

www.mepcafe.com

where

$$V_i = |V_i| \angle \theta_i$$
$$t_{ij} = \text{the transformer tap in branch } ij$$
$$y_{ij} = \text{the branch admittance}$$
$$y_{\text{shunt}_i} = \text{the sum of the branch and bus shunt admittances at bus } i$$

Then, at each bus,

$$P_i\left(|V|,\theta\right) = P_{\text{gen}_i} - P_{\text{load}_i}$$
$$Q_i\left(|V|,\theta\right) = Q_{\text{gen}_i} - Q_{\text{load}_i}$$

The set of equations that represents the first-order approximation of the AC network around the initial point is the same as that generally used in the Newton power flow algorithm. That is,

$$\sum \frac{\partial P_i}{\partial |V_j|} \Delta |V_j| + \sum \frac{\partial P_i}{\partial \theta_j} \Delta \theta_j + \sum \frac{\partial P_i}{\partial t_{ij}} \Delta t_{ij} = \Delta P_{\text{gen}_i}$$
$$\sum \frac{\partial Q_i}{\partial |V_j|} \Delta |V_j| + \sum \frac{\partial Q_i}{\partial \theta_j} \Delta \theta_j + \sum \frac{\partial Q_i}{\partial t_{ij}} \Delta t_{ij} = \Delta Q_{\text{gen}_i}$$

This can be placed in matrix form for easier manipulation:

$$\begin{bmatrix} \dfrac{\partial P_1}{\partial |V_1|} & \dfrac{\partial P_1}{\partial \theta_1} & \cdots \\ \dfrac{\partial Q_1}{\partial |V_1|} & \dfrac{\partial Q_1}{\partial \theta_1} & \cdots \\ & \vdots & \end{bmatrix} \begin{bmatrix} \Delta |V_1| \\ \Delta \theta_1 \\ \vdots \end{bmatrix} = \begin{bmatrix} -\dfrac{\partial}{\partial t_{ij}}(P_i) & 1 & 0 \\ -\dfrac{\partial}{\partial t_{ij}}(Q_i) & 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta t_{ij} \\ \Delta P_{\text{gen}_i} \\ \Delta Q_{\text{gen}_i} \end{bmatrix}$$

This equation will be placed into a more compact format that uses the vectors **x** and **u**, where **x** is the state vector of voltages and phase angles and **u** is the vector of control variables. The control variables are the generator MW, transformer taps, and generator voltage magnitudes (or generator MVAR). Note that at any given generator bus, we can control a voltage magnitude only within the limits of the unit VAR capacity. Therefore, there are times when the role of the state and control are reversed. Note that other controls can easily be added to this formulation. The compact form of this equation is written:

$$\left[J_{px}\right]\Delta\mathbf{x} = \left[J_{pu}\right]\Delta\mathbf{u}$$

Now, we will assume that there are several transmission system-dependent variables, **h**, that represent, for example, MVA flows, load-bus voltages, and line amperes, and

www.mepcafe.com

we wish to find their sensitivity with respect to changes in the control variables. Each of these quantities can be expressed as a function of the state and control variables; that is, for example,

$$\mathbf{h} = \begin{bmatrix} \text{MVA flow}_{nm}(|V|, \theta) \\ |V_k| \end{bmatrix}$$

where $|E_k|$ represents only load-bus voltage magnitude.

As before, we can write a linear version of these variables around the operating point

$$\Delta\mathbf{h} = \begin{bmatrix} \frac{\partial h_1}{\partial |V_1|} & \frac{\partial h_1}{\partial \theta_1} & \cdots \\ \frac{\partial h_2}{\partial |V_1|} & \frac{\partial h_2}{\partial \theta_1} & \cdots \\ \vdots & \cdots & \end{bmatrix} \begin{bmatrix} \Delta |V_1| \\ \Delta\theta_1 \\ \vdots \end{bmatrix} + \begin{bmatrix} \frac{\partial h_1}{\partial u_1} & \frac{\partial h_1}{\partial u_2} & \cdots \\ \frac{\partial h_2}{\partial u_1} & \frac{\partial u_2}{\partial h_2} & \cdots \\ & \vdots & \cdots \end{bmatrix} \begin{bmatrix} \Delta u_1 \\ \Delta u_2 \\ \vdots \end{bmatrix}$$

where

$$h_1 = \text{the line } nm \text{ MVA flow}$$
$$h_2 = \text{the bus } k \text{ voltage magnitude}$$

Again, we can put this into a compact format using the vectors **x** and **u** as before:

$$\Delta\mathbf{h} = [J_{hx}]\Delta\mathbf{x} + [J_{hu}]\Delta\mathbf{u}$$

We will now eliminate the $\Delta\mathbf{x}$ variables; that is,

$$\Delta\mathbf{x} = [J_{px}]^{-1}[J_{pu}]\Delta\mathbf{u}$$

Then, substituting,

$$\Delta\mathbf{h} = [J_{hx}][J_{px}]^{-1}[J_{pu}]\Delta\mathbf{u} + [J_{hu}]\Delta\mathbf{u}$$

This last equation gives the linear sensitivity coefficients between the transmission system quantities, **h**, and the control variables, **u**. This can be used in an *LP* to solve an ACOPF as shown earlier in the chapter.

PROBLEMS

8.1 Solve the following optimization problem using the interior point algorithm:

$$\min f(x_1, x_2) = 0.25x_1^2 + x_2^2$$
$$h(x_1, x_2) = x_2 - 0.05x_1^2 - 0.5x_1 + 2 \le 0$$

www.mepcafe.com

8.2 You are going to use a linear program and a power flow to solve an OPF. The linear program will be used to solve constrained dispatch problems, and the power flow will confirm that you have done the correct thing. For each of the problems, you should use the power flow data for the six-bus problem found in Chapter 4.

The following data on unit cost functions applies to this problem:

$$\begin{aligned}
&\textbf{Unit 1 (bus 1)}: & F(P) &= 600.0 + 6.0P + 0.02P^2 \\
&& P_{\min} &= 70 \text{ MW} \\
&& P_{\max} &= 250.0 \text{ MW} \\
&\textbf{Unit 2 (bus 2)}: & F(P) &= 220.0 + 7.3P + 0.003P^2 \\
&& P_{\min} &= 55 \text{ MW} \\
&& P_{\max} &= 135 \text{ MW} \\
&\textbf{Unit 3 (bus 3)}: & F(P) &= 100.0 + 8.0P + 0.004P^2 \\
&& P_{\min} &= 60 \text{ MW} \\
&& P_{\max} &= 160 \text{ MW}
\end{aligned}$$

When setting up the LP, you should use three straight-line segments with breakpoints as follows:

Unit 1, break points at : 70, 130, 180, and 250 MW

Unit 2, break points at : 55, 75, 95, and 135 MW

Unit 3, break points at : 60, 80, 120, and 160 MW

When using the LP for dispatching, you should ignore losses.

Set up the power flow as follows:

$$\begin{aligned}
\text{Load} &= 300 \text{ MW} \\
\text{Generation on bus 2} &= 100 \text{ MW} \\
\text{Generation on bus 3} &= 100 \text{ MW}
\end{aligned}$$

This should lead to a flow of about 67 MW on line 3–6.

Using the linear program, set up a minimum-cost LP for the three units using the breakpoints earlier and the line flow sensitivity factors, "*a*" factors from Appendix 8C. You are to constrain the system so that the flow on line 3–6 is no greater than 55 MW.

When you obtain an answer from the LP, enter the values for P_2 and P_3 found in the LP into the AC load flow and see if, indeed, the flow on line 3–6 is close to the 55 MW desired. (Be sure the load is still set to 300 MW in the AC power flow.)

www.mepcafe.com

8.3 Using the six-bus power flow example from Chapter 6 with load at 300 MW, try to adjust the MW generated on the three generators and the voltage on each generator to minimize transmission losses. Keep the generators within their economic limits and the voltages at the generators within 0.90 to 1.07 pu volts. Use the following as MVAR limits:

$$\text{Bus 2 generator}: 100 \text{ MVAR max}$$
$$\text{Bus 3 generator}: 60 \text{ MVAR max}$$

8.4 Using the six-bus power flow example from Chapter 6, set up the base case as in Problem 8.2 (300 MW load, 100 MW on generator buses 2 and 3). Solve the base conditions and note that the load voltages on buses 4, 5, and 6 are quite low. Now, drop the line from bus 2 to bus 3 and resolve the power flow. (Note that the VAR limits on buses 2 and 3 should be the same as in Problem 8.2.)

This results in a severe voltage drop at bus 6. Can you correct this voltage so it comes back into normal range (e.g., 0.90 per unit to 1.07 per unit)? Suggested options: Add fixed capacitance to ground at bus 6, raise the voltage at one or more of the generators, reduce the load MW and MVAR at bus 6, etc.

8.5 You are going to solve the following OPF in two different ways. Given a power system with two generators, P_1 and P_2, with their corresponding cost functions $F_1(P_1)$ and $F_2(P_2)$. In addition, the voltage magnitudes on the generator buses are also to be scheduled.

The balance between load and generation will be assumed to be governed by a linear constraint (see Section 6.20 for definition of β):

$$\sum_i \beta_i P = \sum_i \beta_i P_i^0$$

In addition, two constraints have been identified and their sensitivities calculated (see Appendix 8D). The first is a flow constraint where

$$\Delta \text{flow}_{\text{nm}} = \sum_i af_i \Delta P_i + \sum_i av_i \Delta V_i$$

The second constraint involves a voltage magnitude at bus k that is assumed to be sensitive only to the generator voltages:

$$\Delta V_k = \sum_i \gamma_i \Delta V$$

a. Assume that the initial generator outputs are P_1^0 and P_2^0 and that the initial voltage magnitudes are V_1^0 and V_2^0 and that you have obtained the initial flow, flow_{nm}^0, and the initial voltage, V_k^0, from a power flow program.

www.mepcafe.com

Further assume that there are limits to be constrained for flow and voltage: flow_{nm}^{+} and flow_{nm}^{-} and for the voltage V_k^{+} and V_k^{-}.
Express the flow on line *nm* and the voltage on bus *k* as linear functions of the four control variables: P_1, P_2, V_1, V_2.

b. Show the same minimum-cost dispatch solution with an LP where we break each cost function into two segments.

www.mepcafe.com

9

INTRODUCTION TO STATE ESTIMATION IN POWER SYSTEMS

9.1 INTRODUCTION

State estimation is the process of assigning a value to an unknown system state variable based on measurements from that system according to some criteria. Usually, the process involves imperfect measurements that are redundant, and the process of estimating the system states is based on a statistical criterion that estimates the true value of the state variables to minimize or maximize the selected criterion. A commonly used and familiar criterion is that of minimizing the sum of the squares of the differences between the estimated and "true" (i.e., measured) values of a function.

The ideas of least-squares estimation have been known and used since the early part of the nineteenth century. The major developments in this area have taken place in the twentieth century in applications in the aerospace field. In these developments, the basic problems have involved the location of an aerospace vehicle (i.e., missile, airplane, or space vehicle) and the estimation of its trajectory given redundant and imperfect measurements of its position and velocity vector. In many applications, these measurements are based on optical observations and/or radar signals that may be contaminated with noise and may contain system measurement errors. State estimators may be both static and dynamic. Both types of estimators have been developed for power systems. This chapter will introduce the basic development of a static-state estimator.

In a power system, the state variables are the voltage magnitudes and relative phase angles at the system nodes. Measurements are required in order to estimate the system performance in real time for both system security control and constraints on economic

Power Generation, Operation, and Control, Third Edition. Allen J. Wood, Bruce F. Wollenberg, and Gerald B. Sheblé.
© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

www.mepcafe.com

dispatch. The inputs to an estimator are imperfect power system measurements of voltage magnitudes and power, VAR, or ampere-flow quantities. The estimator is designed to produce the "best estimate" of the system voltage and phase angles, recognizing that there are errors in the measured quantities and that there may be redundant measurements. The output data are then used in system control centers in the implementation of the security-constrained dispatch and control of the system as discussed in Chapters 7 and 8.

9.2 POWER SYSTEM STATE ESTIMATION

As introduced in Chapter 7, the problem of monitoring the power flows and voltages on a transmission system is very important in maintaining system security. By simply checking each measured value against its limit, the power system operators can tell where problems exist in the transmission system—and, it is hoped, they can take corrective actions to relieve overloaded lines or out-of-limit voltages.

Many problems are encountered in monitoring a transmission system. These problems come primarily from the nature of the measurement transducers and from communication problems in transmitting the measured values back to the operations control center.

Transducers from power system measurements, like any measurement device, will be subject to errors. If the errors are small, they may go undetected and can cause misinterpretation by those reading the measured values. In addition, transducers may have gross measurement errors that render their output useless. An example of such a gross error might involve having the transducer connected up backward, thus giving the negative of the value being measured. Finally, the telemetry equipment often experiences periods when communication channels are completely out, thus depriving the system operator of any information about some part of the power system network.

It is for these reasons that power system state estimation techniques have been developed. A state estimator, as we will see shortly, can "smooth out" small random errors in meter readings, detect and identify gross measurement errors, and "fill in" meter readings that have failed due to communication failures.

To begin, we will use a simple DC load flow example to illustrate the principles of state estimation. Suppose the three-bus DC load flow of Example 6A were operating with the load and generation shown in Figure 9.1. The only information we have about this system is provided by three MW power flow meters located as shown in Figure 9.2.

Only two of these meter readings are required to calculate the bus phase angles and all load and generation values fully. Suppose we use M_{13} and M_{32} and further suppose that M_{13} and M_{32} give us perfect readings of the flows on their respective transmission lines:

$$M_{13} = 5 \text{ MW} = 0.05 \text{ pu}$$
$$M_{32} = 40 \text{ MW} = 0.40 \text{ pu}$$

www.mepcafe.com

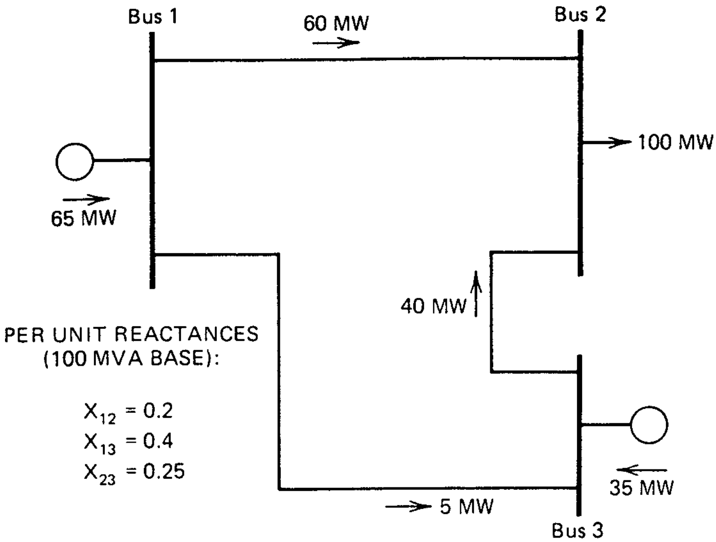


FIGURE 9.1 Three-bus system from Chapter 6, Figures 6.21 and 6.22.

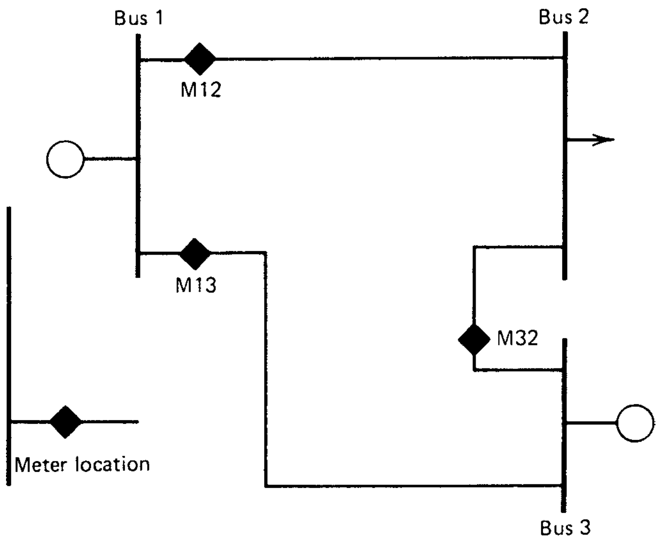


FIGURE 9.2 Meter placement.

Then, the flows on lines 1–3 and 3–2 can be set equal to these meter readings:

$$f_{13} = \frac{1}{x_{13}}(\theta_1 - \theta_3) = M_{13} = 0.05 \text{ pu}$$

$$f_{32} = \frac{1}{x_{23}}(\theta_3 - \theta_2) = M_{32} = 0.40 \text{ pu}$$

www.mepcafe.com

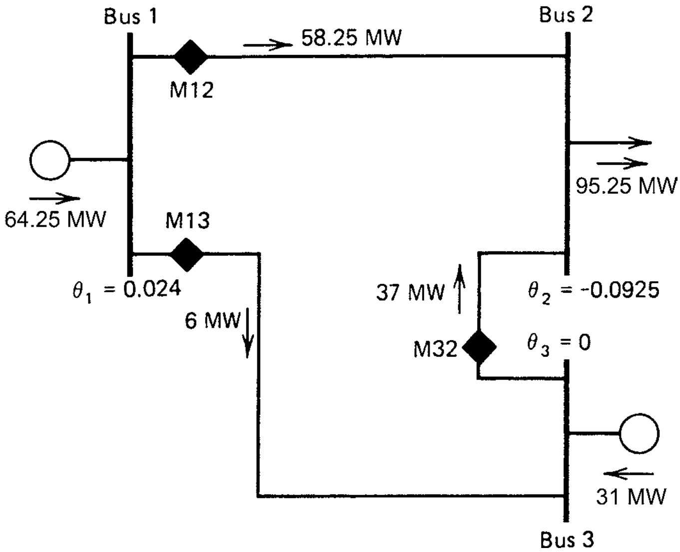


FIGURE 9.3 Flows resulting from use of meters M_{13} and M_{32}.

Since we know that $\theta_3 = 0$ rad, we can solve the f_{13} equation for θ_1 and the f_{32} equation for θ_2, resulting in

$$\theta_1 = 0.02 \text{ rad}$$
$$\theta_2 = -0.10 \text{ rad}$$

We will now investigate the case where all three meter readings have slight errors. Suppose the readings obtained are

$$M_{12} = 62 \text{ MW} = 0.62 \text{ pu}$$
$$M_{13} = 6 \text{ MW} = 0.06 \text{ pu}$$
$$M_{32} = 37 \text{ MW} = 0.37 \text{ pu}$$

If we use only the M_{13} and M_{32} readings, as before, we will calculate the phase angles as follows:

$$\theta_1 = 0.024 \text{ rad}$$
$$\theta_2 = -0.0925 \text{ rad}$$
$$\theta_3 = 0 \text{ rad (still assumed to equal zero)}$$

This results in the system flows as shown in Figure 9.3. Note that the predicted flows match at M_{13} and M_{32}, but the flow on line 1–2 does not match the reading of 62 MW from M_{12}. If we were to ignore the reading on M_{13} and use M_{12} and M_{32}, we could obtain the flows shown in Figure 9.4.

www.mepcafe.com

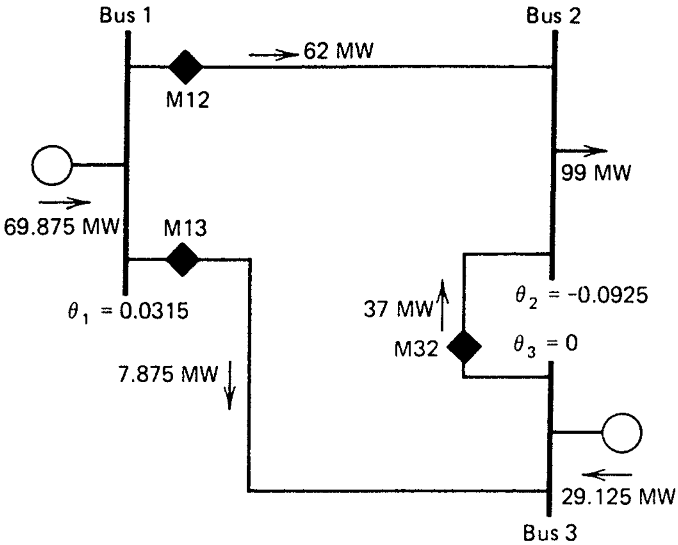


FIGURE 9.4 Flows resulting from use of meters M_{12} and M_{32}.

All we have accomplished is to match M_{12}, but at the expense of no longer matching M_{13}. What we need is a procedure that uses the information available from all three meters to produce the best estimate of the actual angles, line flows, and bus load and generations.

Before proceeding, let us discuss what we have been doing. Since the only thing we know about the power system comes to us from the measurements, we must use the measurements to estimate system conditions. Recall that in each instance the measurements were used to calculate the bus phase angles at bus 1 and 2. Once these phase angles were known, all unmeasured power flows, loads, and generations could be determined. We call θ_1 and θ_2 the *state variables* for the three-bus system, since knowing them allows all other quantities to be calculated. In general, the state variables for a power system consist of the bus voltage magnitude at all buses and the phase angles at all but one bus. The swing or reference-bus phase angle is usually assumed to be zero radians. Note that we could use real and imaginary components of bus voltage if desired. If we can use measurements to estimate the "states" (i.e., voltage magnitudes and phase angles) of the power system, then we can go on to calculate any power flows, generation, loads, and so forth that we desire. This presumes that the network configuration (i.e., breaker and disconnect switch statuses) is known and that the impedances in the network are also known. Automatic load tap-changing autotransformers or phase angle regulators are often included in a network, and their tap positions may be telemetered to the control as a measured quantity. Strictly speaking, the transformer taps and phase angle regulator positions should also be considered as states since they must be known in order to calculate the flows through the transformers and regulators.

www.mepcafe.com

To return to the three-bus DC power flow model, we have three meters providing us with a set of redundant readings with which to estimate the two states θ_1 and θ_2. We say that the readings are redundant since, as we saw earlier, only two readings are necessary to calculate θ_1 and θ_2; the other reading is always "extra." However, the "extra" reading does carry useful information and ought not to be discarded summarily.

This simple example serves to introduce the subject of *static-state estimation*, which is the art of estimating the exact system state given a set of imperfect measurements made on the power system. We will digress at this point to develop the theoretical background for static-state estimation. We will return to our three-bus system in Section 9.3.4.

9.3 MAXIMUM LIKELIHOOD WEIGHTED LEAST-SQUARES ESTIMATION

9.3.1 Introduction

Statistical estimation refers to a procedure where one uses samples to calculate the value of one or more unknown parameters in a system. Since the samples (or measurements) are inexact, the estimate obtained for the unknown parameter is also inexact. This leads to the problem of how to formulate a "best" estimate of the unknown parameters given the available measurements.

The development of the notions of state estimation may proceed along several lines, depending on the statistical criterion selected. Of the many criteria that have been examined and used in various applications, the following three are perhaps the most commonly encountered.

a. *The maximum likelihood criterion*, where the objective is to maximize the probability that the estimate of the state variable, $\hat{\mathbf{x}}$, is the true value of the state variable vector, $\mathbf{x}$ (i.e., maximize $P(\hat{\mathbf{x}}) = \mathbf{x}$)
b. *The weighted least-squares criterion*, where the objective is to minimize the sum of the squares of the weighted deviations of the estimated measurements, $\hat{\mathbf{z}}$, from the actual measurements, $\mathbf{z}$
c. *The minimum variance criterion*, where the object is to minimize the expected value of the sum of the squares of the deviations of the estimated components of the state variable vector from the corresponding components of the true state variable vector

When normally distributed, unbiased meter error distributions are assumed, each of these approaches results in identical estimators. This chapter will utilize the maximum likelihood approach because the method introduces the measurement error weighting matrix $[R]$ in a straightforward manner.

The maximum likelihood procedure asks the following question: "What is the probability (or likelihood) that I will get the measurements I have obtained?" This probability depends on the random error in the measuring device (transducer) as well as the unknown parameters to be estimated. Therefore, a reasonable procedure would be one

www.mepcafe.com

that simply chose the estimate as the value that maximizes this probability. As we will see shortly, the maximum likelihood estimator assumes that we know the probability density function (PDF) of the random errors in the measurement. Other estimation schemes could also be used. The "least-squares" estimator does not require that we know the PDF for the sample or measurement errors. However, if we assume that the PDF of sample or measurement error is a normal (Gaussian) distribution, we will end up with the same estimation formula. We will proceed to develop our estimation formula using the maximum likelihood criterion assuming normal distributions for measurement errors. The result will be a "least-squares" or more precisely a "weighted least-squares" estimation formula, even though we will develop the formulation using the maximum likelihood criteria. We will illustrate this method with a simple electrical circuit and show how the maximum likelihood estimate can be made.

First, we introduce the concept of *random measurement error*. Note that we have dropped the term "sample" since the concept of a measurement is much more appropriate to our discussion. The measurements are assumed to be in error: that is, the value obtained from the measurement device is close to the true value of the parameter being measured but differs by an unknown error. Mathematically, this can be modeled as follows.

Let z^{meas} be the value of a measurement as received from a measurement device. Let z^{true} be the true value of the quantity being measured. Finally, let η be the random measurement error. We can then represent our measured value as

$$z^{\text{meas}} = z^{\text{true}} + \eta \tag{9.1}$$

The random number, η, serves to model the uncertainty in the measurements. If the measurement error is unbiased, the PDF of η is usually chosen as a normal distribution with zero mean. Note that other measurement PDFs will also work in the maximum likelihood method as well. The PDF of η is

$$\text{PDF}(\eta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-\eta^2}{2\sigma^2}\right) \tag{9.2}$$

where σ is called the standard deviation and σ^2 is called the variance of the random number. PDF(η) describes the behavior of η. A plot of PDF(η) is shown in Figure 9.5.

Note that σ, the standard deviation, provides a way to model the seriousness of the random measurement error. If σ is large, the measurement is relatively inaccurate

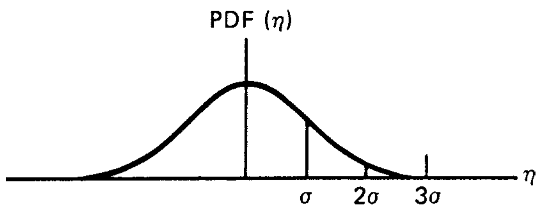


FIGURE 9.5 The normal distribution.

www.mepcafe.com

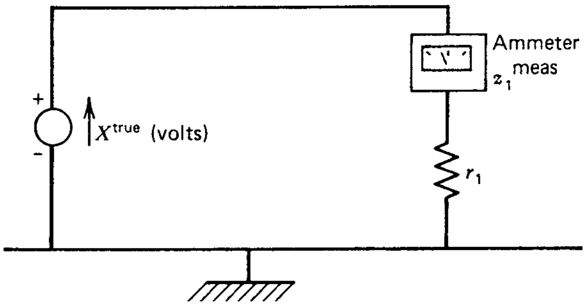


FIGURE 9.6 Simple DC circuit with current measurement.

(i.e., a poor-quality measurement device), whereas a small value of σ denotes a small error spread (i.e., a higher-quality measurement device). The normal distribution is commonly used for modeling measurement errors since it is the distribution that will result when many factors contribute to the overall error.

9.3.2 Maximum Likelihood Concepts

The principle of maximum likelihood estimation is illustrated by using a simple DC circuit example as shown in Figure 9.6. In this example, we wish to estimate the value of the voltage source, x^{true}, using an ammeter with an error having a known standard deviation. The ammeter gives a reading of z_1^{meas}, which is equal to the sum of z_1^{true} (the true current flowing in our circuit) and η_1 (the error present in the ammeter). Then we can write

$$z_1^{meas} = z_1^{true} + \eta_1 \tag{9.3}$$

Since the mean value of η_1 is 0, we then know that the mean value of z_1^{meas} is equal to z_1^{true}. This allows us to write a PDF for z_1^{meas} as

$$\text{PDF}(z_1^{meas}) = \frac{1}{\sigma_1\sqrt{2\pi}}\exp\left[\frac{-(z_1^{meas} - z_1^{true})^2}{2\sigma_1^2}\right] \tag{9.4}$$

where σ_1 is the standard deviation for the random error η_1. If we assume that the value of the resistance, r_1, in our circuit is known, then we can write

$$\text{PDF}(z_1^{meas}) = \frac{1}{\sigma_1\sqrt{2\pi}}\exp\left[\frac{-\left(z_1^{meas} - \frac{1}{r_1}x\right)^2}{2\sigma_1^2}\right] \tag{9.5}$$

Coming back to our definition of a maximum likelihood estimator, we now wish to find an estimate of x (called x^{est}) that maximizes the probability that the observed measurement z_1^{meas} would occur. Since we have the PDF of z_1^{meas}, we can write

www.mepcafe.com

$$\begin{aligned}\text{prob}\,(z_1^{\text{meas}}) &= \int_{z_1^{\text{meas}}}^{z_1^{\text{meas}}+\text{d}z_1^{\text{meas}}} \text{PDF}(z_1^{\text{meas}})\text{d}z_1^{\text{meas}} \text{ as } \text{d}z_1^{\text{meas}} \to 0 \\ &= \text{PDF}(z_1^{\text{meas}})\text{d}z_1^{\text{meas}}\end{aligned} \tag{9.6}$$

The maximum likelihood procedure then requires that we maximize the value of prob(z_1^{meas}), which is a function of x. That is,

$$\max_x \text{prob}(z_1^{\text{meas}}) = \max_x \text{PDF}(z_1^{\text{meas}})\text{d}z_1^{\text{meas}} \tag{9.7}$$

One convenient transformation that can be used at this point is to maximize the natural logarithm of PDF(z_1^{meas}) since maximizing the Ln of PDF(z_1^{meas}) will also maximize PDF(z_1^{meas}). Then we wish to find

$$\max_x \text{Ln}[\text{PDF}(z_1^{\text{meas}})]$$

or

$$\max_x \left[-\text{Ln}(\sigma_1\sqrt{2\pi}) - \frac{(z_1^{\text{meas}} - \frac{1}{r_1}x)^2}{2\sigma_1^2} \right]$$

Since the first term is constant, it can be ignored. We can maximize the function in brackets by minimizing the second term since it has a negative coefficient; that is,

$$\max_x \left[-\text{Ln}(\sigma_1\sqrt{2\pi}) - \frac{(z_1^{\text{meas}} - \frac{1}{r_1}x)^2}{2\sigma_1^2} \right]$$

is the same as

$$\min_x \left[\frac{(z_1^{\text{meas}} - \frac{1}{r_1}x)^2}{2\sigma_1^2} \right] \tag{9.8}$$

The value of x that minimizes the right-hand term is found by simply taking the first derivative and setting the result to 0:

$$\frac{\text{d}}{\text{d}x}\left[\frac{(z_1^{\text{meas}} - \frac{1}{r_1}x)^2}{2\sigma_1^2} \right] = \frac{-(z_1^{\text{meas}} - \frac{1}{r_1}x)}{r_1\sigma_1^2} = 0 \tag{9.9}$$

or

$$x^{\text{est}} = r_1 z_1^{\text{meas}}$$

To most readers this result was obvious from the beginning. All we have accomplished is to declare the maximum likelihood estimate of our voltage as simply the measured current times the known resistance. However, by adding a second measurement circuit, we have an entirely different situation in which the best estimate is not so obvious. Let us now add a second ammeter and resistance as shown in Figure 9.7.

www.mepcafe.com

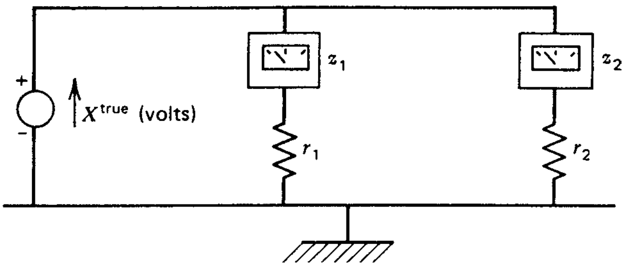


FIGURE 9.7 DC circuit with two current measurements.

Assume that both r_1 and r_2 are known. As before, model each meter reading as the sum of the true value and a random error:

$$\begin{aligned} z_1^{\text{meas}} &= z_1^{\text{true}} + \eta_1 \\ z_2^{\text{meas}} &= z_2^{\text{true}} + \eta_2 \end{aligned} \tag{9.10}$$

where the errors will be represented as independent zero mean, normally distributed random variables with PDFs:

$$\begin{aligned} \text{PDF}\,(\eta_1) &= \frac{1}{\sigma_1\sqrt{2\pi}}\exp\left(\frac{-(\eta_1)^2}{2\sigma_1^2}\right) \\ \text{PDF}(\eta_2) &= \frac{1}{\sigma_2\sqrt{2\pi}}\exp\left(\frac{-(\eta_2)^2}{2\sigma_2^2}\right) \end{aligned} \tag{9.11}$$

and as before we can write the PDFs of z_1^{meas} and z_2^{meas} as

$$\begin{aligned} \text{PDF}\,(z_1^{\text{meas}}) &= \frac{1}{\sigma_1\sqrt{2\pi}}\exp\left[\frac{-\left(z_1^{\text{meas}} - \frac{1}{r_1}x\right)^2}{2\sigma_1^2}\right] \\ \text{PDF}(z_2^{\text{meas}}) &= \frac{1}{\sigma_2\sqrt{2\pi}}\exp\left[\frac{-\left(z_2^{\text{meas}} - \frac{1}{r_2}x\right)^2}{2\sigma_2^2}\right] \end{aligned} \tag{9.12}$$

The likelihood function must be the probability of obtaining the measurements z_1^{meas} and z_2^{meas}. Since we are assuming that the random errors η_1 and η_2 are independent random variables, the probability of obtaining z_1^{meas} and z_2^{meas} is simply the product of the probability of obtaining z_1^{meas} and the probability of obtaining z_2^{meas}.

www.mepcafe.com

$$\begin{aligned}
\text{prob}\,(z_1^{\text{meas}} \text{ and } z_2^{\text{meas}}) &= \text{prob}(z_1^{\text{meas}}) \times (\text{prob}(z_2^{\text{meas}})\\
&= \text{PDF}\,(z_1^{\text{meas}})\text{PDF}(z_2^{\text{meas}})\text{d}z_1^{\text{meas}}\text{d}z_2^{\text{meas}}\\
&= \left[\frac{1}{\sigma_1\sqrt{2\pi}}\exp\left(\frac{-\left(z_1^{\text{meas}}-\frac{1}{r_1}x\right)^2}{2\sigma_1^2}\right)\right]\\
&\quad\times\left[\frac{1}{\sigma_2\sqrt{2\pi}}\exp\left(\frac{-\left(z_2^{\text{meas}}-\frac{1}{r_2}x\right)^2}{2\sigma_2^2}\right)\right]\text{d}z_1^{\text{meas}}\text{d}z_2^{\text{meas}}
\end{aligned} \tag{9.13}$$

To maximize the function, we will again take its natural logarithm:

$$\begin{aligned}
\max_x \text{prob}\,(z_1^{\text{meas}} \text{ and } z_2^{\text{meas}}) &= \max_x\left[-\text{Ln}\left(\sigma_1\sqrt{2\pi}\right)-\frac{\left(z_1^{\text{meas}}-\frac{1}{r_1}x\right)^2}{2\sigma_1^2}\right.\\
&\quad\left.-\text{Ln}\left(\sigma_2\sqrt{2\pi}\right)-\frac{\left(z_2^{\text{meas}}-\frac{1}{r_2}x\right)^2}{2ma_2^2}\right]\\
&= \min_x\left[\frac{\left(z_1^{\text{meas}}-\frac{1}{r_1}x\right)^2}{2\sigma_1^2}+\frac{\left(z_2^{\text{meas}}-\frac{1}{r_2}x\right)^2}{2\sigma_2^2}\right]
\end{aligned} \tag{9.14}$$

The minimum sought is found by

$$\frac{\text{d}}{\text{d}x}\left[\frac{\left(z_1^{\text{meas}}-\frac{1}{r_1}x\right)^2}{2\sigma_1^2}+\frac{\left(z_2^{\text{meas}}-\frac{1}{r_2}x\right)^2}{2\sigma_2^2}\right]=\frac{-\left(z_1^{\text{meas}}-\frac{1}{r_1}x\right)}{r_1\sigma_1^2}-\frac{\left(z_2^{\text{meas}}-\frac{1}{r_2}x\right)}{r_2\sigma_2^2}=0$$

giving

$$x^{\text{est}}=\frac{\left(\frac{z_1^{\text{meas}}}{r_1\sigma_1^2}+\frac{z_2^{\text{meas}}}{r_2\sigma_2^2}\right)}{\left(\frac{1}{r_1^2\sigma_1^2}+\frac{1}{r_2^2\sigma_2^2}\right)} \tag{9.15}$$

If one of the ammeters is of superior quality, its variance will be much smaller than that of the other meter. For example, if $\sigma_2^2 \ll \sigma_1^2$, then the equation for x^{est} becomes

$$x^{\text{est}} \simeq z_2^{\text{meas}} \times r_2$$

www.mepcafe.com

Thus, we see that the maximum likelihood method of estimating our unknown parameter gives us a way to weight the measurements properly according to their quality.

It should be obvious by now that we need not express our estimation problem as a maximum of the product of PDFs. Instead, we can observe a direct way of writing what is needed by looking at Equations 9.8 and 9.14. In these equations, we see that the maximum likelihood estimate of our unknown parameter is always expressed as that value of the parameter that gives the minimum of the sum of the squares of the difference between each measured value and the true value being measured (expressed as a function of our unknown parameter) with each squared difference divided or "weighted" by the variance of the meter error. Thus, if we are estimating a single parameter, x, using N_{m} measurements, we would write the expression

$$\min_{x} J(x) = \sum_{i=1}^{N_{\mathrm{m}}} \frac{\left[z_i^{\mathrm{meas}} - f_i(x)\right]^2}{\sigma_i^2} \tag{9.16}$$

where

f_i = function that is used to calculate the value being measured by the ith measurement

σ_i^2 = variance for the ith measurement

$J(x)$ = measurement residual

N_{m} = number of independent measurements

z_i^{meas} = ith measured quantity

Note that Equation 9.16 may be expressed in per unit or in physical units such as MW, MVAR, or kV.

If we were to try to estimate N_{s} unknown parameters using N_{m} measurements, we would write

$$\min_{\{x_1, x_2, \ldots, x_{N_{\mathrm{s}}}\}} J(x_1, x_2, \ldots, x_{N_{\mathrm{s}}}) = \sum_{i=1}^{N_{\mathrm{m}}} \frac{\left[z_i - f_i(x_1, x_2, \ldots, x_{N_{\mathrm{s}}})\right]^2}{\sigma_i^2} \tag{9.17}$$

The estimation calculation shown in Equations 9.16 and 9.17 is known as a *weighted least-squares* estimator, which, as we have shown earlier, is equivalent to a maximum likelihood estimator if the measurement errors are modeled as random numbers having a normal distribution.

9.3.3 Matrix Formulation

If the functions $f_i(x_1, x_2, \ldots, x_{N_{\mathrm{s}}})$ are linear functions, Equation 9.17 has a closed-form solution. Let us write the function $f_i(x_1, x_2, \ldots, x_{N_{\mathrm{s}}})$ as

$$f_i(x_1, x_2, \ldots, x_{N_{\mathrm{s}}}) = f_i(\mathbf{x}) = h_{i1}x_1 + h_{i2}x_2 + \cdots + h_{iN_{\mathrm{s}}}x_{N_{\mathrm{s}}} \tag{9.18}$$

www.mepcafe.com

Then, if we place all the f_i functions in a vector, we may write

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_{N_\mathrm{m}}(\mathbf{x}) \end{bmatrix} = [H]\mathbf{x} \tag{9.19}$$

where

$[H]$ = an N_m by N_s matrix containing the coefficients of the linear functions $f_i(\mathbf{x})$
N_m = number of measurements
N_s = number of unknown parameters being estimated

Placing the measurements in a vector

$$\mathbf{z}^\mathrm{meas} = \begin{bmatrix} z_1^\mathrm{meas} \\ z_2^\mathrm{meas} \\ \vdots \\ z_{N_\mathrm{m}}^\mathrm{meas} \end{bmatrix} \tag{9.20}$$

we may then write Equation 9.17 in a very compact form:

$$\min_{\mathbf{x}} J(\mathbf{x}) = \left[\mathbf{z}^\mathrm{meas} - \mathbf{f}(\mathbf{x})\right]^\mathrm{T} \left[R^{-1}\right]\left[\mathbf{z}^\mathrm{meas} - \mathbf{f}(\mathbf{x})\right] \tag{9.21}$$

where

$$[R] = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_{N_\mathrm{m}}^2 \end{bmatrix}$$

$[R]$ is called the *covariance matrix of measurement errors*. To obtain the general expression for the minimum in Equation 9.21, expand the expression and substitute $[H]\mathbf{x}$ for $\mathbf{f}(\mathbf{x})$ from Equation 9.19.

$$\begin{aligned} \min_{\mathbf{x}} J(\mathbf{x}) = \{ & \mathbf{z}^{\mathrm{meas}^\mathrm{T}} [R^{-1}]\mathbf{z}^\mathrm{meas} - \mathbf{x}^\mathrm{T} [H]^\mathrm{T} [R^{-1}]\mathbf{z}^\mathrm{meas} \\ & - \mathbf{z}^{\mathrm{meas}^\mathrm{T}} \left[R^{-1}\right][H]\mathbf{x} + \mathbf{x}^\mathrm{T} [H]^\mathrm{T} [R^{-1}][H]\mathbf{x} \} \end{aligned} \tag{9.22}$$

Similar to the procedures of Chapter 3, the minimum of $J(\mathbf{x})$ is found when $\partial J(\mathbf{x})/\partial x_i = 0$, for $i = 1, \ldots, N_\mathrm{s}$; this is identical to stating that the gradient of $J(\mathbf{x})$, $\nabla J(\mathbf{x})$ is exactly 0.

The gradient of $J(\mathbf{x})$ is (see the appendix to this chapter)

www.mepcafe.com

TABLE 9.1 Estimation Formulas

Case	Description	Solution	Comment
$N_s < N_m$		$\mathbf{x}^{est} = [[H]^T[R^{-1}][H]]^{-1} \times \{H]^T[R^{-1}]\mathbf{z}^{meas}$	$\mathbf{x}^{est}$ is the maximum likelihood estimate of x given the measurements $\mathbf{z}^{meas}$
$N_s = N_m$	Completely determined	$\mathbf{x}^{est} = [H]^{-1}\mathbf{z}^{meas}$	$\mathbf{x}^{est}$ fits the measured quantities to the measurements $\mathbf{z}^{meas}$ exactly
$N_s > N_m$	Underdetermined	$\mathbf{x}^{est} = [H]^T[[H][H]^T]^{-1}\mathbf{z}^{meas}$	$\mathbf{x}^{est}$ is the vector of minimum norm that fits the measured quantities to the measurements exactly. (The norm of a vector is equal to the sum of the squares of its components)

$$\nabla J(\mathbf{x}) = -2[H]^T[R^{-1}]\mathbf{z}^{meas} + 2[H]^T[R^{-1}][H]\mathbf{x}$$

Then $\nabla J(\mathbf{x}) = 0$ gives

$$\mathbf{x}^{est} = \left[[H]^T[R^{-1}][H]\right]^{-1}[H]^T[R^{-1}]\mathbf{z}^{meas} \quad (9.23)$$

Note that Equation 9.23 holds when $N_s < N_m$, that is, when the number of parameters being estimated is less than the number of measurements being made.

When $N_s = N_m$, our estimation problem reduces to

$$\mathbf{x}^{est} = [H]^{-1}\mathbf{z}^{meas} \quad (9.24)$$

There is also a closed-form solution to the problem when $N_s > N_m$, although in this case we are not estimating $\mathbf{x}$ to maximize a likelihood function since $N_s > N_m$ usually implies that many different values for $\mathbf{x}^{est}$ can be found that cause $f_i(\mathbf{x}^{est})$ to equal z_i^{meas} for all $i = 1, \ldots, N_m$ exactly. Rather, the objective is to find $\mathbf{x}^{est}$ such that the sum of the squares of x_i^{est} is minimized. That is,

$$\min_{\mathbf{x}} \sum_{i=1}^{N_s} x_i^2 = \mathbf{x}^T\mathbf{x} \quad (9.25)$$

subject to the condition that $\mathbf{z}^{meas} = [H]\mathbf{x}$. The closed-form solution for this case is

$$\mathbf{x}^{est} = [H]^T\left[[H][H]^T\right]^{-1}\mathbf{z}^{meas} \quad (9.26)$$

In power system state estimation, underdetermined problems (i.e., where $N_s > N_m$) are not solved, as shown in Equation 9.26. Rather, "pseudo-measurements" are added to the measurement set to give a completely determined or overdetermined problem. We will discuss pseudo-measurements in Section 9.6.3. Table 9.1 summarizes the results for this section.

www.mepcafe.com

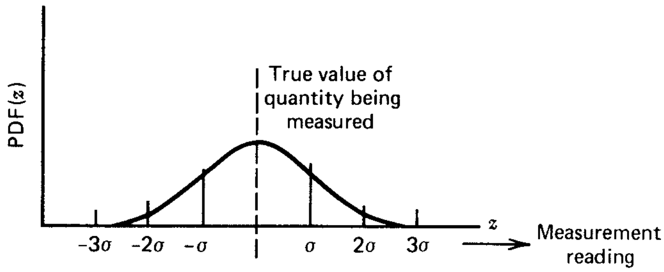


FIGURE 9.8 Normal distribution of meter errors.

9.3.4 An Example of Weighted Least-Squares State Estimation

We now return to our three-bus example. Recall from Figure 9.2 that we have three measurements to determine θ_1 and θ_2, the phase angles at buses 1 and 2. From the development in the preceding section, we know that the states θ_1 and θ_2 can be estimated by minimizing a residual $J(\theta_1,\theta_2)$ where $J(\theta_1,\theta_2)$ is the sum of the squares of individual measurement residuals divided by the variance for each measurement.

To start, we will assume that all three meters have the following characteristics.

Meter full-scale value: 100 MW
Meter accuracy: ±3 MW

This is interpreted to mean that the meters will give a reading within ±3 MW of the true value being measured for approximately 99% of the time. Mathematically, we say that the errors are distributed according to a normal PDF with a standard deviation, σ, as shown in Figure 9.8.

Notice that the probability of an error decreases as the error magnitude increases. By integrating the PDF between -3σ and $+3\sigma$, we come up with a value of approximately 0.99. We will assume that the meter's accuracy (in our case ±3 MW) is being stated as equal to the 3σ points on the PDF. Then ±3 MW corresponds to a metering standard deviation of $\sigma = 1$ MW $= 0.01$ pu.

The formula developed in the last section for the weighted least-squares estimate is given in Equation 9.23, which is repeated here:

$$\mathbf{x}^{\text{est}} = \left[[H]^{\text{T}}\left[R^{-1}\right][H]\right]^{-1}[H]^{\text{T}}\left[R^{-1}\right]\mathbf{z}^{\text{meas}}$$

where

$\mathbf{x}^{\text{est}}$ = vector of estimated state variables

$[H]$ = measurement function coefficient matrixs

$[R]$ = measurement covariance matrix

$\mathbf{z}^{\text{meas}}$ = vector containing the measured values themselves

For the three-bus problem, we have

www.mepcafe.com

$$\mathbf{x}^{\text{est}} = \begin{bmatrix} \theta_1^{\text{est}} \\ \theta_2^{\text{est}} \end{bmatrix} \tag{9.27}$$

To derive the [*H*] matrix, we need to write the measurements as a function of the state variables θ_1 and θ_2. These functions are written in per unit as

$$\begin{aligned} M_{12} &= f_{12} = \frac{1}{0.2}(\theta_1 - \theta_2) = 5\theta_1 - 5\theta_2 \\ M_{13} &= f_{13} = \frac{1}{0.4}(\theta_1 - \theta_2) = 2.5\theta_1 \\ M_{32} &= f_{32} = \frac{1}{0.25}(\theta_3 - \theta_2) = -4\theta_2 \end{aligned} \tag{9.28}$$

The reference-bus phase angle, θ_3, is still assumed to be 0. Then

$$[H] = \begin{bmatrix} 5 & -5 \\ 2.5 & 0 \\ 0 & -4 \end{bmatrix}$$

The covariance matrix for the measurements, [*R*], is

$$[R] = \begin{bmatrix} \sigma_{M12}^2 & & \\ & \sigma_{M13}^2 & \\ & & \sigma_{M32}^2 \end{bmatrix} = \begin{bmatrix} 0.0001 & & \\ & 0.0001 & \\ & & 0.0001 \end{bmatrix}$$

Note that since the coefficients of [*H*] are in per unit, we must also write [*R*] and $\mathbf{z}^{\text{meas}}$ in per unit.

Our least-squares "best" estimate of θ_1 and θ_2 is then calculated as

$$\begin{bmatrix} \theta_1^{\text{est}} \\ \theta_2^{\text{est}} \end{bmatrix} = \left[\begin{bmatrix} 5 & 2.5 & 0 \\ -5 & 0 & -4 \end{bmatrix} \begin{bmatrix} 0.0001 & & \\ & 0.0001 & \\ & & 0.0001 \end{bmatrix}^{-1} \begin{bmatrix} 5 & -5 \\ 2.5 & 0 \\ 0 & -4 \end{bmatrix} \right]^{-1}$$

$$\times \begin{bmatrix} 5 & 2.5 & 0 \\ -5 & 0 & -4 \end{bmatrix} \begin{bmatrix} 0.0001 & & \\ & 0.0001 & \\ & & 0.0001 \end{bmatrix}^{-1} \begin{bmatrix} 0.62 \\ 0.06 \\ 0.37 \end{bmatrix}$$

$$= \begin{bmatrix} 312500 & -250000 \\ -250000 & 410000 \end{bmatrix}^{-1} \begin{bmatrix} 32500 \\ -45800 \end{bmatrix}$$

$$= \begin{bmatrix} 0.028571 \\ -0.094286 \end{bmatrix}$$

www.mepcafe.com

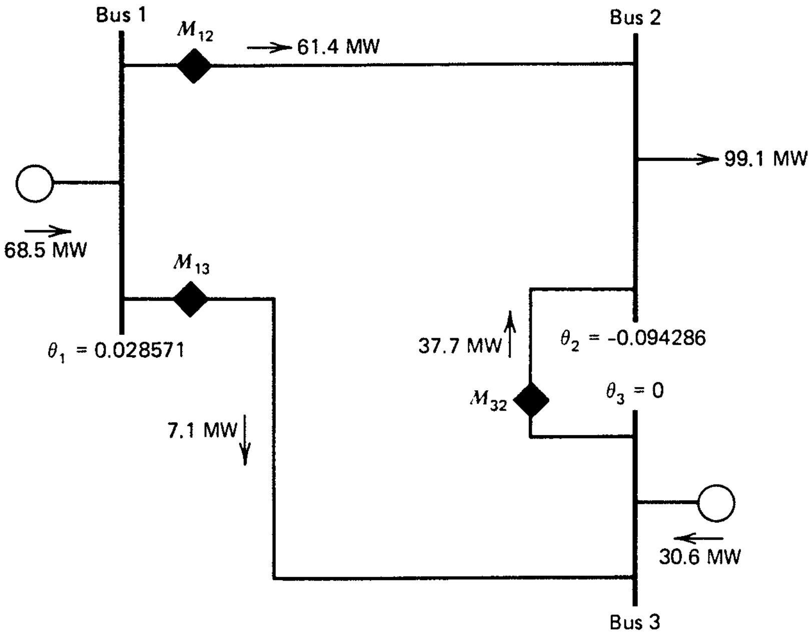


FIGURE 9.9 Three-bus example with best estimates of θ_1 and θ_2.

where

$$z^{\text{meas}} = \begin{bmatrix} 0.62 \\ 0.06 \\ 0.37 \end{bmatrix}$$

From the estimated phase angles, we can calculate the power flowing in each transmission line and the net generation or load at each bus. The results are shown in Figure 9.9.

If we calculate the value of $J(\theta_1,\theta_2)$, the residual, we get

$$\begin{aligned} J\,(\theta_1,\theta_2) &= \frac{[z_{12}-f_{12}(\theta_1,\theta_2)]^2}{\sigma_{12}^2}+\frac{[z_{13}-f_{13}(\theta_1,\theta_2)]^2}{\sigma_{13}^2}+\frac{[z_{32}-f_{32}(\theta_1,\theta_2)]^2}{\sigma_{32}^2} \\ &= \frac{[0.62-(5\theta_1-5\theta_2)]}{0.0001}+\frac{[0.06-(2.5\theta_1)]^2}{0.0001}+\frac{[0.37+(4\theta_2)]^2}{0.0001} \\ &= 2.14 \end{aligned} \tag{9.29}$$

Suppose the meter on the M_{13} transmission line was superior in quality to those on M_{12} and M_{32}. How will this affect the estimate of the states? Intuitively, we can reason that any measurement reading we get from M_{13} will be much closer to the true power flowing on line 1–3 than can be expected when comparing M_{12} and M_{32}

www.mepcafe.com

to the flows on lines 1–2 and 3–2, respectively. Therefore, we would expect the results from the state estimator to reflect this if we set up the measurement data to reflect the fact that M_{13} is a superior measurement. To show this, we use the following metering data.

$$\begin{aligned} \textbf{Meters } M_{12} \textbf{ and } M_{32}: &100 \text{ MW full scale} \\ &\pm 3 \text{ MW accuracy} \\ &(\sigma = 1 \text{ MW} = 0.01 \text{ pu}) \end{aligned}$$

$$\begin{aligned} \textbf{Meter } M_{13}: &100 \text{ MW full scale} \\ &\pm 0.3 \text{ MW accuracy} \\ &(\sigma = 0.1 \text{ MW} = 0.001 \text{ pu}) \end{aligned}$$

The covariance matrix to be used in the least-squares formula now becomes

$$[R] = \begin{bmatrix} \sigma_{M12}^2 & & \\ & \sigma_{M13}^2 & \\ & & \sigma_{M32}^2 \end{bmatrix} = \begin{bmatrix} 1\times10^{-4} & & \\ & 1\times10^{-6} & \\ & & 1\times10^{-4} \end{bmatrix}$$

We now solve with the new $[R]$ matrix:

$$\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \left[\begin{bmatrix} 5 & 2.5 & 0 \\ -5 & 0 & -4 \end{bmatrix} \begin{bmatrix} 1\times10^{-4} & & \\ & 1\times10^{-6} & \\ & & 1\times10^{-4} \end{bmatrix}^{-1} \begin{bmatrix} 5 & -5 \\ 2.5 & 0 \\ 0 & -4 \end{bmatrix} \right]^{-1}$$

$$= \begin{bmatrix} 5 & 2.5 & 0 \\ -5 & 0 & -4 \end{bmatrix} \begin{bmatrix} 1\times10^{-4} & & \\ & 1\times10^{-6} & \\ & & 1\times10^{-4} \end{bmatrix}^{-1} \begin{bmatrix} 5 & -5 \\ 2.5 & 0 \\ 0 & -4 \end{bmatrix}$$

$$= \begin{bmatrix} 6.5\times10^{6} & -2.5\times10^{5} \\ -2.5\times10^{5} & 4.1\times10^{5} \end{bmatrix}^{-1} \begin{bmatrix} 1.81\times10^{5} \\ -0.458\times10^{5} \end{bmatrix}$$

$$= \begin{bmatrix} 0.024115 \\ -0.097003 \end{bmatrix}$$

From these estimated phase angles, we obtain the network conditions shown in Figure 9.10. Compare the estimated flow on line 1–3, as just calculated, to the estimated flow calculated on line 1–3 in the previous least-squares estimate. Setting σ_{M13} to 0.1 MW has brought the estimated flow on line 1–3 much closer to the meter reading of 6.0 MW. Also, note that the estimates of flow on lines 1–2 and 3–2 are now further from the M_{12} and M_{32} meter readings, respectively, which is what we should have expected.

www.mepcafe.com

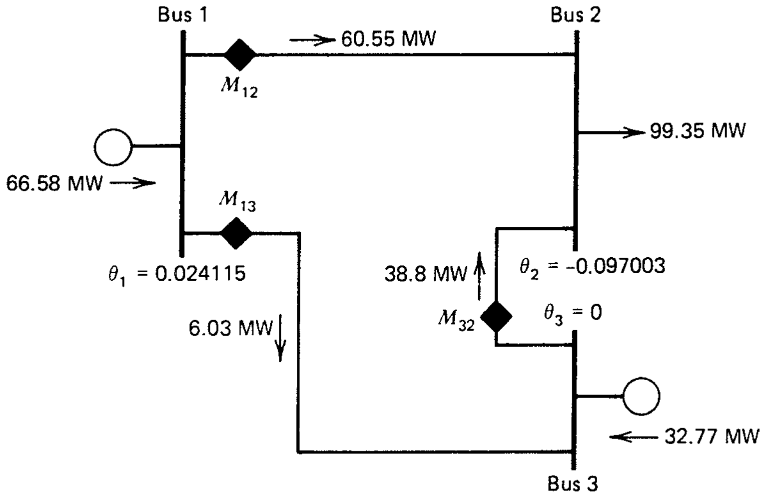


FIGURE 9.10 Three-bus example with better meter at M_{13}.

9.4 STATE ESTIMATION OF AN AC NETWORK

9.4.1 Development of Method

We have demonstrated how the maximum likelihood estimation scheme developed in Section 9.3.2 led to a least-squares calculation for measurements from a linear system. In the least-squares calculation, we are trying to minimize the sum of measurement residuals:

$$\min_{\mathbf{x}} J(\mathbf{x}) = \sum_{i=1}^{N_m} \frac{\left[z_i - f_i(\mathbf{x})\right]^2}{\sigma_i^2} \tag{9.30}$$

In the case of a linear system, the $f_i(\mathbf{x})$ functions are themselves linear and we solve for the minimum of $J(\mathbf{x})$ directly. In an AC network, the measured quantities are MW, MVAR, MVA, amperes, transformer tap position, and voltage magnitude. The state variables are the voltage magnitude at each bus, the phase angles at all but the reference bus, and the transformer taps. The equation for power entering a bus is given in Equation 4.21 and is clearly not a linear function of the voltage magnitude and phase angle at each bus. Therefore, the $f_i(\mathbf{x})$ functions will be nonlinear functions, except for a voltage magnitude measurement where $f_i(\mathbf{x})$ is simply unity times the particular x_i that corresponds to the voltage magnitude being measured. For MW and MVAR measurements on a transmission line from bus i to bus j, we would have the following terms in $J(\mathbf{x})$:

$$\frac{\left\{\mathrm{MW}_{ij}^{\mathrm{meas}} - \left[|E_i|^2\left(G_{ij}\right) - |E_i||E_j|\left(\cos\left(\theta_i - \theta_j\right)G_{ij} + \sin\left(\theta_i - \theta_j\right)B_{ij}\right)\right]\right\}^2}{\sigma_{\mathrm{MW}_{ij}}^2} \tag{9.31}$$

www.mepcafe.com

and

$$\frac{\{\text{MVAR}_{ij}^{\text{meas}} - [-|E_i|^2 (B_{\text{cap}_{ij}} + B_{ij}) - |E_i||E_j|(\sin(\theta_i - \theta_j)G_{ij} - \cos(\theta_i - \theta_j)B_{ij})]\}^2}{\sigma^2_{\text{MVAR}ij}} \tag{9.32}$$

A voltage magnitude measurement would result in the following term in $J(\mathbf{x})$:

$$\frac{\left(|E_i|^{\text{meas}} - |E_i|\right)^2}{\sigma^2_{|E_i|}} \tag{9.33}$$

Similar functions can be derived for MVA or ampere measurements.

If we do not have a linear relationship between the states ($|E|$ values and θ values) and the power flows on a network, we will have to resort to an iterative technique to minimize $J(\mathbf{x})$. A commonly used technique for power system state estimation is to calculate the gradient of $J(\mathbf{x})$ and then force it to 0 using Newton's method, as was done with the Newton load flow in Chapter 6. We will review how to use Newton's method on multidimensional problems before proceeding to the minimization of $J(\mathbf{x})$.

Given the functions $g_i(\mathbf{x})$, $i = 1, \ldots, n$, we wish to find $\mathbf{x}$ that gives $g_i(\mathbf{x}) = g_i^{\text{des}}$, for $i = 1, \ldots, n$. If we arrange the g_i functions in a vector, we can write

$$\mathbf{g}^{\text{des}} - \mathbf{g}(\mathbf{x}) = 0 \tag{9.34}$$

By perturbing $\mathbf{x}$ we can write

$$\mathbf{g}^{\text{des}} - \mathbf{g}(\mathbf{x} + \Delta\mathbf{x}) = \mathbf{g}^{\text{des}} - \mathbf{g}(\mathbf{x}) - [\mathbf{g}'(\mathbf{x})]\Delta\mathbf{x} = 0 \tag{9.35}$$

where we have expanded $\mathbf{g}(\mathbf{x} + \Delta\mathbf{x})$ in a Taylor's series about $\mathbf{x}$ and ignored all higher-order terms. The $[\mathbf{g}'(\mathbf{x})]$ term is the Jacobian matrix of first derivatives of $\mathbf{g}(\mathbf{x})$. Then

$$\Delta\mathbf{x} = \left[\mathbf{g}'(\mathbf{x})\right]^{-1}\left[\mathbf{g}^{\text{des}} - \mathbf{g}(\mathbf{x})\right] \tag{9.36}$$

Note that if $\mathbf{g}^{\text{des}}$ is identically 0, we have

$$\Delta\mathbf{x} = \left[\mathbf{g}'(\mathbf{x})\right]^{-1}\left[-\mathbf{g}(\mathbf{x})\right] \tag{9.37}$$

To solve for $\mathbf{g}^{\text{des}}$, we must solve for $\Delta\mathbf{x}$ using Equation 9.36, then calculate $\mathbf{x}^{\text{new}} = \mathbf{x} + \Delta\mathbf{x}$ and reapply Equation 9.36 until either $\Delta\mathbf{x}$ gets very small or $\mathbf{g}(\mathbf{x})$ comes close to $\mathbf{g}^{\text{des}}$.

Now let us return to the state estimation problem as given in Equation 9.30:

$$\min_{\mathbf{x}} J(\mathbf{x}) = \sum_{i=1}^{N_{\text{m}}} \frac{\left[z_i - f_i(\mathbf{x})\right]^2}{\sigma_i^2}$$

We first form the gradient of $J(\mathbf{x})$ as

www.mepcafe.com

$$\nabla_x \mathbf{J}(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial J(\mathbf{x})}{\partial x_1} \\ \dfrac{\partial J(\mathbf{x})}{\partial x_2} \\ \vdots \end{bmatrix}$$
$$= -2 \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_3}{\partial x_1} & \cdots \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_1} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} \begin{bmatrix} \dfrac{1}{\sigma_1^2} & & \\ & \dfrac{1}{\sigma_1^2} & \\ & & \ddots \end{bmatrix} \begin{bmatrix} z_1 - f_1(\mathbf{x}) \\ z_2 - f_2(\mathbf{x}) \\ \vdots \end{bmatrix} \tag{9.38}$$

If we put the $f_i(\mathbf{x})$ functions in a vector form $\mathbf{f}(\mathbf{x})$ and calculate the Jacobian of $\mathbf{f}(\mathbf{x})$, we would obtain

$$\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \dfrac{\partial f_1}{\partial x_3} & \cdots \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \dfrac{\partial f_2}{\partial x_3} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} \tag{9.39}$$

We will call this matrix $[H]$. Then,

$$[H] = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \dfrac{\partial f_1}{\partial x_3} & \cdots \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \dfrac{\partial f_2}{\partial x_3} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} \tag{9.40}$$

And its transaction is

$$[H]^{\mathrm{T}} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \dfrac{\partial f_1}{\partial x_3} & \cdots \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \dfrac{\partial f_2}{\partial x_3} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} \tag{9.41}$$

Further, we write

$$\begin{bmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \ddots \end{bmatrix} = [R] \tag{9.42}$$

www.mepcafe.com

Equation 9.38 can be written

$$\nabla_x \mathbf{J}(\mathbf{x}) = \left\{ -2[H]^{\mathrm{T}}[R]^{-1} \begin{bmatrix} z_1 - f_1(\mathbf{x}) \\ z_2 - f_2(\mathbf{x}) \\ \vdots \end{bmatrix} \right\} \tag{9.43}$$

To make $\nabla_x \mathbf{J}(\mathbf{x})$ equal 0, we will apply Newton's method as in Equation 9.37; then

$$\Delta \mathbf{x} = \left[\frac{\partial \nabla_x J(\mathbf{x})}{\partial \mathbf{x}} \right]^{-1} \left[-\nabla_x J(\mathbf{x}) \right] \tag{9.44}$$

The Jacobian of $\nabla_x \mathbf{J}(\mathbf{x})$ is calculated by treating $[H]$ as a constant matrix:

$$\begin{aligned} \frac{\partial \nabla_x J(\mathbf{x})}{\partial \mathbf{x}} &= \frac{\partial}{\partial \mathbf{x}} \left\{ -2[H]^{\mathrm{T}}[R]^{-1} \begin{bmatrix} z_1 - f_1(\mathbf{x}) \\ z_2 - f_2(\mathbf{x}) \\ \vdots \end{bmatrix} \right\} \\ &= -2[H]^{\mathrm{T}}[R]^{-1}[-H] \\ &= 2[H]^{\mathrm{T}}[R]^{-1}[H] \end{aligned} \tag{9.45}$$

Then

$$\begin{aligned} \Delta \mathbf{x} &= \frac{1}{2}\left[[H]^{\mathrm{T}}[R]^{-1}[H] \right]^{-1} \left\{ 2[H]^{\mathrm{T}}[R]^{-1} \begin{bmatrix} z_1 - f_1(\mathbf{x}) \\ \vdots \end{bmatrix} \right\} \\ &= \left[[H]^{\mathrm{T}}[R]^{-1}[H] \right]^{-1} [R]^{-1} \begin{bmatrix} z_1 - f_1(\mathbf{x}) \\ z_2 - f_2(\mathbf{x}) \\ \vdots \end{bmatrix} \end{aligned} \tag{9.46}$$

Equation 9.46 is obviously a close parallel to Equation 9.23. To solve the AC state estimation problem, apply Equation 9.46 iteratively as shown in Figure 9.11. Note that this is similar to the iterative process used in the Newton power flow solution.

9.4.2 Typical Results of State Estimation on an AC Network

Figure 9.12 shows our familiar six-bus system with $P + jQ$ measurements on each end of each transmission line and at each load and generator. Bus voltage magnitude is also measured at each system bus. Phase angle measurements are not included in this example.

To demonstrate the use of state estimation on these measurements, the base-case conditions shown in Example 6A were used together with a random number-generating algorithm to produce measurements with random errors. The measurements were obtained by adding the random errors to the base-case flows, loads, generations, and

www.mepcafe.com

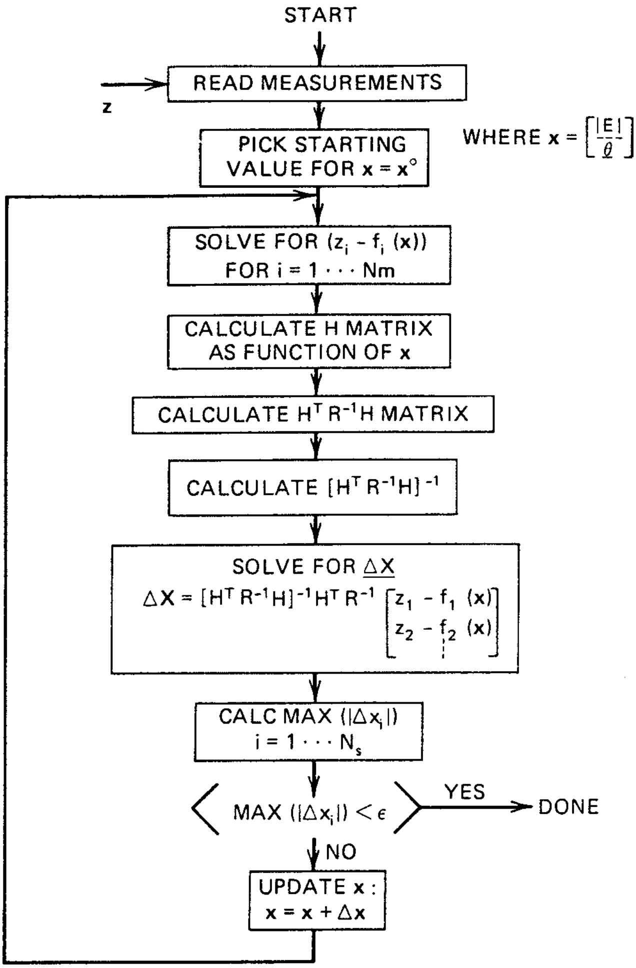


FIGURE 9.11 State estimation solution algorithm.

bus voltage magnitudes. The errors were generated so as to be representative of values drawn from a set of numbers having a normal PDF with zero mean, and variance as specified for each measurement type. The measurement variances used were

$$\begin{aligned} P + jQ \textbf{ measurements}: \sigma &= 3 \text{ MW for the } P \text{ measurement} \\ \sigma &= 3 \text{ MVAR for the } Q \text{ measurement} \\ \textbf{Voltage measurement}: \sigma &= 3.83 \text{ kV} \end{aligned}$$

The base conditions and the measurements are shown in Table 9.2. The state estimation algorithm shown in Figure 9.11 was run to obtain estimates for the bus voltage magnitudes and phase angles given the measurements shown in Table 9.2.

www.mepcafe.com

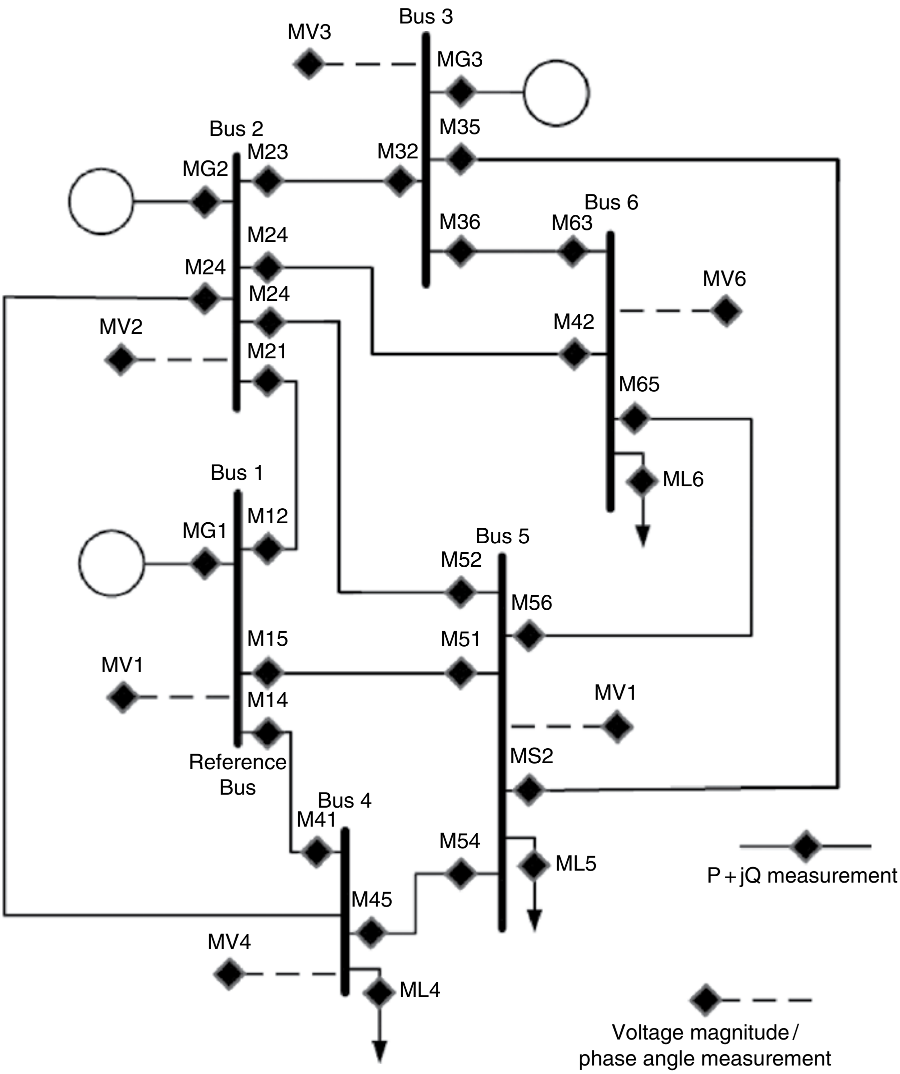


FIGURE 9.12 Six-bus system with measurements.

The procedure took four iterations with $\mathbf{x}^0$ initially being set to 1.0 pu and 0 rad for the voltage magnitude and phase angle at each bus, respectively. At the beginning of each iteration, the sum of the measurement residuals, $J(\mathbf{x})$ (see Equation 9.30), is calculated and displayed. The iterative steps for the six-bus system used here produced the results given in Table 9.3 with the number of active measurements, degree of freedom, and bad data threshold t_j (more on these in Section 9.6.2).

www.mepcafe.com

TABLE 9.2 Base-Case Conditions

Measurement Name	Status	Base Case Value kV	MW	MVAR	Measured Value kV	MW	MVAR
		Bus 1					
MV1	1	246.1			246.1		
MA1	1	0.0			0.0		
M12	1		123.6	-35.6		123.2	-36.0
M14	1		179.9	23.9		180.4	23.8
M15	1		105.0	3.4		105.4	3.8
		Bus 2					
MV2	1	241.5			241.5		
MA2	0	0.0			0.0		
MG2	1		50.0	75.7		51.0	75.5
M21	1		-109.3	59.7		-110.6	60.5
M23	1		11.4	-5.4		11.7	-5.9
M24	1		93.1	17.2		94.1	16.4
M25	1		15.6	5.3		15.3	2.1
M26	1		39.2	-1.1		38.2	-1.3
		Bus 3					
MV3	1	241.5			241.5		
MA3	0	0.0			0.0		
MG3	1		50.0	24.2		49.8	24.9
M32	1		-11.3	-0.9		-12.2	-2.1
M35	1		8.9	7.5		8.4	7.3
M36	1		52.4	17.6		51.0	17.6
		Bus 4					
MV4	1	228.0			228.0		
MA4	0	0.0			0.0		
ML4	1		-100.0	-15.0		-99.4	-16.1
M41	1		-165.5	29.6		-164.5	28.7
M42	1		-89.0	-11.1		-89.2	-12.3
M45	1		-11.0	-3.9		-11.7	-5.5
		Bus 5					
MV5	1	233.3			233.3		
MA5	0	0.0			0.0		
ML5	1		-100.0	-15.0		-101.5	-16.1
M51	1		-97.3	19.1		-96.9	18.3
M52	1		-15.3	-8.7		-16.4	-10.2
M53	1		-8.7	-12.4		-10.3	-12.9
M54	1		11.2	-3.7		12.7	-1.6
M56	1		10.1	-9.2		10.2	-10.2
		Bus 6					
MV6	1	235.4			235.4		
MA6	0	0.0			0.0		
ML6	1		-100.0	-15.0		-101.4	-14.9
M62	1		-38.2	-1.5		-38.6	0.4
M63	1		-51.8	-16.9		-51.0	-16.5
M65	1		-9.9	3.4		-8.7	2.8

The value of $J(\mathbf{x})$ at the end of the iterative procedure would be 0 if all measurements were without error or if there were no redundancy in the measurements. When there are redundant measurements with errors, the value of $J(\mathbf{x})$ will not normally go to 0. Its value represents a measure of the overall fit of the estimated values to the measurement values. The value of $J(\mathbf{x})$ can, in fact, be used to detect the presence of bad measurements.

www.mepcafe.com

TABLE 9.3 Iterative Results of State Estimation Solution

Estimator Iteration Summary

Iteration	Residual J	Number Bad Data Measurements	Active Degrees of Freedom	Threshold t_J	Largest Normalized Residual at	Bad Measurement
1	26569.356					
2	314.437					
3	56.580					
4	56.312					
		62	51	63.440		

The estimated values from the state estimator are shown in Table 9.4, together with the base-case values and the measured values. Notice that, in general, the estimated values do a good job of calculating the true (base-case) conditions from which the measurements were made. For example, measurement M_{23} shows a P flow of 8.6 MW whereas the true flow is 2.9 MW and the estimator predicts a flow of 3.0 MW.

The example shown here started from a base case or "true" state that was shown in Table 9.2. In actual practice, we only have the measurements and the resulting estimate of the state; we never know the "true" state exactly and can only compare measurements with estimates. In the presentations to follow, however, we will leave the base-case or "true" conditions in our illustrations to aid the reader.

The results in Table 9.4 show one of the advantages of using a state estimation algorithm in that, even with measurement errors, the estimation algorithm calculates quantities that are the "best" possible estimates of the true bus voltages and generator, load, and transmission line MW and MVAR values.

There are, however, other advantages to using a state estimation algorithm. First is the ability of the state estimator to detect and identify bad measurements, and second is the ability to estimate quantities that are not measured and telemetered. These are introduced later in the chapter.

9.5 STATE ESTIMATION BY ORTHOGONAL DECOMPOSITION

One problem with the standard least-squares method presented earlier in the chapter is the numerical difficulties encountered with some special state estimation problems. One of these comes about when we wish to drive a state estimator solution to match its measurement almost exactly. This is the case when we have a circuit such as shown in Figure 9.13. All of the actual flows and injections are shown in Figure 9.13 along with the values assumed for the measurements.

In this sample system, the measurement of power at bus 1 will be assumed to be 0 MW. If the value of 0 is dictated by the fact that the bus has no load or generation attached to it, then we know this value of 0 MW with certainty and the concept of an error in its "measured" value is meaningless. Nonetheless, we proceed by setting up

www.mepcafe.com

TABLE 9.4 State Estimation Solution

Measurement Name	Status	Base Case Value kV	MW	MVAR	Measured Value kV	MW	MVAR	Estimated Value kV	MW	MVAR
		Bus 1								
MV1	1	246.1			245.7			247.0		
MA1	1	0.0			0.0			0.0		
MG1	1		228.6	-32.2		225.4	-33.0		226.8	-31.7
M12	1		123.6	-35.6		126.7	-28.3		122.2	-35.2
M14	1		179.9	23.9		181.2	20.8		178.6	22.0
M15	1		105.0	3.4		107.4	0.8		104.6	3.5
		Bus 2								
MV2	1	241.5			246.6			242.3		
MA2	0	-14.5			0.0			-10.5		
MG2	1		50.0	75.7		49.7	71.3		49.7	72.3
M21	1		-109.3	59.7		-106.4	58.8		-108.3	58.5
M23	1		11.4	-5.4		10.4	-7.1		11.1	-5.1
M24	1		93.1	17.2		92.5	12.1		91.8	14.3
M25	1		15.6	5.3		17.7	6.1		16.1	5.2
M26	1		39.2	-1.1		36.2	-2.4		39.1	-0.6
		Bus 3								
MV3	1	241.5			238.5			242.1		
MA3	0	-16.0			0.0			-12.0		
MG3	1		50.0	24.2		50.6	21.7		51.2	23.1
M32	1		-11.3	-0.9		-14.4	-3.6		-11.0	-1.3
M35	1		8.9	7.5		8.1	11.3		9.6	6.9
M36	1		52.4	17.6		52.7	15.6		52.6	17.5
		Bus 4								
MV4	1	228.0			230.7			229.6		
MA4	0	-19.1			0.0			-15.1		
ML4	1		-100.0	-15.0		-100.3	-14.0		-98.2	-12.0
M41	1		-165.5	29.6		-159.9	32.4		-164.5	30.0
M42	1		-89.0	-11.1		-87.2	-11.5		-87.9	-8.6
M45	1		-11.0	-3.9		-8.5	-4.8		-10.4	-3.4
		Bus 5								
MV5	1	233.3			234.4			234.1		
MA5	0	-16.6			0.0			-12.7		
ML5	1		-100.0	-15.0		-102.7	-15.9		-102.1	-15.2
M51	1		-97.3	19.1		-96.3	17.4		-96.9	18.5
M52	1		-15.3	-8.7		-13.8	-13.2		-15.8	-8.6
M53	1		-8.7	-12.4		-7.3	-8.9		-9.4	-11.9
M54	1		11.2	-3.7		9.6	-4.6		10.6	-4.3
M56	1		10.1	-9.2		6.8	-10.7		9.5	-8.9
		Bus 6								
MV6	1	235.4			234.6			236.0		
MA6	0	-18.6			0.0			-14.6		
ML6	1		-100.0	-15.0		-98.9	-20.5		-99.5	-15.9
M62	1		-38.2	-1.5		-37.9	1.9		-38.1	-2.0
M63	1		-51.8	-16.9		-56.3	-16.4		-52.1	-16.9
M65	1		-9.9	3.4		-10.5	3.6		-9.4	3.0

the standard state estimator equations and specifying the value of the measurement σ for M_1 as $\sigma_{M1} = 10^{-2}$. This results in the following solution when using the state estimator equations as shown in Equation 9.23:

P_{flow} estimate on line $1-2 = 30.76$ MW

P_{flow} estimate on line $3-2 = 72.52$

Injection estimate on bus $1 = 0.82$

www.mepcafe.com

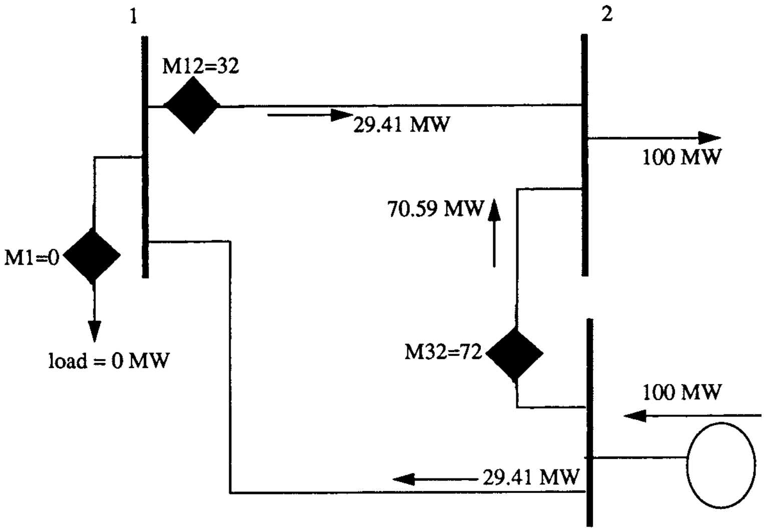


FIGURE 9.13 Zero injection system example.

The estimator has not forced the bus injection to be exactly 0; instead, it reads 0.82 MW. This may not seem like such a big error. However, if there are many such buses (say 100) and they all have errors of this magnitude, then the estimator will have a large amount of load allocated to the buses that are known to be zero.

At first, the solution to this dilemma may seem to be simply forcing the σ value to a very small number for the zero injection buses and rerun the estimator. The problem with this is as follows. Suppose we had changed the zero injection σ to $\sigma_{M1} = 10^{-10}$. Hopefully, this would force the estimator to make the zero injection so dominant that it would result in the correct zero value coming out of the estimator calculation. In this case, the $[H^T R^{-1} H]$ matrix used in the standard least-squares method would look like this for the sample system:

$$[H] = \begin{bmatrix} 5.0 & -5.0 \\ 0 & -4.0 \\ 7.5 & -5.0 \end{bmatrix}$$

$$[R] = \begin{bmatrix} 10^{-4} & & \\ & 10^{-4} & \\ & & 10^{-20} \end{bmatrix}$$

$$[H^T R^{-1} H] = \begin{bmatrix} 56.25\times10^{20} & -37.5\times10^{20} \\ -37.5\times10^{20} & 25.0\times10^{20} \end{bmatrix}$$

www.mepcafe.com

Unfortunately, this matrix is very nearly singular. The reason is that the terms in the matrix are dominated by those terms that are multiplied by the 10^{20} terms from the inverse of the R matrix, and the other terms are so small by comparison that they are lost from the computer (unless one is using an extraordinarily long word length or extra double precision). When the aforementioned is presented to a standard matrix inversion routine or run into a Gaussian elimination solution routine, an error message results and garbage comes out of the estimator.

The solution to this dilemma is to use another algorithm for the least-squares solution. This algorithm is called the orthogonal decomposition algorithm and works as follows.

9.5.1 The Orthogonal Decomposition Algorithm

This algorithm goes under several different names in texts on linear algebra. It is often called the QR algorithm or the Gram–Schmidt decomposition. The idea is to take the state estimation least-squares equation, Equation 9.23, and eliminate the R^{-1} matrix as follows: let

$$[R^{-1}] = R^{-1/2}R^{-1/2} \tag{9.47}$$

where

$$\left[R^{-1/2}\right] = \begin{bmatrix} \dfrac{1}{\sigma_{m1}} & & \\ & \dfrac{1}{\sigma_{m2}} & \\ & & \dfrac{1}{\sigma_{m3}} \end{bmatrix} \tag{9.48}$$

Then

$$[H^{\mathrm{T}}R^{-1}H]^{-1} = [H^{\mathrm{T}}R^{-1/2}R^{-1/2}H]^{-1} = [H'^{\mathrm{T}}H']^{-1} \tag{9.49}$$

with

$$[H'] = [R^{-1/2}][H] \tag{9.50}$$

Finally, Equation 9.23 becomes

$$\mathbf{x}^{\mathrm{est}} = [H'^{\mathrm{T}}H']^{-1}[H'^{\mathrm{T}}]\mathbf{z}'^{\mathrm{meas}} \tag{9.51}$$

where

$$\mathbf{z}'^{\mathrm{meas}} = [R^{-1/2}]\mathbf{z}^{\mathrm{meas}} \tag{9.52}$$

The idea of the orthogonal decomposition algorithm is to find a matrix $[Q]$ such that:

$$[H'] = [Q][U] \tag{9.53}$$

(Note that in most linear algebra textbooks, this factorization would be written as $[H'] = [Q]$ $[R]$; however, we shall use $[Q][U]$ so as not to confuse the identity of the $[R]$ matrix.)

www.mepcafe.com

The matrix [Q] has special properties. It is called an orthogonal matrix so that

$$[Q^{\mathrm{T}}][Q] = [I] \tag{9.54}$$

where [I] is the identity matrix, which is to say that the transpose of [Q] is its inverse. The matrix [U] is now upper triangular in structure, although since the [H] matrix may not be square, [U] will not be square either. Thus,

$$[H'] = \begin{bmatrix} h'_{11} & h'_{12} \\ h'_{21} & h'_{22} \\ h'_{31} & h'_{32} \end{bmatrix} = [Q][U] = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \\ 0 & 0 \end{bmatrix} \tag{9.55}$$

Now, if we substitute [Q][U] for [H'] in the state estimation equation,

$$\mathbf{x}^{\mathrm{est}} = \left[U^{\mathrm{T}}Q^{\mathrm{T}}QU\right]^{-1}\left[U^{\mathrm{T}}\right]\left[Q^{\mathrm{T}}\right]\mathbf{z}' \tag{9.56}$$

or

$$\mathbf{x}^{\mathrm{est}} = \left[U^{\mathrm{T}}U\right]^{-1}U^{\mathrm{T}}\hat{\mathbf{z}} \tag{9.57}$$

since

$$\left[Q^{\mathrm{T}}Q\right] = I$$

and

$$\hat{\mathbf{z}} = \left[Q^{\mathrm{T}}\right]\mathbf{z}' \tag{9.58}$$

Then, by rearranging we get

$$\left[U^{\mathrm{T}}U\right]\mathbf{x}^{\mathrm{est}} = \left[U^{\mathrm{T}}\right]\hat{\mathbf{z}} \tag{9.59}$$

and we can eliminate U^{T} from both sides so that we are left with

$$\left[U\right]\mathbf{x}^{\mathrm{est}} = \hat{\mathbf{z}} \tag{9.60}$$

or

$$\begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1^{\mathrm{est}} \\ x_2^{\mathrm{est}} \end{bmatrix} = \begin{bmatrix} \hat{z}_1 \\ \hat{z}_2 \\ \hat{z}_3 \end{bmatrix} \tag{9.61}$$

This can be solved directly since U is upper triangular:

$$x_2^{\mathrm{est}} = \frac{\hat{z}_2}{u_{22}} \tag{9.62}$$

www.mepcafe.com

and

$$x_1^{\text{est}} = \frac{1}{u_{11}}\left(\hat{z}_1 - u_{12}x_2^{\text{est}}\right) \tag{9.63}$$

The Q matrix and the U matrix are obtained, for our simple two-state–three-measurement problem here, using the Givens rotation method.

For the Givens rotation method, we start out to define the steps necessary to solve:

$$\left[Q^{\text{T}}\right]\left[H'\right] = \left[U\right] \tag{9.64}$$

where $[H]$ is a 2×2 matrix:

$$\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$$

and $[U]$ is

$$\begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix}$$

The $[Q]$ matrix must be orthogonal, and when it is multiplied times $[H]$, it eliminates the h_{21} term. The terms in the $[Q]$ matrix are simply

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix}$$

where

$$c = \frac{h_{11}}{\sqrt{h_{11}^2 + h_{21}^2}} \tag{9.65}$$

and

$$s = \frac{h_{21}}{\sqrt{h_{11}^2 + h_{21}^2}} \tag{9.66}$$

The reader can easily verify that the $[Q]$ matrix is indeed orthogonal and that

$$\begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix} = \begin{bmatrix} 1 & \left(ch_{12} + sh_{22}\right) \\ 0 & \left(-sh_{12} + ch_{22}\right) \end{bmatrix} \tag{9.67}$$

When we solve the $3\times2[H]$ matrix in our three-measurement–two-state sample problem, we apply the Givens rotation three times to eliminate h_{21}, h_{31}, and h_{32}. That is, we need to solve

www.mepcafe.com

$$[Q^{\mathrm{T}}]\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \\ h_{31} & h_{32} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \\ 0 & 0 \end{bmatrix} \tag{9.68}$$

We will carry this out in three distinct steps, where each step can be represented as a Givens rotation. The result is that we represent $[Q^{\mathrm{T}}]$ as the product of three matrices:

$$[Q^{\mathrm{T}}] = [N_3][N_2][N_1] \tag{9.69}$$

These matrices are numbered as shown to indicate the order of application. In the case of the $3 \times 2[H]$ matrix,

$$[N_1] = \begin{bmatrix} c & s & 0 \\ -s & c & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{9.70}$$

where c and s are defined exactly as before. Next, $[N_2]$ must be calculated so as to eliminate the 31 term that results from $[N_1][H]$. The actual procedure loads $[H]$ into $[U]$ and then determines each $[N]$ based on the current contents of $[U]$. The $[N_2]$ matrix will have terms like

$$[N_2] = \begin{bmatrix} c' & 0 & s' \\ 0 & 1 & 0 \\ -s' & 0 & c' \end{bmatrix} \tag{9.71}$$

where c' and s' are determined from $[N_1][H]$. Similarly for $[N_3]$,

$$[N_3] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c'' & s'' \\ 0 & -s'' & c'' \end{bmatrix} \tag{9.72}$$

For our zero injection example, we start with the $[H]$ and $[R]$ matrices as shown before:

$$[H] = \begin{bmatrix} 5.0 & -5.0 \\ 0 & -4.0 \\ 7.5 & -5.0 \end{bmatrix}$$

and

$$[R] = \begin{bmatrix} 10^{-4} & & \\ & 10^{-4} & \\ & & 10^{-20} \end{bmatrix}$$

Then, the $[H']$ matrix is

www.mepcafe.com

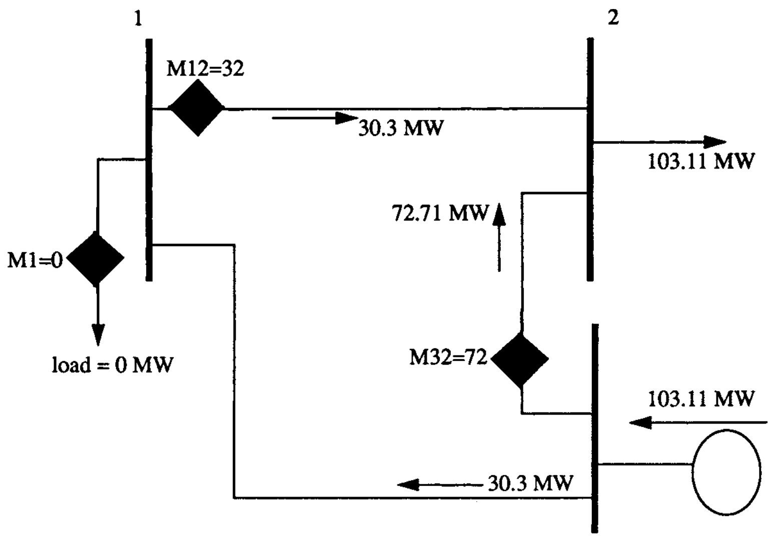


FIGURE 9.14 State estimate resulting from orthogonal decomposition algorithm.

$$[H'] = \begin{bmatrix} 5.0\times10^{2} & -5.0\times10^{2} \\ 0 & -4.0\times10^{2} \\ 7.5\times10^{10} & -5.0\times10^{10} \end{bmatrix}$$

and the measurement vector is

$$\hat{\mathbf{z}} = \begin{bmatrix} 32 \\ 72 \\ 0 \end{bmatrix}$$

The resulting state estimate is shown in Figure 9.14. Note particularly that the injection at bus 1 is estimated to be 0, as we desired.

The orthogonal decomposition algorithm has the advantage that measurement weights can be adjusted to extreme values as demonstrated by the numerical example shown. As such, its robust numerical advantages have made it a useful algorithm for power system state estimators.

9.6 AN INTRODUCTION TO ADVANCED TOPICS IN STATE ESTIMATION

9.6.1 Sources of Error in State Estimation

The quality of the measurements used in a state estimator vary, and the use of a state estimator gives engineers and power system operators a means to cope with errors. A partial list of error sources follows:

www.mepcafe.com

- **Modeling Errors**. These have to do with the fact that the model of the transmission system may be incorrect to begin with. Impedance data for transmission lines and transformers can be calculated wrong or may have drifted due to environmental factors such as ground moisture. Last of all, modeling errors can be present because breaker and disconnect switch status indications can be wrong due to human error or due to telemetry error through the SCADA system.
- **Data Errors**. By these are meant the errors that come with incorrect data specifications that indicate which current transformer is connected to which transducer, which are the positive and the negative sides of measurement devices, which breaker status contacts are wired to which terminal strip connectors in the substation building, etc.
- **Transducer Errors**. These are errors in the values transmitted from current transformers, potential transformers, and transducers that produce signals proportional to P and Q flows. These may be simple bias errors that give a constant offset to the measurement value or they may be random errors giving an error that is distributed around a mean value representing the true measurement value.
- **Sampling Errors**. These are due to the fact that SCADA systems do not sample all measurements simultaneously, nor are all measurements sampled at the same frequency. Thus, we are using measurements of time-varying quantities that were taken at different times and can be in error simply due to the dynamic nature (time-changing value) of the quantity being measured.

9.6.2 Detection and Identification of Bad Measurements

The ability to detect and identify bad measurements is extremely valuable to a power system's operations department. Transducers may have been wired incorrectly, or the transducer itself may be malfunctioning so that it simply no longer gives accurate readings. The statistical theory required to understand and analyze bad measurement detection and identification is straightforward but lengthy. We are going to open the door to the subject in this chapter. The serious student who wishes to pursue this subject should start with the references listed in Chapter 1. For the rest, we present results of these theories and indicate application areas.

To detect the presence of bad measurements, we will rely on the intuitive notion that for a given configuration, the residual, $J(\mathbf{x})$, calculated after the state estimator algorithm converges, will be smallest if there are no bad measurements. When $J(\mathbf{x})$ is small, a vector $\mathbf{x}$ (i.e., voltages and phase angles) has been found that causes all calculated flows, loads, generations, and so forth to closely match all the measurements. Generally, the presence of a bad measurement value will cause the converged value of $J(\mathbf{x})$ to be larger than expected with $\mathbf{x} = \mathbf{x}^{\text{est}}$.

What magnitude of $J(\mathbf{x})$ indicates the presence of bad measurements?

The measurement errors are random numbers so that the value of $J(\mathbf{x})$ is also a random number. If we assume that all the errors are described by their respective

www.mepcafe.com

normal PDFs, then we can show that $J(\mathbf{x})$ has a PDF known as a *chi-squared distribution*, which is written as $\chi^2(K)$. The parameter K is called the degrees of freedom of the chi-squared distribution. This parameter is defined as follows:

$$K = N_{\mathrm{m}} - N_{\mathrm{s}}$$

where

N_{m} = numner of measurements (note that a $p + jQ$ measurement counts as two measurements)

N_{s} = number of states = $(2n–1)$

n = number of buses in the network,

Note that we do not count the phase angle of the reference (slack) bus as a state since it is known and held constant from the start, usually as zero radians.

It can be shown that when $\mathbf{x}=\mathbf{x}^{\mathrm{est}}$, the mean value of $J(\mathbf{x})$ equals K and the standard deviation, $\sigma_{J(\mathbf{x})}$, equals $\sqrt{2K}$.

When one or more measurements are bad, their errors are frequently much larger than the assumed $\pm3\sigma$ error bound for the measurement. However, even under normal circumstances (i.e., all errors within $\pm3\sigma$), $J(\mathbf{x})$ can get to be large—although the chance of this happening is small. If we simply set up a threshold for $J(\mathbf{x})$, which we will call t_J, we could declare that bad measurements are present when $J(\mathbf{x})>t_J$. This threshold test might be wrong in one of two ways. If we set t_J to a small value, we would get many "false alarms." That is, the test would indicate the presence of bad measurements when, in fact, there were none. If we set t_J to be a large value, the test would often indicate that "all is well" when, in fact, bad measurements were present. This can be put on a formal basis by writing the following equation:

$$\begin{array}{c}\text{prob}\,(J(x) > t_J \mid J(\mathbf{x}) \text{ is a chi-squared}) = \alpha \\ \text{with } K \text{ degrees of freedom}\end{array} \tag{9.73}$$

This equation says that the probability that $J(\mathbf{x})$ is greater than t_J is equal to α, given that the probability density for $J(\mathbf{x})$ is chi-squared with K degrees of freedom.

This type of testing procedure is formally known as *hypothesis testing*, and the parameter α is called the *significance level* of the test. By choosing a value for the significance level α, we automatically know what threshold t_J to use in our test. When using a t_J derived in this manner, the probability of a "false alarm" is equal to α. By setting α to a small number, for example $\alpha=0.01$, we would say that false alarms would occur in only 1% of the tests made. A plot of the probability function in Equation 9.73 is shown in Figure 9.15.

In Table 9.3, we saw that the minimum value for $J(\mathbf{x})$ was 56.312. Looking at Figure 9.12 and counting all $P+jQ$ measurements as two measurements, we see that N_{m} is equal to 62 (measurements are 6 V_{mag}+6 P_{inj}+6 Q_{inj}+22 P_{flows}+22 Q_{flows}). Therefore, the degrees of freedom for the chi-square distribution of $J(\mathbf{x})$ in our six-bus sample system are

www.mepcafe.com

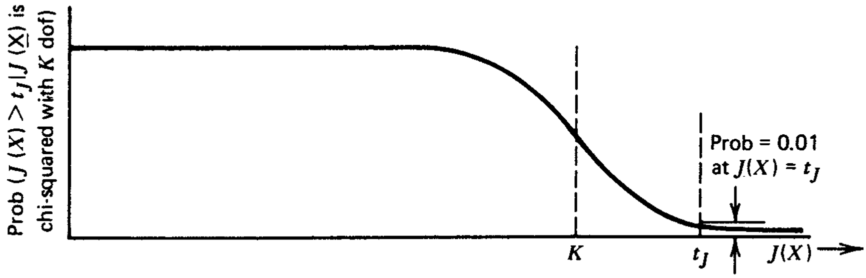


FIGURE 9.15 Threshold test probability function.

$$K = N_m - N_s = N_m - (2n-1) = 51$$

where

$$N_m = 62 \quad \text{and} \quad n = 6$$

If we set our significance level for this test to 0.01 (i.e., $\alpha = 0.01$ in Equation 9.73), we get a t_J of 63.44.[1] Therefore, with a $J(\mathbf{x}) = 56.312$, it seems reasonable to assume that there are no "bad" measurements present.

Now let us assume that one of the measurements is truly bad. To simulate this situation, the state estimation algorithm was rerun with the M_{12} measurement reversed. Instead of $P = 123.6$ and $Q = -35.6$, it was set to $P = -123$ and $Q = +35$ and then noise added.

The value of $J(\mathbf{x})$ for each iteration for this case is given in Table 9.5. The presence of bad data does not prevent the estimator from converging, but it will increase the value of the residual, $J(\mathbf{x})$.

TABLE 9.5 Iterative Results with Bad Measurement

Estimator Iteration Summary

Iteration	Residual J	Number Active Measurements Freedom	Degrees of Bad Data Threshold Largest t_J	Bad Normalized Measurement Residual at
1	26641.718			
2	6706.949			
3	6492.852			
4	6456.845			
5	6455.592			
	62	51	63.440	

[1] Standard tables of $\chi^2(K)$ usually only go up to $K = 30$. For $K > 30$, a very close approximation to $\chi^2(K)$ using the normal distribution can be used. The student should consult any standard reference on probability and statistics to see how this is done.

www.mepcafe.com

The calculated flows and voltages for this situation are shown in Table 9.6. Note that the number of degrees of freedom is still 51 but $J(\mathbf{x})$ is now 6455.592 at the end of our calculation.

Since t_J is 63.44, we would immediately expect bad measurements at our 0.01 significance level. If we had not known ahead of running the estimation algorithm that a bad measurement was present, we would certainly have had good reason to suspect its presence when so large a $J(\mathbf{x})$ resulted.

TABLE 9.6 State Estimation Solution with Measurement M12 Reversed

Measurement Name	Status	Base Case Value kV	MW	MVAR	Measured Value kV	MW	MVAR	Estimated Value kV	MW	MVAR
		Bus 1								
MV1	1	246.1			245.9			247.6		
MA1	1	0.0			0.0			0.0		
MG1	1		228.6	-32.2		226.4	-30.6		179.9	-28.4
M12	1		123.6	-35.6		-125.2	38.7		92.4	-30.1
M14	1		179.9	23.9		184.5	23.8		148.5	20.0
M15	1		105.0	3.4		106.8	3.2		87.5	1.7
		Bus 2								
MV2	1	241.5			243.6			244.0		
MA2	0	-14.5			0.0			-8.5		
MG2	1		50.0	75.7		49.2	77.5		77.3	61.3
M21	1		-109.3	59.7		-108.1	57.5		-84.4	41.6
M23	1		11.4	-5.4		13.2	-5.6		9.4	-3.0
M24	1		93.1	17.2		93.8	14.2		95.1	17.4
M25	1		15.6	5.3		21.0	2.1		19.9	4.1
M26	1		39.2	-1.1		37.1	-2.9		37.3	1.1
		Bus 3								
MV3	1	241.5			241.9			242.9		
MA3	0	-16.0			0.0			-9.7		
MG3	1		50.0	24.2		51.8	17.6		57.3	15.2
M32	1		-11.3	-0.9		-14.6	-2.8		-9.3	-3.5
M35	1		8.9	7.5		9.8	6.4		14.4	3.1
M36	1		52.4	17.6		52.7	19.3		52.2	15.6
		Bus 4								
MV4	1	228.0			234.3			230.4		
MA4	0	-19.1			0.0			-13.1		
ML4	1		-100.0	-15.0		-104.0	-19.3		-99.8	-16.6
M41	1		-165.5	29.6		-160.6	28.3		-138.8	14.5
M42	1		-89.0	-11.1		-86.1	-13.0		-91.0	-11.2
M45	1		-11.0	-3.9		-8.7	-5.9		-8.9	-5.4
		Bus 5								
MV5	1	233.3			234.4			235.9		
MA5	0	-16.6			0.0			-11.3		
ML5	1		-100.0	-15.0		-98.8	-10.6		-102.1	-12.9
M51	1		-97.3	19.1		-103.3	16.1		-82.2	11.5
M52	1		-15.3	-8.7		-14.7	-13.3		-19.5	-7.3
M53	1		-8.7	-12.4		-6.3	-10.0		-14.2	-7.9
M54	1		11.2	-3.7		9.4	-3.1		9.0	-2.5
M56	1		10.1	-9.2		9.2	-9.6		4.7	-6.6
		Bus 6								
MV6	1	235.4			241.6			237.2		
MA6	0	-18.6			0.0			-12.2		
ML6	1		-100.0	-15.0		-101.0	-12.6		-92.7	-18.8
M62	1		-38.2	-1.5		-37.0	1.3		-36.4	-4.1
M63	1		-51.8	-16.9		-50.6	-19.9		-51.7	-15.1
M65	1		-9.9	3.4		-8.2	6.6		-4.7	0.4

www.mepcafe.com

So far, we can say that by looking at $J(\mathbf{x})$, we can detect the presence of bad measurements. But if bad measurements are present, how can one tell which measurements are bad? Without going into the statistical theory, we give the following explanation of how this is accomplished.

Suppose we are interested in the measurement of megawatt flow on a particular line. Call this measured value z_i. In Figure 9.16a, we have a plot of the normal PDF of z_i. Since we assume that the error in z_i is normally distributed with zero mean value, the PDF is centered on the true value of z_i. Since the errors on all the measurements are assumed normal, we will assume that the estimate $\mathbf{x}^{\text{est}}$ is approximately normally distributed and that any quantity that is a function of $\mathbf{x}^{\text{est}}$ is also an approximately normally distributed quantity. In Figure 9.16b, we show the PDF for the calculated megawatt flow, f_i, which is a function of the estimated state, $\mathbf{x}^{\text{est}}$. We have drawn the density function of f_i as having a smaller deviation from its mean than the measurement z_i to indicate that, due to redundancy in measurements, the estimate is more accurate.

The difference between the estimate, f_i, and the measurement, z_i, is called the *measurement residual* and is designated y_i. The PDF for y_i is also normal and is shown in Figure 9.16c as having a zero mean and a standard deviation of σ_{y_i}. If we divide the difference between the estimate f_i and the measurement z_i by σ_{y_i}, we obtain what is called a *normalized measurement residual*. The normalized measurement residual is designated y_i^{norm} and is shown in Figure 9.16d along with its PDF, which is normal and has a standard deviation of unity. If the absolute value of y_i^{norm} is greater than 3, we have good reason to suspect that z_i is a bad measurement value. The usual procedure in identifying bad measurements is to calculate all f_i values for the N_{m} measurements once $\mathbf{x}^{\text{est}}$ is available from the state estimator. Using the z_i values that were used in the estimator and the f_i values, a measurement residual y_i can be calculated for each measurement. Also, using information from the state estimator, we can calculate σ_{y_i} (see references in Chapter 1 for details of this calculation). Using y_i and σ_{y_i}, we can calculate a normalized residual for each measurement. Measurements having the largest absolute normalized residual are labeled as prime suspects. These prime suspects are removed from the state estimator calculation one at a time, starting with the measurement having the largest normalized residual. After a measurement has been removed, the state estimation calculation (see Figure 9.11) is rerun. This results in a different $\mathbf{x}^{\text{est}}$ and therefore a different $J(\mathbf{x})$. The chi-squared PDF for $J(\mathbf{x})$ will have to be recalculated, assuming that we use the same significance level for our test. If the new $J(\mathbf{x})$ is now less than the new value for t_J, we can say that the measurement that was removed has been identified as bad. If, however, the new $J(\mathbf{x})$ is greater than the new t_J, we must proceed to recalculate $f_i(\mathbf{x}^{\text{est}})$, σ_{y_i}, and then y_i^{norm} for each of the remaining measurements. The measurement with the largest absolute y_i^{norm} is then again removed and the entire procedure repeated successively until $J(\mathbf{x})$ is less than t_J. This is a problem that the identification process may encounter, wherein several measurements may need to be removed to eliminate one "bad" measurement. That is, the identification procedure often cannot pinpoint a single bad measurement but instead identifies a group of

www.mepcafe.com

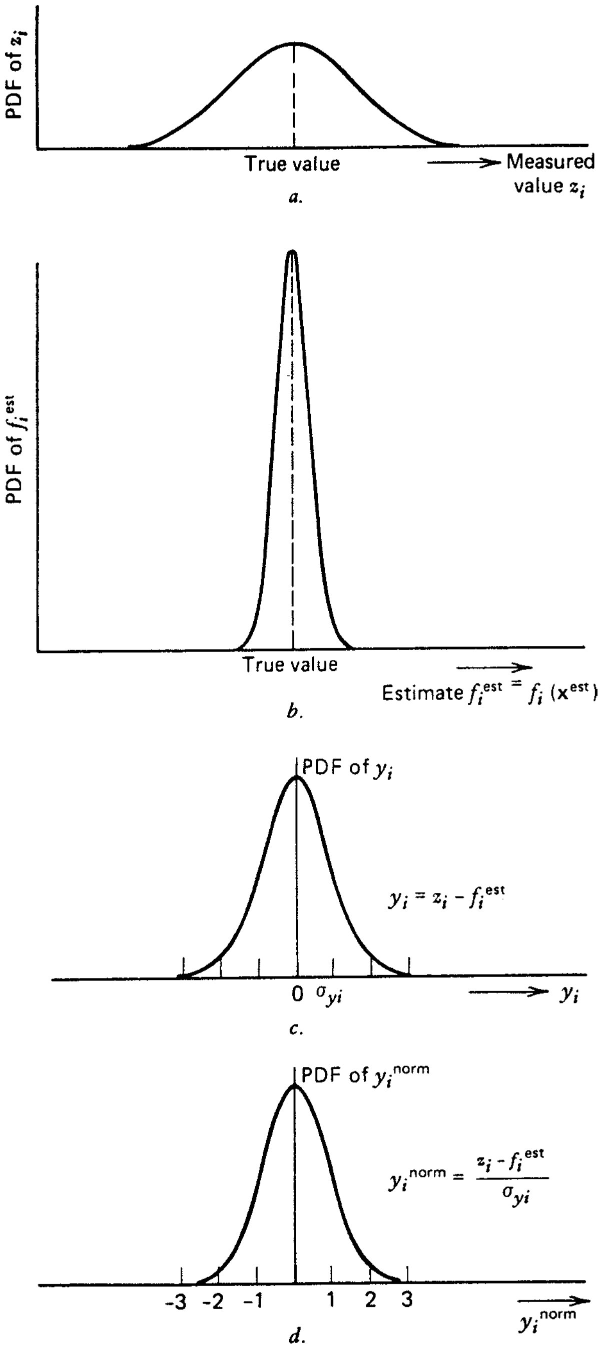


FIGURE 9.16 PDF of the normalized measurement residual.

www.mepcafe.com

measurements, one of which is bad. In such cases, the groups must be eliminated to eliminate the bad measurement.

The example given in Table 9.2 and Table 9.3 is now re-executed with the estimator's bad data detection logic activated. The results are given in Table 9.7 and Table 9.8 on the next page, with the M_{12} measurement value reversed. When $J(\mathbf{x}) > t_J$, M_{12} was found to have the largest normalized residual and removed.

TABLE 9.7 State Estimation Solution After Removal of Bad Data

Measurement Name	Measurement Status	Base Case Value kV	Base Case Value MW	Base Case Value MVAR	Measured Value kV	Measured Value MW	Measured Value MVAR	Estimated Value kV	Estimated Value MW	Estimated Value MVAR
		Bus 1								
MV1	1	246.1			245.9			248.9		
MA1	1	0.0			0.0			0.0		
MG1	1		228.6	-32.2		226.4	-30.6		228.7	-32.4
M12	0		123.6	-35.6		-125.2	38.7		123.3	-35.3
M14	1		179.9	23.9		184.5	23.8		179.6	24.5
M15	1		105.0	3.4		106.8	3.2		105.4	3.0
		Bus 2								
MV2	1	241.5			243.6			244.1		
MA2	0	-14.5			0.0			-11.9		
MG2	1		50.0	75.7		49.2	77.5		50.4	75.7
M21	1		-109.3	59.7		-108.1	57.5		-109.4	58.6
M23	1		11.4	-5.4		13.2	-5.6		11.4	-4.8
M24	1		93.1	17.2		93.8	14.2		93.2	18.4
M25	1		15.6	5.3		21.0	2.1		15.9	4.7
M26	1		39.2	-1.1		37.1	-2.9		39.4	-1.2
		Bus 3								
MV3	1	241.5			241.9			243.7		
MA3	0	-16.0			0.0			-13.4		
MG3	1		50.0	24.2		51.8	17.6		50.1	20.2
M32	1		-11.3	-0.9		-14.6	-2.8		-11.3	-1.7
M35	1		8.9	7.5		9.8	6.4		9.0	6.2
M36	1		52.4	17.6		52.7	19.3		52.5	15.7
		Bus 4								
MV4	1	228.0			234.3			230.5		
MA4	0	-19.1			0.0			-16.4		
ML4	1		-100.0	-15.0		-104.0	-19.3		-100.1	-17.1
M41	1		-165.5	29.6		-160.6	28.3		-165.5	27.5
M42	1		-89.0	-11.1		-86.1	-13.0		-89.2	-12.4
M45	1		-11.0	-3.9		-8.7	-5.9		-10.9	-4.7
		Bus 5								
MV5	1	233.3			234.4			236.3		
MA5	0	-16.6			0.0			-14.0		
ML5	1		-100.0	-15.0		-98.8	-10.6		-101.1	-12.5
M51	1		-97.3	19.1		-103.3	16.1		-97.7	18.9
M52	1		-15.3	-8.7		-14.7	-13.3		-15.6	-8.2
M53	1		-8.7	-12.4		-6.3	-10.0		-8.8	-11.2
M54	1		11.2	-3.7		9.4	-3.1		11.2	-3.1
M56	1		10.1	-9.2		9.2	-9.6		9.9	-8.8
		Bus 6								
MV6	1	235.4			241.6			238.0		
MA6	0	-18.6			0.0			-15.9		
ML6	1		-100.0	-15.0		-101.0	-12.6		-100.2	-14.0
M62	1		-38.2	-1.5		-37.0	1.3		-38.4	-1.6
M63	1		-51.8	-16.9		-50.6	-19.9		-51.9	-15.2
M65	1		-9.9	3.4		-8.2	6.6		-9.8	2.8

www.mepcafe.com

TABLE 9.8 Iterative Results Showing Bad Data Removal

Number Active Measurements	Degrees of Freedom	Bad Data Threshold t_J
62	51	63.440

Iteration	Residual J
1	26641.718
2	6706.949
3	6492.852
4	6456.845
5	6455.592

Bad Data Detected — Largest Normalized Residual = 76.947
Bad Measurement at M12
Bad Measurement Removed

Number Active Measurements	Degrees of Freedom	Bad Data Threshold t_J
60	49	61.182

Iteration	Residual J
1	809.343
2	37.555
3	37.481

The ability to detect (using the chi-squared statistic) and identify (using normalized residuals) is an extremely useful feature of a state estimator. Without the state estimator calculation using the system measurement data, those measurements whose values are not obviously wrong have little chance of being detected and identified. With the state estimator, the operations personnel have a greater assurance that quantities being displayed are not grossly in error.

9.6.3 Estimation of Quantities Not Being Measured

The other useful feature of a state estimator calculation is the ability to calculate (or estimate) quantities not being telemetered. This is most useful in cases of failure of communication channels connecting operations centers to remote data-gathering equipment or when the remote data-gathering equipment fails. Often data from some network substations are simply unavailable because no transducers or data-gathering equipment was ever installed.

An example of this might be the failure of all telemetry from buses 3, 4, 5, and 6 in our six-bus system. Even with the loss of these measurements, we can run the state estimation algorithm on the remaining measurements at buses 1 and 2, calculate the bus voltage magnitudes and phase angles at all six buses, and then calculate all network generations, loads, and flows. The results of such a calculation are given in Table 9.9. Notice that the estimate of quantities at the untelemetered buses is close to the base case obtained when using the full set of measurements (i.e., compare Table 9.9 to Table 9.4).

www.mepcafe.com

TABLE 9.9 State Estimator Solution with Measurements at Buses 1 and 2 Only

Measurement		Base Case Value			Measured Value			Estimated Value		
Name	Status	kV	MW	MVAR	kV	MW	MVAR	kV	MW	MVAR
		Bus 1								
MV1	1	246.1			245.3			246.6		
MA1	1	0.0			0.0			0.0		
MG1	1		228.6	-32.2		229.9	-34.0		229.8	-31.5
M12	1		123.6	-35.6		123.9	-30.4		124.1	-34.5
M14	1		179.9	23.9		174.7	26.6		178.1	26.0
M15	1		105.0	3.4		107.6	3.2		105.8	3.0
		Bus 2								
MV2	1	241.5			242.3			241.4		
MA2	0	-14.5			0.0			-13.6		
MG2	1		50.0	75.7		50.8	77.5		53.2	76.0
M21	1		-109.3	59.7		-111.1	57.5		-109.8	58.6
M23	1		11.4	-5.4		14.0	-4.9		11.8	-3.5
M24	1		93.1	17.2		94.0	17.5		90.2	20.1
M25	1		15.6	5.3		16.9	5.6		15.7	4.0
M26	1		39.2	-1.1		47.6	-4.6		45.3	-3.2
		Bus 3								
MV3	0	241.5						240.3		
MA3	0	-16.0						-15.1		
MG3	0		50.0	24.2					60.1	11.4
M32	0		-11.3	-0.9					-11.7	-2.8
M35	0		8.9	7.5					7.9	4.5
M36	0		52.4	17.6					63.9	9.8
		Bus 4								
MV4	0	228.0						227.5		
MA4	0	-19.1						-18.0		
ML4	0		-100.0	-15.0					-96.8	-20.0
M41	0		-165.5	29.6					-164.0	26.4
M42	0		-89.0	-11.1					-86.3	-14.4
M45	0		-11.0	-3.9					-10.5	-5.5
		Bus 5								
MV5	0	233.3						234.1		
MA5	0	-16.6						-15.8		
ML5	0		-100.0	-15.0					-96.5	-8.7
M51	0		-97.3	19.1					-97.9	19.7
M52	0		-15.3	-8.7					-15.4	-7.5
M53	0		-8.7	-12.4					-7.8	-9.5
M54	0		11.2	-3.7					10.7	-2.1
M56	0		10.1	-9.2					14.0	-9.5
		Bus 6								
MV6	0	235.4						235.4		
MA6	0	-18.6						-18.4		
ML6	0		-100.0	-15.0					-121.0	-2.6
M62	0		-38.2	-1.5					-44.0	1.5
M63	0		-51.8	-16.9					-63.2	-8.0
M65	0		-9.9	3.4					-13.8	3.9

9.6.4 Network Observability and Pseudo-measurements

What happens if we continue to lose telemetry so that fewer and fewer measurements are available? Eventually, the state estimation procedure breaks down completely. Mathematically, the matrix

$$\left[[H]^{\mathrm{T}}\left[R^{-1}\right][H]\right]$$

in Equation 9.46 becomes singular and cannot be inverted. There is also a very interesting engineering interpretation of this phenomenon that allows us to alter the situation so that the state estimation procedure is not completely disabled.

If we take the three-bus example used in the beginning of Section 9.2, we note that when all three measurements are used, we have a redundant set and we can use a least-squares fit to the measurement values. If one of the measurements is lost, we have just enough measurements to calculate the states. If, however, two measurements

www.mepcafe.com

are lost, we are in trouble. For example, suppose M_{13} and M_{32} were lost leaving only M_{12}. If we now apply Equation 9.23 in a straightforward manner, we get

$$M_{12} = f_{12} = \frac{1}{0.2}(\theta_1 - \theta_2) = 5\theta_1 - 5\theta_2$$

Then

$$[H] = [5 - 5]$$

$$[R] = \left[\sigma^2_{M12}\right] = [0.0001]$$

and

$$\begin{bmatrix} \theta_1^{est} \\ \theta_2^{est} \end{bmatrix} = \left[\begin{bmatrix} 5 \\ -5 \end{bmatrix}[0.0001]^{-1}[5 - 5]\right]^{-1} [5 - 5][0.0001]^{-1}(0.55)$$
$$= \begin{bmatrix} 2500 & -2500 \\ -2500 & 2500 \end{bmatrix}^{-1} [5 - 5][0.0001]^{-1}(0.55) \quad (9.74)$$

The matrix to be inverted in Equation 9.74 is clearly singular and, therefore, we have no way of solving for θ_1^{est} and θ_2^{est}. Why is this? The reasons become quite obvious when we look at the one-line diagram of this network as shown in Figure 9.17.

With only M_{12} available, all we can say about the network is that the phase angle across line 1–2 must be 0.11 rad, but with no other information available, we cannot

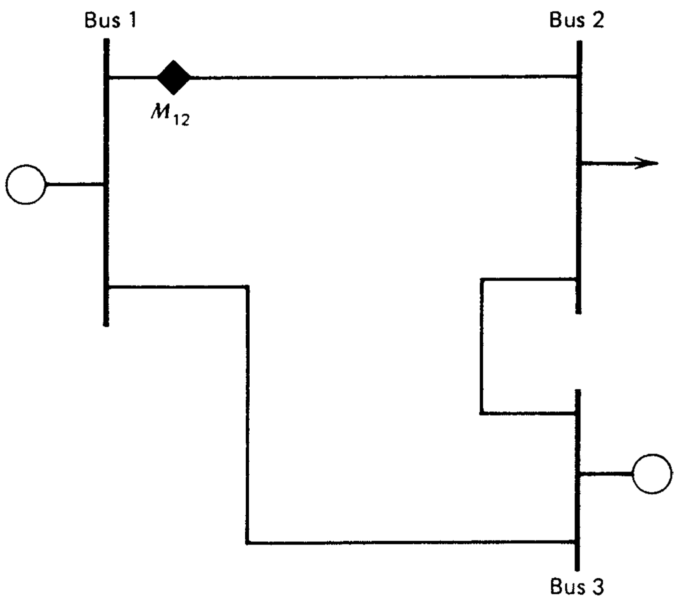


FIGURE 9.17 "Unobservable" measurement set.

www.mepcafe.com

tell what relationship a_1 or θ_2 has to θ_3, which is assumed to be 0 rad. If we write down the equations for the net injected power at bus 1 and bus 2, we have

$$\begin{aligned} P_1 &= 7.5\theta_1 - 5\theta_2 \\ P_2 &= -5\theta_1 + 9\theta_2 \end{aligned} \tag{9.75}$$

If measurement M_{12} is reading 55 MW (0.55 pu), we have

$$\theta_1 - \theta_2 = 0.11 \tag{9.76}$$

and by substituting Equation 9.17 into Equation 9.55 and eliminating θ_1, we obtain

$$P_2 = 1.6P_1 - 1.87 \tag{9.77}$$

Furthermore,

$$P_3 = -P_1 - P_2 = -0.6P_1 + 1.87 \tag{9.78}$$

Equations 9.77 and 9.78 give a relationship between P_1, P_2, and P_3, but we still do not know their correct values. The technical term for this phenomenon is to say that the network is *unobservable*; that is, with only M_{12} available, we cannot observe (calculate) the state of the system.

It is very desirable to be able to circumvent this problem. Often a large power system network will have missing data that render the network unobservable. Rather than just stop the calculations, a procedure is used that allows the estimator calculation to continue. The procedure involves the use of what are called *pseudo-measurements*. If we look at Equations 9.77 and 9.78, it is obvious that θ_1 and θ_2 could be estimated if the value of any one of the bus injections (i.e., P_1, P_2, or P_3) could be determined by some means other than direct measurement. This value, the pseudo-measurement, is used in the state estimator just as if it were an actual measured value.

To determine the value of an injection without measuring it, we must have some knowledge about the power system beyond the measurements currently being made. For example, it is customary to have access to the generated MW and MVAR values at generating stations through telemetry channels (i.e., the generated MW and MVAR would normally be measurements available to the state estimator). If these channels are out and we must have this measurement for observability, we can probably communicate with the operators in the plant control room by telephone and ask for the MW and MVAR values and enter them into the state estimator calculation manually. Similarly, if we needed a load-bus MW and MVAR for a pseudo-measurement, we could use historical records that show the relationship between an individual load and the total system load. We can estimate the total system load fairly accurately by knowing the total power being generated and estimating the network losses. Finally, if we have just experienced a telemetry failure, we could use the most recently estimated values from the estimator (assuming that it is run periodically) as pseudo-measurements. Therefore, if needed, we can provide the state estimator with a reasonable value to use as a pseudo-measurement at any bus in the system.

www.mepcafe.com

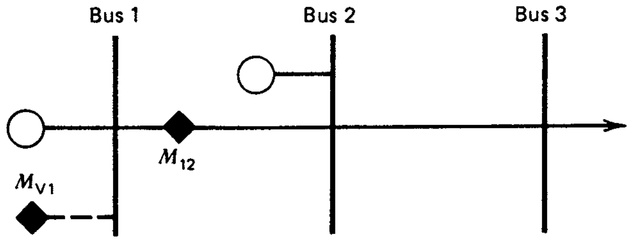


FIGURE 9.18 Unobservable system showing importance of location of pseudo-measurements.

The three-bus sample system in Figure 9.18 requires one pseudo-measurement. Measurement M_{12} allows us to estimate the voltage magnitude and phase angle at bus 2 (bus 1's voltage magnitude is measured and its phase angle is assumed to be 0). But without knowing the generation output at the generator unit on bus 2 or the load on bus 3, we cannot tell what voltage magnitude and phase angle to place on bus 3; hence, the network is unobservable. We can make this three-bus system observable by adding a pseudo-measurement of the net bus injected MW and MVAR at bus 2 or bus 3, but not at bus 1. That is, a pseudo-measurement at bus 1 will do no good at all because it tells nothing about the relationship of the phase angles between bus 2 and bus 3.

When adding a pseudo-measurement to a network, we simply write the equation for the pseudo-measurement injected power as a function of bus voltage magnitudes and phase angles as if it were actually measured. However, we do not wish to have the estimator treat the pseudo-measurement the same as a legitimate measurement, since it is often quite inaccurate and is little better than a guess. To circumvent this difficulty, we assign a large standard deviation to this measurement. The large standard deviation allows the estimator algorithm to treat the pseudo-measurement as if it were a measurement from a very poor-quality metering device.

To demonstrate the use of pseudo-measurements on our six-bus test system, all measurements were removed from buses 2, 3, 4, 5, and 6 so that bus 1 had all remaining measurements. This rendered the network unobservable and required adding pseudo-measurements at buses 2, 3, and 6. In the case, the pseudo-measurements were just taken from our base-case power flow. The results are shown in Table 9.10. Notice that the resulting estimates are quite close to the measured values for bus 1 but that the remaining buses have large measurement residuals. The net injections at buses 2, 3, and 6 do not closely match the pseudo-measurements since the pseudo-measurements were weighted much less than the legitimate measurements.

9.7 THE USE OF PHASOR MEASUREMENT UNITS (PMUS)

A **p**hasor **m**easurement **u**nit (PMU) has been defined by the IEEE as "a device that produces Synchronized Phasor, Frequency, and Rate of Change of Frequency (ROCOF) estimates from voltage and/or current signals and a time synchronizing signal."[2]

[2] IEEE C37.118-2011 Standard for Synchrophasor

www.mepcafe.com

TABLE 9.10 State Estimator Solution with Measurements at Bus 1 and Pseudo-measurements at Buses 2, 3, and 6

Measurement		Base Case Value			Measured Value			Estimated Value		
Name	Status	kV	MW	MVAR	kV	MW	MVAR	kV	MW	MVAR
		Bus 1								
MV1	1	246.1			246.3			246.3		
MA1	1	0.0			0.0			0.0		
MG1	1		228.6	-32.2		228.7	-25.5		228.3	-26.1
M12	1		123.6	-35.6		119.9	-34.6		108.4	-39.3
M14	1		179.9	23.9		175.9	20.8		181.2	26.3
M15	1		105.0	3.4		109.0	2.2		119.9	13.2
		Bus 2								
MV2	0	241.5						245.0		
MA2	0	-14.5						-12.9		
MG2	1		50.0	75.7	Pseudo	50.0	75.7		46.2	74.1
M21	0		-109.3	59.7					-97.0	57.6
M23	0		11.4	-5.4					-8.4	-11.8
M24	0		93.1	17.2					128.8	23.6
M25	0		15.6	5.3					43.1	12.6
M26	0		39.2	-1.1					-20.4	-7.8
		Bus 3								
MV3	0	241.5						250.5		
MA3	0	-16.0						-12.0		
MG3	1		50.0	24.2	Pseudo	50.0	24.0		47.9	23.4
M32	0		-11.3	-0.9					8.5	5.1
M35	0		8.9	7.5					59.2	14.8
M36	0		52.4	17.6					-19.9	3.5
		Bus 4								
MV4	0	228.0						227.2		
MA4	0	-19.1						-19.2		
ML4	0		-100.0	-15.0					-121.8	-14.4
M41	0		-165.5	29.6					-166.5	28.2
M42	0		-89.0	-11.1					-121.3	-10.5
M45	0		-11.0	-3.9					-0.5	-3.9
		Bus 5								
MV5	0	233.3						227.4		
MA5	0	-16.6						-19.1		
ML5	0		-100.0	-15.0					-260.0	-22.4
M51	0		-97.3	19.1					-109.7	18.8
M52	0		-15.3	-8.7					-41.3	-11.3
M53	0		-8.7	-12.4					-55.4	-11.8
M54	0		11.2	-3.7					0.5	-3.9
M56	0		10.1	-9.2					-54.2	-14.2
		Bus 6								
MV6	0	235.4						250.4		
MA6	0	-18.6						-11.0		
ML6	1		-100.0	-15.0	Pseudo	100.0	15.0		97.9	14.4
M62	0		-38.2	-1.5					20.6	2.8
M63	0		-51.8	-16.9					19.9	-5.5
M65	0		-9.9	3.4					57.3	17.1

By metering a common point on the sine wave of the voltage a two parts of a power system with high precision (error within 1 ms), we can determine the phase angle difference between the two voltages using the time difference as shown in the Figure 9.19 that follows.

Phasor Measurement Units have many useful features. They can be used to see system dynamic effects and to help engineers measure the system's response to

www.mepcafe.com

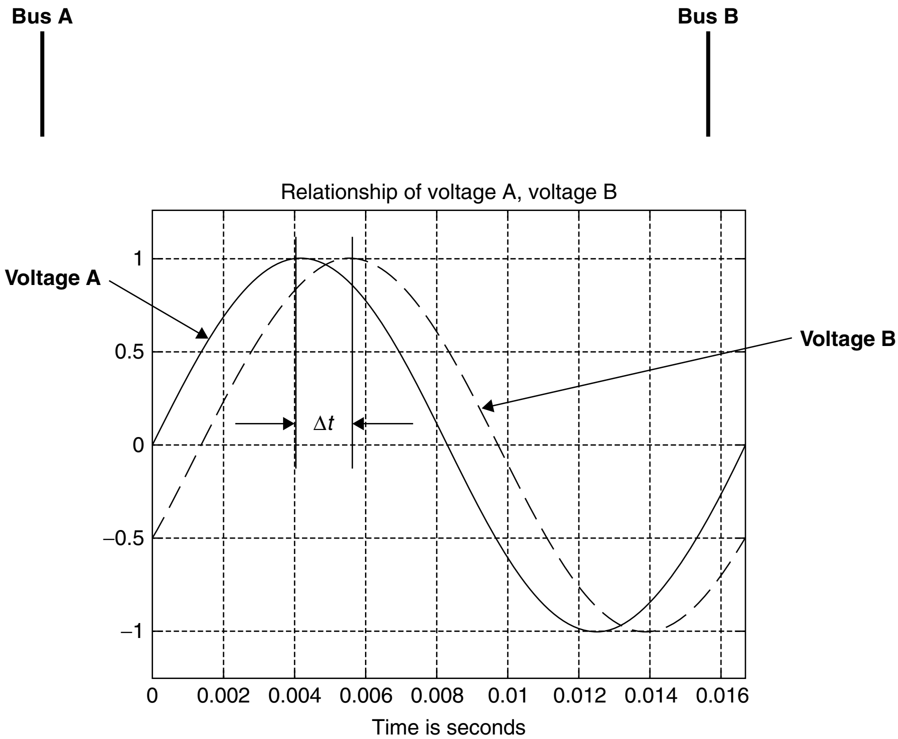


FIGURE 9.19 Phase difference as a function of time difference.

disturbances so as to correctly calculate dynamic parameters. PMUs require that the measurements be taken virtually simultaneously so that a GPS signal is required at the substation of the measurement. In addition, accurate determinations of the peak of the voltage waveform can be determined through the use of signal processing techniques on the fast sampled voltage signal. These have good effects when the PMU phase and voltage magnitude are used in a state estimator. In addition to the voltage magnitude and phase angle determination, PMUs also usually include measurements of the line or transformer current magnitude and angle that can also be used to advantage in a state estimator.

In this chapter, we assume that the state estimator is there only to provide a means of providing an accurate power flow solution to a model that is parallel to the actual power system itself and to detect and identify bad measurements. We do not deal with the use of PMUs to monitor system dynamics. The PMUs do improve the state estimator accuracy through the following means[3]:

[3] *PMU Impact on State Estimation Reliability for Improved Grid Security, Hongxia Wu, and Jay Giri*, IEEE Transactions on Power Systems, 2006

Recent Experience with a Hybrid SCADA/PMU On-line State Estimator, Rene Avila-Rosales, Mark J. Rice, Jay Giri, Lisa Beard, and Floyd Galvan, IEEE Transactions on Power Systems, 2009

www.mepcafe.com

TABLE 9.11 State Estimator with Voltage and Phase Angle Measurements Set to Very High Accuracy

Measurement Name	Status	Base Case Value kV	MW	MVAR	Measured Value kV	MW	MVAR	Estimated Value kV	MW	MVAR
		Bus 1								
MV1	1	246.1			246.1			246.1		
MA1	1	0.0			0.0			0.0		
MG1	1		228.6	-32.2		227.3	-33.4		230.8	-32.7
M12	1		123.6	-35.6		129.6	-32.7		124.5	-36.0
M14	1		179.9	23.9		178.9	26.0		181.3	23.9
M15	1		105.0	3.4		108.1	7.8		106.3	3.4
		Bus 2								
MV2	1	241.5			241.7			241.6		
MA2	1	-14.5			-14.4			-14.4		
MG2	1		50.0	75.7		53.0	74.8		51.8	76.5
M21	1		-109.3	59.7		-110.6	61.7		-110.0	60.6
M23	1		11.4	-5.4		11.8	-9.1		11.8	-5.3
M24	1		93.1	17.2		95.2	18.2		93.8	17.2
M25	1		15.6	5.3		16.0	0.4		16.2	5.2
M26	1		39.2	-1.1		40.8	-1.1		40.0	-1.2
		Bus 3								
MV3	1	241.5			241.4			241.5		
MA3	1	-16.0			-16.0			-16.0		
MG3	1		50.0	24.2		53.4	22.6		50.3	23.4
M32	1		-11.3	-0.9		-17.9	-1.9		-11.7	-1.0
M35	1		8.9	7.5		6.9	5.5		9.2	7.3
M36	1		52.4	17.6		53.6	20.2		52.9	17.1
		Bus 4								
MV4	1	228.0			228.3			228.1		
MA4	1	-19.1			-19.1			-19.1		
ML4	1		-100.0	-15.0		-97.1	-16.6		-100.4	-15.0
M41	1		-165.5	29.6		-161.3	24.8		-166.6	30.5
M42	1		-89.0	-11.1		-87.8	-12.9		-89.6	-11.0
M45	1		-11.0	-3.9		-8.9	-0.4		-10.8	-4.0
		Bus 5								
MV5	1	233.3			233.2			233.4		
MA5	1	-16.6			-16.6			-16.6		
ML5	1		-100.0	-15.0		-99.5	-12.1		-102.2	-13.7
M51	1		-97.3	19.1		-97.1	24.3		-98.4	19.8
M52	1		-15.3	-8.7		-17.6	-11.2		-15.9	-8.6
M53	1		-8.7	-12.4		-6.1	-12.1		-9.0	-12.2
M54	1		11.2	-3.7		12.7	0.6		11.0	-3.6
M56	1		10.1	-9.2		10.2	-9.4		10.0	-9.2
		Bus 6								
MV6	1	235.4			235.6			235.4		
MA6	1	-18.6			-18.6			-18.6		
ML6	1		-100.0	-15.0		-101.2	-16.3		-101.1	-14.4
M62	1		-38.2	-1.5		-41.7	-1.5		-39.0	-1.3
M63	1		-51.8	-16.9		-51.3	-15.3		-52.3	-16.4
M65	1		-9.9	3.4		-4.8	1.9		-9.9	3.4

- The time accuracy of PMU measurements means that time differences now taken for granted in SCADA systems are largely eliminated.
- Providing direct measurement of a voltage magnitude and phase allows the state estimator to have additional measurements for the same number of states (voltage magnitudes and voltage phase angles).

www.mepcafe.com

Experience with the use of adding PMU data to existing state estimators is reported in the engineering literature (see footnote) and includes:

- Improved reliability and robustness of the state estimator.
- PMUs can improve the estimation of inaccurate active power measurements close to the PMU substation.
- If only phase angle is used in the state estimator, then the PMU has little effect on improving accuracy of voltage magnitude and reactive power.
- The best performance is obtained when PMUs are evenly distributed over the power system.
- By observing data trends from the PMUs, a circuit breaker or disconnect switch change can be easily detected and identified.
- The use of PMU data from an external system (i.e., a neighboring power system) can improve the performance of the estimator.
- Using current phasors as well as voltage phasors improves the ability to detect and identify errors in nearby substations.
- The observability is improved with PMU use and the number of critical measurements is reduced.
- The normalized residuals used to detect problems with topology and parameter estimation are reduced.
- Last of all, the use of PMUs in a state estimator makes the detection and identification of errors in the PMU itself easier.

In the final AC state estimation example in Table 9.11, the accuracy of both voltage magnitude measurements and phase angle measurements was set so that σ was one tenth of that used in the same data set shown in Figure 9.2.

9.8 APPLICATION OF POWER SYSTEMS STATE ESTIMATION

In this last section, we will try to present the "big picture" showing how state estimation, contingency analysis, and generator corrective action fit together in a modern operations control center. Figure 9.20 is a schematic diagram showing the information flow between the various functions to be performed in an operations control center computer system. The system gets its information about the power system from remote terminal units that encode measurement transducer outputs and opened/closed status information into digital signals that are transmitted to the operations center over communication circuits. In addition, the control center can transmit control information such as raise/lower commands to generators and open/close commands to circuit breakers and switches. We have broken down the information coming into the control center as breaker/switch status indications and analog measurements. The analog measurements of generator output must be used directly by the automatic generation control (AGC) program (see Chapter 10), whereas all other data will be processed by the state estimator before being used by other programs.

www.mepcafe.com

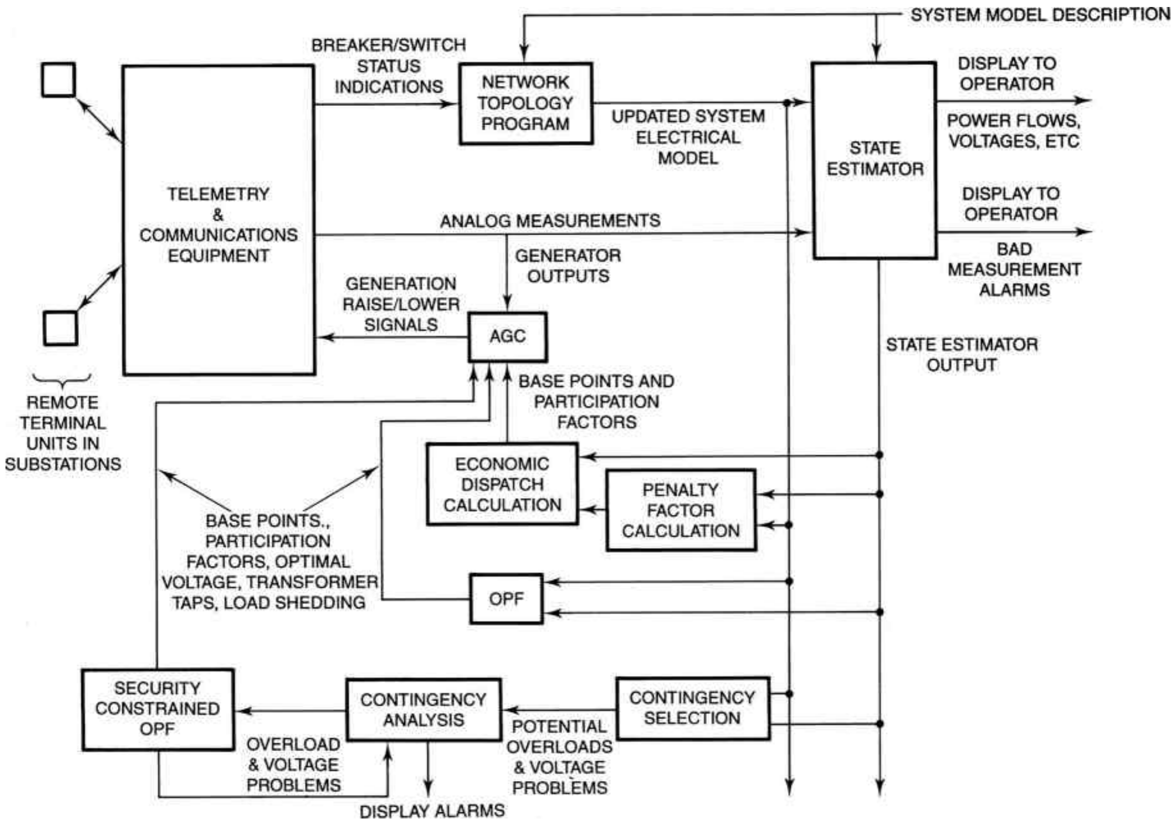


FIGURE 9.20 Energy control center system security schematic.

www.mepcafe.com

In order to run the state estimator, we must know how the transmission lines are connected to the load and generation buses. We call this information the *network topology*. Since the breakers and switches in any substation can cause the network topology to change, a program must be provided that reads the telemetered breaker/switch status indications and restructures the electrical model of the system. An example of this was shown in Figure 6.9 and repeated here in Figure 9.21, where the opening of four breakers requires two electrical buses to represent the substation instead of one electrical bus. We have labeled the program that reconfigures the

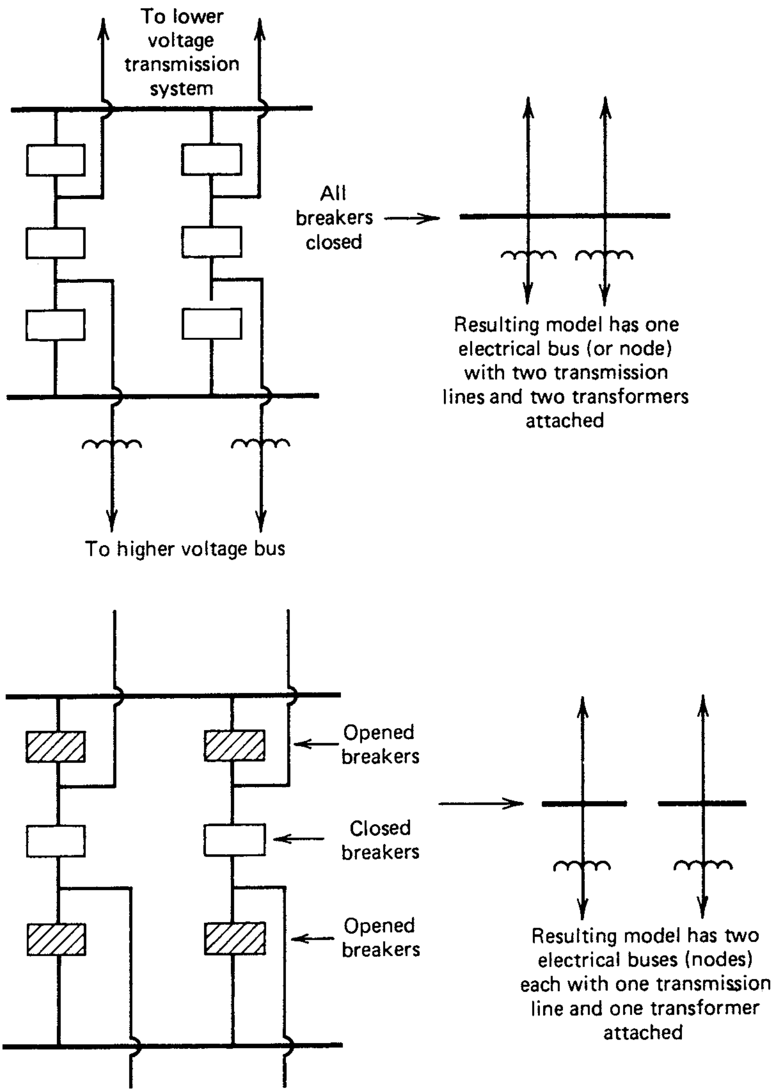


FIGURE 9.21 Example of network topology updating.

www.mepcafe.com

electrical model as the *network topology program*.[4] The network topology program must have a complete description of each substation and how the transmission lines are attached to the substation equipment. Bus sections that are connected to other bus sections through closed breakers or switches are designated as belonging to the same electrical bus. Thus, the number of electrical buses and the manner in which they are interconnected can be changed in the model to reflect breaker and switch status changes on the power system itself.

As seen in Figure 9.20, the electrical model of the power system's transmission system is sent to the state estimator program together with the analog measurements. The output of the state estimator consists of all bus voltage magnitudes and phase angles, transmission line MW and MVAR flows calculated from the bus voltage magnitude and phase angles, and bus loads and generations calculated from the line flows. These quantities, together with the electrical model developed by the network topology program, provide the basis for the economic dispatch program, contingency analysis program, and generation corrective action program. Note that since the complete electrical model of the transmission system is available, we can directly calculate bus penalty factors as shown in Chapter 6.

9.9 IMPORTANCE OF DATA VERIFICATION AND VALIDATION

When a state estimator is first installed, engineers working on the installation usually find that many of the measurements used by the operations office are poorly calibrated, some are connected backward, and some measurements are connected so that they are measuring the wrong quantity. All of this adds up to a very difficult period of time while the engineers go measurement by measurement to determine if measurements are calibrated properly, connected in the correct polarity, and connected to the proper transducers so as to measure what they are supposed to measure. Engineers should plan for a good deal of time to deal with these problems so that the state estimator can be trusted. This will mean that tests are run to verify measurements by comparison with other measurements, for example, that are at the opposite end of a transmission line and that both measurements make sense. The bad measurement detection logic can also be used to great advantage in these periods to find measurements that are off calibration or not connected properly.

When measurement are verified and used in the state estimator, the results can be trusted and used for all other applications (OPF, security analysis, etc.) that depend on the solved power flow results from the state estimator.

9.10 POWER SYSTEM CONTROL CENTERS

Figure 9.22 shows the power system control center design delivered in the early 1970s and commonly labeled as an Energy Management System (EMS). The remote terminal units (RTUs) sampled the data and transmitted the data every 2 s for critical

[4] Alternative names that are often used for this program are "system status processor" and "network configurator."

www.mepcafe.com

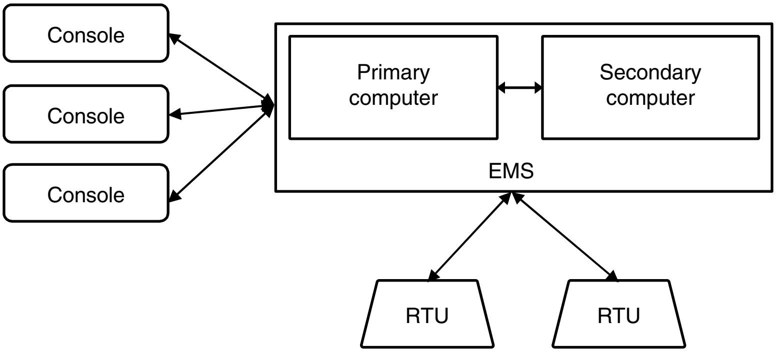


FIGURE 9.22 Energy management system.

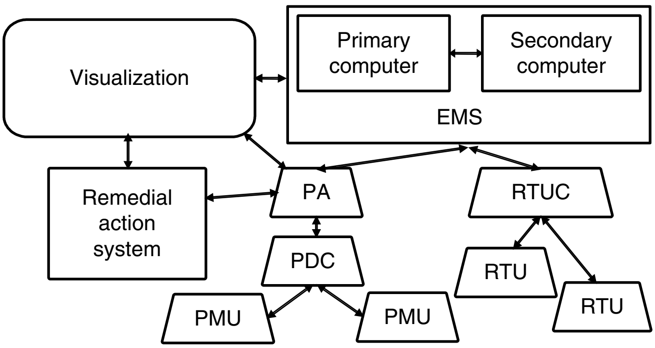


FIGURE 9.23 Energy management system architecture augmented with phasor measurements.

data and every minute for noncritical data. Critical data included generation, frequency, tie line flows, and breaker status data. Note that angle data was not included as time stamping data was not feasible. Data included voltage magnitude, real power, reactive power, and breaker status (open or closed). The consoles were the person machine interface for the power system operators to review the data and the output of the applications discussed in this text. There were one or more RTUs in each substation. The RTU concentrators were used to reduce the communication costs and to provide the initial data analysis to verify transmission quality.

Figure 9.23 shows the inclusion of phasor measurement unit (PMU) data. Phasor data is the voltage and current magnitude and angle since global positioning satellite (GPS) time synchronization enables time stamping of data accurate to find the phase of each variable. Global positioning satellite technology enables digital relays to determine the phase angles of each variable using waveform curve fitting techniques. Phasor data is presently sampled up to 200 times/s and sent to phasor data concentrators (PDCs). Phasor measurement units are added components of digital relays or are stand-alone versions of digital relays to acquire the data and provide the GPS time stamp. The PDCs gather data to synchronize data of the same time stamp and

www.mepcafe.com

perform analysis of data quality. State Estimator algorithms are now split at multiple levels, from the substation through the PDC to the EMS to provide faster data validation at the power flow level. It is expected that PMUs will replace RTUs as the cost of both systems is not feasible. Phasor data enables a new set of algorithms to augment the traditional algorithms presented in this book at the phasor analysis (PA) level. The visualization system is an extended person machine interface which enables operators to understand the power system state (secure, alert, etc.) with extensive graphing and topology presentations at enhanced consoles and projected displays. The remedial action system (RAS) is a collection of schemes to extend the relay protection system from the substation level to the regional level.

APPENDIX 9A
Derivation of Least-Squares Equations

One is often confronted with problems wherein data have been obtained by making measurements or taking samples on a process. Furthermore, the quantities being measured are themselves functions of other variables that we wish to estimate. These other variables will be called the state variables and designated **x**, where the number of state variables is N_s. The measurement values will be called **z**. We will assume here that the process we are interested in can be modeled using a linear model. Then we say that each measurement z_i is a linear function of the states x_i; that is,

$$z_i = h_i(\mathbf{x}) = h_{i1}x_1 + h_{i2}x_2 + \cdots + h_{iN_s}x_{N_s} \tag{9A.1}$$

We can also write this equation as a vector equation if we place the h_{ij} coefficients into a vector **h**; that is,

$$\mathbf{h}_i = \begin{bmatrix} h_{i1} \\ h_{i2} \\ \vdots \\ h_{iN_s} \end{bmatrix} \tag{9A.2}$$

Then Equation 9A.1 becomes

$$z_i = \mathbf{h}_i^{\mathrm{T}}\mathbf{x} \tag{9A.3}$$

where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N_s} \end{bmatrix} \tag{9A.4}$$

www.mepcafe.com

Finally, we can write all the measurement equations in a compact form

$$\mathbf{z} = [H]\mathbf{x}$$

where

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{N_\mathrm{m}} \end{bmatrix}$$

$$[H] = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1N_\mathrm{s}} \\ h_{21} & h_{22} & \cdots & \\ \vdots & & & \\ h_{N_\mathrm{m}1} & & \cdots & h_{N_\mathrm{m}N_\mathrm{s}} \end{bmatrix}$$

where row i of $[H]$ is equal to vector $\mathbf{h}_i^\mathrm{T}$ (see Equation 9A.2).

With N_m measurements we can have three possible cases to solve. That is, N_s, the number of states, is either less than N_m, equal to N_m, or greater than N_m. We will deal with each case separately.

9A.1 THE OVERDETERMINED CASE (N_m>N_s)

In this case, we have more measurements or samples than state variables; therefore, we can write more equations, $h_i(\mathbf{x})$, than we have unknowns x_j. One way to estimate the x_i values is to minimize the sum of the squares of difference between the measurement values z_i and the estimate of z_i that is, in turn, a function of the estimates of x_i. That is, we wish to minimize

$$J(\mathbf{x}) = \sum_{i=1}^{N_\mathrm{m}} [z_i - h_i(x_1, x_2, \ldots, x_{N_\mathrm{s}})]^2 \tag{9A.5}$$

Equation 9A.5 can be written as

$$J(\mathbf{x}) = \sum_{i=1}^{N_\mathrm{m}} (z_i - \mathbf{h}_i^\mathrm{T}\mathbf{x})^2 \tag{9A.6}$$

and this can be written in a still more compact form as

$$J(\mathbf{x}) = (\mathbf{z} - [H]\mathbf{x})^\mathrm{T}(\mathbf{z} - [H]\mathbf{x}) \tag{9A.7}$$

www.mepcafe.com

If we wish to find the value of **x** that minimizes $J(\mathbf{x})$, we can take the first derivative of $J(\mathbf{x})$ with respect to each $x_j (j = 1, \ldots, N_s)$ and set these derivatives to 0. That is,

$$\frac{\partial J(\mathbf{x})}{\partial x_j} = 0 \quad \text{for} \quad j = 1 \ldots N_s \tag{9A.8}$$

If we place these derivatives into a vector, we have what is called the gradient of $J(\mathbf{x})$, which is written $\nabla_x \mathbf{J}(\mathbf{x})$. Then,

$$\nabla_x \mathbf{J}(\mathbf{x}) = \begin{bmatrix} \frac{\partial J(\mathbf{x})}{\partial x_1} \\ \frac{\partial J(\mathbf{x})}{\partial x_2} \\ \vdots \\ \end{bmatrix} \tag{9A.9}$$

Then the goal of forcing each derivative to 0 can be written as

$$\nabla_x \mathbf{J}(\mathbf{x}) = \mathbf{0} \tag{9A.10}$$

where **0** is a vector of N_s elements, each of which is 0. To solve this problem, we will first expand Equation 9A.7:

$$\begin{aligned} J(x) &= (z - [H]x)^{\mathrm{T}}(\mathbf{z} - [H]\mathbf{x}) \\ &= \mathbf{z}^{\mathrm{T}}\mathbf{z} - \mathbf{x}^{\mathrm{T}}[H]^{\mathrm{T}}\mathbf{z} - \mathbf{z}^{\mathrm{T}}[H]\mathbf{x} + \mathbf{x}^{\mathrm{T}}[H]^{\mathrm{T}}[H]\mathbf{x} \end{aligned} \tag{9A.11}$$

The second and third terms in Equation 9A.11 are identical, so that we can write

$$J(\mathbf{x}) = \mathbf{z}^{\mathrm{T}}\mathbf{z} - 2\mathbf{z}^{\mathrm{T}}[H]\mathbf{x} + \mathbf{x}^{\mathrm{T}}[H]^{\mathrm{T}}[H]\mathbf{x} \tag{9A.12}$$

Before proceeding, we will derive a few simple relationships.

The gradient is always a vector of first derivatives of a scalar function that is itself a function of a vector. Thus, if we define $F(\mathbf{y})$ to be a scalar function, then its gradient $\nabla_y \mathbf{F}$ is

$$\nabla_y \mathbf{F} = \begin{bmatrix} \frac{\partial F}{\partial y_1} \\ \frac{\partial F}{\partial y_2} \\ \vdots \\ \frac{\partial F}{\partial y_n} \end{bmatrix} \tag{9A.13}$$

www.mepcafe.com

Then, if we define F as follows,

$$F = \mathbf{y}^{\mathrm{T}}\mathbf{b} = [y_1 \quad y_2 \quad \cdots]\begin{bmatrix} b_1 \\ b_2 \\ \vdots \end{bmatrix} \tag{9A.14}$$

where $\mathbf{b}$ is a vector of constants $b_i, i = 1, \ldots, n$, then F can be expanded as

$$F = y_1 b_1 + y_2 b_2 + y_3 b_3 + \cdots \tag{9A.15}$$

and the gradient of F is

$$\nabla_y \mathbf{F} = \begin{bmatrix} \dfrac{\partial F}{\partial y_1} \\ \dfrac{\partial F}{\partial y_2} \\ \vdots \\ \dfrac{\partial F}{\partial y_n} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \mathbf{b} \tag{9A.16}$$

It ought to be obvious that writing F with y and $\mathbf{b}$ reversed makes no difference. That is,

$$F = \mathbf{b}^{\mathrm{T}}\mathbf{y} = \mathbf{y}^{\mathrm{T}}\mathbf{b} \tag{9A.17}$$

$$\text{and, therefore, } \nabla_y(\mathbf{b}^{\mathrm{T}}\mathbf{y}) = \mathbf{b}$$

Suppose we now write the vector $\mathbf{b}$ as the product of a matrix $[A]$ and a vector $\mathbf{u}$:

$$\mathbf{b} = [A]\mathbf{u} \tag{9A.18}$$

Then, if we take F as shown in Equation 9A.14,

$$F = \mathbf{y}^{\mathrm{T}}\mathbf{b} = \mathbf{y}^{\mathrm{T}}[A]\mathbf{u} \tag{9A.19}$$

we can say that

$$\nabla_y \mathbf{F} = [A]\mathbf{u} \tag{9A.20}$$

Similarly, we can define

$$\mathbf{b}^{\mathrm{T}} = \mathbf{u}^{\mathrm{T}}[A] \tag{9A.21}$$

www.mepcafe.com

If we can take F as shown in Equation 9A.17,

$$F = \mathbf{b}^{\mathrm{T}}\mathbf{y} = \mathbf{u}^{\mathrm{T}}[A]\mathbf{y}$$

then

$$\nabla_{y}\mathbf{F} = [A]^{\mathrm{T}}\mathbf{u} \tag{9A.22}$$

Finally, we will look at a scalar function F that is quadratic, namely,

$$\begin{aligned} F &= \mathbf{y}^{\mathrm{T}}[A]\mathbf{y} \\ &= \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & \vdots & \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} y_i a_{ij} y_j \end{aligned} \tag{9A.23}$$

Then

$$\nabla_{y}\mathbf{F} = \begin{bmatrix} \dfrac{\partial F}{\partial y_1} \\ \dfrac{\partial F}{\partial y_2} \\ \vdots \\ \dfrac{\partial F}{\partial y_n} \end{bmatrix} = \begin{bmatrix} 2a_{11}y_1 + 2a_{12}y_2 + \cdots \\ 2a_{21}y_1 + 2a_{22}y_2 + \cdots \\ \vdots \end{bmatrix} = 2[A]\mathbf{y} \tag{9A.24}$$

Then, in summary,

$$\begin{array}{|ll|} \hline 1.\ \mathrm{F} = \mathrm{y}^{\mathrm{T}}\mathbf{b} & \nabla_{y}\mathbf{F} = \mathbf{b} \\ 2.\ F = \mathbf{b}^{\mathrm{T}}\mathbf{y} & \nabla_{y}\mathbf{F} = \mathbf{b} \\ 3.\ F = \mathbf{y}^{\mathrm{T}}[A]\mathbf{u} & \nabla_{y}\mathbf{F} = [A]\mathbf{u} \\ 4.\ F = \mathbf{u}^{\mathrm{T}}[A]\mathbf{y} & \nabla_{y}\mathbf{F} = [A]^{T}\mathbf{u} \\ 5.\ F = \mathbf{y}^{\mathrm{T}}[A]\mathbf{y} & \nabla_{y}\mathbf{F} = 2[A]\mathbf{u} \\ \hline \end{array} \tag{9A.25}$$

www.mepcafe.com

We will now use Equation 9A.5 to derive the gradient of $J(\mathbf{x})$, that is, $\nabla_x \mathbf{J}$, where $J(\mathbf{x})$ is shown in Equation 9A.12. The first term, $\mathbf{z}^T\mathbf{z}$, is not a function of $\mathbf{x}$, so we can discard it. The second term is of the same form as (4) in Equation 9A.25, so that

$$\nabla_x(-2\mathbf{z}^T[H]\mathbf{x}) = -2[H]^T\mathbf{z} \tag{9A.26}$$

The third term is the same as (5) in Equation 9A.25 with $[H]^T[H]$ replacing $[A]$; then,

$$\nabla_x(\mathbf{x}^T[H]^T[H]\mathbf{x}) = 2[H]^T[H]\mathbf{x} \tag{9A.27}$$

Then from Equations 9A.26 and 9A.27, we have

$$\nabla_x \mathbf{J} = -2[H]^T\mathbf{z} + 2[H]^T[H]\mathbf{x} \tag{9A.28}$$

But, as stated in Equation 9A.10, we wish to force $\nabla_x \mathbf{J}$ to 0. Then

$$-2[H]^T\mathbf{z} + 2[H]^T[H]\mathbf{x} = 0$$

or

$$\mathbf{x} = [[H]^T[H]]^{-1}[H]^T\mathbf{z} \tag{9A.29}$$

If we had wanted to put a different weight, w_i, on each measurement, we could have written Equation 9A.6 as

$$J(\mathbf{x}) = \sum_{i=1}^{N_m} w_i(z_i - \mathbf{h}_i^T\mathbf{x})^2 \tag{9A.30}$$

which can be written as

$$J(\mathbf{x}) = (\mathbf{z} - [H]\mathbf{x})^T[W](\mathbf{z} - [H]\mathbf{x})$$

where $[W]$ is a diagonal matrix. Then

$$J(\mathbf{x}) = \mathbf{z}^T[W]\mathbf{z} - \mathbf{x}^T[H]^T[W]\mathbf{z} - \mathbf{z}^T[W][H]\mathbf{x} + \mathbf{x}^T[H]^T[W][H]\mathbf{x}$$

If we once again use Equation 9A.25, we would obtain

$$\nabla_x \mathbf{J} = -2[H]^T[W]\mathbf{z} + 2[H]^T[W][H]\mathbf{x}$$

www.mepcafe.com

and

$$\nabla_x \mathbf{J} = 0$$

gives

$$\mathbf{x} = ([H]^{\mathrm{T}}[W][H])^{-1}[H]^{\mathrm{T}}[W]\mathbf{z} \tag{9A.31}$$

9A.2 THE FULLY DETERMINED CASE ($N_{\mathrm{m}} = N_{\mathrm{s}}$)

In this case, the number of measurements is equal to the number of state variables, and we can solve for **x** directly by inverting [H]:

$$\mathbf{x} = [H]^{-1}\mathbf{z} \tag{9A.32}$$

9A.3 THE UNDERDETERMINED CASE ($N_{\mathrm{m}} < N_{\mathrm{s}}$)

In this case, we have fewer measurements than state variables. In such a case, it is possible to solve for many solutions $\mathbf{x}^{\mathrm{est}}$ that cause $J(\mathbf{x})$ to equal 0. The usual solution technique is to find $\mathbf{x}^{\mathrm{est}}$ that minimizes the sum of the squares of the solution values. That is, we find a solution such that

$$\sum_{j=1}^{N_{\mathrm{s}}} x_j^2 \tag{9A.33}$$

is minimized while meeting the condition that the measurements will be solved for exactly. To do this, we treat the problem as a constrained minimization problem and use Lagrange multipliers as shown in Appendix 3A.

We formulate the problem as

$$\begin{aligned} &\textbf{Minimize}: \sum_{j=1}^{N_{\mathrm{s}}} x_j^2 \\ &\textbf{Subject to}: z_i = \sum_{j=1}^{N_{\mathrm{s}}} h_{ij} x_j \quad \text{for} \quad i = 1,\ldots,N_{\mathrm{m}} \end{aligned} \tag{9A.34}$$

This optimization problem can be written in vector–matrix form as

$$\begin{aligned} &\min \quad \mathbf{x}^{\mathrm{T}}\mathbf{x} \\ &\text{subject to} \quad \mathbf{z} = [H]\mathbf{x} \end{aligned} \tag{9A.35}$$

www.mepcafe.com

The Lagrangian for this problem is

$$\mathrm{L} = \mathbf{x}^{\mathrm{T}}\mathbf{x} + \lambda^{\mathrm{T}}(\mathbf{z} - [H]\mathbf{x}) \tag{9A.36}$$

Following the rules set down in Appendix 3A, we must find the gradient of $\mathcal{L}$ with respect to **x** and with respect to λ. Using the identities found in Equation 9A.25, we get

$$\nabla_x \mathrm{L} = 2\mathbf{x} - [H]^{\mathrm{T}}\lambda = 0$$

which gives

$$\mathbf{x} = \frac{1}{2}[H]^{\mathrm{T}}\lambda$$

and

$$\nabla_\lambda \mathrm{L} = \mathbf{z} - [H]\mathbf{x} = 0$$

which gives

$$\mathbf{z} = [H]\mathbf{x}$$

Then

$$\mathbf{z} = \frac{1}{2}[H][H]^{\mathrm{T}}\lambda$$

or

$$\lambda = 2[[H][H]^{\mathrm{T}}]^{-1}\mathbf{z}$$

and, finally,

$$\mathbf{x} = [H]^{\mathrm{T}}[[H][H]^{\mathrm{T}}]^{-1}\mathbf{z} \tag{9A.37}$$

The reader should be aware that the matrix inversion shown in Equations 9A.29, 9A.32, and 9A.37 may not be possible. That is, the $[[H]^{\mathrm{T}}[H]]$ matrix in Equation 9A.32 may be singular, or $[H]$ may be singular in Equation 9A.37, or $[[H][H]^{\mathrm{T}}]$ may be singular in Equation 9A.37. In the overdetermined case ($N_{\mathrm{m}} > N_{\mathrm{s}}$) whose solution is Equation 9A.29 and the fully determined case ($N_{\mathrm{m}} = N_{\mathrm{s}}$) whose solution is Equation 9A.32, the singularity implies what is known as an "unobservable" system. By unobservable we mean that the measurements do not provide sufficient information to allow a determination of the states of the system. In the case of the underdetermined case ($N_{\mathrm{m}} < N_{\mathrm{s}}$) whose solution is Equation 9A.37, the singularity simply implies that there is no unique solution to the problem.

www.mepcafe.com

PROBLEMS

9.1 Using the three-bus sample system shown in Figure 9.1, assume that the three meters have the following characteristics.

Meter	Full Scale (MW)	Accuracy (MW)	σ (pu)
M_{12}	100	±6	0.02
M_{13}	100	±3	0.01
M_{32}	100	±0.6	0.002

a. Calculate the best estimate for the phase angles θ_1 and θ_2 given the following measurements.

Meter	Measured Value (MW)
M_{12}	60.0
M_{13}	4.0
M_{32}	40.5

b. Calculate the residual $J(\mathbf{x})$. For a significance level, α, of 0.01, does $J(\mathbf{x})$ indicate the presence of bad data? Explain.

9.2 Given a single transmission line with a generator at one end and a load at the other, two measurements are available as shown in Figure 9.24. Assume that we can model this circuit with a DC load flow using the line reactance shown. Also, assume that the phase angle at bus 1 is 0 rad. Given the meter characteristics and meter readings telemetered from the meters, calculate the best estimate of the power flowing through the transmission line.

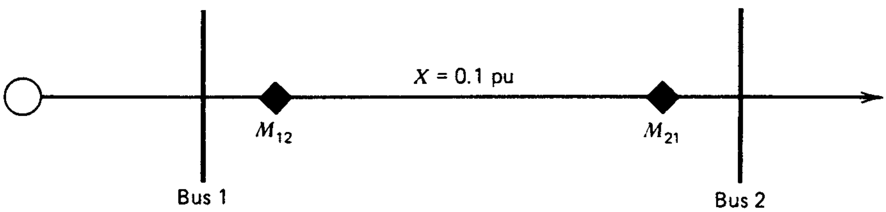


FIGURE 9.24 Measurement configuration for Problem 9.2.

Meter	Full Scale (MW)	Meter Standard Deviation (σ) in Full Scale	Meter Reading (MW)
M_{12}	200	1	62
M_{21}	200	5	−52

Note: M_{12} measures power flowing from bus 1 to bus 2; M_{21} measures power flowing from bus 2 to bus 1. Use 100 MVA as base.

www.mepcafe.com

9.3 You are given in the following network with meters at locations as shown in Figure 9.25.

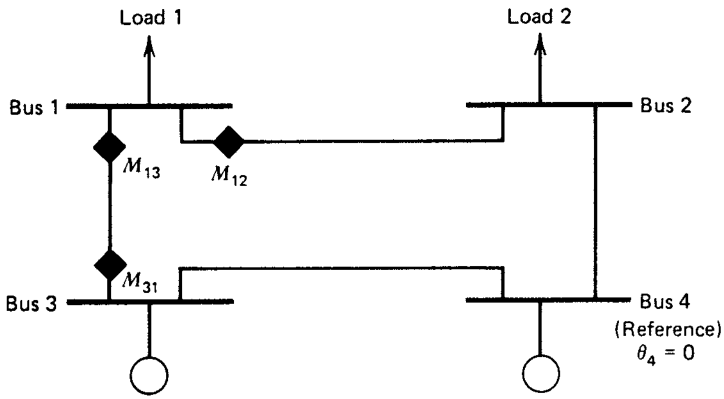


FIGURE 9.25 Four-bus system with measurements for Problem 9.3.

Branch Impedances (pu)
$X_{12}=0.25$
$X_{13}=0.50$
$X_{24}=0.40$
$X_{34}=0.10$

Measurement Values	Measurement Errors
$M_{13}=-70.5$	$\sigma_{13}=0.01$
$M_{31}=72.1$	$\sigma_{31}=0.01$
$M_{12}=21.2$	$\sigma_{12}=0.02$

a. Is this network observable? Set up the least-squares equations and try to invert $[H^T R^{-1} H]$.

b. Suppose we had a measurement of generation output at bus 3 and included it in our measurement set. Let this measurement be the following:

$$M_{3\text{gen}} = 92 \text{ MW with } \sigma = 0.015$$

Repeat part a including this measurement.

9.4 Given the network shown in Figure 9.26, the network is to be modeled with a DC power flow with line reactances as follows (assume 100-MVA base):

www.mepcafe.com

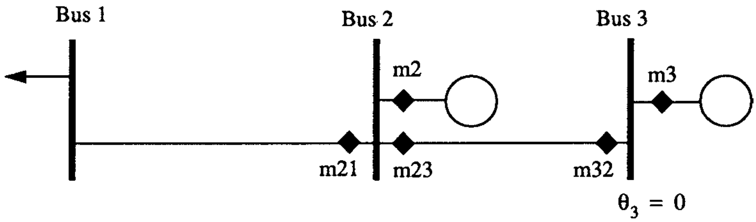


FIGURE 9.26 Network for Problem 9.4.

$$x_{12} = 0.1 \text{ pu}$$

$$x_{23} = 0.25 \text{ pu}$$

The meters are all of the same type with a standard deviation of σ=0.01 pu for each. The measured values are

$$M_3 = 105 \text{ MW}$$

$$M_{32} = 98 \text{ MW}$$

$$M_{23} = -135 \text{ MW}$$

$$M_2 = 49 \text{ MW}$$

$$M_{21} = 148 \text{ MW}$$

a. Find the phase angles that result in a best fit to the measured values.

b. Find the value of the residual function J.

c. Calculate estimated generator output of each generator and the estimated power flow on each line.

d. Are there any errors in the measurements? If you think so, explain which meters are apt to be in error and why. Remove the suspected bad measurement and try to resolve the state estimator.

9.5 You are to purchase and install a set of programs that are to act as a monitor for the system security of a major utility company. You have solicited bids from major manufacturers of computer systems and are responsible for reviewing the technical contents of each bid. One of the manufacturers proposes to install a system with the flowchart and description given in Figure 9.27.

www.mepcafe.com

Bidders design:

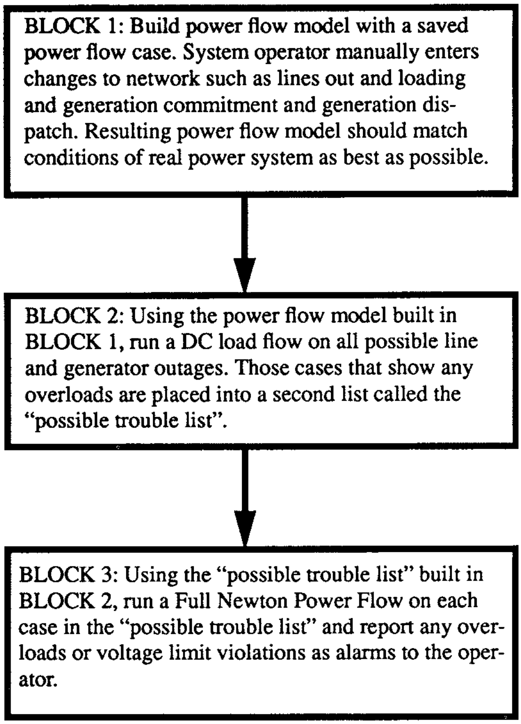


FIGURE 9.27 Diagram for Problem 9.5.

Bidders design:

a. Write down as many of the design flaws that you can find in this bidder's design.

b. Create a new design that you think will be a state-of-the-art system.

www.mepcafe.com

10

CONTROL OF GENERATION

10.1 INTRODUCTION

So far, this text has concentrated on methods of establishing optimum dispatch and scheduling of generating units. It is important to understand that the early success of electric generation was the inherent detection of demand changes without any communication and control infrastructure. The first response mechanism to keep the production and the consumption operating at the speed of light, it the ability of the generator to satisfy conservation of energy instantaneously by matching mechanical power to the electric power consumed. Indeed, the control of generator units was the first problem faced in early power-system design. The methods developed for control of individual generators and eventually control of large interconnections play a vital role in modern energy control centers.

A generator driven by a steam turbine can be represented as a large rotating mass with two opposing torques acting on the rotation. As shown in Figure 10.1, T_{mech}, the mechanical torque, acts to increase rotational speed whereas T_{elec}, the electrical torque, acts to slow it down. When T_{mech} and T_{elec} are equal in magnitude, the rotational speed, ω, will be constant. If the electrical load is increased so that T_{elec} is larger than T_{mech}, the entire rotating system will begin to slow down. This inherent balancing of energy enabled electrical systems to operate without prior knowledge of demand increases or decreases. It is the inertia of the system which maintains the security of the system as demand changes. Since it would be damaging to let the equipment slow down too far, something must be done to increase the mechanical

Power Generation, Operation, and Control, Third Edition. Allen J. Wood, Bruce F. Wollenberg, and Gerald B. Sheblé.
© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

www.mepcafe.com

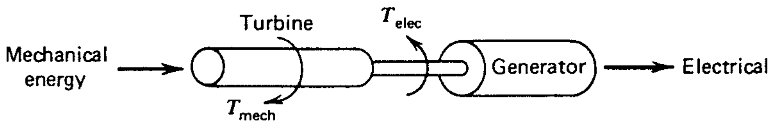


FIGURE 10.1 Mechanical and electrical torques in a generating unit.

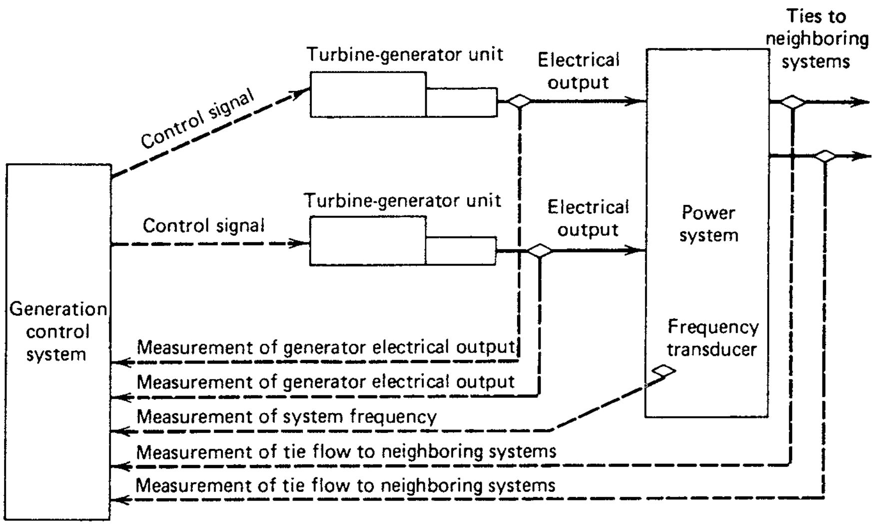


FIGURE 10.2 Overview of generation control problem.

torque T_{mech} to restore equilibrium; that is, to bring the rotational speed back to an acceptable value and the torques to equality so that the energy balance is restored at the desired speed.

This process must be repeated constantly on a power system because the loads change constantly. Furthermore, because there are many generators supplying power into the transmission system, some means must be provided to allocate the load changes to the generators. To accomplish this, multiple layers of control systems are connected to the generator units. A governor on each unit maintains its speed while supplementary control, usually originating at a remote control center, acts to allocate generation. Figure 10.2 shows an overview of the generation control problem. There is also a boiler control system to increase steam production to follow the governor action to maintain the energy balance. The coordination of the boiler control system with the governor is beyond the scope of this book.

It is the realm of transient stability that the first response of the system is analyzed. Transient stability to demand changes is treated in an appendix. This chapter assumes that the system is stable to any changes such that the governors have time to respond.

www.mepcafe.com

10.2 GENERATOR MODEL

Before starting, it will be useful for us to define electrical–mechanical system terms:

$$\begin{aligned}
\omega &= \text{rotational speed (rad/sec)} \\
\alpha &= \text{rotational acceleration} \\
\delta &= \text{phase angle of a rotating machine} \\
T_{\text{net}} &= \text{net accelerating torque in a machine} \\
T_{\text{mech}} &= \text{mechanical torque exerted on the machine by the turbine} \\
T_{\text{elec}} &= \text{electrical torque exerted on the machine by the generator} \\
P_{\text{net}} &= \text{net accelerating power}
\end{aligned}$$

$$\begin{aligned}
P_{\text{mech}} &= \text{mechanical power input} \\
P_{\text{elec}} &= \text{electrical power output} \\
I &= \text{moment of inertia for the machine} \\
M &= \text{angular momentum of the machine}
\end{aligned}$$

where all quantities (except phase angle) will be in per unit on the machine base, or, in the case of ω, on the standard system frequency base. Thus, for example, M is in per unit power/per unit frequency/s.

In the development to follow, we are interested in deviations of quantities about steady-state values. All steady-state or nominal values will have a "0" subscript (e.g., ω_0, T_{net_0}), and all deviations from nominal will be designated by a "Δ" (e.g., $\Delta\omega$, ΔT_{net}). Some basic relationships are

$$I\alpha = T_{\text{net}} \tag{10.1}$$

$$M = \omega I \tag{10.2}$$

$$P_{\text{net}} = \omega T_{\text{net}} = \omega(I\alpha) = M\alpha \tag{10.3}$$

To start, we will focus our attention on a single rotating machine. Assume that the machine has a steady speed of ω_0 and phase angle δ_0. Due to various electrical or mechanical disturbances, the machine will be subjected to differences in mechanical and electrical torque, causing it to accelerate or decelerate. We are chiefly interested in the deviations of speed, $\Delta\omega$, and deviations in phase angle, $\Delta\delta$, from nominal.

The phase angle deviation, $\Delta\delta$, is equal to the difference in phase angle between the machine as subjected to an acceleration of α and a reference axis rotating at exactly ω_0. If the speed of the machine under acceleration is

$$\omega = \omega_0 + \alpha t \tag{10.4}$$

www.mepcafe.com

then

$$\Delta\delta = \underbrace{\int(\omega_0 + \alpha t)\,\mathrm{d}t}_{\substack{\text{Machine absolute}\\ \text{phase angle}}} - \underbrace{\int\omega_0\,\mathrm{d}t}_{\substack{\text{Phase angle of}\\ \text{reference axis}}}$$
$$= \omega_0 t + \frac{1}{2}\alpha t^2 - \omega_0 t \tag{10.5}$$
$$= \frac{1}{2}\alpha t^2$$

The deviation from nominal speed, $\Delta\omega$, may then be expressed as

$$\Delta\omega = \alpha t = \frac{\mathrm{d}}{\mathrm{d}t}(\Delta\delta) \tag{10.6}$$

The relationship between phase angle deviation, speed deviation, and net accelerating torque is

$$T_{\text{net}} = I\alpha = I\frac{\mathrm{d}}{\mathrm{d}t}(\Delta\omega) = I\frac{\mathrm{d}^2}{\mathrm{d}t^2}(\Delta\delta) \tag{10.7}$$

Next, we will relate the deviations in mechanical and electrical power to the deviations in rotating speed and mechanical torques. The relationship between net accelerating power and the electrical and mechanical powers is

$$P_{\text{net}} = P_{\text{mech}} - P_{\text{elec}} \tag{10.8}$$

which is written as the sum of the steady-state value and the deviation term,

$$P_{\text{net}} = P_{\text{net}_0} + \Delta P_{\text{net}} \tag{10.9}$$

where

$$P_{\text{net}_0} = P_{\text{mech}_0} - P_{\text{elec}_0}$$
$$\Delta P_{\text{net}} = \Delta P_{\text{mech}} - \Delta P_{\text{elec}}$$

Then

$$P_{\text{net}} = (P_{\text{mech}_0} - P_{\text{elec}_0}) + (\Delta P_{\text{mech}} - \Delta P_{\text{elec}}) \tag{10.10}$$

Similarly for torques,

$$T_{\text{net}} = (T_{\text{mech}_0} - T_{\text{elec}_0}) + (\Delta T_{\text{mech}} - \Delta T_{\text{elec}}) \tag{10.11}$$

Using Equation 10.3, we can see that

$$P_{\text{net}} = P_{\text{net}_0} + \Delta P_{\text{net}} = (\omega_0 + \Delta\omega)(T_{\text{net}_0} + \Delta T_{\text{net}}) \tag{10.12}$$

www.mepcafe.com

Substituting Equations 10.10 and 10.11, we obtain

$$\left(P_{\text{mech}_0} - P_{\text{elec}_0}\right) + \left(\Delta P_{\text{mech}} - \Delta P_{\text{elec}}\right) = \left(\omega_0 + \Delta\omega\right)\left[\left(T_{\text{mech}_0} - T_{\text{elec}_0}\right) + \left(\Delta T_{\text{mech}} - \Delta T_{\text{elec}}\right)\right] \tag{10.13}$$

Assume that the steady-state quantities can be factored out since

$$P_{\text{mech}_0} = P_{\text{elec}_0}$$

and

$$T_{\text{mech}_0} = T_{\text{elec}_0}$$

and further assume that the second-order terms involving products of $\Delta\omega$ with ΔT_{mech} and ΔT_{elec} can be neglected. Then

$$\Delta P_{\text{mech}} - \Delta P_{\text{elec}} = \omega_0\left(\Delta T_{\text{mech}} - \Delta T_{\text{elec}}\right) \tag{10.14}$$

As shown in Equation 10.7, the net torque is related to the speed change as follows:

$$\left(T_{\text{mech}_0} - T_{\text{elec}_0}\right) + \left(\Delta T_{\text{mech}} - \Delta T_{\text{elec}}\right) = I\frac{\mathrm{d}}{\mathrm{d}t}(\Delta\omega) \tag{10.15}$$

Then, since $T_{\text{mech}_0} = T_{\text{elec}_0}$, we can combine Equations 10.14 and 10.15 to get

$$\begin{aligned}\Delta P_{\text{mech}} - \Delta P_{\text{elec}} &= \omega_0 I\frac{\mathrm{d}}{\mathrm{d}t}(\Delta\omega)\\ &= M\frac{\mathrm{d}}{\mathrm{d}t}(\Delta\omega)\end{aligned} \tag{10.16}$$

This can be expressed in Laplace transform operator notation as

$$\Delta P_{\text{mech}} - \Delta P_{\text{elec}} = Ms\Delta\omega \tag{10.17}$$

This is shown in block diagram form in Figure 10.3.

The units for M are watts per radian per second. We will always use per unit power over per unit speed per second where the per unit refers to the machine rating as the base (see Example 10A).

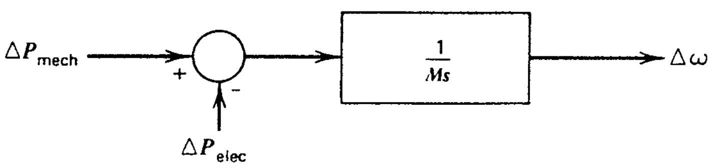


FIGURE 10.3 Relationship between mechanical and electrical power and speed change.

www.mepcafe.com

10.3 LOAD MODEL

The loads on a power system consist of a variety of electrical devices. Some of them are purely resistive, some are motor loads with variable power–frequency characteristics, and others exhibit quite different characteristics. Since motor loads are a dominant part of the electrical load, there is a need to model the effect of a change in frequency on the net load drawn by the system. The relationship between the change in load due to the change in frequency is given by

$$\Delta P_{\mathrm{L(freq)}} = D\Delta\omega \quad \text{or} \quad D = \frac{\Delta P_{\mathrm{L(freq)}}}{\Delta\omega}$$

where D is expressed as percent change in load divided by percent change in frequency. For example, if load changed by 1.5% for a 1% change in frequency, then D would equal 1.5. However, the value of D used in solving for system dynamic response must be changed if the system base MVA is different from the nominal value of load. Suppose the D referred to here was for a net connected load of 1200 MVA and the entire dynamics problem were to be set up for a 1000-MVA system base. Note that $D = 1.5$ tells us that the load would change by 1.5 pu for 1 pu change in frequency. That is, the load would change by 1.5×1200 MVA or 1800 MVA for a 1 pu change in frequency. When expressed on a 1000-MVA base, D becomes

$$D_{1000-\mathrm{MVA\ base}} = 1.5 \times \left(\frac{1200}{1000}\right) = 1.8$$

The net change in P_{elec} in Figure 10.3 (Equation 10.15) is

$$\Delta P_{\mathrm{elec}} = \underbrace{\Delta P_{\mathrm{L}}}_{\substack{\text{Nonfrequency-sensitive} \\ \text{load change}}} + \underbrace{D\Delta\omega}_{\substack{\text{Frequency-sensitive} \\ \text{load change}}} \tag{10.18}$$

Including this in the block diagram results in the new block diagram shown in Figure 10.4.

Example 10A: We are given an isolated power system with a 600-MVA generating unit having an M of 7.6 pu MW/pu frequency/sec on a machine base. The unit is supplying a load of 400 MVA. The load changes by 2% for a 1% change in frequency. First, we will set up the block diagram of the equivalent generator load system. Everything will be referenced to a 100 MVA base.

$$M = 7.6 \times \frac{600}{1000} = 4.56 \text{ on a 1000-MVA base}$$

$$D = 2 \times \frac{400}{1000} = 0.8 \text{ on a 1000-MVA base}$$

Then the block diagram is as shown in Figure 10.5.

www.mepcafe.com

Suppose the load suddenly increases by 10 MVA (or 0.01 pu); that is,

$$\Delta P_L(s) = \frac{0.01}{s}$$

then

$$\Delta \omega(s) = -\frac{0.01}{s}\left(\frac{1}{4.56s + 0.8}\right)$$

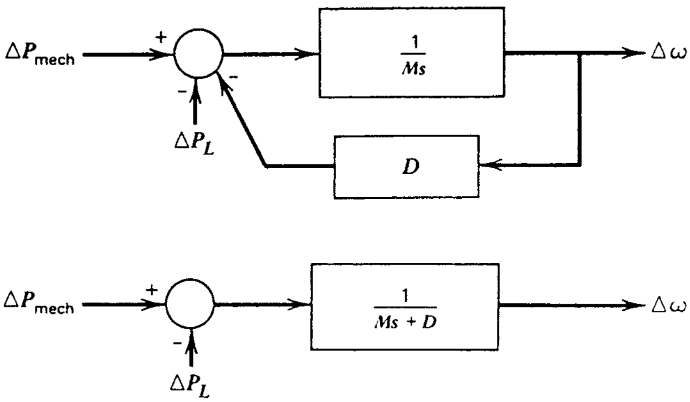


FIGURE 10.4 Block diagram of rotating mass and load as seen by prime-mover output.

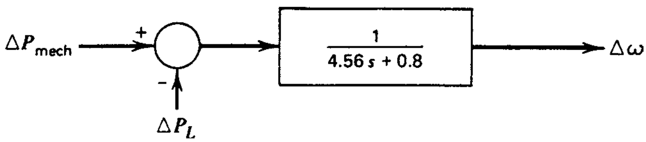


FIGURE 10.5 Block diagram for system in Example 10A.

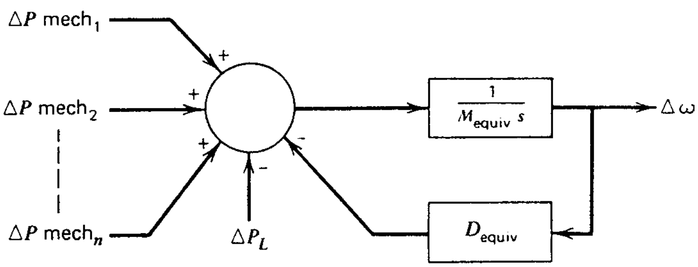


FIGURE 10.6 Multi-turbine-generator system equivalent.

www.mepcafe.com

or taking the inverse Laplace transform,

$$\begin{aligned}\Delta\omega(t) &= (0.01/0.8)e^{-(0.8/4.56)t} - (0.01/0.8)\\ &= 0.0125e^{-0.175t} - 0.0125\end{aligned}$$

The final value of $\Delta\omega$ is -0.0125 pu, which is a drop of 0.75 Hz on a 60-Hz system.

When two or more generators are connected to a transmission system network, we must take account of the phase angle difference across the network in analyzing frequency changes. However, for the sake of governor analysis, which we are interested in here, we can assume that frequency will be constant over those parts of the network that are tightly interconnected. When making such an assumption, we can then lump the rotating mass of the turbine generators together into an equivalent that is driven by the sum of the individual turbine mechanical outputs. This is illustrated in Figure 10.6, where all turbine generators were lumped into a single equivalent rotating mass, M_{equiv}.

Similarly, all individual system loads were lumped into an equivalent load with damping coefficient, D_{equiv}.

10.4 PRIME-MOVER MODEL

The prime mover driving a generator unit may be a steam turbine or a hydro turbine. The models for the prime mover must take account of the steam supply and boiler control system characteristics in the case of a steam turbine, or the penstock characteristics for a hydro turbine. Throughout the remainder of this chapter, only the simplest prime-mover model, the nonreheat turbine, will be used. The models for other more complex prime movers, including hydro turbines, are developed in the references (see Reference List).

The model for a nonreheat turbine, shown in Figure 10.7, relates the position of the valve that controls emission of steam into the turbine to the power output of the turbine, where

$$\begin{aligned}T_{\text{CH}} &= \text{“charging time” time constant}\\ \Delta P_{\text{valve}} &= \text{per unit change in valve position from nominal}\end{aligned}$$

The combined prime-mover–generator-load model for a single generating unit can be built by combining Figure 10.4 and Figure 10.7, as shown in Figure 10.8.

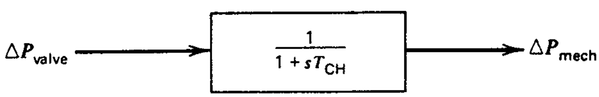


FIGURE 10.7 Prime-mover model.

www.mepcafe.com

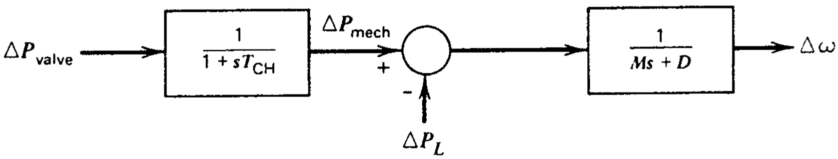


FIGURE 10.8 Prime-mover–generator-load model.

10.5 GOVERNOR MODEL

Suppose a generating unit is operated with fixed mechanical power output from the turbine. The result of any load change would be a speed change sufficient to cause the frequency-sensitive load to exactly compensate for the load change (as in Example 10A). This condition would allow system frequency to drift far outside acceptable limits. This is overcome by adding a governing mechanism that senses the machine speed, and adjusts the input valve to change the mechanical power output to compensate for load changes and to restore frequency to nominal value. The earliest such mechanism used rotating "flyballs" to sense speed and to provide mechanical motion in response to speed changes. Modern governors use electronic means to sense speed changes and often use a combination of electronic, mechanical, and hydraulic means to effect the required valve position changes. The simplest governor, called the *isochronous governor*, adjusts the input valve to a point that brings frequency back to nominal value. If we simply connect the output of the speed-sensing mechanism to the valve through a direct linkage, it would never bring the frequency to nominal. To force the frequency error to 0, one must provide what control engineers call reset action. Reset action is accomplished by integrating the frequency (or speed) error, which is the difference between actual speed and desired or reference speed.

We will illustrate such a speed-governing mechanism with the diagram shown in Figure 10.9. The speed-measurement device's output, ω, is compared with a reference, ω_{ref}, to produce an error signal, $\Delta\omega$. The error, $\Delta\omega$, is negated and then amplified by a gain K_G and integrated to produce a control signal, ΔP_{valve}, which causes the main steam supply valve to open (ΔP_{valve} position) when $\Delta\omega$ is negative. If, for example, the machine is running at reference speed and the electrical load increases, ω will fall below ω_{ref} and $\Delta\omega$ will be negative. The action of the gain and integrator will be to open the steam valve, causing the turbine to increase its mechanical output, thereby increasing the electrical output of the generator and increasing the speed ω. When ω exactly equals ω_{ref}, the steam valve stays at the new position (further opened) to allow the turbine generator to meet the increased electrical load.

The isochronous (constant speed) governor of Figure 10.9 cannot be used if two or more generators are electrically connected to the same system since each generator would have to have precisely the same speed setting or they would "fight" each other, each trying to pull the system's speed (or frequency) to its own setting. To be able to run two or more generating units in parallel on a generating system, the

www.mepcafe.com

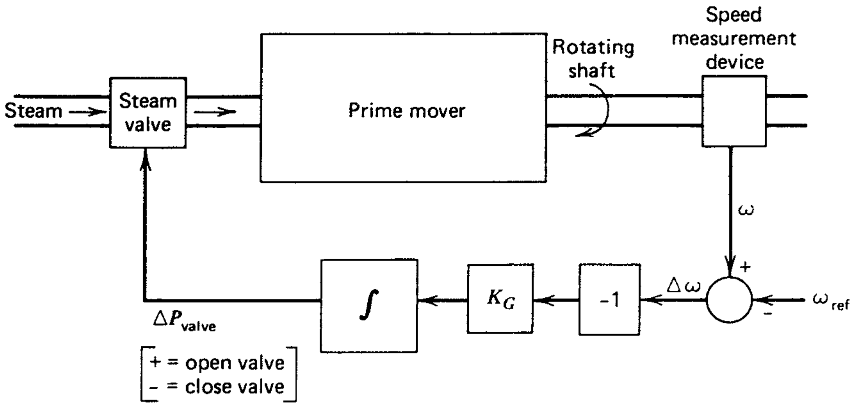


FIGURE 10.9 Isochronous governor.

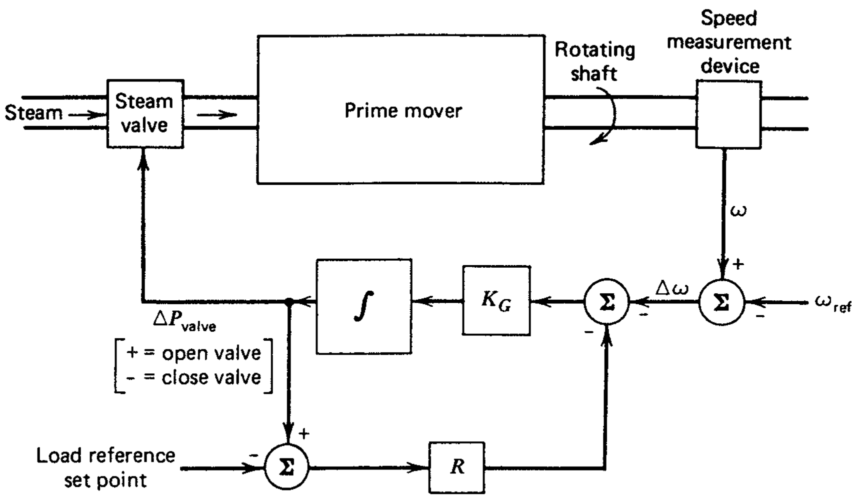


FIGURE 10.10 Governor with speed-droop feedback loop.

governors are provided with a feedback signal that causes the speed error to go to 0 at different values of generator output.

This can be accomplished by adding a feedback loop around the integrator as shown in Figure 10.10. Note that we have also inserted a new input, called the *load reference*, which we will discuss shortly. The block diagram for this governor is shown in Figure 10.11, where the governor now has a net gain of $1/R$ and a time constant T_G.

The result of adding the feedback loop with gain R is a governor characteristic as shown in Figure 10.12. The value of R determines the slope of the characteristic. That is, R determines the change on the unit's output for a given change in frequency. Common practice is to set R on each generating unit so that a change from 0 to 100%

www.mepcafe.com

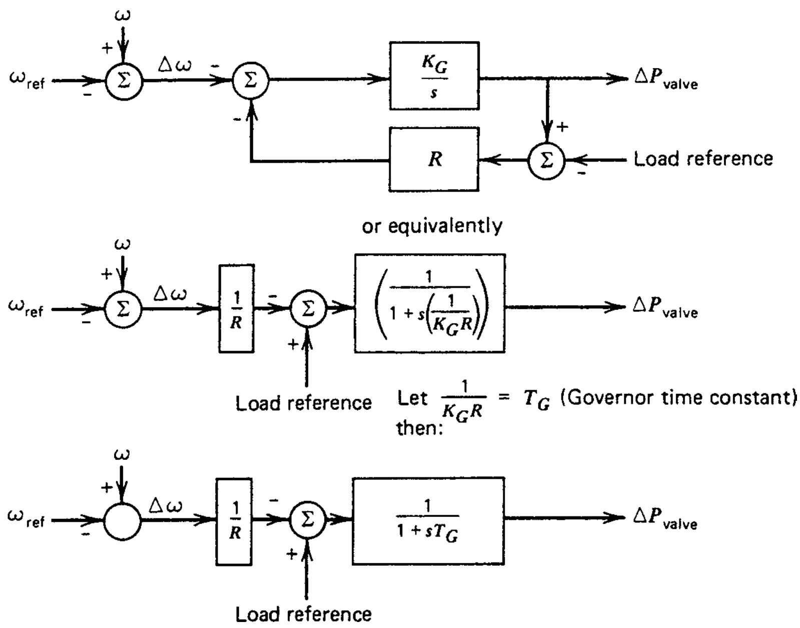


FIGURE 10.11 Block diagram of governor with droop.

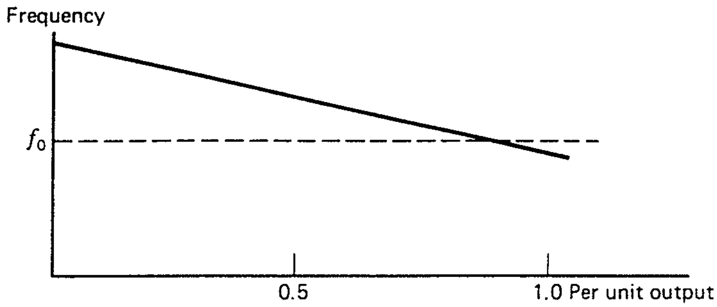


FIGURE 10.12 Speed-droop characteristic.

(i.e., rated) output will result in the same frequency change for each unit. As a result, a change in electrical load on a system will be compensated by generator unit output changes proportional to each unit's rated output.

If two generators with drooping governor characteristics are connected to a power system, there will always be a unique frequency, at which they will share a load change between them. This is illustrated in Figure 10.13, showing two units with drooping characteristics connected to a common load.

As shown in Figure 10.13, the two units start at a nominal frequency of f_0. When a load increase, ΔP_L, causes the units to slow down, the governors increase output until the units seek a new, common operating frequency, f'. The amount of load pickup on each unit is proportional to the slope of its droop characteristic. Unit 1

www.mepcafe.com

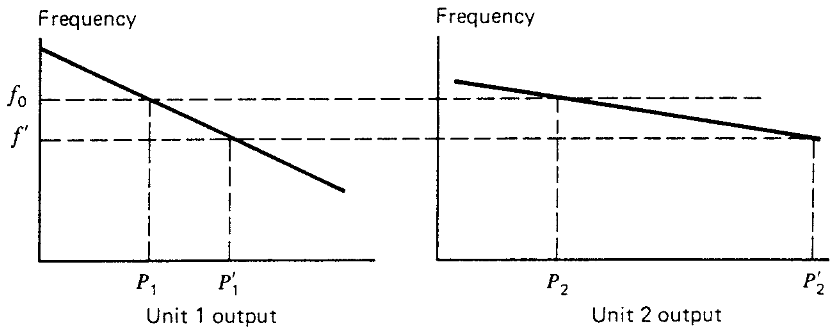


FIGURE 10.13 Allocation of unit outputs with governor droop.

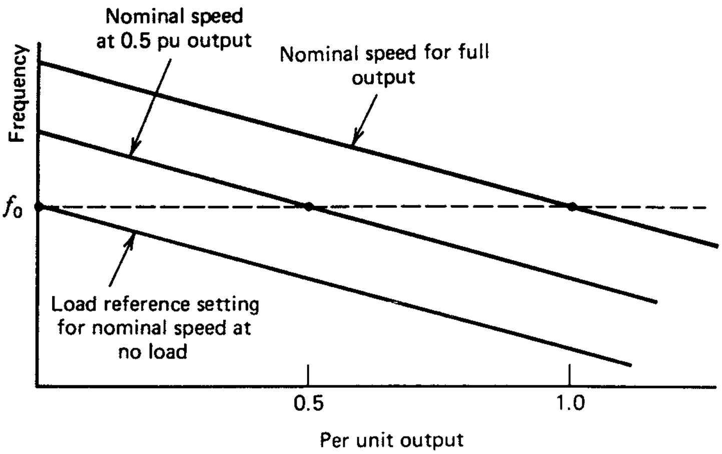


FIGURE 10.14 Speed-changer settings.

increases its output from P_1 to P_1', unit 2 increases its output from P_2 to P_2' such that the net generation increase, $P_1' - P_1 + P_2' - P_2$, is equal to ΔP_L. Note that the actual frequency sought also depends on the load's frequency characteristic as well. The use of power electronic controls in factories and in residences has decreased the inertia contribution from the demand. Thus, the system inertia has decreased. Also, newer turbines are reduced in weight, further reducing the system inertia.

Figure 10.10 shows an input labeled "load reference set point." By changing the load reference, the generator's governor characteristic can be set to give reference frequency at any desired unit output. This is illustrated in Figure 10.14. *The basic control input to a generating unit as far as generation control is concerned is the load reference set point*. By adjusting this set point on each unit, a desired unit dispatch can be maintained while holding system frequency close to the desired nominal value. This is the boundary between response capability and economic allocation by economic dispatch.

Note that a steady-state change in ΔP_{valve} of 1.0 pu requires a value of R pu change in frequency, $\Delta\omega$. One often hears unit regulation referred to in percent. For instance, a 3% regulation for a unit would indicate that a 100% (1.0 pu) change in valve position

www.mepcafe.com

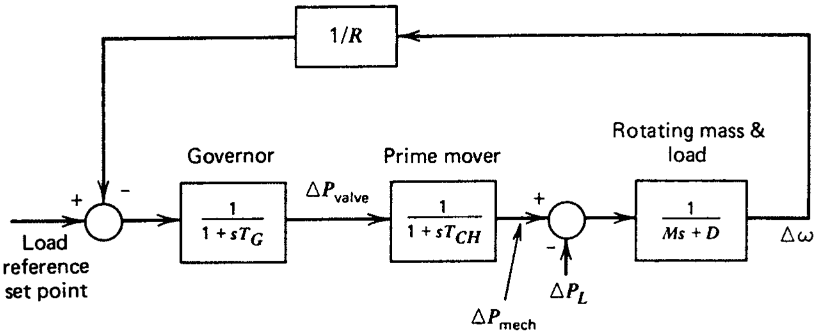


FIGURE 10.15 Block diagram of governor, prime mover, and rotating mass.

(or equivalently a 100% change in unit output) requires a 3% change in frequency. Therefore, R is equal to pu change in frequency divided by pu change in unit output. That is,

$$R = \frac{\Delta\omega}{\Delta P} \text{ pu}$$

At this point, we can construct a block diagram of a governor–prime–mover–rotating mass/load model as shown in Figure 10.15. Suppose that this generator experiences a step increase in load,

$$\Delta P_L(s) = \frac{\Delta P_L}{s} \tag{10.19}$$

The transfer function relating the load change, ΔP_L, to the frequency change, $\Delta\omega$, is

$$\Delta\omega(s) = \Delta P_L(s)\left[\frac{\dfrac{-1}{Ms+D}}{1+\dfrac{1}{R}\left(\dfrac{1}{1+sT_G}\right)\left(\dfrac{1}{1+sT_{CH}}\right)\left(\dfrac{1}{Ms+D}\right)}\right] \tag{10.20}$$

The steady-state value of $\Delta\omega(s)$ may be found by

$$\begin{aligned}\Delta\omega \text{ steady state} &= \lim_{s\to 0}[s\Delta\omega(s)] \\ &= \frac{-\Delta P_L\left(\dfrac{1}{D}\right)}{1+\left(\dfrac{1}{R}\right)\left(\dfrac{1}{D}\right)} = \frac{-\Delta P_L}{\dfrac{1}{R}+D}\end{aligned} \tag{10.21}$$

Note that if D were 0, the change in speed would simply be

$$\Delta\omega = -R\Delta P_L \tag{10.22}$$

www.mepcafe.com

If several generators (each having its own governor and prime mover) were connected to the system, the frequency change would be

$$\Delta\omega = \frac{-\Delta P_L}{\dfrac{1}{R_1} + \dfrac{1}{R_2} + \cdots + \dfrac{1}{R_n} + D} \tag{10.23}$$

10.6 TIE-LINE MODEL

The power flowing across a transmission line can be modeled using the DC load flow method shown in Chapter 6.

$$P_{\text{tie flow}} = \frac{1}{X_{\text{tie}}}(\theta_1 - \theta_2) \tag{10.24}$$

This tie flow is a steady-state quantity. For purposes of analysis here, we will perturb Equation 10.24 to obtain deviations from nominal flow as a function of deviations in phase angle from nominal.

$$\begin{aligned} P_{\text{tie flow}} + \Delta P_{\text{tie flow}} &= \frac{1}{X_{\text{tie}}}[(\theta_1 + \Delta\theta_1) - (\theta_2 + \Delta\theta_2)] \\ &= \frac{1}{X_{\text{tie}}}(\theta_1 - \theta_2) + \frac{1}{X_{\text{tie}}}(\Delta\theta_1 - \Delta\theta_2) \end{aligned} \tag{10.25}$$

Then

$$\Delta P_{\text{tie flow}} = \frac{1}{X_{\text{tie}}}(\Delta\theta_1 - \Delta\theta_2) \tag{10.26}$$

where $\Delta\theta_1$ and $\Delta\theta_2$ are equivalent to $\Delta\delta_1$ and $\Delta\delta_2$ as defined in Equation 10.6. Then, using the relationship of Equation 10.6,

$$\Delta P_{\text{tie flow}} = \frac{T}{s}(\Delta\omega_1 - \Delta\omega_2) \tag{10.27}$$

where $T = 377 \times 1/X_{\text{tie}}$ (for a 60-Hz system).

Note that $\Delta\theta$ must be in radians for ΔP_{tie} to be in per unit megawatts, but $\Delta\omega$ is in per unit speed change. Therefore, we must multiply $\Delta\omega$ by 377 rad/s (the base frequency in rad/s at 60 Hz). T may be thought of as the "tie-line stiffness" coefficient.

Suppose now that we have an interconnected power system broken into two areas each having one generator. The areas are connected by a single transmission line. The power flow over the transmission line will appear as a positive load to one area and an equal but negative load to the other, or vice versa, depending on the direction of flow. The direction

www.mepcafe.com

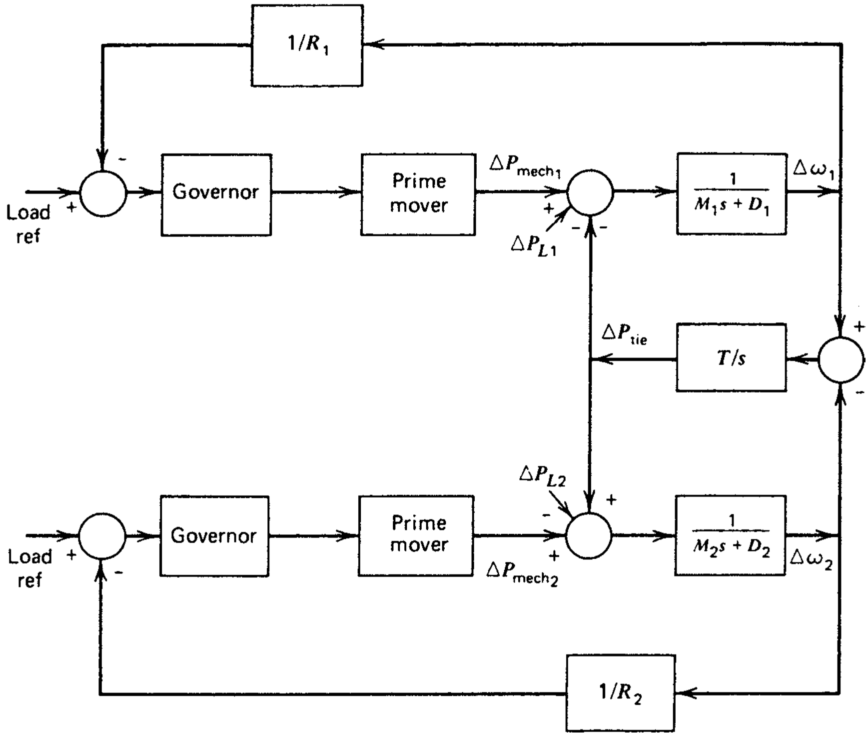


FIGURE 10.16 Block diagram of interconnected areas.

of flow will be dictated by the relative phase angle between the areas, which is determined by the relative speed deviations in the areas. A block diagram representing this interconnection can be drawn as in Figure 10.16. Note that the tie power flow was defined as going from area 1 to area 2; therefore, the flow appears as a load to area 1 and a power source (negative load) to area 2. If one assumes that mechanical powers are constant, the rotating masses and tie line exhibit damped oscillatory characteristics known as synchronizing oscillations. (See Problem 10.3 at the end of this chapter.)

It is quite important to analyze the steady-state frequency deviation, tie-flow deviation, and generator outputs for an interconnected area after a load change occurs. Let there be a load change ΔP_{L_1} in area 1. In the steady state, after all synchronizing oscillations have damped out, the frequency will be constant and equal to the same value on both areas. Then

$$\Delta\omega_1 = \Delta\omega_2 = \Delta\omega \quad \text{and} \quad \frac{d(\Delta\omega_1)}{dt} = \frac{d(\Delta\omega_2)}{dt} = 0 \tag{10.28}$$

and

$$\Delta P_{\text{mech}_1} - \Delta P_{\text{tie}} - \Delta P_{L_1} = \Delta\omega D_1$$

www.mepcafe.com

$$\Delta P_{\text{mech}_2} + \Delta P_{\text{tie}} = \Delta\omega D_2$$

$$\Delta P_{\text{mech}_1} = \frac{-\Delta\omega}{R_1} \tag{10.29}$$

$$\Delta P_{\text{mech}_2} = \frac{-\Delta\omega}{R_2}$$

By making appropriate substitutions in Equation 10.29,

$$\begin{aligned} -\Delta P_{\text{tie}} - \Delta P_{\text{L}_1} &= \Delta\omega\left(\frac{1}{R_1} + D_1\right) \\ +\Delta P_{\text{tie}} &= \Delta\omega\left(\frac{1}{R_2} + D_2\right) \end{aligned} \tag{10.30}$$

or, finally,

$$\Delta\omega = \frac{-\Delta P_{\text{L}_1}}{\frac{1}{R_1} + \frac{1}{R_2} + D_1 + D_2} \tag{10.31}$$

from which we can derive the change in tie flow:

$$\Delta P_{\text{tie}} = \frac{-\Delta P_{\text{L}_1}\left(\frac{1}{R_2} + D_2\right)}{\frac{1}{R_1} + \frac{1}{R_2} + D_1 + D_2} \tag{10.32}$$

Note that the conditions described in Equations 10.28 through 10.32 are for the new steady-state conditions after the load change. The new tie flow is determined by the net change in load and generation in each area. We do not need to know the tie stiffness to determine this new tie flow, although the tie stiffness will determine how much difference in phase angle across the tie will result from the new tie flow.

Example 10B: You are given two system areas connected by a tie line with the following characteristics:

Area 1	Area 2
$R = 0.01$ pu	$R = 0.02$ pu
$D = 0.8$ pu	$D = 1.0$ pu
Base MVA = 500	Base MVA = 500

www.mepcafe.com

A load change of 100 MW (0.2 pu) occurs in area 1. What is the new steady-state frequency and what is the change in tie flow? Assume both areas were at nominal frequency (60 Hz) to begin.

$$\Delta\omega = \frac{-\Delta P_{L_1}}{\frac{1}{R_1}+\frac{1}{R_2}+D_1+D_2} = \frac{-0.2}{\frac{1}{0.01}+\frac{1}{0.02}+0.8+1} = -0.00131752 \text{ pu}$$

$$f_{new} = 60 - 0.00132(60) = 59.92 \text{ Hz}$$

$$\Delta P_{tie} = \Delta\omega\left(\frac{1}{R_2}+D_2\right) = -0.00131752\left(\frac{1}{0.02}+1\right) = -0.06719368 \text{ pu}$$

$$= -33.6 \text{ MW}$$

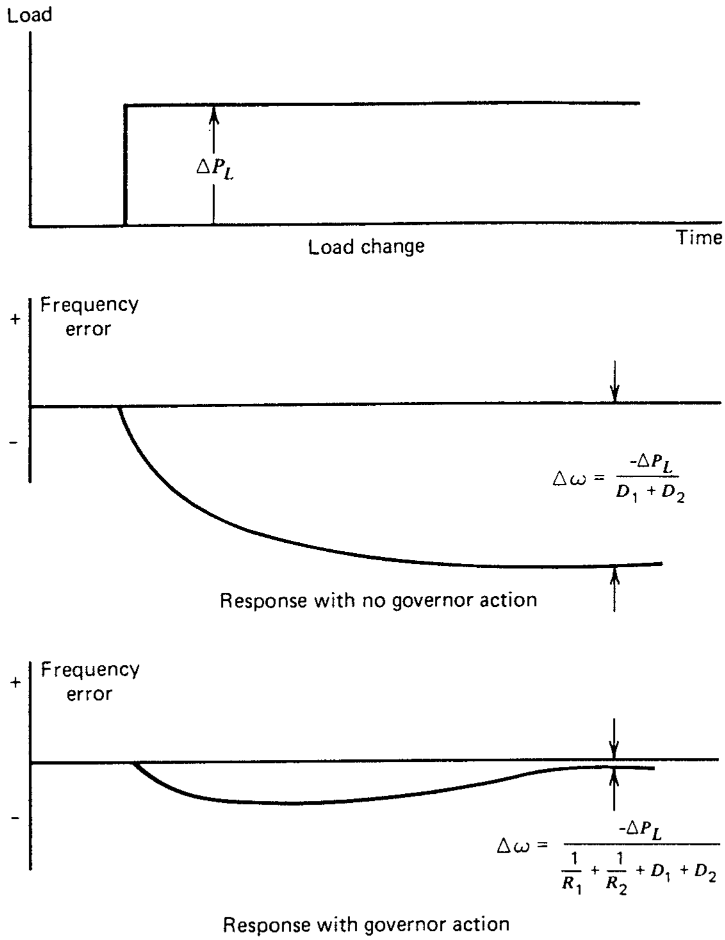


FIGURE 10.17 Frequency response to load change.

www.mepcafe.com

The change in prime-mover power would be

$$\Delta P_{\text{mech}_1} = \frac{-\Delta\omega}{R_1} = -\left(\frac{-0.00131752}{0.01}\right) = 0.13175231 \text{ pu} = 65.876\,\text{MW}$$
$$\Delta P_{\text{mech}_2} = \frac{-\Delta\omega}{R_2} = -\left(\frac{-0.00131752}{0.02}\right) = 0.06587615 \text{ pu} = 32.938\,\text{MW}$$
$$= 98.814\,\text{MW}$$

The total changes in generation is 98.814 MA, which is 1.186 MW short of the 100 MW load change. The change in total area load due to frequency drop would be

$$\text{For area } 1 = \Delta\omega D_1 = -0.0010540\,\text{pu} = -0.527\,\text{MW}$$
$$\text{For area } 2 = \Delta\omega D_2 = -0.00131752\,\text{pu} = -0.6588\,\text{MW}$$

Therefore, the total load change = 1.186 MW, which accounts for the difference in total generation change and total load change. (See Problem 10.2 for further variations on this problem.)

If we were to analyze the dynamics of the two-area systems, we would find that a step change in load would always result in a frequency error. This is illustrated in Figure 10.17, which shows the frequency response of the system to a step-load change. Note that Figure 10.17 only shows the average frequency (omitting any high-frequency oscillations).

10.7 GENERATION CONTROL

Automatic generation control (AGC) is the name given to a control system having three major objectives:

1. To hold system frequency at or very close to a specified nominal value (e.g., 50 or 60 Hz).
2. To maintain the correct value of interchange power between control areas.
3. To maintain each unit's generation at the most economic value.

10.7.1 Supplementary Control Action

To understand each of the three objectives just listed, we may start out assuming that we are studying a single generating unit supplying load to an isolated power system. As shown in Section 10.5, a load change will produce a frequency change with a magnitude that depends on the droop characteristics of the governor and the frequency characteristics of the system load. Once a load change has occurred, a supplementary control must act to restore the frequency to nominal value. This

www.mepcafe.com

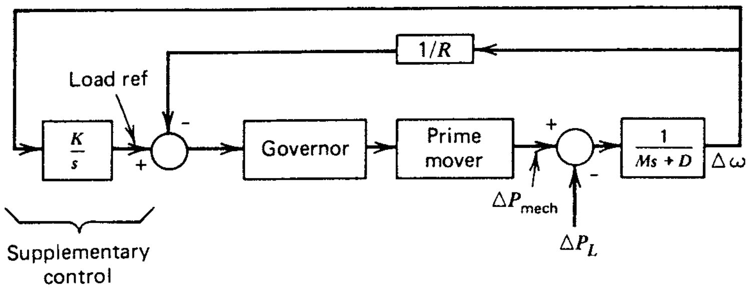


FIGURE 10.18 Supplementary control added to generating unit.

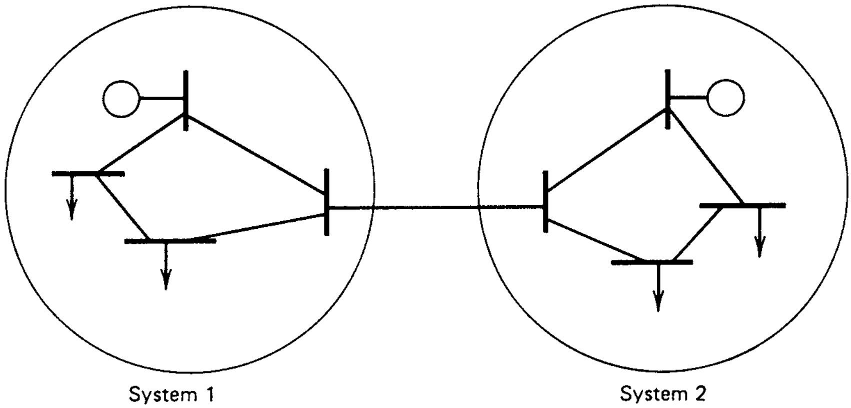


FIGURE 10.19 Two-area system.

can be accomplished by adding a reset (integral) control to the governor, as shown in Figure 10.18. This book examines the use of integral control while a more complete use of proportional, integral, and derivative (PID) is often applied. It is noted that adaptive state space controllers have been research but not commonly implemented.

The reset control action of the supplementary control will force the frequency error to 0 by adjustment of the speed reference set point. For example, the error shown in the bottom diagram of Figure 10.17 would be forced to 0.

10.7.2 Tie-Line Control

When two utilities interconnect their systems, they do so for several reasons. One is to be able to buy and sell power with neighboring systems whose operating costs make such transactions profitable as analyzed in Chapter 11. Further, even if no power is being transmitted over ties to neighboring systems, if one system has a sudden loss of a generating unit, the units throughout all the interconnection will experience a frequency change and can help in restoring frequency.

www.mepcafe.com

Interconnections present a very interesting control problem with respect to allocation of generation to meet load. The hypothetical situation in Figure 10.19 will be used to illustrate this problem. Assume both systems in Figure 10.19 have equal generation and load characteristics ($R_1 = R_2$, $D_1 = D_2$) and, further, assume system 1 was sending 100 MW to system 2 under an interchange agreement made between the operators of each system. Now, let system 2 experience a sudden load increase of 30 MW. Since both units have equal generation characteristics, they will both experience a 15 MW increase, and the tie line will experience an increase in flow from 100 MW to 115 MW. Thus, the 30 MW load increase in system 2 will have been satisfied by a 15 MW increase in generation in system 2, plus a 15 MW increase in tie flow into system 2. This would be fine, except that system 1 contracted to sell only 100 MW, not 115 MW, and its generating costs have just gone up without anyone to bill the extra cost as presented in Chapter 11. What is needed at this point is a control scheme that recognizes the fact that the 30 MW load increase occurred in system 2 and, therefore, would increase generation in system 2 by 30 MW while restoring frequency to nominal value. It would also restore generation in system 1 to its output before the load increase occurred.

Such a control system must use two pieces of information: the system frequency and the net power flowing in or out over the tie lines. Such a control scheme would, of necessity, have to recognize the following:

1. If frequency decreased and net interchange power leaving the system increased, a load increase has occurred outside the system.
2. If frequency decreased and net interchange power leaving the system decreased, a load increase has occurred inside the system.

This can be extended to cases where frequency increases. We will make the following definitions:

$$
\begin{aligned}
P_{\text{net int}} &= \text{total actual net interchange} \\
&\quad (+\text{for power leaving the system}; -\text{for power entering}) \qquad (10.33) \\
P_{\text{net int sched}} &= \text{scheduled or desired value of interchange}
\end{aligned}
$$

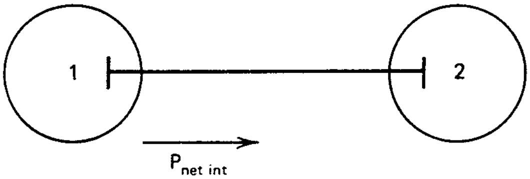
$$\Delta P_{\text{net int}} = P_{\text{net int}} - P_{\text{net int sched}}$$

Then, a summary of the tie-line frequency control scheme can be given as in the table in Figure 10.20.

We define a *control area* to be a part of an interconnected system within which the load and generation will be controlled as per the rules in Figure 10.20. The control area's boundary is simply the tie-line points where power flow is metered. All tie lines crossing the boundary must be metered so that total control area net interchange power can be calculated.

www.mepcafe.com

$\Delta\omega$	$\Delta P_{\text{net int}}$	Load change		Resulting control action
−	−	ΔP_{L_1}	+	Increase P_{gen} in system 1
		ΔP_{L_2}	0	
+	+	ΔP_{L_1}	−	Decrease P_{gen} in system 1
		ΔP_{L_2}	0	
−	+	ΔP_{L_1}	0	Increase P_{gen} in system 2
		ΔP_{L_2}	+	
+	−	ΔP_{L_1}	0	Decrease P_{gen} in system 2
		ΔP_{L_2}	−	



ΔP_{L_1} = Load change in area 1
ΔP_{L_2} = Load change in area 2

FIGURE 10.20 Tie-line frequency control actions for two-area system.

The rules set forth in Figure 10.20 can be implemented by a control mechanism that weighs frequency deviation, $\Delta\omega$, and net interchange power, $\Delta P_{\text{net int}}$. The frequency response and tie flows resulting from a load change, ΔP_{L_1}, in the two-area system of Figure 10.16 are derived in Equations 10.28 through 10.32. These results are repeated here.

Load Change	Frequency Change	Change in Net Interchange
ΔP_{L_1}	$\Delta\omega = \dfrac{-\Delta P_{L_1}}{\dfrac{1}{R_1}+\dfrac{1}{R_2}+D_1+D_2}$	$\Delta P_{\text{net int}_1} = \dfrac{-\Delta P_{L_1}\left(\dfrac{1}{R_2}+D_2\right)}{\dfrac{1}{R_1}+\dfrac{1}{R_2}+D_1+D_2}$

(10.34)

This corresponds to the first row of the table in Figure 10.20; we would therefore require that

$$\Delta P_{\text{gen}_1} = \Delta P_{L_1}$$
$$\Delta P_{\text{gen}_2} = 0$$

www.mepcafe.com

The required change in generation, historically called the *area control error* or ACE, represents the shift in the area's generation required to restore frequency and net interchange to their desired values. The equations for ACE for each area are

$$
\begin{aligned}
\text{ACE}_1 &= -\Delta P_{\text{net int}_1} - B_1 \Delta\omega \\
\text{ACE}_2 &= -\Delta P_{\text{net int}_2} - B_2 \Delta\omega
\end{aligned}
\tag{10.35}
$$

where B_1 and B_2 are called *frequency bias factors*. We can see from Equation 10.34 that setting bias factors as follows:

$$
\begin{aligned}
B_1 &= \left(\frac{1}{R_1} + D_1\right) \\
B_2 &= \left(\frac{1}{R_2} + D_2\right)
\end{aligned}
\tag{10.36}
$$

results in

$$
\begin{aligned}
\text{ACE}_1 &= \left(\frac{+\Delta P_{L_1}\left(\dfrac{1}{R_2} + D_2\right)}{\dfrac{1}{R_1} + \dfrac{1}{R_2} + D_1 + D_2}\right) - \left(\frac{1}{R_1} + D_1\right)\left(\frac{-\Delta P_{L_1}}{\dfrac{1}{R_1} + \dfrac{1}{R_2} + D_1 + D_2}\right) = \Delta P_{L_1} \\
\text{ACE}_2 &= \left(\frac{-\Delta P_{L_1}\left(\dfrac{1}{R_2} + D_2\right)}{\dfrac{1}{R_1} + \dfrac{1}{R_2} + D_1 + D_2}\right) - \left(\frac{1}{R_2} + D_2\right)\left(\frac{-\Delta P_{L_1}}{\dfrac{1}{R_1} + \dfrac{1}{R_2} + D_1 + D_2}\right) = 0
\end{aligned}
$$

This control can be carried out using the scheme outlined in Figure 10.21. Note that the values of B_1 and B_2 would have to change each time a unit was committed or decommitted, in order to have the exact values as given in Equation 10.36. Actually, the integral action of the supplementary controller will guarantee a reset of ACE to 0 even when B_1 and B_2 are not adjusted.

10.7.3 Generation Allocation

If each control area in an interconnected system had a single generating unit, the control system of Figure 10.21 would suffice to provide stable frequency and tie-line interchange. However, power systems consist of control areas with many generating units with outputs that must be set according to economics. That is, we must couple an economic dispatch calculation to the control mechanism so it will know how much of each area's total generation is required from each individual unit.

www.mepcafe.com

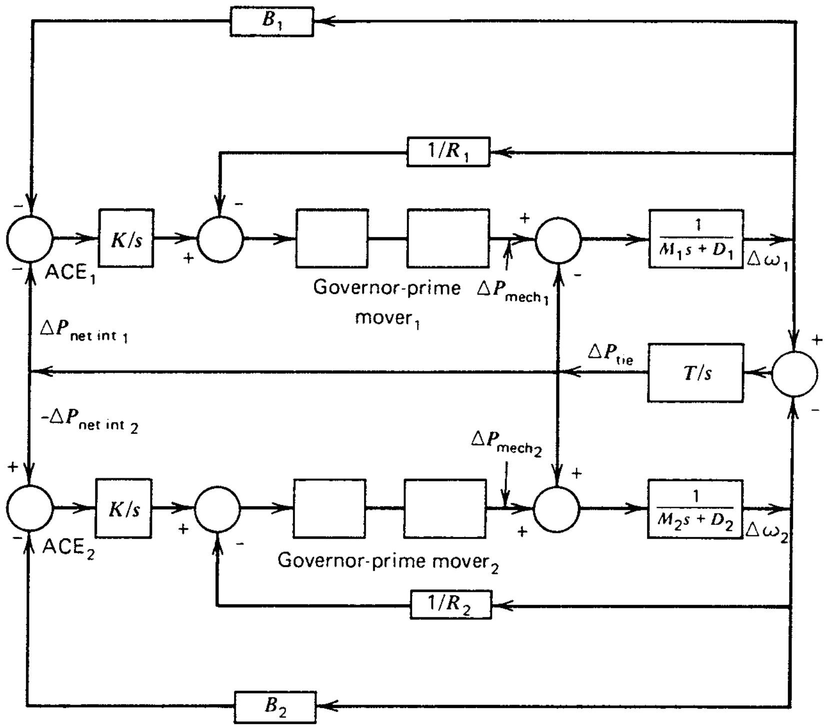


FIGURE 10.21 Tie-line bias supplementary control for two areas.

One must remember that a particular total generation value will not usually exist for a very long time, since the load on a power system varies continually as people and industries use individual electric loads. Therefore, it is impossible to simply specify a total generation, calculate the economic dispatch for each unit, and then give the control mechanism the values of megawatt output for each unit—unless such a calculation can be made very quickly. Until the widespread use of digital computer-based control systems, it was common practice to construct control mechanisms such as we have been describing using analog computers. Although analog computers are not generally proposed for new control-center installations today, there are some in active use. An analog computer can provide the economic dispatch and allocation of generation in an area on an instantaneous basis through the use of function generators set to equal the units' incremental heat rate curves. *B* matrix loss formulas were also incorporated into analog schemes by setting the matrix coefficients on precision potentiometers. The use of analog computer diagrams is just as important now as state space techniques are used for analysis and for design. It is noted that digital computer control uses state space equations to emulate the control implemented by analog computers, but with far more sophistication.

When using digital computers, it is desirable to be able to carry out the economic-dispatch calculations at intervals of one to several minutes. Either the output of the

www.mepcafe.com

economic-dispatch calculation is fed to an analog computer (i.e., a "digitally directed analog" control system) or the output is fed to another program in the computer that executes the control functions (i.e., a "direct digital" control system). Whether the control is analog or digital, the allocation of generation must be made instantly when the required area total generation changes. Since the economic-dispatch calculation is to be executed every few minutes, a means must be provided to indicate how the generation is to be allocated for values of total generation other than that used in the economic-dispatch calculation.

The allocation of individual generator output over a range of total generation values is accomplished using base points and participation factors as presented in Chapter 3. The economic-dispatch calculation is executed with a total generation equal to the sum of the present values of unit generation as measured. The result of this calculation is a set of base-point generations, $P_{i_{\text{base}}}$, which is equal to the most economic output for each generator unit. The rate of change of each unit's output with respect to a change in total generation is called the unit's *participation factor*, *pf*. The base point and participation factors are used as follows:

$$P_{i_{\text{des}}} = P_{i_{\text{base}}} + pf_i \times \Delta P_{\text{total}} \tag{10.37}$$

where

$$\Delta P_{\text{total}} = P_{\text{newtotal}} - \sum_{\substack{\text{all} \\ \text{gen}}} P_{i_{\text{base}}} \tag{10.38}$$

and

$$\begin{aligned} P_{i_{\text{des}}} &= \text{new desired output from unit } i \\ P_{i_{\text{base}}} &= \text{base-point generation for unit } i \\ pf_i &= \text{participation factor for unit } i \\ \Delta P_{\text{total}} &= \text{change in total generation} \\ P_{\text{newtotal}} &= \text{new total generation} \end{aligned}$$

Note that by definition (e.g., see Equation 3.8) the participation factors must sum to unity. In a direct digital control scheme, the generation allocation would be made by running a computer code that was programmed to execute according to Equations 10.37 and 10.38.

10.7.4 Automatic Generation Control (AGC) Implementation

Modern implementation of AGC schemes usually consists of a central location where information pertaining to the system is telemetered. Control actions are determined in a digital computer and then transmitted to the generation units via the same telemetry

www.mepcafe.com

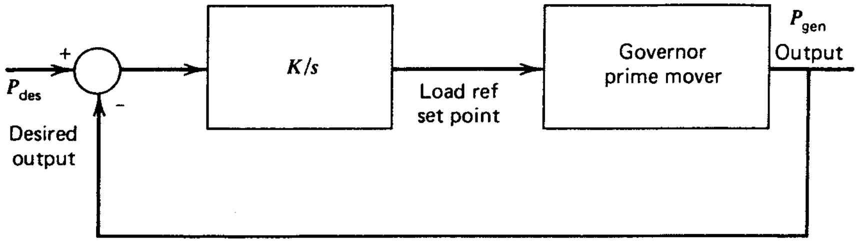


FIGURE 10.22 Basic generation control loop.

channels. To implement an AGC system, one would require the following information at the control center:

1. Unit megawatt output for each committed unit.
2. Megawatt flow over each tie line to neighboring systems.
3. System frequency.

The output of the execution of an AGC program must be transmitted to each of the generating units. The original controls were transmitted as raise or lower pulses of varying lengths to the unit. Control equipment then changes the unit's load reference set point up or down in proportion to the pulse length. The "length" of the control pulse was encoded in the bits of a digital word that is transmitted over a digital telemetry channel. The use of digital telemetry is commonplace in modern systems wherein supervisory control (opening and closing substation breakers), telemetry information (measurements of MW, MVAR, MVA voltage, etc.), and control information (unit raise/lower) is all sent via the same channels.

The basic reset control loop for a unit consists of an integrator with gain K as shown in Figure 10.22. The control loop is implemented as shown in Figure 10.23. The P_{des} control input used in Figure 10.22 and Figure 10.23 is a function of system frequency deviation, net interchange error, and each unit's deviation from its scheduled economic output.

The overall control scheme we are going to develop starts with ACE, which is a measure of the error in total generation from total desired generation. ACE is calculated according to Figure 10.24. ACE serves to indicate when total generation must be raised or lowered in a control area. However, ACE is not the only error signal that must "drive" our controller. The individual units may deviate from the economic output as determined by the base point and participation-factor calculation.

The AGC control logic must also be driven by the errors in unit output so as to force the units to obey the economic dispatch. To do this, the sum of the unit output errors is added to ACE to form a composite error signal that drives the entire control system. Such a control system is shown schematically in Figure 10.25, where we have combined the ACE calculation, the generation allocation calculation, and the unit control loop into a centralized computer system, often called as Energy Management System (EMS).

www.mepcafe.com

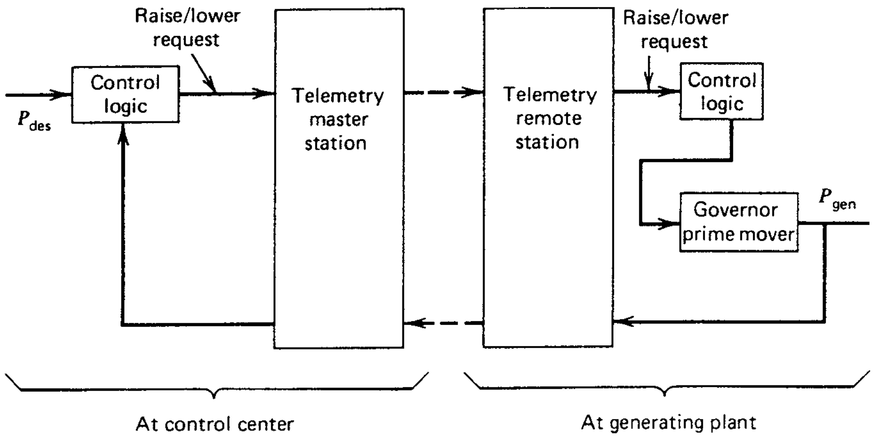


FIGURE 10.23 Basic generation control loop via telemetry.

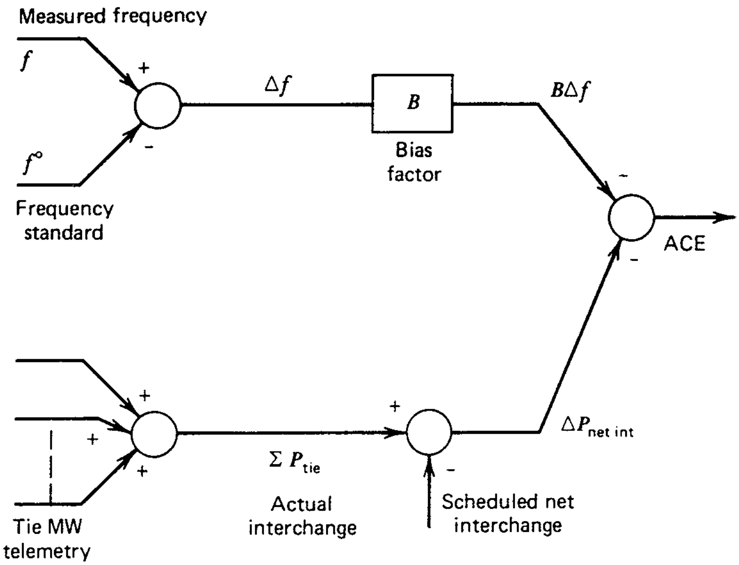


FIGURE 10.24 ACE calculation.

Investigation of Figure 10.25 shows an overall control system that will try to drive ACE to 0 as well as driving each unit's output to its required economic value. Readers are cautioned that there are many variations to the control execution shown in Figure 10.25. This is especially true of digital implementations of AGC where great sophistication can be programmed into an AGC digital implementation of the equivalent state space controller.

www.mepcafe.com

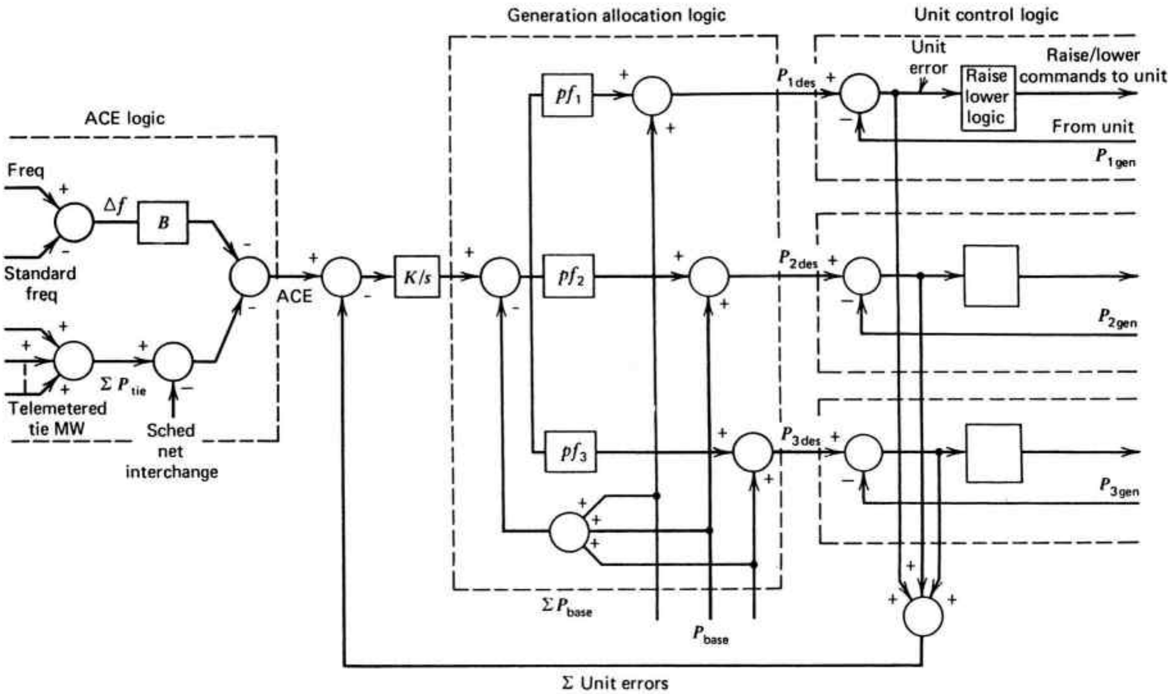


FIGURE 10.25 Overview of AGC logic.

www.mepcafe.com

Often the question is asked as to what constitutes "good" AGC design. This is difficult to answer, other than in a general way, since what is "good" for one system may be different in another. Three general criteria can be given:

1. The ACE signal should ideally be kept from becoming too large. Since ACE is directly influenced by random load variations, this criterion can be treated statistically by saying that the standard deviation of ACE should be small.
2. ACE should not be allowed to "drift." This means that the integral of ACE over an appropriate time should be small. "Drift" in ACE has the effect of creating system time errors or what are termed *inadvertent interchange errors*.
3. The amount of control action called for by the AGC should be kept to a minimum. Many of the errors in ACE, for example, are simply random load changes that need not cause control action. Trying to "chase" these random load variations will only wear out the unit speed-changing hardware.

To achieve the objectives of good AGC, many features are added, as described briefly in the next section.

10.7.5 AGC Features

This section will serve as a simple catalog of some of the features that can be found in most AGC systems.

Assist action: Often the incremental heat rate curves for generating units will give trouble to an AGC when an excessive ACE occurs. If one unit's participation factor is dominant, it will take most of the control action and the other units will remain relatively fixed. Although it is the proper thing to do as far as economics are concerned, the one unit that is taking all the action will not be able to change its output fast enough when a large ACE calls for a large change in generation. The assist logic then comes into action by moving more of the units to correct ACE. When the ACE is corrected, the AGC then restores the units back to economic output.

Filtering of ACE: As indicated earlier, much of the change in ACE may be random noise that need not be "chased" by the generating units. Most AGC programs use elaborate, adaptive nonlinear filtering schemes to try to filter out random noise from true ACE deviations that need control action.

Telemetry failure logic: Logic must be provided to insure that the AGC will not take wrong action when a telemetered value it is using fails. The usual design is to suspend all AGC action when this condition happens.

Unit control detection: Sometimes a generating unit will not respond to control actions. For the sake of overall control, the AGC ought to take this into account. Such logic will detect a unit that is not following control actions and suspend control to it, thereby causing the AGC to reallocate control action among the other units on control.

Ramp control: Special logic allows the AGC to ramp a unit from one output to another at a specified rate of change in output. This is most useful in bringing units on line and up to full output.

www.mepcafe.com

Rate limiting: All AGC designs must account for the fact that units cannot change their output too rapidly. This is especially true of thermal units where mechanical and thermal stresses are limiting. The AGC must limit the rate of change such units will be called on to undergo during fast load changes.

Unit control modes: Many units in power systems are not under full AGC control. Various special control modes must be provided such as manual, base load, and base load and regulating. For example, base load and regulating units are held at their base load value—but are allowed to move as assist action dictates, and are then restored to base-load value.

10.7.6 NERC Generation Control Criteria

NERC, the North-American Electric Reliability Corporation designs and administers standards for reliable operation in North America. NERC has written two generation control performance standards:

1. Control Performance Standard
2. Disturbance Control Standard

Both of these are described in the NERC document listed in reference 1. In addition, an excellent description of an implementation of the NERC AGC performance standards is given in a paper in reference 2 and a discussion of that paper in reference 3.

The NERC standards set criteria that measure the performance of each control area's efforts to control frequency and interchange flow and give specific targets that must be met to be in compliance. This is expressed in two inequalities:

$$\text{CPS1}: \text{AVG}_{\text{Period}}\left[\left(\frac{\text{ACE}_i}{-10B_i}\right) * \Delta F_1\right] \le \varepsilon_1^2$$

This is the rolling average over a 12 month period of the 1 min average, index *i*, where

ACE_i is the 1 min average of ACE for the minute indexed as i

B_i is the average frequency bias in effect over that minute

ΔF_1 is the 1 min average of the interconnection's frequency error

ε_1^2 is a constant determined by each interconnection to meet a specified frequency bound

$$CSP2: \text{AVG}_{\text{10-minute}}(\text{ACE}_i) \le L_{10}$$

This is the average ACE for at least 90% of the 10 min periods during a calendar month,
where

$$L_{10} = 1{,}65\varepsilon_{10}\sqrt{(-10B_i)(-10B_s)}$$

www.mepcafe.com

ε_{10} is the RMS of the 10 min average frequency error over a given year and is the same for all control areas in the interconnection.

According to the discussion in reference 3, the CSP1 and CSP2 standards were to allow control areas to minimize generator control actions or maneuvering by allowing maximum allowance for the value of ACE. The larger the allowed value of ACE, the less do generation units need to reverse and maneuver.

PROBLEMS

10.1 Suppose that you are given a single area with three generating units as shown in Figure 10.26.

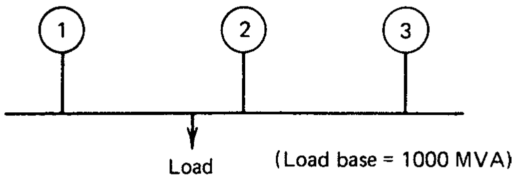


FIGURE 10.26 Three-generator system for Problem 10.1.

Three-generator system for Problem 10.1.

Unit	Rating (MVA)	Speed Droop R (Per Unit on Unit Base)
1	100	0.01
2	500	0.015
3	500	0.015

The units are initially loaded as follows:

$$P_1 = 80\,\text{MW}$$
$$P_2 = 300\,\text{MW}$$
$$P_3 = 400\,\text{MW}$$

Assume $D = 0$; what is the new generation on each unit for a 50-MW load increase? Repeat with $D = 1.0$ pu (i.e., 1.0 pu on load base). Be careful to convert all quantities to a common base when solving.

10.2 Using the values of R and D in each area, for Example 10B, resolve for the 100-MW load change in area 1 under the following conditions:

$$\text{Area 1: base MVA} = 2000\ \text{MVA}$$
$$\text{Area 2: base MVA} = 500\ \text{MVA}$$

www.mepcafe.com

Then solve for a load change of 100 MW occurring in area 2 with R values and D values as in Example 10B and base MVA for each area as before.

10.3 Given the block diagram of two interconnected areas shown in Figure 10.27 (consider the prime-mover output to be constant, i.e., a "blocked" governor):

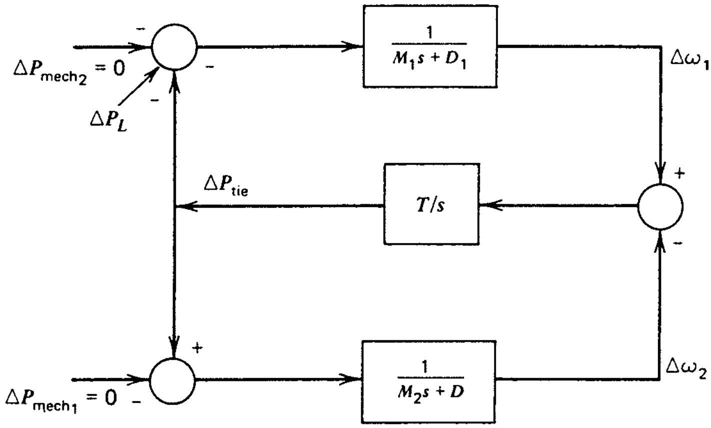


FIGURE 10.27 Two-area system for Problem 10.3.

a. Derive the transfer functions that relate $\Delta\omega_1(s)$ and $\Delta\omega_2(s)$ to a load change $\Delta P_L(s)$.

b. For the following data (all quantities refer to a 1000-MVA base),

$$M_1 = 3.5\,\text{pu} \quad D_1 = 1.00$$
$$M_2 = 4.0\,\text{pu} \quad D_2 = 0.75$$
$$T = 377 \times 0.02\text{pu} = 7.54\,\text{pu}$$

calculate the final frequency for load-step change in area 1 of 0.2 pu (i.e., 200 MW). Assume frequency was at nominal and tie flow was 0 pu.

c. Derive the transfer function relating tie flow, $\Delta P_{tie}(s)$ to $\Delta P_L(s)$. For the data of part b calculate the frequency of oscillation of the tie power flow. What happens to this frequency as tie stiffness increases (i.e., $T \to \infty$)?

10.4 Given two generating units with data as follows:

Unit 1: Fuel cost: $F_1 = 1.0\$/\text{MBtu}$

$$H_1(P_1) = 500 + 7P_1 + 0.002P_1^2 \text{ MBtu/h}$$
$$150 < P_1 < 600 \text{ Rate limit} = 2\,\text{MW/min}$$

Unit 2: Fuel cost: $F_2 = 0.98\$/\text{MBtu}$

$$H_2(P_2) = 200 + 8P_2 + 0.0025P_2^2 \text{ MBtu/h}$$
$$125 \le P_2 \le 500\,\text{MW Rate limit} = 2\,\text{MW/min}$$

www.mepcafe.com

a. Calculate the economic base points and participation factors for these two units supplying 500 MW total. Use Equation 3.8 to calculate participation factors.

b. Assume a load change of 10 MW occurs and that we wish to clear the ACE to 0 in 5 min. Is this possible if the units are to be allocated by base points and participation factors?

c. Assume the same load change as in part b, but assume that the rate limit on unit 1 is now 0.5 MW/min.

This problem demonstrates the flaw in using Equation 10.35 to calculate the participation factors. An alternate procedure would generate participation factors as follows:

Let t be the time in minutes between economic-dispatch calculation executions. Each unit will be assigned a range that must be obeyed in performing the economic dispatch.

$$\begin{aligned} P_i^{\max} &= P_i^0 + t \times \text{rate limit}_i \\ P_i^{\min} &= P_i^0 - t \times \text{rate limit}_i \end{aligned} \tag{10.39}$$

The range thus defined is simply the maximum and minimum excursion the unit could undergo within t minutes. If one of the limits described is outside the unit's normal economic limits, the economic limit would be used. Participation factors can then be calculated by resolving the economic dispatch at a higher value and enforcing the new limits described previously.

a. Assume $T=5$ min and that the perturbed economic dispatch is to be resolved for 510 MW. Calculate the new participation factors as

$$pf_i = \frac{P_i^{\Delta} - P_{i_{\text{basept}}}}{\Delta P_{\text{total}}}$$

where

$$\begin{aligned} P_{i_{\text{basept}}} &= \text{base economic solution} \\ P_{1_{\text{base}}} + P_{2_{\text{base}}} &= 500 \text{ MW} \\ P_i^{\Delta} &= \text{perturbed solution} \\ P_1^{\Delta} + P_2^{\Delta} &= 510 \text{ MW} \end{aligned}$$

with limits as calculated in Equation 10.35.

Assume the initial unit generations P_i^0 were the same as the base points found in part a. And assume the rate limits were as in part c (i.e., unit 1 rate $lim=0.5$ MW/min, unit 2 rate $lim=2$ MW/min). Now check to see if 1 part c gives a different result.

www.mepcafe.com

10.5 The interconnected systems in the eastern United States and Canada have a total capacity of about 5×10^5 MW. The equivalent inertia and damping constants are approximately

$$M = 8 \text{ pu MW / pu frequency / sec}$$

and

$$D = 1.5$$

both on the system capacity base. It is necessary to correct for time errors every so often. The electrical energy involved is not insignificant.

a. Assume that a time error of 1 s is to be corrected by deliberately supplying a power unbalance of a constant amount for a period of 1 h. Find the power unbalance required. Express the amount in MWH.

b. Is this energy requirement a function of the power unbalance? Assume a power unbalance is applied to the system of a duration "delta *T*". During this period, the unbalance of power is constant; after the period it is 0. Does it make any difference if the length of time is long or short? Show the response of the system. The time deviation is the integral of the frequency deviation.

10.6 In Figure 10.16 assume that system 2 represents a system so large that it is effectively an "infinite bus." M_2 is much greater than M_1 and the frequency deviation in system 2 is 0.

a. Draw the block diagram including the tie line between areas 1 and 2. What is the transfer function for a load change in area 1 and the tie flow?

b. The reactance of the tie is 1 pu on a 1000-MW base. Initially, the tie flow is 0. System 1 has an inertia constant (M_1) of 10 on the same base. Load damping and governor action are neglected. Determine the equation for the tie-line power flow swings for a sudden short in area 1 that causes an instantaneous power drop of 0.02 pu (2%), which is restored instantly. Assume that $\Delta P_{L_i}(s) = -0.02$, and find the frequency of oscillation and maximum angular deviation between areas 1 and 2.

REFERENCES

1. Policy 1 – Generation Control and Performance, version 2, October 8, 2002, NERC.
2. Yao, M., Shoults, R. R., Kelm, R., "AGC Logic Based on NERC's New Control Performance Standard and Disturbance Control Standard," *IEEE Transactions on Power Systems*, Vol. 15, No. 2, May 2000, pp. 852–857.
3. Jaleeli, N., VanSlyck, L. S., Discussion of "AGC Logic Based on NERC's New Control Performance Standard and Disturbance Control Standard", published with ref 2 above. *IEEE Transactions on Power Systems*, Vol. 15 No. 4, 2000, pp. 1455–1456.

www.mepcafe.com

11

INTERCHANGE, POOLING, BROKERS, AND AUCTIONS

11.1 INTRODUCTION

This chapter reviews the exchange of energy and of power, primarily the practices in the United States, Canada, and other countries where there are numerous, major electric utilities operating in parallel in multiple AC interconnections, some augmented with DC Links. In many parts of the world, similar commercial structures of the electric power industry have been copied or evolved in a slightly different manner. Many countries have one to two major generation and transmission utilities with local distribution utilities. As these countries interconnect, the industry structure is important in discussing the interchange of energy since the purchase and sale of energy is a commercial business where the parties to any transaction expect to enhance their own economic positions or conform to regulations. The advent of micro-grids increase the role of interchange for economic and reliability issues.

The business environment chapter outlined the economic benefit of buying energy instead of making it under various conditions. This chapter centers on the exchange of energy over periods of time under the make it or buy it framework. In North America, the "marketplace" is large, geographically separated into three areas, and the transmission networks in the major interconnections are owned and operated by multiple entities. This has led to the development of a number of common practices in the interchange of power and energy between electric utilities. Where the transmission network is (or was) owned by a single entity, the past and developing practices regarding transactions have evolved. We confine the discussions of the commercial aspects of the electric energy markets to the practices in North America, circa early 2012.

Power Generation, Operation, and Control, Third Edition. Allen J. Wood, Bruce F. Wollenberg, and Gerald B. Sheblé.
© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

www.mepcafe.com

It is recognized in this work that the exchange of power often occurs under abnormal or emergency situations. While these situations are critical for the heightened reliability of the system, the probability calculations are beyond the scope of this work. These are considered contingent contracts since they are dependent on the occurrence of future events. Instead, the amount of resources is identified as a given in the following examples.

Many industrial firms had a need for process heat or steam and developed internal generation (i.e., cogeneration plants) to supply steam and electric power. Some developed electric power beyond the internal needs of the plant so that they could arrange for sale of the excess to the connected utility systems. The earlier competitive markets only involved "wholesale transactions," the sale and purchase of electric energy to utilities for ultimate delivery to the consumer. Competitive markets are now required to include Demand Response Management (DRM) as noted in the business environment Chapter 2. The introduction of Combined Heat and Power (CHP) has expanded the co-generation concepts to every industrial development.

Because of multiple transmission system ownerships in North America and the absence of a single entity charged with the control of the entire (or even regional) bulk power system, there are many unresolved issues for operational and for reliability decisions. These concern generation control, control of flows on the transmission system circuits, and establishment of schemes for setting "fair and equitable" rates for the use of the transmission network by parties beyond the utility owner of the local network to maintain reliability. Commonwealth Edison (COMED) had implemented a wholesale transmission access tariff that maintained the reliability of the grid at the distribution system connection.

The movement toward more open access participation continues and more entities are becoming involved in the operation of the interconnected systems. Almost all of the nonutility participants are involved in supplying power and energy to utilities or large industrial firms. The use of a transmission system by parties other than its owner involves "wheeling" arrangements. Wheeling contracts are an arrangement to use the transmission system owned by another party to deliver the energy or power to the counterparty. There have been wheeling arrangements as long as there have been interconnections between more than two utilities. In most cases, the development of transmission service (i.e., wheeling) rates has been based on simplified physical models designed to facilitate commercial arrangements. As long as the market was restricted to a few parties, these arrangements were usually simple. With the introduction of many nonutility participants, there is a need for the development of rate structures based on more realistic models of the power system. Since the need to maintain flow limitations, various contracts involving flowgates or financial transmission rights have been developed. These approaches are outlined later in the chapter.

The growth of the number and size of energy transactions has emphasized the need for intersystem agreements on power flows over "parallel" transmission circuits. Two neighboring utilities may engage in the purchase or sale of a large block of power. They may have more than enough unused transmission capacity in the direct interconnections between the systems to carry the power. But, since the systems are interconnected in an AC network that includes a large number of utilities, when the transaction takes place,

www.mepcafe.com

a large portion of the power may actually flow over circuits owned by other systems. The flow pattern is determined by physical laws, not commercial arrangements. The problems caused by these parallel path flows have been handled (at least in North America) by mutual agreements between interconnected utility systems. In the past, there was a general, if unspoken, agreement to attempt to accommodate the transactions. But, as the numbers and sizes of the transactions have increased, there have been more incidences of local circuit overloads caused by remote transactions. Most recently, a tagging system has been implemented to document the actual flows.

We emphasize these points because utilities in other parts of the world do not operate in the same manner. Many of the problems associated with transmission system use, transmission access, and parallel path issues, are a consequence of multiple ownership of the transmission network. They are *structural problems*, not physical problems. On the other hand, when a formerly nationalized grid is deregulated and turned into a single, privatized network there are problems, primarily due to the intended use of the network has changed from a regulated point of support design from a new resource to a given demand area to an open access design to connect the most sellers with the most buyers.

Energy transfers are due to economies of scale, timing of unit additions, jointly owned units, distributed generation, combined heat and power (CHP), Independent Power Producers (IPP) (especially gas turbines), as well as renewable generation farms of wind generation or solar cells, and other smart grid developments. The Economic Dispatch, Unit Commitment, and Limited Energy Supply chapters provide the basic analysis tools for interchange evaluation.

Jointly owned units (JOU) lead to dynamic interchange of energy to follow area control error or to follow remote demand. Chapter 10 detailed the implemented algorithms for such operation to maintain a given interchange schedule of contracts.

As noted in Chapter 10, inadvertent flows occur since the control is on the area net only with inadequate control to force flows as per contract. This causes errors and power wheeling. Such inadvertent flows have to be scheduled for payback at the same cost or price as when the flow occurred.

Flows can be controlled if enough devices were added to the system such as flexible AC transmission (FACTS) devices, high voltage direct current (HVDC) Links, as well as quadrature phase shifting transformers. However, these devices are installed at a high cost. There are also high operating and infrastructure costs associated with coordinated control. The Transmission System Effects, Security Analysis, and Optimal Power Flow chapters provide the analysis tools for flow control issues.

This chapter reviews the practices that have evolved in all-utility interchange arrangements. This leads to a brief discussion of power pools and other commercial arrangements designed to facilitate economic interchange. Many of the issues raised by the use of the transmission system are unresolved issues that await the full and mature development of new patterns for coordinating bulk power system operations and defining, packaging, and pricing transmission services. We only discuss possible outcomes based on engineering and economic analysis.

Where there is regulation of utility charges to consumers, prices are usually based on costs as the commissions have placed a fixed rate of return for utility profits.

www.mepcafe.com

There may be some dispute over what costs should be included and how they should be allocated to each consumer class, but, generally, the notion of cost-based pricing is firmly established.

Prices are based on market action rather than being administered by governments in most competitive markets. There is usually a stated principle that utilities may recover no more than a given margin above "cost." This is called a price ceiling. There is also often a price floor to prevent inadequate return on investment. Where utilities are dealing with each other or with nonutility entities, there may, or may not, be an obligation to base prices on costs. In many situations, market forces will set price levels. Transactions will be negotiated when both parties can agree upon terms that each considers advantageous, or at least satisfactory.

11.2 INTERCHANGE CONTRACTS

The interchange types introduced within this chapter include the following contract categories: energy, dynamic energy, contingent, market based, transmission line use, and reliability.

11.2.1 Energy

Economy A

This type of interchange was defined to clearly show that the economic dispatch program alone was needed to solve for the costs. The economic advantage is due to lower cost production in another area.

Economy B

This type of interchange was defined to clearly show that the unit commitment program alone was needed to solve for the costs. The economic advantage is due to lower cost of production in another area.

Unit Sales

This type of interchange was defined to show the sale of a percentage of the unit over a given time frame. The economic advantage is due to the capital and to the operation production cost of a newer unit. This is an alternative to JOU operation as the period is finite and not for the life of the plant.

Diversity

Daily diversity interchange arrangements may be made between two large systems covering operating areas that span different time zones. Under such circumstances, one system may experience its peak load at a different time of the day than the other system simply because the second system is 1 h or more behind. If the two systems experience such a phenomenon, they can help each other by interchanging power during the peak. The system that peaked first would buy power from the other and then pay it back when the other system reached its peak load. Note that there are four time zones across the United States.

This type of interchange can also occur between systems that peak at different seasons of the year. Typically, one system will peak in the summer due to air-conditioning

www.mepcafe.com

load and the other will peak in winter due to heating load. The winter-peaking system would buy power during the winter months from the summer-peaking system whose system load is presumably lower at that time of year. Then in the summer, the situation is reversed and the summer-peaking system buys power from the winter-peaking system. Several continents have countries with northern and southern states with a difference in environmental conditioning equipment.

Energy Banking

Energy-banking agreements usually occurred when a high storage capability hydro system is interconnected to a run of river hydro or thermal system. During high water runoff periods, the hydro system may have energy to spare and will sell it to the thermal system. Conversely, the hydro system may also need to import energy during periods of low runoff. The prices for such arrangements are usually set by negotiations between the specific systems involved in the agreement.

Instead of accounting for the interchange and charging each other for the transactions on the basis of hour-by-hour operating costs, it is common practice in some areas for utilities to agree to a banking arrangement, whereby one of the systems acts as a bank and the other acts as a depositor. The depositor would "deposit" energy whenever it had a surplus and only the MWh "deposited" would be accounted for. Then, whenever the depositor needed energy, it would simply withdraw the energy up to MWh it had in the account with the other system. Which system is "banker" or "depositor" depends on the exchange contract. It may be that the roles are reversed as a function of the time of year.

Distributed Generation, such as wind and solar cells, are now just as important an impact on energy storage to virtually schedule the use of energy to coincide with the availability of renewable energy.

Inadvertent

The AGC systems of utilities are not perfect devices with the result that there are regularly occurring instances where the error in controlling interchange results in a significant, accumulated amount of energy. This is known as *inadvertent interchange*. Under normal circumstances, system operators will "pay back" the accumulated inadvertent interchange energy megawatt-hour for megawatt-hour, usually during similar time periods in the next week. Differences in cost rates should not be ignored.

Occasionally, utilities will suffer prolonged shortages of fuel or water, and the inadvertent interchange energy may grow beyond normal practice. If done deliberately, this is known as "leaning on the ties." When this occurs, systems will normally agree to pay back the inadvertent energy at the same time of day that the errors occurred. The intent is to equalize the economic transfer. In severe fuel shortage situations, interconnected utilities may agree to compensate each other by paying for the inadvertent interchange at market based price levels that reflect the real cost of generating the exchange energy.

Take or Pay

Take or pay contracts are similar to the take or pay fuel contracts presented previously. The energy is sold on a monthly or a multi-monthly basis sometimes in the same pattern as a futures contract. The solution is found in a similar fashion to take or pay fuel contracts. Whatever energy that is not used is lost at the end of the period.

www.mepcafe.com

Other Conditions
The above contracts can be classified as firm on non-firm depending on the conditions for contract termination. The most common reason for contract termination is lack of flow capability or emergency operation.

11.2.2 Dynamic Energy

The dynamic energy contracts that change as a function of other flows include the following.

Jointly Owned Units
Many fossil-fuel-fired power plants require economies of scale to justify implementation. When neighboring utilities find that the demand growth does not justify a new power plant on their own, a pool of owners is formed to build the new plant. The plant is subsequently operated by the agreed shares of the unit with dynamic interchange across all control areas to deliver the generation from that JOU to the owner as if that plant was in the owners control area. The Intermountain Power Project (IPP) in Utah is an example of this operation combined with a Jointly Owned Transmission line that is an HVDC link from Lynndyll, Utah, the location of the coal mine, to Los Angeles, California. There are 36 participants of this project using AC and DC lines.

Remote Demand
After mergers or purchases of other utilities, a demand center can be isolated within another control area. Such demand is the responsibility of the serving utility. Since the demand and changes in demand are the responsibility of a utility not in the control area, a dynamic interchange agreement is needed for the responsible utility to provide for the demand as it changes. This demand is an isolated control area that is linked to the control area of the owning company.

11.2.3 Contingent

The contingent contracts that are implemented based on events include the following.

Capacity
Normally, a power system will add generation to make sure that the available capacity of the units it has equals its predicted peak load plus a reserve to cover unit outages. If for some reason this criterion cannot be met, the system may enter into a capacity agreement with a neighboring system, provided that neighboring system has surplus capacity beyond what it needs to supply its own peak load and maintain its own reserves. In selling capacity, the system that has a surplus agrees to cover the reserve needs of the other system. This may require running an extra unit during certain hours, which represents a cost to the selling system. The advantage of such agreements is to let each system schedule generation additions at longer intervals by buying capacity when it is short and selling capacity when a large unit has just been brought on line and it has a surplus. Pure capacity reserve interchange agreements do not entitle the purchaser to any energy other than emergency energy requirements.

www.mepcafe.com

Emergency
It is very likely that at some future time a power system will have a series of generation failures that require it to import power or shed load. Under such emergencies, it is useful to have agreements with neighboring systems that commit them to supply power so that there will be time to shed demand. This may occur at times that are not convenient or economical from an incremental cost point of view. Therefore, such agreements often stipulate that emergency power be priced very high. Such agreements are based on the probability of such events.

Spinning Reserve
Spinning reserve is the available capacity to cover the forced outage of another unit. Spinning reserve applies to capability already synchronized and able to respond within 5 min or some other time frame. The amount of spinning reserve needed is based on the probability of a forced unit or interchange outage event.

Ready Reserve
Ready reserve is the available capacity to cover the forced outage of another unit. Spinning reserve applies to capability not yet synchronized but able to respond within 15 min or some other time frame. The amount of ready reserve needed is based on the probability of a forced unit or interchange outage event.

11.2.4 Market Based

Interchange contracts have expanded dramatically within the United States. Several control centers have been enforcing thousands of contracts each hour. Many federal and state commissions found large economic gains in the use of these contracts.

A new growth within the United Sates is the evolution of Demand Side Management (DSM) as demonstrated by Florida Power and Light (FP&L) during the 1970s oil crisis. FP&L placed air-conditioning units under remote control to turn them off for a 5 min interval each hour. This is essentially a miniature, localized rolling blackout. After the state commission mandated a price ceiling, FP&L used the DSM to control frequency instead of the generation and to keep the total demand below the price ceiling. The success demonstrated that such a response enables price control by buyers. This is critical as costs continue to climb.

Demand Response Management (DRM) is now required within the United States due to Federal Energy Regulatory Commission (FERC) order 745. Such power and energy reductions by demand are now to be handled as interchange or unit contracts in a competitive market environment on a static or dynamic basis depending on the capability of the demand to meet contractual requirements.

The number of contracts implemented has become hard to maintain. The solution on the east coast was to form a power pool where all interchange could be dynamically dispatched as if multiple companies operated as one company. The Pennsylvania New Jersey Maryland (PJM) pool is a prime example of this interchange growth. Others include Southern Company that operated as a holding company coordinating the various utility subsidiaries. The American Electric Power Company was a similar umbrella holding company approach to integrating many utilities to operate as one.

www.mepcafe.com

The oil crisis of the middle 1970s brought forward a centralized market as ordered by the state utility commission. The Florida Energy Broker (FEB) was established to find interchange opportunities on an hourly basis (Economy A). This auction mechanism was very successful in finding interchange opportunities. The economics were so strong that many of the repeating interchanges found each hour became firm interchange contracts on a long-term basis. This diminished the number of opportunities found but did result in a higher level of economic efficiency. The FEB was the first commodity markets established in the regulated environment.

"Re-regulation" of the U.S. electric industry was an outgrowth of the "deregulation" of the telephone, natural gas, and airlines industries. The establishment of markets was forced in the manner outlined in Chapter 2. The emulation of such markets is explained in the later part of this chapter. This work outlines how to simulate the market based on Linear Programming and extensions to the Lagrangian Unit Commitment algorithm.

11.2.5 Transmission Use

The contracts that are tariffs for use of facilities to support the energy or power flow most generally as implemented as a postage stamp rate, \$/MW rate, or the Locational Marginal Price (LMP) differences most recently.

Another phenomenon that can take place with multiple neighbors is called "wheeling." This occurs when a system's transmission system is simply being used to transmit power from one neighbor, through an intermediate system, to a third system. The intermediate system's AGC will keep net interchange to a specified value, regardless of the power being passed through it. The power being passed through will change the transmission losses incurred in the intermediate system. When the losses are increased, this can represent an unfair burden on the intermediate system, since if it is not part of the interchange agreement, the increased losses will be supplied by the intermediate system's generation. As a result, systems often assess a "wheeling" charge for such power passed through its transmission network.

The determination of an appropriate (i.e., "acceptable") wheeling charge involves both engineering and economics. Utilities providing a wheeling service to other utilities are enlarging the scope of the market for interchange transactions. Past practices amongst utilities have been established by mutual agreement amongst interconnected systems in a region. A transaction between two utilities that are not directly interconnected may also be arranged by having each intermediate utility purchase and resell the power until it goes from the original generator of the sale power to the utility ultimately purchasing it. This is known (in the United States, at least) as *displacement*.

The notion of selling transmission service is not new. A number of different pricing schemes have been proposed and used. Most are based upon simplified models that allow such fictions as the "contract path." Some are based on an attempt to mimic a power flow, in that they would base prices on incremental power flows

www.mepcafe.com

determined in some cases by using DC power flow models. The very simplest rates are a charge per MWh transferred, and ignore any path considerations.

11.2.5.1 Postage Stamp. The postage stamp is the simplest tariff for charges as it replicates the cost of sending a letter from one post office to another within the United States. The total cost of operating the transmission system is calculated for the year. Then the contract flows for each hour are divided by the expected contract flows for the annual period using the transmission system. This percentage is multiplied by the total annual cost to find the cost for each contract.

11.2.5.2 Dollar Per MW Mile. The Dollar per megawatts (MW) Mile is a more complex tariff as the cost of each transmission line is calculated, then the interchange contracts using this line are charged based on the MW flow each hour multiplied by the length of the line. The difficulty is to identify which contracts use which transmission lines. Elaborate tagging systems have been implemented to keep track of the use of each piece of equipment by each interchange contract.

11.2.5.3 Locational Margin Pricing (LMP). Rates for transmission service have a great deal of influence on transactions when wheeling is involved. We have previously considered energy transaction prices based on split-savings concepts. Where wheeling services are involved, this same idea might be carried over so that the selling and wheeling utilities would split the savings with the purchaser on some agreed-upon basis. Both the seller and wheeling systems would want to recover their costs and would wish to receive a profit by splitting the savings of the purchaser. Some would argue that transmission services should be offered on the basis of a "cost plus" price. A split-savings arrangement involving four or five utility systems might lose its economic attractiveness to the buyer by the time the potential savings were redistributed.

Locational Marginal Price (LMP) is dealt with in Section 3.11. Strictly speaking, LMP results from running an auction such as we show in Section 11.11 wherein a security-constrained optimal power flow (SCOPF) is used to implement the auction, then as shown in Chapter 8, the Lagrange multipliers are the LMP at each bus. They are "prices" since they come from an auction where generation is offered at a price to the auction participants. If, however, we are dealing with an optimal power flow (OPF) or SCOPF based on generator cost functions, then the same Lagrange multipliers are referred to as the Bus Incremental Cost (BIC) at each bus. We shall adopt LMP for discussions of bus prices or bus costs because LMP is the most frequently heard term today.

More complex transmission pricing schemes are based on the "marginal cost" or "marginal price" of transmission which is based on the use of LMPs. The numerical evaluation of LMP is straightforward for a system in economic dispatch. In that case, the bus penalty factor times the incremental cost of power at the bus is equal to the system incremental cost (see Section 8.11), except for generator buses that are at upper or lower limits. This is true for load buses as well as generator buses. Note that the cost of operation has to include the cost of generation, the cost of transportation

www.mepcafe.com

losses, and the use of transmission and distribution lines. This text assumes that the cost of distribution is handled separately.
Consider any power system in economic dispatch.

1. If we have a single generator, then the cost to deliver an additional small increment of power at the generator bus is equal to the incremental cost of power for that generator. The cost of the transmission and distribution is normally embedded in this cost.
2. If we have more than one generator attached to a bus and this is the only source of power, and the generators have been dispatched economically (i.e., equal incremental system cost (λ)), then the cost to deliver an additional small increment of power at this bus is equal to the system incremental cost. The cost of the transmission and distribution is normally embedded in this cost.
3. If there are multiple generators at different buses throughout the power system, and they have been dispatched economically, that is, accurate penalty factors have been calculated and used in the economic dispatch—then the cost of delivery of an additional small increment of power at any individual generator bus will be that generator's own incremental cost. This cost will not be equal across the system due to the fact that each generator's incremental cost is multiplied by its penalty factor.

It is important to stress that we are talking of an "additional small increment" of power at a bus and not a large increment. The application of derivatives requires small incremental changes if the derivatives are to be assumed linear. If the power increase is very small, the aforementioned three statements hold. If we are talking of a large increment in power delivered anywhere, the optimal dispatch must be recalculated and the cost is not equal to the incremental cost in any of the earlier three cases.

If we have the case shown in Figure 11.1, the power is all delivered to a load bus that is separated from either generator by a transmission line. In this case, the incremental cost of delivery of power to the load is not equal to the incremental cost of delivery at either generator bus. The exact value of the incremental cost at the load bus can be calculated, however, using the techniques developed in Section 8.11. We are calling the incremental in cost to deliver power at a bus the LMP, which plays a very important role in the operation of modern power systems. For a power system without any transmission limitations, the LMP at any bus in the system will usually be fairly close to the LMP at other buses. However, when there is a transmission constraint, this no longer holds.

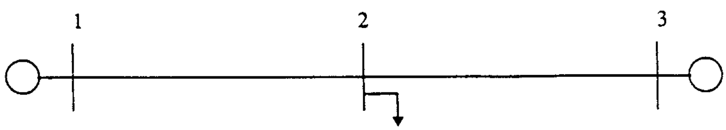


FIGURE 11.1 Three-bus system.

www.mepcafe.com

Suppose the following situation were to arise in the system in Figure 11.1:

1. Generator 1 has high incremental cost and is at its low limit.
2. Generator 3 has low incremental cost and is not at either limit.

In such a case, the LMP at the load bus will be very close to the low incremental cost of the generator at bus 3.

Now let there be a limit to the power flowing on the transmission line from bus 3 to bus 2 so that no further power can be generated at bus 3. When the load is increased at bus 2, the increase must come entirely from the generator at bus 1 and its LMP will be much higher, reflecting the incremental cost of the bus 1 generator. Thus, the LMPs are very useful to show when loading of the transmission system shifts the cost of delivery at certain buses in the network.

Next, let us consider how the LMPs can be used to calculate the *short-run marginal costs* (SRMC) of wheeling. Figure 11.2 shows three systems, A, B, and C, with A selling P_w MW to system C and system B wheeling that amount. The figure shows a single point for injecting the power (bus 1) and a single point for delivery to system C (bus 2). The operators of the wheeling system, B, can determine the incremental cost of power at both buses by using an OPF. If these operators were to purchase the

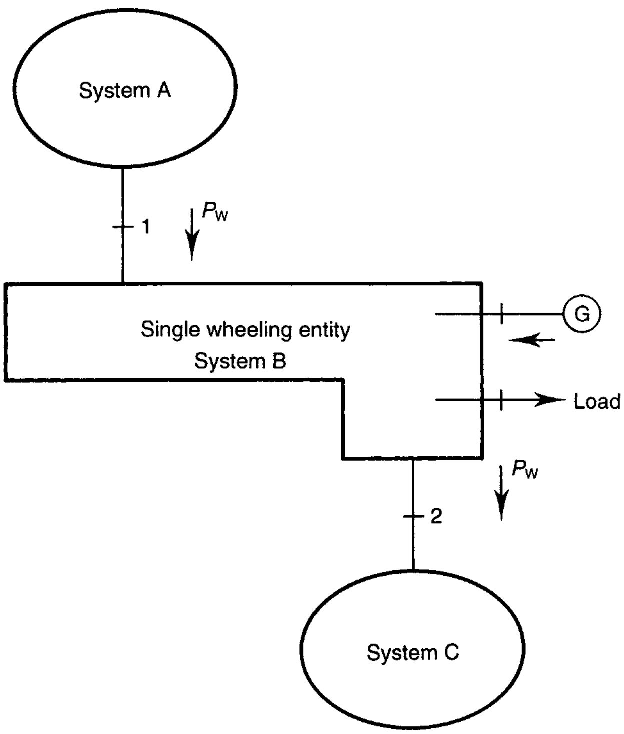


FIGURE 11.2 Simple wheeling example.

www.mepcafe.com

block of wheeled power at bus 1 at the incremental cost there, and sell it to system C at the incremental cost of power at bus 2, they would recover their (short-run) marginal cost of transmission. Many engineers and economists have suggested that transmission service prices should be based upon these marginal costs since they include the cost of incremental transmission losses and network constraints.

The equation to determine this marginal cost is

$$\Delta F = \left(\frac{\partial F}{\partial P_{\mathrm{w}}}\right)\Delta P_{\mathrm{w}} = \left[\frac{\partial F}{\partial P_i} - \frac{\partial F}{\partial P_j}\right]\Delta P_{\mathrm{w}}$$

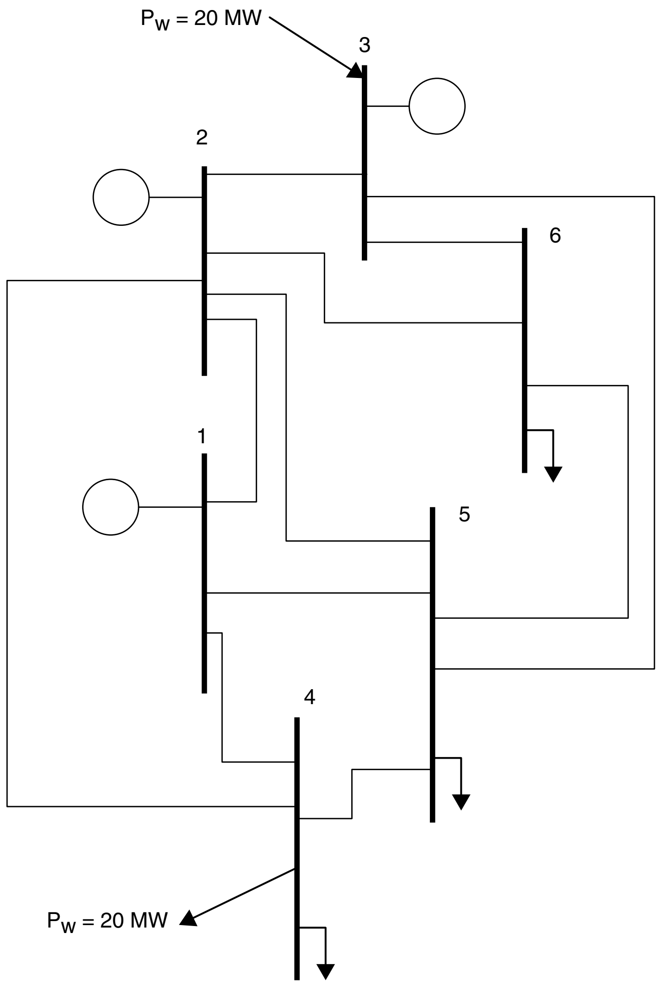


FIGURE 11.3 Six-bus case with 20 MW being wheeled between bus 3 and bus 4.

www.mepcafe.com

where the power ΔP_w is injected at bus i and withdrawn at bus j. Various implementations of the OPF may be programmed to determine the rate of change of the objective function with respect to independent variables and constraints. These computations may be used to evaluate the marginal transmission cost directly.

The six-bus case previously introduced in Chapter 6 may be used to illustrate these ideas. Two separate wheeling examples were run. In both examples, 20 MW was injected at bus 3 and withdrawn at bus 4 as indicated in Figure 11.3.

Three cases were run, Case 1 shows results for an OPF with no transaction and no line limit checking. Note that we here assumed that the line from bus 3 to 6 has been doubled in capacity to 120 MW with the addition of a second circuit. Case 2 shows the same system with the 20 MW transaction added but no line limits checked, Case 3 shows results with the 20 MW transaction and checking line limits. Results are presented in the following.

From	To	P_{flow}	$P_{flowMax}$	P_{flow}
Bus	**Bus**			**Lambda**
2	4	60.03	60	5.7068

The short-run marginal transmission cost rates (in $/MWh) found were

13.2494 − 12.376 = 0.8734 for the unconstrained case and
16.4711 − 12.2015 = 4.2696 for the constrained case.

In the unconstrained example, the marginal cost reflects the effects of the incremental losses. The system dispatch is altered a slight amount to accommodate the additional losses caused by the 20-MW wheeling transfer. No major generation shifts are required. When the flow on the direct line, 2–4, is constrained, the generation pattern is shifted in the OPF solution to reduce the MW flow on that circuit. In doing so, the marginal cost of wheeling is increased to reflect that change.

The effect of a constraint can be illustrated by considering the three-system wheeling situation shown on Figure 11.2. Suppose the transmission system is lossless. With no constraints on power flows, the marginal cost of power will be the same throughout the system. (It will be equal to the incremental cost of the next MWh generated in system B.) Now suppose that there is a constraint in system B such that *before the wheeled power is injected*, no more power may flow from the area near bus 1 to loads near bus 2. See Figure 11.4 which shows a cut labeled "Transmission bottleneck." Then, when the power to be wheeled is injected at bus 1 and withdrawn at bus 2, the schedule in system B will be adjusted so that the delivered power is absorbed near bus 1 and generated by units near bus 2. The difference in marginal costs will now increase, reflecting the marginal cost of the constraint. With no constraint violations, marginal costs of wheeling rise gradually to reflect incremental losses. When constraints are reached, the marginal wheeling costs are more volatile and change rapidly.

It is a strategy to take advantage of a poorly designed transmission system within a competitive market system. This also gives rise to capital market hurdle for competition.

www.mepcafe.com

Case 1: No transaction and no line limits

Bus	P_{min}	P_{gen}	P_{max}	Q_{min}	Q_{gen}	Q_{max}	P_{load}	Q_{load}	V_{min}	V_{bus}	V_{max}	kV	Lambda P
1	50	78.1	200	−100	2.5	150	0	0	0.95	1.07UL	1.07	246.1	12.5083
2	37.5	118.7	150	−100	1	150	0	0	0.95	1.07UL	1.07	246.1	12.4572
3	45	110.2	180	−100	5.1	120	0	0	0.95	1.07UL	1.07	246.1	12.4653
4							100	15	0.95	1.04	1.07	238.6	13.1214
5							100	15	0.95	1.04	1.07	238.5	13.1902
6							100	15	0.95	1.05	1.07	240.5	12.8995

Total operating cost = 4232.4286 $/h

Case 2: With 20 MW transaction from 3 to 4, no line limits checked

Bus	P_{min}	P_{gen}	P_{max}	Q_{min}	Q_{gen}	Q_{max}	P_{load}	Q_{load}	V_{min}	V_{bus}	V_{max}	kV	Lambda P
1	50	83.5	200	−100	4.8	150	0	0	0.95	1.07UL	1.07	246.1	12.5249
2	37.5	119.2	150	−100	5	150	0	0	0.95	1.07UL	1.07	246.1	12.4505
3	45	105.5	180	−100	2.1	120	−20	0	0.95	1.07UL	1.07	246.1	12.376
4							120	15	0.95	1.03	1.07	237.3	13.2494
5							100	15	0.95	1.04	1.07	238.3	13.1763
6							100	15	0.95	1.05	1.07	240.5	12.835

Total operating cost = 4247.8355 $/h
Line flow 2 TO 4 = 77 mw

www.mepcafe.com

Case 3: With transaction and with line limits checked, limit hit on line 2–4

Bus	P_{min}	P_{gen}	P_{max}	Q_{min}	Q_{gen}	Q_{max}	P_{load}	Q_{load}	V_{min}	V_{bus}	V_{max}	kV	Lambda P
1	50	139.8	200	−100	−1.8	150	0	0	0.95	1.07UL	1.07	246.1	13.1639
2	37.5	77.7	150	−100	2.3	150	0	0	0.95	1.06	1.07	243.6	11.731
3	45	91.4	180	−100	15.6	120	−20	0	0.95	1.07UL	1.07	246.1	12.2015
4							120	15	0.95	1.03	1.07	236.2	16.4177
5							100	15	0.95	1.03	1.07	237.5	13.5237
6							100	15	0.95	1.04	1.07	239.6	12.6114

Total operating cost = 4296.7396 $/h
Line flow on line 2–4 now at limit

www.mepcafe.com

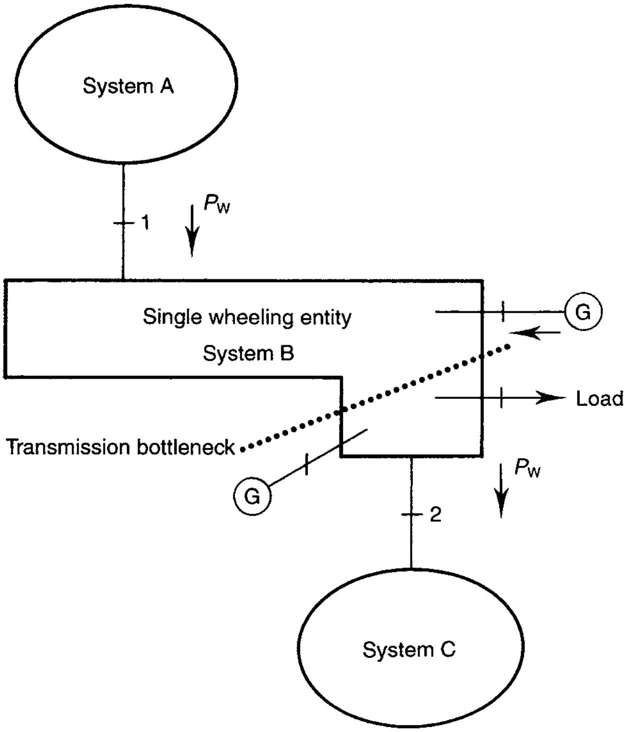


FIGURE 11.4 Simple wheeling example with a “transmission bottleneck.”

Marginal-cost-based pricing for transmission services has a theoretical appeal. Not everyone is in agreement that transmission services should be priced this way. If the entire transaction is priced at the marginal cost rate after the transaction is in place, the wheeling utility may over- or under-recover its changes in operating costs. Perhaps more importantly, short-run marginal operating costs do not reflect the revenue required to pay the costs related to the investments in the wheeling system’s facilities. Note that the capital charges are dropped when the derivative is found as they are of constant value. It is noted that the market price or fair market value normally includes the fixed or capital cost.

These facilities make it possible to wheel the power. (It is quite possible that short-run marginal wheeling costs could be negative if a transaction were to result in incremental power flows that reduced the losses in the wheeling system.) Any pricing structure for transmission service needs to incorporate some means of generating the funds required to install and support any new facilities that are needed in order to accommodate growing demands for service. These are the long-run marginal costs. If the transmission network is to be treated as a separate entity, the price structure for transmission service needs to include the long-run costs as well as short-run operating costs. It is best to include the real option value in the market based rate to estimate the value of any added asset.

www.mepcafe.com

11.2.5.4 Jointly Owned Transmission (JOT). These projects are typified by the delivery of power and/or energy over large distances such as the Intermountain Power Project HVDC Link. These projects require control over the flows between control areas to accurately reflect and to implement the interchange contract as designed.

11.2.6 RELIABILITY

Electric power systems interconnect because the interconnected system is more reliable, it is operated in a more secure fashion, and it may be operated at less cost than if left as separate parts. We saw in a previous chapter that interconnected systems have better regulating characteristics. A load change in any of the systems is taken care of by all units in the interconnection, not just the units in the control area where the load change occurred. This fact also makes interconnections more reliable, since the loss of a generating unit in one of them can be made up from spinning reserve among units throughout the interconnection. Thus, if a unit is lost in one control area, governing action from units in all connected areas will increase generation outputs to make up the deficit until standby units can be brought on line. If a power system were to run in isolation and lose a large unit, the chance of the other units in that isolated system being able to make up the deficit is greatly reduced. Extra units would have to be run as spinning reserve, and this would mean less-economic operation. Furthermore, a generation system will generally require a smaller installed generation capacity reserve if it is planned as part of an interconnected system. It is noted that it is not the number of generators but the amount of inertia that is connected. Inertia can be provided by flywheels, batteries with fast inverters, and other energy storage devices. Inertia requirements lead to the establishment of markets for this ancillary service.

Operational Planning groups have been established in many countries to coordinate the various commodity markets needed for electric power and energy production from the fuel resources through the transportation of the electric power. These details are outside the scope of this work.

11.3 ENERGY INTERCHANGE BETWEEN UTILITIES

One of the most important reasons for interconnecting with neighboring systems centers on the better economics of operation that can be attained when utilities are interconnected. This opportunity to improve the operating economics arises any time two power systems are operating with different incremental costs. If there is a sufficient difference in the incremental cost between the systems, it will pay both systems to exchange power at an equitable price. To see how this can happen, one need merely reason as follows. Given the following situation:

- Utility A is generating at a lower incremental cost than utility B.
- If utility B were to buy the next megawatt of power for its load from utility A at a price less than if it generated that megawatt from its own generation, it would save money in supplying that increment of load.

www.mepcafe.com

- Utility A would benefit economically from selling power to utility B, as long as utility B is willing to pay a price that is greater than utility A's cost of generating that block of power.

The key to achieving a mutually beneficial transaction is in establishing an agreed price for the economy interchange sale. Regulated business environment traditionally used difference in costs. The competitive business environment uses differences in prices.

There are other, longer-term interchange transactions that are economically advantageous to interconnected utilities. One system may have a surplus of power and energy and may wish to sell it to an interconnected company on a long-term firm-supply basis. It may, in other circumstances, wish to arrange to see this excess only on a "when, and if available" basis. The purchaser would probably agree to pay more for a firm supply (the first case) than for the interruptible supply of the second case. Nonfirm contracts require the use of reliability measures to properly address the pricing. These are beyond the present scope of this work.

Note that power systems which operate under regulated rules use differences in costs and those operating under competitive rules use differences in prices.

In all these transactions, the question of a "fair and equitable price" enters into the arrangement. The economy interchange examples that follow are all based on an equal division of the operating costs that are saved by the utilities involved in the interchange. This is not always the case since "fair and equitable" is a very subjective concept; what is fair and equitable to one party may appear as grossly unfair and inequitable to the other. The 50–50 split of savings in the examples in this chapter should not be taken as advocacy of this particular price schedule. It is used since it has been quite common in interchange practices in the United States under the noncompetitive business environment. Pricing arrangements for long-term interchange between companies vary widely and may include "take-or-pay" contracts, split savings, or fixed price schedules.

Before we look at the pricing of interchange power, we will present an example showing how the interchange power affects production costs.

Two utility operating areas are adjacent with transmission connections of sufficient capability. Data giving the composite heat rates and fuel costs in both areas are given. The construction of the composite cost curve for a utility is defined previously. Only one linear segment is used in the following examples to clarify the economic situation (Figure 11.5).

Example 11A

		Cost Coefficients			Unit Limits	
					$P_{min} < P < P_{max}$	
Company Composite Curve	Fuel Cost f_i (It/MBtu)	$F_i(P_i) = f_i(a_i + b_iP_i + c_iP_i^2)$			$P_{min} < P < pr_i nax$	
		a_i	b_i	c_i	P_{min} (MW)	P_{max} (MW)
1	2.0	561	7.92	0.001562	300	1200
2	2.0	310	7.85	0.00194	360	1470

www.mepcafe.com

Area 1:	Load = 700 MW
	Max total generation = 1200 MW
	Min total generation = 300 MW
Area 2:	Load = 1100 MW
	Max total generation = 1470 MW
	Min total generation = 360 MW

First, we will assume that each area operates independently; that is, each will supply its own load from its own generation. This necessitates performing a separate economic dispatch calculation for each area. The results of an independent economic dispatch are given here.

Area 1:	P = 700 MW
	Incremental cost = 20.2136 $/MW
	Production cost = 13,741 $/h
Area 2	PB = 1100 MW
	Incremental cost = 24.236 $/MW
	Production cost = 22,585 $/h
Combined	Total operating cost for both areas = 36,326 $/h

Now suppose the two areas are interconnected by several transmission circuits such that the two areas may be thought of, and operated, as one system without any flow limitations. If we now dispatch them as one system, considering the loads in each area to be the same as just shown, we get a different dispatch for the units.

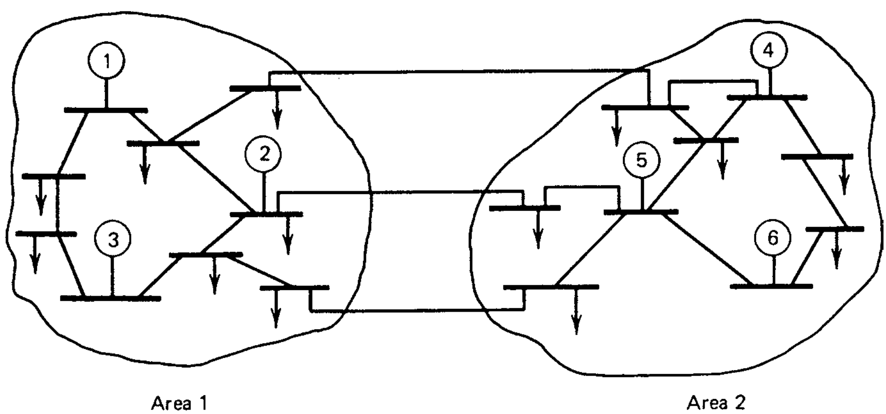


FIGURE 11.5 Interconnected areas for Example 11A.

www.mepcafe.com

Area 1:	PA = 986 MW Incremental cost = 22.00 \$/MW Production cost = 19,777 \$/h
Area 2	PB = 814 MW Incremental cost = 22.016 \$/MW Production cost = 15,971 \$/h
Combined	Interchange power = 286 MW from area 1 to area 2 Total generation for entire system = 1800.0 MW Total operating cost for both areas = 35,748 \$/h

The *interchange power* corresponds to the net power flow over the interconnecting circuits. Note also that the overall cost of operating both systems is now less than the sum of the costs to operate the areas when each supplied its own load.

Example 11A has shown that interconnecting two power systems can have a marked economic advantage when power can be interchanged. If we look at the net change in operating cost for each area, we will discover that area 1 had a decrease in operating cost while area 2 had an increase. Obviously, area 1 should pay area 2 for the power transmitted over the interconnection, but how much should be paid? This question can be, and is, approached differently by each party.

Assume that the two systems did interchange the 286 MW for 1 h. Analyzing the effects of this interchange gives the following:

Area 1 costs:	Without the interchange	\$13,741
	with the interchange	\$19,777
	Increased	\$6,036
Area 2 costs:	Without interchange	\$22,585
	with interchange	\$15,971
	Reduced	\$6,614
	Combined, net savings	\$578

Area 1 can argue that area 2 had a net decrease in operating cost of \$6614 and therefore area 2 ought to pay area 1 this amount. Note that if this were agreed to, area 1 would have a net decrease in its operating costs when the revenues from the sale are included.

Area 2 can argue that area 1 had a net operating cost increase of \$6036 and therefore area 2 ought to pay area 1 this amount. Note that if this were agreed to, area 2 should reduce its net operating cost by \$15,971 minus the amount paid \$6,036 when the cost of the purchase is included.

The problem with each of these approaches is, of course, that there is no agreement concerning a mutually acceptable price. In both cases, one party to the transaction gets all the economic benefits while the other gains nothing. A common practice in such cases is to price the sale at the cost of generation plus one-half the savings in operating costs of the purchaser. This splits the savings equally between the two operating areas. This means that area 2 would pay area 1 coming close to the same savings if both areas were dispatched as one economic dispatch area.

www.mepcafe.com

Such transactions are usually not carried out if the net savings are very small. In such a case, the errors in measuring interchange flows might cause the transaction to be uneconomic. The transaction may also appear to be uneconomic to a potential seller if the utility is concerned with conserving its fuel resources to serve its own customers or if the price of fuel is volatile. The additional costs for long-term life cycle changes are often a concern that is not addressed in many contracts. Fuel volatility and customer requirements may prohibit the exchange of interchange across state boundaries if the sates are competing for customers.

11.4 INTERUTILITY ECONOMY ENERGY EVALUATION

All economy energy valuations are based on the difference before the contract and after the contract. The noncompetitive environment first did with and without study, then split benefit to each company by factor of 2. The competitive environment is similar except that the prices are based on the negotiated market price at the time of the exchange. This approach requires excellent price forecasting to take advantage of increased profits. The future value of fuel is a major consideration that is beyond the scope of this work.

In Example 11A, we saw how two power systems could operate interconnected for less money than if they operated separately. We obtained a dispatch of the interconnected systems by assuming that we had all the information necessary (input–output curves, fuel costs, unit limits, on-line status, etc.) in one location and could calculate the overall dispatch as if the areas were part of the same system. However, unless the two power systems have formed a power pool or transmit this information to each other, or a third party, who will arrange the transaction; this assumption is incorrect. The most common situation involves system operations personnel, located in offices within each of the control areas, who can communicate to each other by telephone. We can assume that each office has the data and computation equipment needed to perform an economic dispatch calculation for its own power system and that all information about the neighboring system must come over some other communications network. How should the two operations offices coordinate their operations to obtain best economic operation of both systems?

The simplest way to coordinate the operations of the two power systems is to note that if someone were performing an economic dispatch for both systems combined, the most economic way to operate would require the incremental cost to be the same at each generating plant, assuming that losses are ignored. The two operations offices can achieve the same result by taking the following steps:

1. Assume there is no interchange power being transmitted between the two systems.
2. Each system operations office runs an economic dispatch calculation for its own system.
3. The offices can determine which system has the lower incremental cost. The operations office in the system with lower incremental cost then runs a series of economic dispatch calculations, each one having a greater total demand (i.e., the total load is increased at each step). Similarly, the operations office in

www.mepcafe.com

the system having higher incremental cost runs a series of economic dispatch calculations, each having a lower total demand.

4. Each increase in total demand on the system with lower incremental cost will tend to raise its incremental cost, and each decrease in demand on the high incremental cost system will tend to lower its incremental cost. By running the economic dispatch steps and conversing over the telephone, the two operations offices can determine the level of interchange energy that will bring the two systems toward most economic operation.

Under idealized "free market" conditions where both utilities are attempting to minimize their respective operating costs, and assuming no physical limitations on the transfer, their power negotiations (or bartering) will ideally lead to the same economic results as a pool dispatch performed on a single area basis. These assumptions, however, are critical. In many practical situations, there are both physical and institutional constraints that prevent interconnected utility systems from achieving optimum economic operation.

The Florida Energy Brokerage system of the mid 1970s automated this evaluation process. Each utility was required to enter two ascending and two descending bids/offers. The interchange was evaluated based on interconnection capability. The solution was then communicated to each utility for implementation. After continuous matches were made, the contracts were committed to on a continuous basis. Thus, the matching eventually failed since all economic transactions were implemented.

11.5 INTERCHANGE EVALUATION WITH UNIT COMMITMENT

A unit commitment calculation would allow the valuation to be adjusted for the ramp rates, the start-up and shutdown times, and other unit operating rules to take more effective advantage of the interchange power economics over a longer time period. The solution methods of Chapter 4 are the key analysis methods for these transactions.

In Examples 11A, there was an implicit assumption that conditions remained constant on the two power systems as the interchange was evaluated. Usually, this assumption is a good one if the interchange is to take place for a period of up to 1 h. However, there may be good economic reasons to transmit interchange power for periods extending from several hours to several days. Obviously, when studying such extended periods, we will have to take into account many more factors than just the relative incremental costs of the two systems.

Extended interchange transactions require that a model of the load to be served in each system (i.e., the expected load levels as a function of time) be included, as well as the unit commitment schedule for each. The procedure for studying interchange of power over extended periods of time is as follows:

1. Each system must run a base-unit commitment study extending over the length of the period in question. These base-unit commitment studies are run without the interchange, each system serving its own load as given by a load forecast extending over the entire time period.

www.mepcafe.com

2. Each system then runs another unit commitment, one system having an increase in load, the other a decrease in load over the time the interchange is to take place.
3. Each system then calculates a total production cost for the base-unit commitment and for the unit commitment reflecting the effect of the interchange. The difference in cost for each system represents the cost of the interchange power (a positive change in cost for the selling system and a negative change in cost for the buying system). The price for the interchange can then be negotiated. If the agreed-on pricing policy is to "split the savings," the price will be set by splitting the savings of the purchaser and adding the change in the cost for the selling system. If the savings are negative, it obviously would not pay to carry out the interchange contract.

It may pay for one system to leave an uneconomical unit off-line entirely during a peak in load and buy the necessary interchange power instead.

The use of energy storage between the two areas is one common economic advantage. Another is the difference between the weekday demand levels and the weekend demand levels. Pumped energy storage, even if accomplished by batteries or other newer devices, results in large economic savings when the weekend demand reduction is included in the analysis.

11.6 MULTIPLE UTILITY INTERCHANGE TRANSACTIONS—WHEELING

The term "wheeling" has a number of definitions; we will stay with a simple one. Wheeling is the use of some party's (or parties') transmission system(s) for the benefit of other parties. Wheeling occurs on an AC interconnection that contains more than two utilities (more properly, two parties) whenever a transaction takes place. (If there are only two parties, there is no third party to perform wheeling.) As used here, the term "parties" includes both utility and nonutility organizations.

Consider the six interconnected control areas shown in Figure 11.6. Suppose areas *A* and *C* negotiate the sale of 100 MW by *A*–*C*. Area *A* will increase its scheduled net interchange by 100 MW and *C* will reduce its net interchange schedule by the same amount. (We ignore losses.) The generation in *A* will increase by the 100 MW sale and that in *C* will decrease by the 100 MW purchase. Figure 11.6b shows the resulting *changes* in power flows, obtained by finding the difference between power flows before and after the transaction. Note that not all of the transaction flows over the direct interconnections between the two systems. The other systems are all wheeling some amount of the transaction. (In the United States, these are called "*parallel path or loop flows*.")

The number of possibilities for transactions is very large, and the power flow pattern that results depends on the configuration and the purchase–sale combination plus the schedules in all of the systems. In the United States, various arrangements have been worked out between the utilities in different regions to facilitate interutility

www.mepcafe.com

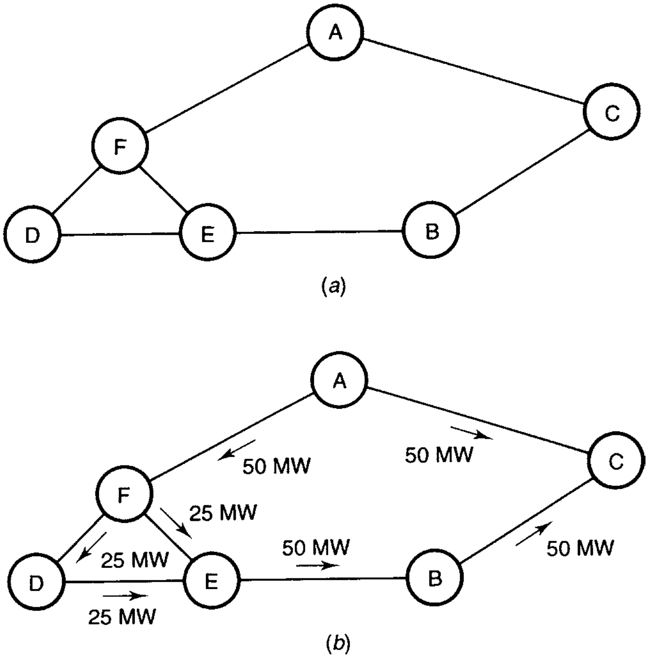


FIGURE 11.6 Six interconnected control areas. (**a**) Configuration; (**b**) incremental power flows when area A sells 100 MW to area C.

transactions that involve wheeling. These past arrangements would generally ignore flows over parallel paths where the two systems were contiguous and owned sufficient transmission capacity to permit the transfer. (This capacity is usually calculated on the basis of nominal or nameplate ratings) In that case, wheeling was *not* taking place, by mutual agreement. The extension of this arrangement to noncontiguous utilities led to the artifice known as the "contract path." In making arrangements for wheeling, the two utilities would rent the capability needed on any path that would interconnect the two utilities. Thus, on Figure 11.6, a 100-MW transaction between systems A and D might involve arranging a "contract path" between them that would have 100 MW available. Flows over any parallel paths are ignored. As artificial as these concepts may appear, they are commercial arrangements that have the merit of facilitating mutually beneficial transactions between systems.

Difficulties arise when wheeling increases power losses in the intervening systems and when the parallel path flows utilize capacity that is needed by a wheeling utility. Increased transmission losses may be supplied by the seller so that the purchaser in a transaction receives the net power that was purchased. In other cases, the transaction cost may include a payment to the wheeling utility to compensate it for the incremental losses. The relief of third-party network element loading caused by wheeling is a more difficult problem to resolve. If it is a situation that involves overloading a third party's system on a recurring basis, the utilities engaged in the transaction may be

www.mepcafe.com

required to cease the transfer or pay for additional equipment in someone else's system. Both approaches have been used in the past.

Loop flows and arrangements for parallel path compensation become more important as the demand for transmission capacity increases at a faster rate than actual capability does. This is the situation in most developed countries. New, high-voltage transmission facilities are becoming more difficult to construct. Another unresolved issue has to do with the participation of organizations that are basically consumers. Should they be allowed access to the power transmission network so that they may arrange for energy supplies from nonlocal resources? In the deregulated natural gas industry in the United States, this has been done.

Expansion of Phase Angle Regulating (PAR) transformers, Flexible AC transmission devices, back to back HVDC links, and long distance HVDC Links enables more control. An HVDC system is presently proposed to overlay the present AC systems within the United States. The Tres Amigas back to back triangular HVDC superstation is to connect three systems that have limited flow capability. The benefits include increased transactions, especially for renewable generation. European Union nations and China have planned an extensive HVDC system to overlay the AC system and provide more control for renewable resources.

Wheeling is inherent if utilities can contract beyond immediate connections as implemented in the western half of the United States through the Western Electricity Coordinating Council (WECC). East coast interchange was only between interconnected immediate neighbors. However, the density of the transmission system and the lack of flow control has shown that wheeling is inherent also in the eastern half of the United States.

Most power systems are interconnected with all their immediate neighboring systems. This may mean that one system will have interchange power being bought and sold simultaneously with several neighbors. In this case, the price for the interchange must be set while taking account of the other interchanges. For example, if one system were to sell interchange power to two neighboring systems in sequence, it would probably quote a higher price for the second sale, since the first sale would have raised its incremental cost. On the other hand, if the selling utility was a member of a power pool, the sale price might be set by the power and energy pricing portions of the pool agreement to be at a level such that the seller receives the cost of the generation for the sale plus one-half the total savings of all the purchasers. In this case, assuming that a pool control center exists, the sale prices would be computed by this center and would differ from the prices under multiple interchange contracts. The order in which the interchange transaction agreements are made is very important in costing the interchange where there is no central pool dispatching office. Such analysis is required by federal and state agencies in most nations.

For example, consider a three-party transaction. **CoA** locates power and energy in **CoC** and makes an arrangement with an intervening system **CoB** for transmission. Then **CoC** sells to **CoB** and **CoB** sells to **CoA**. The price level to **CoA** may be set as the cost of **CoC**'s generation plus the wheeling charges of **CoB** plus one-half of **CoA**'s savings. It may also be set at **CoB**'s net costs plus one-half of **CoA**'s savings. Price is a matter of negotiation in this type of transaction, when prior agreements on pricing policies are absent.

www.mepcafe.com

Often, utility companies will enter into interchange agreements that give the amount and schedule of the interchange power but leave the final price out. Instead of agreeing on the price, the contract specifies that the systems will operate with the interchange and then decide on its cost after it has taken place. By doing so, the systems can use the actual load on the systems and the actual unit commitment schedules rather than the predicted load and commitment schedules. Even when the price has been negotiated prior to the interchange, utilities will many times wish to verify the economic gains projected by performing after-the-fact production costs.

Sequence of calculation impacts is long and subject to interpretation if the contracts are not all inclusive procedures. Assumptions on relative costs for fuels must be avoided as well as assumed costs of transmission system use.

Power systems are often interconnected with many neighboring systems and interchange may be carried out with each one. When carrying out the after-the-fact production costs, the operations offices must be careful to duplicate the order of the interchange agreements.

When several two-party interchange agreements are made, the pricing must follow the proper sequence. The central dispatch of a pool can avoid this problem by developing a single cost rate for every transaction that takes place in a given interval.

11.7 POWER POOLS

Interchange of power between systems can be economically advantageous, as has been demonstrated previously. However, when a system is interconnected with many neighbors, the process of setting up one transaction at a time with each neighbor can become very time consuming and will rarely result in the optimum production cost. To overcome this burden, several utilities may form a power pool that incorporates a central dispatch office. The power pool is administered from a central location that has responsibility for setting up interchange between members, as well as other administrative tasks. The pool members relinquish certain responsibilities to the pool operating office in return for greater economies in operation.

The agreement the pool members sign is usually very complex. The complexity arises because the members of the pool are attempting to gain greater benefits from the pool operation and to allocate these benefits equitably among the members. In addition to maximizing the economic benefits of interchange between the pool members, pools help member companies by coordinating unit commitment and maintenance scheduling, providing a centralized assessment of system security at the pool office, calculating better hydro-schedules for member companies, and so forth. Pools provide increased reliability by allowing members to draw energy from the pool transmission grid during emergencies as well as covering each other's reserves when units are down for maintenance or on forced outage. Ideally, the power pool enables the parties to be operated as if one company resulting in economies of scale.

Some of the difficulties in setting up a power pool involving nonaffiliated companies or systems arise because the member companies are independently owned and for the most part independently operated. Therefore, one cannot just make the assumption that the pool is exactly the same entity as a system under one ownership.

www.mepcafe.com

If one member's transmission system is heavily loaded with power flows that chiefly benefit that member's neighbors, then the system that owns the transmission is entitled to a reimbursement for the use of the transmission facilities. If one member is directed to commit a unit to cover a reserve deficiency in a neighboring system, that system is also likewise entitled to a reimbursement.

These reimbursement arrangements are built into the agreement that the members sign when forming the pool. The more the members try to push for maximum economic operation, the more complicated such agreements become. Nevertheless, the savings obtainable are quite significant and have led many interconnected utility systems throughout the world to form centrally dispatched power pools when feasible.

A list of operating advantages for centrally dispatched power pools, ordered by greatest expected economic advantage, might look as follows:

1. Minimize operating costs (maximize operating efficiency).
2. Perform a system-wide unit commitment.
3. Minimize the reserves being carried throughout the system.
4. Coordinate maintenance scheduling to minimize costs and maximize reliability by sharing reserves during maintenance periods.
5. Maximize the benefits of emergency procedures.

There are disadvantages that must be weighed against these operating and economic advantages. Although it is generally true that power pools with centralized dispatch offices will reduce overall operating costs, some of the individual utilities may perceive the pool requirements and disciplines as disadvantageous. Factors that have been cited include the following:

1. The complexity of the pool agreement and the continuing costs of supporting the interutility structure required to manage and administer the pool.
2. The operating and investment costs associated with the central dispatch office and the needed communication and computation facilities.
3. The relinquishing of the right to engage in independent transactions outside of the pool by the individual companies to the pool office and the requirement that any outside transactions be priced on a split-saving basis based on pool members' costs.
4. The additional complexity that may result in dealing with regulatory agencies if the pool operates in more than one state.
5. The feeling on the part of the management of some utilities that the pool structure is displacing some of an individual system's management responsibilities and restricting some of the freedom of independent action possible to serve the needs of its own customers.
6. Strategy.

Power pools without central dispatch control centers can be administered through a central office that simply acts as a brokerage house to arrange transactions among members. In the opposite extreme, the pool can have a fully staffed central office

www.mepcafe.com

with real-time data telemetered to central computers that calculate the best pool-wide economic dispatch and provide control signals to the member companies.

By far, the most difficult task of pool operation is to decide who will pay what to whom for all the economic transactions and special reimbursements built into the pool agreement. There are several ways to solve this problem, so it is critical to define the process as the pool is formed or reorganized.

The PJM interchange is one of the oldest power pools. The implementation was to generate a composite curve at the pool level, select the system lambda that would meet the total demand, and transmit that lambda to each member. Each member would dispatch its native generation based on that system incremental cost. The lack of communication systems at that time demonstrated the use of a single variable to set unit output across a large area. Such an implementation can be interpreted as an auction system of either the English or Dutch variety as discussed in the economic dispatch chapter.

All methods of allocating the savings achieved by a central pool dispatch are based on the premise that no pool member should have higher generation production expenses than it could achieve by dispatching its own generation to meet its own load.

We saw previously in the pool broker system that one of the ways to allocate pool savings is simply to split them in proportion to each system's net interchange during the interval. In the broker method of matching buyers and sellers based on their incremental and decremented costs, calculations of savings are relatively easy to make since the agreed incremental costs and amounts of energy must be transmitted to the broker at the start. When a central economic dispatch is used, it is easier to act as if the power were sold to the pool by the selling systems and then bought from the pool by the buying systems. In addition, allowances may be made for the fact that one system's transmission system is being used more than others in carrying out the pool transactions.

There are two general types of allocation schemes that have been used in U.S. pool control centers. One may be performed in a real-time mode with cost and savings allocations made periodically using the incremental and decremented costs of the systems. In this scheme, power is sold to and purchased from the pool and participants' accounts are updated currently. In the other approach, the allocation of costs and savings is done after the fact using total production costs.

The second type of savings allocation method is based on after-the-fact computations of individual pool member costs as if each were operating strictly so as to serve their own individual load. In this type of calculation, the unit commitment, hydro-schedules, and economic dispatch of each individual pool member are recomputed for an interval after the pool load has been served. This "own load dispatch" is performed with each individual system's generating capacity, including any portions of JOU, to achieve maximum operating economy for the individual system.

The costs for these computed individual production costs are then summed and the total pool savings are computed as the difference between this cost and the actual cost determined by the central pool dispatch.

These savings are then allocated among the members of the pool according to the specific rules established in the pool agreement. One method could be based on rules similar to those illustrated previously. That is, any interval for which savings are being distributed, buyers and sellers will split the savings equally.

www.mepcafe.com

Specific computational procedures may vary from pool to pool. Those members of the pool supplying energy in excess of the needs of their own loads will be compensated for their increased production expenses and receive a portion of the overall savings due to a pool-wide dispatch. The process is complicated because of the need to perform individual system production cost calculations. Pool agreements may contain provisions for compensation to members supplying capacity reserves as well as energy to the pool. A logical question that requires resolution by the pool members involves the fairness of comparing an after-the-fact production cost analysis that utilizes a known load pattern with a pool dispatch that was forced to use load forecasts. With the load pattern known with certainty, the internal unit commitment may be optimized to a greater extent that was feasible by the pool control center.

Computing power today enables hourly or 5 min calculations as desired. Note that the New Zealand market uses 5 min intervals to simplify the number of contracts needed over the course of an hour.

Complete own-load dispatch computations for cost and savings allocations are usually performed for a weekly period. The implementation may be complex since hourly loads and unit status data are required. An on-line, real-time allocation scheme avoids these complications.

No matter how these savings allocations are performed, students should appreciate that any estimates of "savings" involves finding the difference between actual, known costs and *costs as they might have been*. There is a great deal of room for disagreement about how to estimate these second, hypothetical costs.

11.8 THE ENERGY-BROKER SYSTEM

As with sales and purchases of various commodities or financial instruments (e.g., stock), it is often advantageous for interconnected power systems to deal through a broker who sets up sales and purchases of energy instead of dealing directly with each other. The advantage of this arrangement is that the broker can observe all buy-and-sell offers at one time and achieve better economy of operation. When utilities negotiate exchanges of power and energy in pairs, the "marketplace" is somewhat haphazard like a bazaar. The introduction of a central broker to accept quotations to sell and quotations to purchase creates an orderly marketplace where supply, demand, and prices are known simultaneously.

The simplest form of "broker" scheme is the "bulletin board." In this type of scheme, the utility members post offers to buy or sell power and energy at regular, frequent intervals. Members are free to access the bulletin board (via some sort of data exchange network) at all times. Members finding attractive offers are free to contact those posting the offers and make direct arrangements for the transaction. Like any such informally structured market, many transactions will be made outside the marketplace. More complex brokers are those set up to arrange the matching of buyers and sellers directly, and, perhaps, to set transaction prices.

The Florida Energy Group was one power broker scheme that was successfully used and achieved significant economic benefits for customers. The companies that were members of the broker system send hourly buy-and-sell offers for energy to the

www.mepcafe.com

broker who matches them according to certain rules. Hourly, each member transmits an incremental cost and the number of megawatt-hours it is willing to sell or its decremental cost and the number of megawatt-hours it will buy. The broker sets up the transactions by matching the lowest cost seller with the highest cost buyer, proceeding in this manner until all offers are processed. The matched buyers and sellers will price the transaction on the basis of rules established in setting up the power broker scheme. A common arrangement is to compensate the seller for this incremental generation costs and split the savings of the buyer equally with the seller. The pricing formula for this arrangement is as follows. Let

$$F_s = \text{incremental cost of the selling utility}(\$/\text{MWh})$$
$$F_b = \text{decremental cost of the buying utility}(\$/\text{MWh})$$
$$F_c = \text{cost rate of the transaction}(\$/\text{MWh})$$

Then,

$$\begin{aligned} F_c &= F_s + \frac{1}{2}(F_b - F_s) \\ &= \frac{1}{2}(F_s + F_b) \end{aligned}$$

In words, the transaction's cost rate is the average of the seller's incremental cost and the purchaser's decremental cost. In this text, decremental cost is the reduction in operating cost when the generation is reduced a small amount. Example 11B illustrates the power brokerage process. The buy/sell offers are often called bids to buy and ask to sell. The bid/ask is equivalent to buy/sell.

Example 11B: In this example, two power systems have sent their bid/ask offers to the broker. In the table that follows, these are tabulated and the maximum pool savings possible is calculated. The data of Example 11A is used to show the connection to economic dispatch and interchange analysis.
The following bids and asks are entered to the broker.

Company	Incremental Cost ($/MWh)	Selling Energy (MWh)	Total Increase in Cost ($)
Bids to sell (offer)			
A	20.526	50	14,759 – 13,741 = 1,018
B	24.624	50	23,806 – 22,585 = 1,221

Company	Incremental Cost ($/MWh)	Buying Energy (MWh)	Total Decrease in Cost ($)
Asks to buy			
A	19.90	50	12,738 – 13,741 = –1,003
B	23.848	50	21,383 – 22,584 = –1,201

www.mepcafe.com

The broker sets up transactions as shown in the following table. Note that the transaction results in a savings since the incremental cost increases between the seller and the buyer. The rates and total payments are easily computed under the split-savings arrangement.

Transaction	Price Computation	Total Transaction Savings ($)
A sells 50 MWh to B	(20.526 + 19.90)/2 = 20.213 \$/MWh 20.213 × 50 =	1018 − 1201 = 183
No match		
No match		
	Total	183

A receives \$1010.65 from B. Note that each participant benefits: A receives payment above its costs; B receives savings due to avoided costs.

The example should be repeated until the same solution is found as in the interchange examples. Note that the amount to bid/ask has to be reduced to converge to the same optimal Economic Dispatch (ED) solution. Additionally, the supply and demand curves should be drawn from the iterative solution process. Note that each solution is equivalent to a single iteration of ED.

The chief advantage of a broker system is its simplicity. All that is required to get a broker system into operation is a communications circuit to each member's operations office and some means of setting up the transactions. The transactions can be set up manually or, in the case of more modern brokerage arrangements, by a computer program that is given all the buy/sell offers and automatically sets up the transactions. With this type of broker, the quoting systems are commonly only informed of the "match" suggested by the broker and are free to enter into the transaction or not as each see fit.

Competition calls for discriminatory pricing requiring a decision for each contract for most commodities. Such pricing reduces the revenue for sellers who are aggressive to sell at a lower price. Discriminatory pricing is not used in most electric markets.

Economists have sometimes argued that the broker pricing scheme should set one single "clearing price" for energy each time period, this is uniform pricing (one price). The logic behind this is that the market-determined price level should be based on the participants' needs and willingness to buy or sell. This removes the absolute need for quoting cost-based prices. Utilities would be free to quote offers at whatever price level they wished, but would be (under most rules that have been suggested) obligated to deliver or purchase the energy quoted at the market clearing price. The transactions market would be similar to the stock exchange. Objections raised have been that in times of shortage, price levels could rise dramatically and uncontrollably.

This would alter the price for the aforementioned broker since the same price would be used from the last set of matches instead of the price from each iteration of the process.

www.mepcafe.com

Uniform pricing removes the risk for most sellers and buyers as only the last matched bid and ask have to be accurate. If the auction is a single iteration method as used in most power pools, this is appropriate. However, if trading on a continual basis, as is done with most commodities, then discriminatory pricing will lead to faster solutions or consensus of market price.

Power broker schemes can be extended to handle long-term economy interchange and to arrange other interchange contracts, even capacity sales. This enables brokers to assist in minimizing costs for spinning reserves and coordinate unit commitments in interconnected systems. The inclusion of capital costs is possible but not normally explicitly computed. The capital cost may be included in the heat rate curve implicitly.

Example 11C: In this example, four power systems have sent their buy/sell offers to the broker. In the table that follows, these are tabulated and the maximum pool savings possible is calculated using uniform pricing.

Selling Energy	Incremental Cost ($/MWh)	MWh for Sale	Seller's Total Increase in Cost ($)
Utilities			
A	25	100	2500
B	30	100	3000
Buying Energy	**Decremental Cost ($/MWh)**	**MWh for Purchase**	**Buyer's Total Decrease in Cost ($)**
Utilities			
C	35	50	1750
D	45	150	6750

$$\text{Net pool savings} = (1750\$ + 6750\$) - (2500\$ + 3000\$)$$
$$= 8500\$ - 5500\$ = 3000\$$$

The broker sets up transactions as shown in the following table.

Transaction	Savings Computation		Total Transaction Savings ($)
1. A sells 100 MWh to D	100 MWh (45–25)$/MWh =		2000
2. B sells 50 MWh to D	50 MWh (45–30)$/MWh =		750
3. B sells 50 MWh to C	50 MWh (35–30)$/MWh=		250
		Total	3000

The rates and total payments are easily computed under the uniform pricing arrangement at the price of the last match, the highest value. These are shown in the

www.mepcafe.com

following table. Note the benefit to the lower discriminatory priced auctions match of the previous example.

Transaction	Price ($/MWh)	Total Cost ($)
1. A sells 100 MWh to D	37.5	3750
2. B sells 50 MWh to D	37.5	1875
3. B sells 50 MWh to C	37.5	1875
	Total	7500

A receives $3750 from D; B receives $1875 from D and C. Note that each participant benefits: A receives more than its costs and more above the split savings arrangement.

11.9 TRANSMISSION CAPABILITY GENERAL ISSUES

The available interchange capability has been researched extensively since the 1970s Commonwealth Edison (COMED) was one of the first utilities to calculate the limits of interchange to maximize the amount of power transferred into and across their transmission system.

The original interchange capability was based on the ability of the transmission system to move real power from one company to another. This was based on the thermal power limitations of the grid. First contingency cases were then added to ensure that the transmission could support the interchange after the first contingency had occurred. This work was extended by several researchers, especially Pai and Sauer, who included transient stability limitations within the calculation. Their concepts were recently expanded by Sheblé to include the probability of each outage such that the loss of load probability (LOLP) and expected unserved energy (EUE) could be maintained at specified levels as each transaction was added.

This section describes the calculation processes as expanded from these previous concepts into an interchange contract evaluation and approval service. The described process is based to a large extent on the Midwest Independent System Operator (ISO) implementation.

Transmission enables interchange and also limits the amount that can be exchanged. This topic involves both technical and structural considerations. There are some technical issues that transcend the organizational market structure issues, but many of these arise only because of the multiple ownership of interconnected power transmission networks. There are basic technical issues of defining a network's capability to transfer power that involve physical capacity to handle power flows reliably (or *securely*). Even here, nontechnical matters are involved in defining *acceptable levels of network unreliability*. In an economic environment where capital and financing is available to develop multiple parallel paths in a transmission network, transmission capability may be restricted by the desire of the utilities and involved governmental agencies to insure very high levels of system security. Widespread blackouts and prolonged power shortages are to be avoided. Networks

www.mepcafe.com

are designed with large capacity margins so that elements tend to be loaded conservatively. Normal failures of single major elements will not cause loss of load. Even simultaneous occurrences of two failures of major elements will not cause load curtailment. In most foreseeable circumstances, there will not be cascading outages that spread across the interconnected system. *Cascading outages* can occur where the loss of a transmission circuit, due to a prolonged fault, would result in the overloading of parallel circuits. These, in turn, might be opened in time by the action of protective relaying systems. Thus, the single event could cascade into a regional series of events that could result in a blackout.

In economic climates where capital and financing are difficult to obtain, and in areas where environmental restrictions prevent adding transmission capacity, power transmission networks may be designed using less-stringent reliability standards and operated in a fashion such that loads are expected to be curtailed when major transmission elements suffer outages. Security and reliability standards may be similar to the previous situation, with the exception that controlled load disconnection is not considered to be a "failure" event. Even in systems where "defensive operational scheduling" practices are normally followed (i.e., loss of single or two major system elements does not result in cascading outages), there are occasions where it is more economical to resort to using special system controls. These might drop load automatically when a remote generation source loses one of its transmission links to the system. This is a simple example; there are more complex arrangements that have been used. When a variety of specialized system control schemes are used, it is necessary to keep track of the various systems and keep every interconnected system abreast of changes and new developments. Several U.S. utilities have added series and shunt compensation to the high voltage AC network. Such complex systems require more extensive operating and fault protection using Remedial Action Schemes (RAS). Such schemes are implemented by extensions to the local relay system as regional computer systems which operate without human intervention.

In any interconnected system, there is a need to define in quantitative terms the maximum amount of power that may be transferred without violating whatever system reliability and security criteria are in place. Therefore, it is necessary to consider the types of operating limitations that exist in AC power networks. These include thermal limits sets by the capability of the lines and apparatus to absorb and dissipate the heat created by the current flowing in the various elements. These limits are usually expressed as a maximum allowable temperature rise above specified ambient conditions. The intent is to prevent the extreme, sustained temperatures that might cause lines to sag and equipment to be damaged. Even with these straightforward thermal limitations, there are variable ambient conditions that make actual danger points occur in the summer at lower power transfers than in the colder months. Next are limits set by the interplay of system limitations, equipment limitations, economics, and service reliability ("security") standards. These include voltage–VAR-related conditions and stability considerations.

Voltage and VAR conditions arise because voltage magnitudes within the system must remain within a bandwidth that is set by the voltage tolerances of both system and consumer equipment. Large high-voltage equipment and consumer equipment

www.mepcafe.com

(motors, transformers, etc.) are generally limited to excursions of about ±5% of their rated voltage. The voltage magnitude bandwidth tolerance on the system is affected (and generally enlarged) by the ability of various voltage-correcting devices to restore voltages to a bandwidth acceptable to the apparatus. Key control devices include tap changing transformers and various types of VAR-supply devices. At shorter transmission distances (say 50 miles or 100 km), the thermal limits and voltage–VAR limitations generally are the restricting system conditions. Of course, it is theoretically possible to add additional circuits and VAR-support equipment, but economic considerations generally set a practical limit on what is done to increase transmission capacity.

Transmission capability limits can be imposed by voltage instability, steady-state stability, and transient stability. In all cases, the network has to be able to survive possible conditions that can lead to unstable situations. These instability-inducing conditions usually become more intense as the system loading increases. The need to avoid these operating regimes then places a practical limit on the power that can be transmitted. At longer distances it is usually transient stability that sets the limits. The various limits are found by testing the network under increasingly heavy loading conditions and seeking ways to alleviate or prevent the instabilities. At some point, it becomes impossible or uneconomic to increase the limits further. Besides economic considerations, the actual power transfer limitations found will depend upon the testing criteria utilized. Is it sufficient to test the network's ability to survive a single-phase fault that is successfully cleared and the line reclosed, or should the network be tested using a bolted three-phase fault that requires switching a line segment?

11.10 AVAILABLE TRANSFER CAPABILITY AND FLOWGATES

How important is Available Transfer Capability (ATC)? It is the point where power system reliability constrains electricity market efficiency. ATC has a huge impact on market outcomes and system reliability, so the results of ATC are of great financial interest to all involved.

Commonwealth Edison (COMED) operates a large capacity of nuclear electric generating stations. Even BPA operates a Hugh capacity of hydroelectric generation stations. The Grand Coulee Dam alone has a capacity of 6,765 MW (just one station). Most of BPA's capacity is along the Columbia River which starts in Canada. As a result, how Canada utilizes its part of the Columbia River has a huge impact on the ability of BPA to utilize its Hydro Stations on the river. The United States and Canada operate the Columbia River under a Treaty Agreement. To explain the Treaty in oversimplified terms, Canada has built and operates Columbia River Dams to the benefit of the United States (i.e., BPA's hydro units). BPA must make all attempts to give Canada access to markets in the United States (especially energy-hungry California). This means BPA is always trying to ship power across its system between California and Canada. There is also an agreement that energy that had to be produced in the United States can be stored in Canada for later use by the United States if the water does not spill in the interim.

www.mepcafe.com

Large amounts of money are at stake in both cases. COMED has sold over $800 million in electricity to neighbors in one month. COMED has also wheeled large amount of electric energy from the east to the west. During the first month of 2000, Hydro Quebec sold over $350 billion in electricity to California! Thus, any savings would result in a much lower price to the consumers as nuclear and water generated electricity is the least expensive.

ISOs have elected to utilize a flow-based approach for the evaluation of transmission service requests (“TSRs”). The Available Flowgate Capacity Methodology used requires the calculation of AFC in accordance with the requirements of North American Electric Reliability Council (NERC) Reliability Standard MOD-030.

11.10.1 Definitions

Flowgates (see Section 7.5) are a predetermined set of constraints on the transmission system that are expected to experience loading to the limit most of the time. In the flow-based process, these constraints are used to measure the commitments on the transmission system. Flowgates are classified as two types: Power Transfer Distribution Factor (“PTDF”) and Outage Transfer Distribution Factor (“OTDF”). A PTDF flowgate is a flowgate that monitors the flow on single or multiple transmission elements without a contingency. An OTDF flowgate is a flowgate that monitors flow on single or multiple transmission elements for the loss of other transmission elements. A transmission element is a transformer, a transmission line segment, or a generator.

The **Available Flowgate Capability (AFC)** on a flowgate refers to the amount of MW transfer capacity on a flowgate that remains available for additional transmission service above and beyond the already committed and approved uses of the transmission system. Existing uses of the transmission system include the generation to load impacts on the flowgates and transmission service that has already been sold. AFC values are time and service type dependent. An **ISO*** calculates Firm and Non-Firm AFC values for up to 3 years into the future from the next hour.

The distribution factor quantifies the incremental impact of a power transfer on a flowgate for a particular transmission path and is expressed as a percentage. Distribution factors are calculated as shown in the Chapter 7.

An **AFC zone** is the smallest granular area that is modeled in AFC calculations. All included transmission service reservations are mapped to these AFC zones for modeling impacts and evaluation. In general, the AFC zones are the balancing authority areas (“BAAs”) modeled in the Interconnection Reliability Assessment Group (“IRAG”) Multiregional Modeling Working Group (“MMWG”) models.

An **AFC path** is defined by a unique source and sink that are modeled as AFC zones. Since the ISO OASIS† uses a list of sources and sinks that may be different than the granularity in the AFC calculation, an electrical equivalent table is used to

*In the United States, the markets and the transmission system are simultaneously managed by an Independent System Operator.

†OASIS is the Open Access Same-Time Information System. This is a web page service where registered users wanting to place transactions on a power network can obtain information as to transmission capability between points on the network.

www.mepcafe.com

map the source-and-sink combinations to AFC zones. The electrical equivalent table can also be used to map the source and sink to a study zone that is a combination of multiple AFC Zones to evaluate TSRs more accurately.

The Capacity Benefit Margin (CBM) is the amount of firm transmission transfer capability preserved by the ISO for Load Serving Entities ("LSEs") within a BAA to enable access by the LSEs to get generation from the interconnected systems to meet generation reliability requirements. The CBM is applied as a Total Flowgate Capability ("TFC") reduction to the limiting element(s) of the associated flowgate.

The Transmission Reliability Margin (TRM) is the amount of transmission transfer capability necessary to provide reasonable assurance that the interconnected transmission network will be secure. TRM accounts for the inherent uncertainty in system conditions and the need for operating flexibility to ensure reliable system operations as system conditions change.

The Total Flowgate Capability (TFC) is the maximum amount of power that can flow across the flowgate without overloading (either on an actual or contingency basis) any element of the flowgate. An MW proxy can substitute to ensure sufficient margin against adverse voltage or stability conditions. Flowgate TFCs represent the Interconnection Reliability Operating Limit ("IROL") or System Operating Limit ("SOL") of the defined flowgates per NERC Reliability Standard MOD-030, R2.4.

The **Share of Total Flowgate Capability ("STFC")** is the share of the TFC that a particular entity is allocated and that is determined using the various rules laid out in the CMP under the various seams agreements. It is used to limit the firm transmission service in future time periods.

The **Available Share of Total Flowgate Capability ("ASTFC")** is the available share of the TFC for an entity that is available on a flowgate for use as Firm transmission service. It is calculated as the share of the TFC minus the existing firm commitments that the entity has on the flowgate.

Local Balancing Authority ("LBA") is an operational entity or a Joint Registration Organization which is (i) responsible for compliance to NERC for the subset of NERC Balancing Authority Reliability Standards defined in the Balancing Authority Agreement for their local area within the ISO Balancing Authority Area, (ii) a Party to Balancing Authority Agreement, excluding ISO, and (iii) other requirements as specified by FERC, NERC, and state public utility commissions. See Section 2.2.1 for discussion of regulatory agencies such as FERC and NERC.

ISO and its Transmission Owners utilize an "$N - 1$" criteria to determine additions, deletions, or modifications to the list of flowgates. The ISO uses the following criteria to determine if a flowgate should be included in its AFC process:

- Once per year, the ISO completes a first contingency transfer analysis to identify flowgates. This first contingency transfer analysis is performed for ATC paths internal to ISO's market footprint up to 10,000 MW, which is beyond the maximum capability of any path to be studied and, further, is performed such that the first three limiting elements and their worst associated contingency combinations with an OTDF of at least 5% and within ISO's market footprint are included as flowgates. The ISO also includes the results from all adjacent

www.mepcafe.com

> Balancing Authority ("BA") source-and-sink combinations for such analysis. The results of this analysis are reviewed by the operations planning personnel to ensure the first contingency criteria used are consistent with those that are used in planning of operations including use of Special Protection System.

Furthermore, if any limiting element is kept within its limit for its associated worst contingency by operating within the limits of another flowgate, then no new flowgate needs to be established for such limiting elements or contingencies. Likewise, only the most limiting element in a series configuration needs to be included as a flowgate.

The ISO includes any limiting element/contingency combination within its Transmission model that has been requested to be included by any other TSP using the flowgate methodology. The requested flowgate will be included in the ISO AFC process, if it has at least 5% PTDF or OTDF impact from any generator within ISO's market footprint, or at least 5% PTDF or OTDF impact from any BAA within ISO's market footprint to an adjacent BAA.

The ISO also includes any flowgate within its Reliability Coordinator's area that has been subjected to an interconnection-wide congestion management procedure within the last 12 months unless it was created to address temporary operating conditions. If a flowgate has not been subjected to an interconnection-wide congestion management procedure within the last 24 months, the ISO can remove the flowgate from its AFC process.

The ISO also includes any limiting element/contingency combination within its market footprint that has been requested to be included by the ISO real-time operations personnel to manage congestion. The requested flowgate will be included in the ISO AFC process, if it has at least 5% PTDF or OTDF impact from any generator within the ISO's market footprint, or at least 5% PTDF or OTDF impact from any BAA within ISO's market footprint to an adjacent BAA.

The ISO uses the facility ratings provided by Transmission Owners/Operators to establish flowgate TFCs per the ISO SOL/IROL Methodology document. Flowgate definitions and TFCs are provided by the Transmission Owners/Operators (SOL flowgates), or the ISO Reliability Coordinator (IROL flowgates), which are determined based upon the summer and winter peak ambient conditions and by the flowgate selection criteria described previously. These flowgate TFCs represent the IROL or SOL of the defined flowgates. A list of flowgates along with their definitions and data are stored in the ISO web-based database. It is managed by the ISO and is updated at least once per calendar year or as requested within 30 calendar days from a flowgate request. Any addition to the flowgate list due to this annual update will be communicated to the ISO Transmission Owners/Operators for validation.

Furthermore, the ISO conducts an annual review of TFC for each flowgate and establishes them once per calendar year based on the inputs from its Transmission Owners/Operators. Besides the annual review, it is up to the Transmission Owners/Operators to notify ISO of any changes in the Rating that would affect the TFC of a flowgate used in the AFC process. If the change is related to temporary system conditions, the Transmission Owners/Operators need to specify the length of time

www.mepcafe.com

for the temporary TFC. The ISO will make updates to TFC changes within seven calendar days of the notification.

Additionally, the ISO uses assumptions no more limiting than those used in the planning of operations for the corresponding time period. More specifically, the ISO performs studies and provides daily reports for the next day operating condition. Any IROL and SOL changes from these daily reports that are related to existing flowgates in the AFC process are then implemented into the AFC process.

The ISO updates the list of flowgates used in the AFC calculations once a month due to computer system limitations. Sometimes, there is a need to add a flowgate during the month for an outage in the system or a new permanent flowgate that cannot wait until the monthly flowgate update process. Thus, the ISO has allowance for and models temporary flowgates in the AFC process. When a temporary flowgate is inactive, the TFC is set to infinity and the definition is set to a small transformer far outside the ISO footprint to prevent any response factor for any transaction path. When a flowgate is active, the definition (monitored and any contingent elements) is updated to reflect the constraint being modeled and the TFC is set to that of the flowgate for the time period that the constraint is active.

11.10.2 Process

Interchange is entered as a transmission service request (TSR) through an Internet-based database, the ISO OASIS site, for evaluation. TSR requests received are evaluated by the ISO Tariff Administration and the evaluation results are posted to the OASIS for the customer.

A transmission service request is an interchange contract defined by a Point of Receipt ("POR"), a Point of Delivery ("POD"). The POR and POD are the source and sink that identify the transmission path for the service requested. The valid list of PORs and PODs includes all qualified participants. The valid list of sources and sinks includes all entities and commercial nodes (generator and load areas) within the ISO business area and non-ISO business areas (BAs) as needed. The Service Increment describes the time increment of the request. The service increment can be hourly, daily, weekly, monthly, or yearly. The Service Class describes the class of service (firm or non-firm). The Service Type describes the type of service (network or point to point service) that can be specified.

TSRs are classified by status. These include, but are not limited to "STUDY," "ACCEPTED," "REFUSED," and "CONFIRMED." Once a TSR is received and under evaluation by an ISO, the TSR is considered to be in the STUDY status phase. If the TA team evaluates the request and determines that the requested can be granted, the request then achieves ACCEPTED status. If the request cannot be granted, the request is evaluated to determine the availability of partial service. If no partial service is available, the TSR is REFUSED. If the request achieves the ACCEPTED status and the customer wishes to take the service, the customer CONFIRMS the request. A TSR can also have path-related attributes with respect to the ISO footprint.

www.mepcafe.com

Each firm TSR request goes through an ATC check. If a firm TSR is constrained on one reciprocal flow and the ATC value is lower than the allowed value, the ATC value should be utilized for the purpose of approving/denying service. In this case, while the Allocation process might indicate that the entity has rights to a particular flow through the Allocation process; current conditions on that flow indicate that selling those rights would result in reliability problem due to insufficient ATC.

If a Reciprocal Entity uses all of its firm Allocation and desires to obtain additional capacity from another Reciprocal Entity who has remaining capacity, that additional capacity may be obtained through a request to share or transfer unused allocation.

The TSRs received are evaluated separately by class of service. If the request is for Non-Firm service, it is evaluated against the Non-Firm AFC on the limiting flowgates for the requested service. If the request is for Firm service, it is evaluated against the Firm AFC on the limiting flowgates for the requested service and the ASTFC on the limiting flowgates for the requested service.

11.10.3 Calculation ATC Methodology

ISO uses a power-flow-based approach to determine the ATC for evaluating requests for transmission service. For each transmission path defined by a source and sink, the flow-based approach identifies a set of Most Limiting Flows that impact this path. The incremental impact of a transaction on a flowgate is quantified by a distribution factor expressed by a percentage. For a TSR to be granted on a path, the incremental effect of the MW amount of the request must be smaller than the ATC on all Most Limiting flowgates impacted by this path.

The ISO uses up to a specified number of flows in the list of Most Limiting Flows for each transmission path while evaluating TSRs. These flowgates are determined based upon their available incremental transfer using the amount of capacity available and the response factor of the transmission path.

The ISO builds a power flow model for a peak condition for each time interval per service increment. The base flows from these models are used to compute AFC values for Firm and Non-Firm transmission service. Each calculation cycle is referred to as a horizon.

The power flow case is built from a base network topology model, load forecast, generation and transmission outages, and net interchange for each of the ISO areas and first tier areas. The individual loads are scaled to match with the load forecast value for that time point. The generation and transmission outages are applied using a topology processor. The generation dispatch order is determined based on the ISO market history.

A power flow is performed on the base case by enforcing net interchange. The power flow solution provides base flows. The distribution factors are computed from the solved case as well. Distribution factors represent the percent of a transaction from a specified source to a specified sink that appears on a point of impact. The distribution factors are topology dependent and can change due to system conditions. The Transmission Provider uses the thresholds defined subsequently to decide whether a transaction has a significant impact.

www.mepcafe.com

In order to produce credible constrained facility ATCs, the Transmission Provider must consider the effects of system conditions and transmission service that has been sold by other transmission providers.

ISO utilizes a bus branch representation of the power system for all horizons. Special treatment for DC lines and phase shifters, which regulate MW flows through them, are necessary to facilitate accurate calculation of ATCs and distribution factors.

Net interchange represents the MW amount of power that is either exported from or imported to an area. Net interchange values are expressed in MW and are signed numbers typically rounded to the first decimal. A positive number indicates that the area is exporting power and a negative number indicates that the area is importing power.

The solved power flow process generates a set of base flows and distribution factors for every flow in the ATC process for each time increment in each horizon run. Positive Impact (PIRULE) and Counterflow Impact (CIRULE) Rules are needed as it cannot be assumed that a Reservation is always scheduled for its full MW amount in real time. An example would be where a yearly reservation is scheduled during various seasons with a smaller MW amount than the original reservation MW amount or situations where an hourly or daily reservation could not be scheduled.

Similarly, while selling transmission service on a path, it is assumed that the Reservations in the reverse direction that impact the flowgates on the requested path are not flowing at the same time. In reality, some of these counterflow Reservations will be scheduled and flowing at the same time.

ATC is defined in this work as

$$\text{ATC} = \text{TTC} - \text{CBM} - \text{TRM} - \text{ETC}$$

This work will not cover all of these terms except to state that CBM, TRM, and ETC are margins required for stability, reliability, and previous commitments. We will only address the calculation of "TTC" in this work. Broadly define ATC as "The maximum amount of additional MW transfer possible between two parts of a power system." Note that additional recognizes that existing transfers are considered part of the "base case." Typically these two points of delivery are in different control areas. The points of delivery can really be any group of power injections.

What defines the Maximum injections? The rules include the following:

- No overloads should occur in the system as the transfer is implemented.
- No overloads should occur in the system *during contingencies* as the transfer is implemented.

Assume we want to calculate the ATC by incrementing the transfer, resolving the power flow, and iterating in this manner. This is similar to the solutions in the security analysis chapter.

Let us first gauge the computer requirements to achieve this number. Assume that 10 is a reasonable guess for the number of iterations that it will take to determine the ATC. This process has to be done under each contingency. Assume we have 1000

www.mepcafe.com

contingencies. This means we have 10 × 1000 power flows to solve. If it takes 10 sec to solve each power flow (a reasonable estimate as of this writing), then it will take 28 h to complete the computation for ONE transfer request!

An overview of the underlying mathematics of the power flow has been given in Chapter 6. An explanation of the linearized analysis techniques come from the security analysis chapter. Remember that the full AC Power Flow Equations Solution requires iteration of several thousand equations. Remember that the large matrix (**J**) is called the Jacobian, which are the sensitivities of the flow to the system state. The Real Power derivative equations and the Reactive Power derivative equations are stated in the power flow chapter. Remember that the derivatives simplify to a fixed set of "decoupled" equations following the assumptions of Brian Stott. This accelerates the iterations immensely even when control equipment is included, such as transformers and HVDC links. Remember that the B′ and B″ can be simply stated as follows:

- B′ is the imaginary part of the Y-Bus with all the "shunt terms" removed
- B″ is the imaginary part of the Y-Bus with all the "shunt terms" double counted

Also, as another possible simplification, the "DC Power Flow" equations are simply the real part of the decoupled power flow equations when voltages and reactive power are ignored. Then only angles and real power are solved for by iterating.

Note that the Full AC Power flow was solved by iterating

$$\begin{bmatrix} \Delta\delta \\ \Delta V \end{bmatrix} = [\mathbf{J}]^{-1} \begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix}$$

As shown in the security analysis chapter, one can model the transfer as a change in the injections ΔP for the Buyer and the Seller:

$$\Delta T_{\mathrm{B}} = [0 \quad 0 \quad \mathrm{PF}_{\mathrm{B}f} \quad 0\, \mathrm{PF}_{\mathrm{B}g} \quad 0]^{\mathrm{T}} \qquad \sum_{h=1}^{N} \mathrm{PF}_{\mathrm{B}h} = 1$$

$$\Delta T_{\mathrm{S}} = [0 \quad \mathrm{PF}_{\mathrm{S}x} \quad 0 \quad \mathrm{PF}_{\mathrm{S}y}\, 0 \quad 0]^{\mathrm{T}} \qquad \sum_{z=1}^{N} \mathrm{PF}_{\mathrm{S}z} = 1$$

As an example, assume buyer consists of—85% from bus 3 and 15% from bus 5. Then assume seller consists of—65% from bus 2 and 35% from bus 4, then the vectors would be

$$\Delta T_{\mathrm{B}} = [0 \quad 0 \quad .65 \quad 0 \quad .35 \quad 0]^{\mathrm{T}}$$
$$\Delta T_{\mathrm{S}} = [0 \quad .85 \quad 0 \quad .15 \quad 0 \quad 0]^{\mathrm{T}}$$

Then solve for the voltage and angle sensitivities by solving an iteration of the Full AC Power Flow:

www.mepcafe.com

$$\begin{bmatrix} \Delta\delta_S \\ \Delta V_S \end{bmatrix} = [J]^{-1} \begin{bmatrix} \Delta T_S \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} \Delta\delta_B \\ \Delta V_B \end{bmatrix} = [J]^{-1} \begin{bmatrix} \Delta T_B \\ 0 \end{bmatrix}$$

These are the sensitivities of the Buyer and Seller "sending power to the slack bus".

If we assume the total sensitivity to the transfer is the seller minus the buyer sensitivity, then

$$\Delta\delta = \Delta\delta_S - \Delta\delta_B$$
$$\Delta V = \Delta V_S - \Delta V_B$$

Implicitly, this assumes that ALL the change in losses shows up at the slack bus. Instead assign the change to the BUYER instead by defining

$$k = \frac{\Delta \text{Slack}_S}{\Delta \text{Slack}_B} = \frac{\text{Change in slack bus generation for seller sending power to slack}}{\text{Change in slack bus generation for buyersending power to slack}}$$

Then

$$\begin{bmatrix} \Delta\delta \\ \Delta V \end{bmatrix} = \begin{bmatrix} \Delta\delta_S \\ \Delta V_S \end{bmatrix} - k \begin{bmatrix} \Delta\delta_B \\ \Delta V_B \end{bmatrix}$$

Change in slack bus generation for buyer sending power to slack is equal to the change in slack bus generation for seller sending power to slack.

If we use the DC Power Flow equations

$$\Delta\delta = [B']^{-1} \Delta P$$

then determine angle sensitivities

$$\Delta\delta_S = [B']^{-1} \Delta T_S$$
$$\Delta\delta_B = [B']^{-1} \Delta T_B$$

Since the DC Power Flow ignores losses:

$$\Delta\delta = \Delta\delta_S - \Delta\delta_B$$

DC Power Flow equations:

$$\Delta\delta = \left[B'\right]^{-1} \Delta P$$

www.mepcafe.com

need to be augmented to include an equation that describes the change in flow on a phase shifter controlled branch as being 0. Similar to a line outage compensation in the security analysis chapter.

Line flow change is found:

$$P_{k,m} = B_{\delta}\Delta\delta + B_{\alpha}\Delta\alpha$$

Thus, instead of DC power flow equations we use the control augmented equations:

$$\begin{bmatrix}\Delta\delta \\ \Delta\alpha\end{bmatrix} = \begin{bmatrix}B' & 0 \\ B_{\delta} & B_{\alpha}\end{bmatrix}^{-1} \begin{bmatrix}\Delta P \\ 0\end{bmatrix}$$

Otherwise, the process is the same.

Phase Shifters are often on lower voltage paths (230 kV or less) with relatively small limits due to the construction costs of these devices. They are put there in order to manage the flow on a path that would otherwise commonly see overloads. Without including them in the sensitivity calculation, they constantly show up as "overloaded" when using Linear ATC tools.

Remember that PTDF measures the sensitivity of line MW flows to a MW transfer. Line flows are simply a function of the voltages and angles at its terminal buses. Using the Calculus Chain Rule, the PTDF is simply a function of these voltage and angle sensitivities. If Full AC equations are used, then $P_{k,m}$ is the flow from bus k to bus m:

$$\text{PTDF} = \Delta P_{k,m} = \left[\frac{\partial P_{k,m}}{\partial V_k}\right]\Delta V_k + \left[\frac{\partial P_{k,m}}{\partial V_m}\right]\Delta V_m + \left[\frac{\partial P_{k,m}}{\partial \delta_k}\right]\Delta \delta_k + \left[\frac{\partial P_{k,m}}{\partial \delta_m}\right]\Delta \delta_m$$

As a simplification, the Lossless DC approximations yield

$$\left[\frac{\partial P_{k,m}}{\partial V_k}\right] = 0$$

$$\left[\frac{\partial P_{k,m}}{\partial V_m}\right] = 0$$

$$\left[\frac{\partial P_{k,m}}{\partial \delta_k}\right] = b_{k,m}$$

$$\left[\frac{\partial P_{k,m}}{\partial \delta_m}\right] = -b_{k,m}$$

Define the Line Outage Distribution Factors (LODFs) as a function of the flow change to the injection. Then $\text{LODF}_{l,k}$ is the percent of the pre-outage flow on Line K that will show up on Line L after the outage of Line K:

$$\text{LODF}_{l,m} = \frac{\Delta P_{l.k}}{P_k}$$

www.mepcafe.com

Linear impact of an outage is determined by modeling the outage as a "transfer" between the terminals of the line similar to the example in the security analysis chapter.

The transfer request can be defined as transfer defined by change at each point of delivery.

Assume

$$\tilde{P} = \Delta P_n = \Delta P_m$$

Then the flow on the Switches is ZERO; thus, Opening Line K is equivalent to the "transfer."

Modeling an LODF as a Transfer is accomplished by setting up a transfer of P_k MW from Bus n to Bus m as a linear equivalent to outaging the transmission line. Let us assume we know what P_k is equal to, so then we can calculate the values relevant to the LODF. Calculating the relevant values by using PTDFs for a "transfer" from Bus n to Bus m as an estimate of post-outage flow on Line l:

$$\Delta P_{l.k} = \text{PTDF}_l \ \tilde{P}_k$$

And an estimate of flow on Line k after transfer:

$$\tilde{P}_k = P_k + \text{PTDF}_k \ \tilde{P}_k$$

$$\tilde{P}_k = \frac{P_k}{1 - \text{PTDF}_k}$$

Thus, one can show:

$$\text{LODF}_{l,k} = \frac{\Delta P_{l,k}}{P_k} = \frac{\text{PTDF}_k \tilde{P}_k}{P_k} = \frac{\text{PTDF}_l (P_k / (1 - \text{PTDF}_k))}{P_k}$$

$$\text{LODF}_{l,k} = \frac{\text{PTDF}_l}{1 - \text{PTDF}_k}$$

$$\tilde{P}_k = \frac{P_k}{1 - \text{PTDF}_k}$$

We have a simple function of PTDF values.

Line Closure Distribution Factors (LCDFs) are similarly found. $\text{LCDF}_{l,k}$ is the percent of the post-closure flow on Line k which will show up on Line l after the closure of Line k:

$$\text{LODF}_{l,k} = \frac{\Delta P_{l,k}}{\tilde{P}_k}$$

Linear impact of a closure is determined by modeling the closure as a "transfer" between the terminals of the line just as a line outage was modeled in security analysis.

www.mepcafe.com

Modeling the LCDF as a Transfer is accomplished by creating a "transfer" defined by the change in injections at point of delivery.

Assume

$$\tilde{P}_k = \Delta P_n = \Delta P_m$$

Then the net flow to and from the rest of the system are both 0; thus, closing line k is equivalent to the "transfer." Thus, setting up a transfer of $-P_k$ MW from Bus n to Bus m is linearly equivalent to outaging the transmission line.

Let us assume we know what $-P_k$ is equal to, then we can calculate the values relevant to the LODF. Note: The negative sign is used so that the notation is consistent with the LODF "transfer" direction. The estimate of post-closure flow on Line l

$$\Delta P_{l.k} = \text{PTDF}_l \left(-\tilde{P}_k\right)$$

Thus, one can write:

$$\text{LCDF}_{l,k} = \frac{\Delta P_{l,k}}{\tilde{P}_k} = \frac{-\text{PTDF}_k \tilde{P}_k}{\tilde{P}_k} = -\text{PTDF}_l$$

Thus, the LCDF is exactly equal to the PTDF for a transfer between the terminals of the line.

Modeling Linear Impact of a Contingency can be estimated by OTDFs. These are the percent of a transfer that will flow on a branch M after the contingency occurs. The Outage Flows (OMWs) are the estimated flow on a branch M after the contingency occurs.

OTDFs and OMWs can be shown for a Single Line Outage:

$$\text{OTDF}_{\text{m},\ell} = \text{PTDF}_{\text{m}} + \text{LODF}_{\text{m},\ell}\text{PTDF}_{\ell}$$
$$\text{OMW}_{\text{m},\ell} = \text{MW}_{\text{m}} + \text{LODF}_{\text{m},\ell}\text{MW}_{\ell}$$

as well as for a Multiple Line Outage:

$$\text{OTDF}_{\text{m,c}} = \text{PTDF}_{\text{m}} + \sum_{k=1}^{n_c}[\text{LODF}_{\text{m},k}\text{NetPTDF}_k]$$
$$\text{OMW}_{\text{m,c}} = \text{MW}_{\text{m}} + \sum_{k=1}^{n_c}[\text{LODF}_{\text{m},k}\text{NetMW}_k]$$

We need to define NetPTDF_k and NetMW_k. Each NetPTDF_k is a function of all the other NetPTDFs because the change in status of a line affects all other lines (including other outages). Assume we know all NetPTDFs except for the first one, NetPTDF_ℓ. Then one can write:

$$\text{NetPTDF}_\ell = \text{PTDF}_\ell + \sum_{k=2}^{n_c}[\text{LODF}_{\ell,k}\text{NetPTDF}_k]$$

www.mepcafe.com

In general for each Contingent Line *N*, define

$$\mathrm{PTDF}_N = \mathrm{NetPTDF}_N - \sum_{\substack{k=1\\k\neq N}}^{n_c} [\mathrm{LODF}_{N,k}\mathrm{NetPTDF}_k]$$

Thus, we have a set of n_c equations and n_c unknowns where n_c is the number of contingent lines:

$$\begin{bmatrix} 1 & -\mathrm{LODF}_{1,2} & \dots & -\mathrm{LODF}_{1,n_c} \\ -\mathrm{LODF}_{2,1} & 1 & \dots & -\mathrm{LODF}_{2,n_c} \\ \dots & \dots & 1 & \dots \\ -\mathrm{LODF}_{n_c,1} & -\mathrm{LODF}_{n_c,2} & \dots & 1 \end{bmatrix} \begin{bmatrix} \mathrm{NetPDTF}_1 \\ \mathrm{NetPDTF}_2 \\ \dots \\ \mathrm{NetPDTF}_{n_c} \end{bmatrix} = \begin{bmatrix} \mathrm{PTDF}_1 \\ \mathrm{PTDF}_2 \\ \dots \\ \mathrm{PTDF}_{n_c} \end{bmatrix}$$

Thus, the previous set of equations is similarly found by Gaussian Elimination:

$$\mathrm{NetPTDF}_C = [\mathrm{LODF}_{CC}]^{-1}\mathrm{PTDF}_C$$

Same type of derivation shows that

$$\mathrm{NetMW}_C = [\mathrm{LODF}_{CC}]^{-1}\mathrm{MW}_C$$

Given the high number of transfer requests, the goal of fast ATC analysis is to avoid power flow solutions. When completely solving ATC, the number of power flow solutions required is equal to the product of

- The number of contingencies
- The number of iterations required to determine the ATC (this is normally smaller than the number of contingencies)
- The number of control devices that will react to return the system to a more desired "normal state"

Consider three methods of which two are linearized.

First, the Single Linear Step (fully linearized) which performs a single power flow, then all linear solutions that are *extremely* fast. The Iterated Linear Step (mostly linear), requires iterations of power flow to ramp out to the maximum transfer level, but no power flows solved for any contingencies.

Second, use the IL, then Full AC solution that requires iterations of power flow and full solution of contingencies.

Third, consider a Single Linear Step ATC.

For each line in the system determine a Transfer Limiter Value T_M

www.mepcafe.com

$$T_M = \begin{bmatrix} \dfrac{\text{Limit}_M - \text{MW}_M}{\text{PTDF}_M}; & \text{PTDF}_M > 0 \\ \infty(\text{infinite}); & \text{PTDF}_M = 0 \\ \dfrac{-\text{Limit}_M - \text{MW}_M}{\text{PTDF}_M}; & \text{PTDF}_M < 0 \end{bmatrix}$$

Then, for each line during each contingency determine another Transfer Limiter Value.

The linear estimates of OTDF and OMW are quite accurate (usually within 2%) based on present network structures and values. However, this can lead to big errors in ATC estimates.

Consider the following:

- Assume a line's present flow is 47 MW and its limit is 100 MW
- Assume OTDF = 0.5%; assume OMW = 95 MW
- Then ATC = (100 − 95)/0.005 = 1000 MW
- Assume 2% error in OMW (1 MW out of 50 MW change estimate)
- Actual OMW is 96 MW
- Assume 0% error in OTDF
- Actual ATC is then (100–96)/0.005 = 800 MW
- 2% error in OMW estimate results in a 25% overestimate of the ATC

The transfer limit can then be calculated to be the minimum value of TM or TM,C for all lines and contingencies.
A robust simulator saves several values with each Transfer Limiter:

- TM or TM,C [Transfer Limit]
- Line being monitored [Limiting Element]
- Contingency [Limiting Contingency]

This is needed for filtering errors

- OTDF or PTDF value [%PTDF_OTDF]
- OMW or MW value [Pre-Transfer Flow Estimate]
- Limit Used (negative Limit if PTDF_OTDF <0)
- MW value initially [Initial Value]

The pros and cons of linear step ATC is that the Single Linear Step ATC is *extremely* fast. But it is dependent on the linearization being quite accurate in modeling the impact of contingencies and transfers. It only uses derivatives around the present operating point. Thus, control changes as you ramp out to the transfer limit are NOT

www.mepcafe.com

sufficiently small for the chain rule to apply as the derivatives are only valid for small changes in operating point. Thus, special arrangements for Phase Shifters and HVDC links must be included. The possibility of generators participating in the transfer hitting limits is NOT always modeled but should be to include the actual future state of the power system.

The Iterated Linear Step ATC that takes into account these control changes, performs the following:

1. Step Size = ATC using Single Linear Step
2. If [abs(Step Size) < Tolerance] then stop
3. Ramp transfer out an additional amount of Step Size
4. Resolve Power Flow (slow part, but takes into account all controls)
5. At new operating point, Step Size = ATC using Single Linear Step
6. Go to step 2

This is a reasonably fast procedure to address all transfer requests. It is unfortunately on the order of 10 times slower than Single Linear Step.

Takes into account all control changes because a full AC Power Flow is solved to ramp the transfer. Including OPF constraints in (IL) to enforce Interface Flows. When ramping out the transfer, Simulator can be set to enforce a specified flow on an interface.

This introduces a radical change in control variables that is best modeled by completely resolving using the OPF. The objective of the OPF is to minimize the total controller changes (sum of generator output changes).

Why would this be necessary? It is needed to represent a normal operating guideline that is obeyed when transfers are changed.

Consider BPA responsibilities to flows in the Pacific North West. Operating procedures for BPA require maintenance of the "interface" flows into Seattle in specific ranges. These are stability constraints!

IL, then Full AC Method performs the following:

1. Run Iterated Linear Step and ramp transfer out ATC Value found
2. Step Size = 10% of the initial Linear Step Size saved during the (IL) method, or 50 MW, whichever is larger.
3. Run Full Contingency Analysis on the ramped transfer state
4. If there are violations then change the sign of Step Size
5. If [abs(Step Size) < Tolerance] then Stop
6. Ramp transfer out an additional amount of Step Size and resolve Power Flow
7. At new operating point, Run Full Contingency Analysis
8. If [(Step Size > 0) and (There are Violation)] OR [(Step Size < 0) and (There are NO Violations)] THEN Step Size:= −Step Size/2
9. Go to step 5

www.mepcafe.com

This is an extremely slow process, so few transfer requests would be authorized. It is the "Number of Contingencies" times slower than the iterated linear. If you have 1000 contingencies, then this is 1000 times slower.

Given the need to process a large number of request, the following recommendations handle the process in line with the given data accuracy.

Single Linear Step

– Use for all preliminary analysis, and most analysis in general.

Iterated Linear Step

– Only use if you know that important controls change as you ramp out to the limit.

Full AC should not be used presently as it is just too slow.

Note that the marginal gain in accuracy compared to (IL) (less than 2%) does not justify the time requirements. Remember that ATC numbers probably are not any more than 2% accurate anyway! Note what limits were chosen, what generation participates in the transfer, how accurate is the demand forecast for the period of the request, etc.

Clearly, an on-line auction scheme that immediately clears transactions is needed to provide open access to all participants in a fully automated process. The Wholesale Market Analysis implemented by COMED did just this with a probability approach to contingency analysis. This tariff was put into place in 1995.

11.11 SECURITY CONSTRAINED UNIT COMMITMENT (SCUC)

The Security Constrained Unit Commitment is a means of building a unit commitment schedule where each hour's set of committed generating units is tested for contingencies (n-1 outages) to be sure that the commitment does not leave the system vulnerable to outages of transmission lines or generators.

11.11.1 Loads and Generation in a Spot Market Auction

An open market where generator companies and load companies participate by entering bids to buy and sell could be quite simple if they did not have to connect through a transmission system that had to be operated securely. That is, without the transmission system such a market could be accomplished, as many commodity markets are, by having representatives in a room with offers to buy and sell and matching them. However, such a scheme might very well result in a set of generation/load patterns on the transmission system that would not pass basic security analysis and would be rejected. The best means of accommodating transmission system security and having an open market is to set up an OPF as an auction where generator asking price, or bid, is entered as a function $F_i(P_i)$ which is presented to the market

www.mepcafe.com

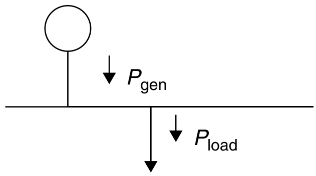


FIGURE 11.7 Simple market example.

by the generation company as a \$/MWh function. Note this is not a cost function as we have used in Chapters 3 and 8 but a function representing the price being asked by the generation company.

If the loads in a power system are assumed to be constant, we say they are price-inelastic and that they will not change the load drawn from the bus regardless of the price being charged to them. If all load is price-inelastic, then the OPF has an objective function that contains only the generator bid functions. For loads that are price-elastic, we use a "worth" function $W_j(P_j)$ which represents the price the entity operating that load is willing to pay to purchase P_j amount of load.

The OPF objective is then to maximize the worth of load served and minimize the cost of generation. Since we usually approach the OPF as a minimization problem, our objective function then becomes

$$\min \sum_i F_i(P_i) - \sum_j W_j(P_j)$$

In the field of economics, the generation would be a supplier and $F_i(P_i)$ is the "supply curve." Similarly, $W_j(P_j)$ is the "demand curve." We can set up a very simple market as shown subsequently with one generator and one load which send their respective functions to the market coordinator who decides how much power is to be produced by the generator and consumed by load (Figure 11.7).

We now state the optimization as

$$\min F_{\text{gen}}(P_{\text{gen}}) - W_{\text{load}}(P_{\text{load}})$$

Subject to $P_{\text{gen}} - P_{\text{load}} = 0$

The Lagrange function for this simple optimization is

$$\mathcal{L} = F_{\text{gen}}(P_{\text{gen}}) - W_{\text{load}}(P_{\text{load}}) + \lambda(P_{\text{gen}} - P_{\text{load}})$$

The solution to this optimization is

$$\frac{d\mathcal{L}}{dP_{\text{gen}}} = \frac{dF_{\text{gen}}(P_{\text{gen}})}{dP_{\text{gen}}} + \lambda = 0$$
$$\frac{d\mathcal{L}}{dP_{\text{load}}} = \frac{-dW_{\text{load}}(P_{\text{load}})}{dP_{\text{load}}} - \lambda = 0$$

www.mepcafe.com

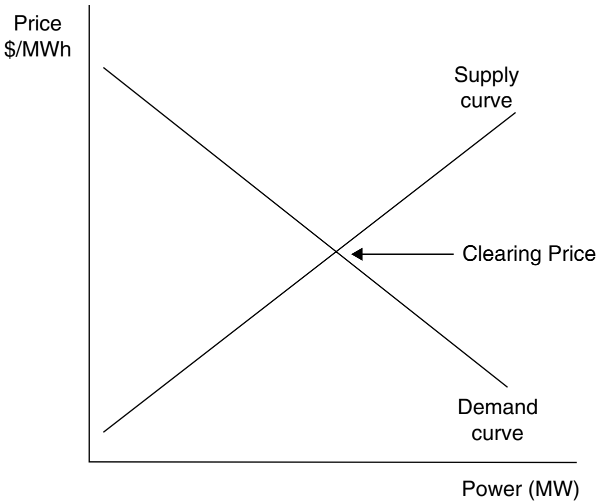


FIGURE 11.8 Supply–demand equilibrium price.

with the result

$$\frac{dF_{gen}(P_{gen})}{dP_{gen}} = \frac{dW_{load}(P_{load})}{dP_{load}}$$

The derivative or marginal functions ($dF_{gen}(P_{gen})/dP_{gen}$) and ($dW_{load}(P_{load})/dP_{load}$) are the supply price curve and the demand price curve, respectively. Usually, economists draw these as a rising curve for supply and a falling curve for demand, and this makes sense here as well. Generation will have high marginal cost (price) as generation goes up, and similarly, demand will be highest when the price is lowest and then go down as the price falls. The result is a graph like this seen in any elementary economics text (Figure 11.8). The point at which the marginal price for supply and the marginal price for demand are equal is called the clearing price.

11.11.2 Shape of the Two Functions

We have been using a quadratic function to express the generator cost or pricing function throughout the book as

$$F_{gen}(P_{gen}) = a + bP_{gen} + cP_{gen}^2$$

We shall also represent the demand as a quadratic function with an important note:

$$W_{load}(P_{load}) = A + BP_{load} + CP_{load}^2$$

www.mepcafe.com

The values of a and A in these functions are not important to this discussion since we are only interested in the first derivatives. Then the derivative functions are

$$\frac{\mathrm{d}F_{\text{gen}}(P_{\text{gen}})}{\mathrm{d}P_{\text{gen}}} = b + 2cP_{\text{gen}}$$

$$\frac{\mathrm{d}W_{\text{load}}(P_{\text{load}})}{\mathrm{d}P_{\text{load}}} = B + 2CP_{\text{load}}$$

We note that for generators b and c are both positive, while for the load function, B is positive and C must be negative, and $B > b$ to result in the supply demand curves shown in the figure.

11.11.3 Meaning of the Lagrange Multipliers

As shown in Chapter 8 on OPF, we have one Lagrange multiplier, λ, for each equality constraint in the OPF. Thus, there is a Lagrange multiplier for each bus in the power flow, and if the OPF is solving the AC network there is one Lagrange multiplier for the real power and one for the reactive power at the load. The Lagrange multiplier is the instantaneous price of the next small increment of load—making it the zone price in \$/MWh or the LMP for that bus (see Section 8.11).

If one is using the OPF with the DC power flow model then there are no incremental losses and the LMPs for the system will all be equal in the absence of any binding line or contingency constraints. With binding line and contingency constraints, the bus LMPs are different as shown in Example 8C.

The LMPs are considered prices at each load bus and are the price to be charged to any load on the bus, similarly, the LMP at a generator bus is the price paid to the generator supplying power at the generator bus. If again, we are using the DC power flow in the OPF, then, without congestion, all LMPs are equal and the total dollars collected from the loads equals exactly the dollars paid out to the generators. If there is congestion, the dollars collected will be greater than that paid out and this surplus is the basis of making payments to holders of contracts for transmission access.

Example 11C: In this example we shall use our six-bus system, but instead of the usual generator "cost" functions, we will use generator bid price functions and load worth functions, as given in the following table:

TABLE 11.1 Table of Generator and Load Functions for the Market Simulation

	a	b	c	$P_{\min}$	$P_{\max}$
Gen on bus 1	213.1	11.669	0.00533	50	200
Gen on bus 2	200	10.333	0.00889	37.5	150
Gen on bus 3	240	10.833	0.00741	45	180
Gen on bus 4	0	16	–0.015	0	100
Gen on bus 5	0	16	–0.015	0	100
Gen on bus 6	0	16	–0.015	0	100

www.mepcafe.com

The result with no line limits being checked is shown in the next table.

TABLE 11.2 Base Case Dispatch with no Limit Checking

Bus Number	Generation (MW′)	Load (MW)	Bus Lambda
1	72.6	0	12.44
2	118.7	0	12.44
3	108.7	0	12.44
4	0	100	12.44
5	0	100	12.44
6	0	100	12.44

Note that bus lambdas (LMP) are the same reflecting no limits hit.

With line limit checking there is a difference is shown in the next table.

TABLE 11.3 Results with Limit Checking

Bus Number	Generation (MW)	Load (MW)	Bus Lambda
1	98.6	0	12.72
2	120.7	0	12.48
3	77.4	0	11.98
4	0	98.9	13.03
5	0	100	12.62
6	0	97.9	13.06

Note that the bus lambdas are now different and small amounts of load were not served at bus 4 and bus 5. Lines 2–4 and lines 3–6 were both at their limit of 60 MW.

11.11.4 The Day-Ahead Market Dispatch

Markets usually start with receiving selling bids from generators and purchasing bids from loads for each hour of a 24 h period starting at 12:00 a.m. and ending at 11:59 p.m. the next day. Typically, these bids must be submitted by a deadline during the day and the 24 h ahead calculation is done and results released before the day being scheduled starts.

To perform the day-ahead calculation, a security constrained OPF is built into a unit commitment program, that is, the SCUC so that each candidate combination of commitment of generators is subjected to an SCOPF to be sure that the resulting commitment is secure and to generate LMPs for each bus for that hour.

Upon conclusion, the SCUC reports the schedule of generation commitment for the day ahead as well as the LMPs that will be used each hour to pay generators and to charge loads.

www.mepcafe.com

11.12 AUCTION EMULATION USING NETWORK LP

The auction emulation in economic dispatch used heuristic methods to match the suppliers and the buyers. That presentation did not include flow limits. Flow limits can be added as point-to-point real power flow limitations simply by adding flow limitations from each buyer to each seller. Instead, either a DC- or AC-based power flow calculation can be implemented to include heat-limited transfer capability, transient and dynamic stability limitations, as well as voltage stability limitations.

11.13 SEALED BID DISCRETE AUCTIONS

There are many markets in every society. Farmer's markets, fish markets, grocery markets, and most stores are examples. A market is where a number of sellers meet with a number of buyers to exchange products. There are many types of markets. One-sided markets have a representative for one of the sides, such as sellers, who deals with a number of buyers. Two-sided markets have buyers and sellers in one common forum, normally a trading ring. Some markets are open outcry, where prices are posted (yelled) for all to see. Some markets are sealed bid, where the bids are submitted and then opened to determine the winner of the auction. The auctioneer is the representative for one of the two sides who decides what the price and quantity are for each sale. All of the contracts discussed in this section assume that all are bilateral contracts.

Consider the market situation when "Sellers" have an amount of product to sell:

$$a_i > 0 \quad \text{for} \quad i = 1, \ldots, m$$

and when "Buyers" have an amount to buy:

$$b_i > 0 \quad \text{for} \quad i = 1, \ldots, n$$

Assume for now that there is one homogeneous product offered by each seller. Thus, there is a common product from each seller. Assume that there are heterogeneous products amongst sellers, thus there is some difference between sellers. The price would have to be the same if all sellers offered the same product with the same quality, unless some of the information was hidden. There are different values for the product from each seller for the following discussion.

The following model assumes that a buyer j tenders a sealed bid for the product from a seller i:

$$c_{ij} \geq 0$$

This is the maximum amount the buyer is willing to pay for one unit of product from seller i's based on the buyers value of the product.

www.mepcafe.com

Designate the selling of the product as seller *i* sells to buyer *j* an amount *x*:

$$x_{ij} \geq 0$$

at a price based on the auction, designated by the Lagrangian multiplier for the seller's constraint at the optimum:

$$u_i \geq 0$$

The total bid value is found by multiplying the bid times the quantity exchanged. The total cost is the product of the price times the quantity exchanged. Thus, there is a need to find the prices and the amounts.

Seller *i* receives a price and quantity set by the auctioneer or the rules of the exchange.

The buyer's surplus is determined by the KKT conditions at the optimal solution, just as the transportation problem is solved:

$$c_{ij} - u_i - v_j = 0$$

Note that the Lagrangian dual variable *v* is the buyer's surplus. Note that we should require the price to be positive:

$$v_j \geq 0$$

That the aforementioned should be a solution to the auction problem can be presented as a set of fairness requirements that both the buyers and the sellers would agree to enter the auction:
The rational seller should be required to:

- Not sell more than is available
- Get the same price for all units sold
- If the price is positive, then more units are sold
- If the price drops to 0, then no more units are sold

Since we are using the simplex, one should notice that a buyer and a seller are matched as each vertex is reached. The simplex will continue to match buyer and seller until the entire buyer surplus is consumed. The equivalent LP formulas are

$$\begin{aligned}
&\sum x_{ij} \leq a_i \\
&u_i \text{ constant } \forall x_{ij} \text{ sold} \\
&u_i > 0 \Rightarrow x_{ij} = a_i \\
&x_{ij} < a_i \Rightarrow u_i = 0
\end{aligned}$$

www.mepcafe.com

The rational buyer should be required to

- Not buy more than is needed
- Get the same surplus for all units bought
- If the surplus is positive, then more units are bought
- If the surplus drops to 0, then no more units are bought

The equivalent LP formulas are

$$\begin{aligned}
&\sum x_{ij} \le b_j \\
&v_j \text{ constant } \forall x_{ij} \text{ bought} \\
&v_j > 0 \Rightarrow x_{ij} = b_j \\
&x_{ij} < b_j \Rightarrow v_j = 0
\end{aligned}$$

Note that at this optimal solution that the price and the surplus will be of zero value. This shows how a market could collapse when supply does not equal demand. We will show an additional condition later that raises the final price above 0.

The LP form based on the aforementioned requirements is presented to determine how this formulation fits the desired auction solution. The objective is to maximize the number of transactions to obtain all of the buyer's positive surplus:

$$\begin{aligned}
\text{Maximize} \quad & \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \\
\text{subject to:} \quad & \sum_{j=1}^{n} x_{ij} \le a_i, \quad i = 1,2,\ldots,m \\
& \sum_{i=1}^{m} x_{ij} \ge b_j, \quad j = 1,2,\ldots,n \\
& x_{ij} \ge 0, \quad i = 1,2,\ldots,m, \quad j = 1,2,\ldots,n
\end{aligned}$$

Each of these constraints should be checked with the aforementioned requirements.

The objective function embodies the auction goal that the maximum total bid value is found. An economist would refer to this as the "invisible hand." This is also an economic potential function. This potential function is at a maximum at the optimum.

As determined by the LP conditions for optimum, and for the transportation algorithm, the complementary slackness conditions for the seller are the following:

$$\begin{aligned}
&\sum_{j=1}^{n} x_{ij}^{*} - a_i \le 0 && \forall i \\
&u_i^{*} \ge 0 && \forall i \\
&u_i^{*}\left(\sum_{j=1}^{n} x_{ij}^{*} - a_i\right) = 0 && \forall i
\end{aligned}$$

www.mepcafe.com

The complementary slackness conditions for the buyer follow

$$\sum x_{ij} - b_j \le 0$$
$$v_j^* \ge 0$$
$$v_j^* \left(\sum x_{ij}^* - b_j \right) = 0$$

These conditions fit the auction rules desired as stated previously. The dual variables satisfy the dual complementary slackness at the optimum:

$$u_i^* + v_j^* - c_{ij} \ge 0$$
$$x_{ij}^* \ge 0 \qquad \forall_{i,j}$$
$$x_{ij}^* \left(u_i^* + v_j^* - c_{ij} \right) = 0$$

As required by the KKT conditions, just as the aforesaid hold for the individual transactions, they hold for the sum of all transactions. Thus, the value distributed at the optimum is spread to the buyers and the sellers.

$$\sum\sum c_{ij} x_{ij}^* = \sum\sum u_i^* x_{ij}^* + \sum\sum v_j^* x_{ij}^*$$

The first problem with this solution is that the dual variables are not unique at the optimum. An artificial value "*k*" can be added to the first dual variable (instead of setting it to 0). Thus, the prices are arbitrary. It depends on the rules of the auction, the availability of other auctions, the oversupply or the undersupply of this market, and other conditions. The problem with this formulation for the sellers is that the price drops to 0, a condition that the sellers would not agree to enter without an additional option. This additional option is to set a minimum price for the sale of any units. This minimum price is a reservation price.

Example D: Consider the following problem as matching the following sellers and buyers where the prices are an indication of the transmission charges plus operational cost.

TABLE 11.4 Bid/Ask Offers

Seller	Buyer	Price	Amount
1	1	25	100
1	2	30	100
1	3	50	200
2	1	30	100
2	2	35	100
2	3	40	200
3	1	50	100
3	2	40	100
3	3	30	200

www.mepcafe.com

The limits of transmission are given in the following table from each seller to each buyer.

TABLE 11.5 Inter Agent Flow Limits

Seller/Buyer	1	2	3
1	200	200	200
2	100	100	100
3	100	200	200

The LP solution is found using either the Simplex or Transportation Network Flows method.

TABLE 11.6 Matched Bid/Ask Offers

Seller	Buyer	Price	Amount
1	1	25	0
2	1	30	0
3	1	15	100
1	2	30	200
2	2	35	100
3	2	20	0
1	3	50	5
2	3	40	100
3	3	30	200

This technique can be modified to use ATCs and reservation prices.

Reservation Price

The reservation price is the lowest price for seller to be willing to sell the commodity. It shows price rigidity.

The formulation is to change the LP formulation to include the reservation price:

$$\text{Max} \quad \ldots + \sum \Pi_i y_i$$

$$\sum x_{ij} + y_i = a_i$$

The corresponding change to the dual constraints follows:

$$u_i^* + v_j^* \geq c_{ij} \qquad \forall_{i,j}$$

$$u_i^* \geq \Pi_i \qquad \forall_i$$

The additional complementary slackness conditions follow:

$$\left(u_i^* - \Pi_i\right) \; y_i^* = 0 \quad \forall_i$$

www.mepcafe.com

Now it is required at the optimal solution that all quantities either are sold or are bought back by the seller:

$$y_i^* > 0 \Rightarrow u_i^* = \Pi_i$$
$$u_i^* > 0 \Rightarrow y_i^* = 0$$

PROBLEMS

11.1 Four areas are interconnected as shown in Figure 11.12. Each area has a total generation capacity of 700 MW currently on-line. The minimum loading of these units is 140 MW in each area. Area loads for a given hour are as shown in Figure 11.12. The transmission lines are each sufficient to transfer any amount of power required (Figure 11.9).

The composite input–output production cost characteristics of each area are as follows:

$$F_1 = 200 + 2P_1 + 0.005P \text{ (g/h)}$$
$$F_2 = 325 + 3P_2 + 0.002143P; \text{ (ft/h)}$$
$$F_3 = 275 + 2.6P_3 + 0.003091/1 \text{ (g/h)}$$
$$F_4 = 190 + 3.1P_4 + 0.00233/1 \text{ (ft/h)}$$

In all cases, 140 << 700 MW. Find the cost of each area if each independently supplies its own load, and the total cost for all four areas.

11.2 Assume that area 1 of Problem 10.1 engages in two transactions:

a. Area 1 buys 190 MW from area 2.

b. Area 1 sells 120 MW to area 3.

For each of these transactions, the price is based upon a 50–50 split-savings agreement. Find the price of each transaction, the net generation costs for each area including the sum it pays or receives under the

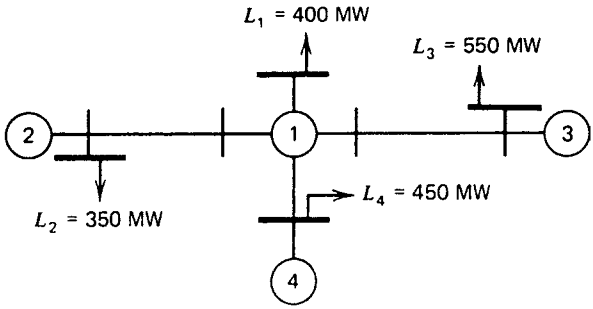


FIGURE 11.9 Four-area system for Problem 11.1.

www.mepcafe.com

split-savings agreement, with the order of the transactions (as given earlier) being as follows:

i. **a** then **b**.

ii. **b** then **a**. In both instances, find the total cost for the four-area pool.

11.3 Assume that the four areas of Problem 11.1 are centrally dispatched by a pool control center.

a. Find the generation and production cost in each area.

b. Assume a split-savings pool agreement such that each area exporting receives its increased costs of production plus its proportionate share of 50% of the pool savings. Find the cost per MWh of transfer energy (i.e., "pool energy") and the net production cost of each area.

11.4 Assume that the four areas of Problem 11.1 are members of a "power broker." Previous to the hour shown in Problem 11.1, each area submits quotations to the broker to sell successive blocks of 25 or 50 MW and bids to purchase blocks of 25 or 50 MW. In furnishing these data to the broker, assume that the prices quoted are the average incremental costs for the block. The broker matching rules are as follows:

Rule 1. Quotations to sell and bids to buy are matched only wherever there is a direct connection between the quoting and bidding company.

Rule 2. Transactions are arranged in a priority order where the lowest remaining incremental cost for the quoting area is matched with the highest decremental cost for the bidding areas. [That is, lowest available incremental cost energy available for sale is matched with the area with the greatest available potential incremental cost savings (= decremental cost).]

Rule 3. "Matches" may be made for all or part of a block. The remainder of the block must be used, if possible, before the next block is utilized. Matching will cease when the absolute value of the difference between the incremental and decremental cost drops below 0.33 Il/MWh.

Rule 4. No area may be both a buyer and a seller in any given hour to the same other company.

Rule 5. The price per MWh for each matched transaction is one-half the sum of the absolute values of the incremental and decremental costs.

For this problem, assume that quotes and bids are supplied to the broker by each area as follows:

Area	Quotes to Sell	Quotes to Buy
	100 MW in 25 MW blocks	100 MW in 25 MW blocks
2	200 MW in 50 MW blocks	None
3	None	200 MW in 50 MW blocks
4	25 MW	25 MW

www.mepcafe.com

a. Set up the power broker matching system and establish the transactions that can take place and the price of each.

b. Assume that all feasible transactions take place and find the net production cost to each area and the pool.

11.5 Repeat Problem 11.4 with the following assumptions simultaneously taken in place of those in Problem 11.4:

a. Each area is interconnected with every other area and transfers may take place directly between all pairs of areas.

b. The matched transactions will proceed until the difference between decremental costs is 0 instead of 0.33 g/MWh.

11.6 Repeat Problem 11.5 with one "clearing price" that applies to all transactions and is equal to the price determined for the last matched transaction.

11.7 Use the cost data for the six-bus base case in Chapter 6, and the power flow and generator output data presented in Figures 11.3 (Cases 1, 2, and 3) that illustrate the wheeling of 20 MW between bus 3 and bus 4. We want to compute an estimate of the utility's net costs under all three cases. Let

Net cost = total production cost for all generators—charges for wheeling
Note that the total production cost for each case is given in the tables just after Figure 11.3.
Produce a table that shows the power generation for each unit and the total system operating cost in $/h for the three cases: the base case and the two wheeling cases.
For the two cases with 20 MW being wheeled, compute the charges for wheeling as (20 MW × the SRMC) for wheeling given in the chapter. These are

$$13.2494 - 12.376 = 0.8734 \text{ for the unconstrained case and}$$
$$16.4711 - 12.2015 = 4.2696 \text{ for the constrained case.}$$

These charges represent income to the utility and reduce the total operating cost. (The question is really: "Does the use of the SRMC for wheeling only recover additional operating costs for the wheeling, or does it make an added profit for the utility?" Remember, this is only one example.)

11.8 Find the ATCs for all generation injection and demand buses for Example 8D.

11.9 Find the solution for Example 8D using the network flows auction method.

11.10 You are asked to analyze the bilateral transaction versus a market auction for the three-bus system (Figure 11.10):

www.mepcafe.com

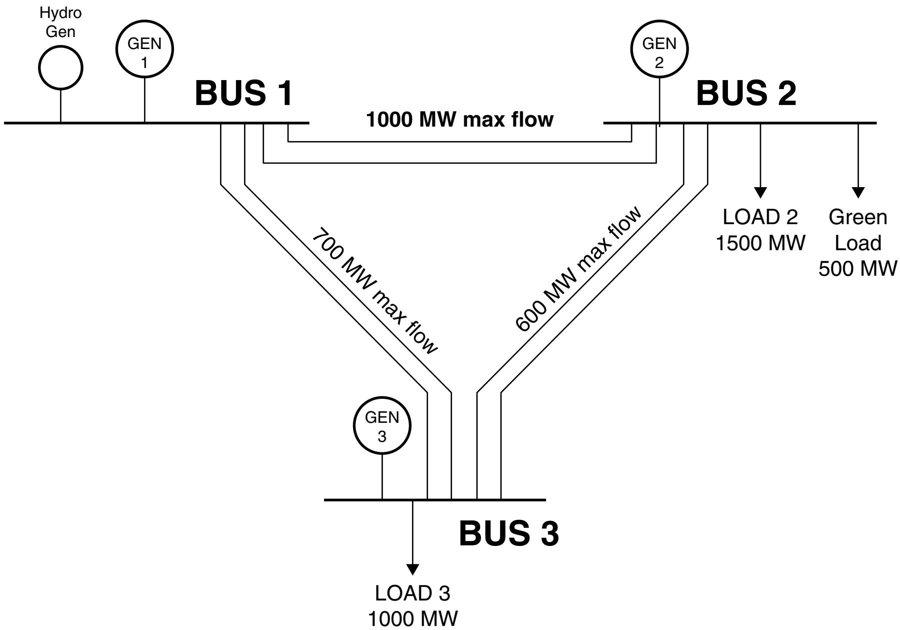


FIGURE 11.10 Three-bus system for Problem 11.10.

The Green Load at bus 2 is trying to buy 500 MW and would like to buy from the Hydro Generator on bus 1. The Green Load trading group calls up the Hydro Generator trading group and asks for a price for 500 MW. The Hydro Generator Trading Group asks a price of 60 $/MWh. The hydro group has to pay the transmission charges for any contract.
The traders at Green Load trading think that this price is too high and say "no thanks—we will take our chances in the Market". Green Trading then puts in a bid. Green trading analyzes past market results and put in a bid that will result, it believes, in 500 MW total load delivered at a price below the bilateral price it got from the Hydro group. So Green trading puts in a bid of 50 $/MWh for the 500 MW.

Bilateral Transaction Case: In this case the transaction is done completely outside the market and only the MW transaction is superimposed onto the transmission system. The price for the MW sold to Green Trading Group by the Hydro Generator is 60 $/MWh. The generator has to pay any transmission charges determined by the results of the auction.

The results are (Figure 11.11):

Market case: Here the Hydro Generator trading group places a bid to sell power. This is done by giving a cost function to the auction operator of

$$C(P_{\text{hydro}}) = 0 + 60P_{\text{hydro}} + 0.002P_{\text{hydro}}^2$$
$$\text{for} \quad 0 \le P_{\text{hydro}} \le 500 \text{ MW}$$

www.mepcafe.com

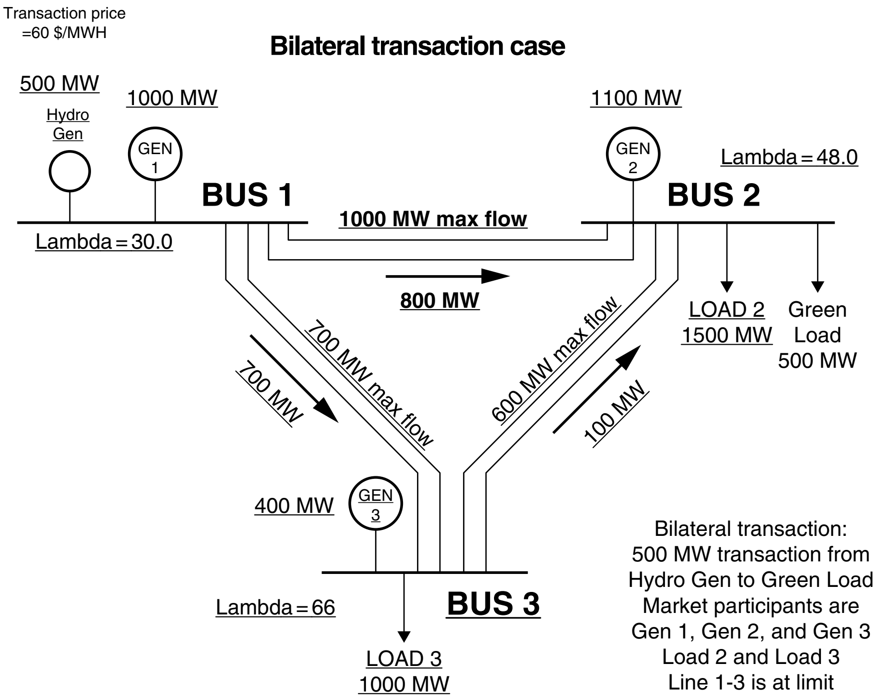


FIGURE 11.11 Results of the Bilateral Transaction Case for Problem 11.10.

The Green Load submits a bid to buy function of

$$C(P_{\text{green}}) = 0 - 50P_{\text{green}} \quad \text{for} \quad 0 \le P_{\text{green}} \le 500\,\text{MW}$$

The other generator cost functions and the cost functions of the loads are as follows:

Bid	A	B	C	Max (MW)
Gen 1	0.0	20	0.005	1100
Gen 2	0.0	5	0.0025	1100
Gen 3	0.0	50	0.02	1100
Load 2	0.0	−100	0	1500
Load 3	0.0	−100	0	1000

The market auction results with the Hydro Generator and the Green Load as part of the market are (Figure 11.12):

www.mepcafe.com

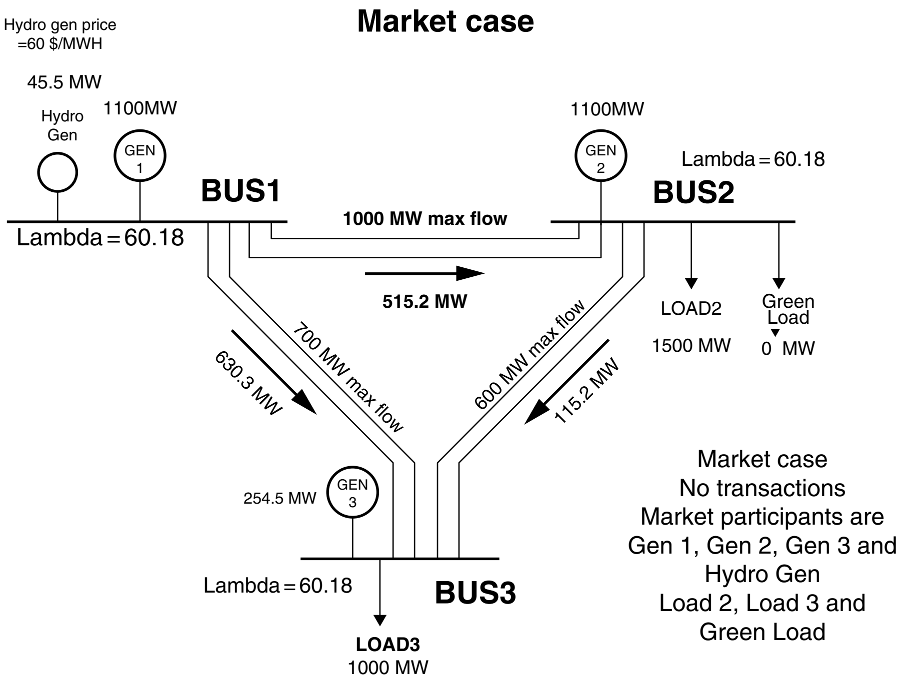


FIGURE 11.12 Results of the Market dispatch for Problem 11.10.

Note that the Hydro Generator only sells 45.5 MW and the Green Load receives nothing.

You are to find the following:

1. The student should plot the demand and supply curves for the market case. Does this help to see why the result is as shown?
2. For the Bilateral case: The dollar income received by the Hydro Generator, the transmission charges, and the net income of the Hydro Generator.
3. For the Bilateral case: The dollars paid by the Green Load.
4. For the Market case: The dollar income received by the Hydro Generator.
5. In this case, we know that the Green Load had the opportunity to make the transaction with the Hydro Generator and chose to go into the Market. Was the Green Load's decision good for the Hydro Generator?

www.mepcafe.com

12

SHORT-TERM DEMAND FORECASTING

12.1 PERSPECTIVE

Operational Planning is based on Inventory and Production Theory in all industries. Inventory Theory is based on storage policies. These policies are derived based on objectives of supply subject to transportation constraints, production, and demand. Inventory policies are dependent upon some forecast of demand, as well as present and future prices of resources. The forecast is an essential component of Production Management when finding optimal manufacturing policies. These policies are derived from production and demand models and are dependent upon forecast of demand (sales) or consumption of the services. Forecasting is an essential process of any successful company, especially when production oriented or trading oriented. Forecasting is often associated in operations research with inventory management problems. Other industries where forecasting plays a critical role include marketing, financial planning, and service industries. It is extremely rare to think of examples where managerial decisions are made in the absence of some form of forecasting. It is important to note that a forecast is not a final product; it is often recursive in nature as the need for refined solutions are required. The forecast is often revisited to determine the economic drivers for the product, the price anomalies (spikes), or the volatility of the forecast based on exogenous events.

Predictions of future events and conditions are called forecasts, and the act of making such predictions is called forecasting. Forecasting is a key element of decision making. Its purpose is to reduce the risk in decision making and reduce unexpected profit or cost.

Power Generation, Operation, and Control, Third Edition. Allen J. Wood, Bruce F. Wollenberg, and Gerald B. Sheblé.
© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

www.mepcafe.com

A major aim of an electric power utility is to accurately forecast demand requirements. In broad terms, power system demand forecasting can be categorized into long-term, medium-term, and short-term functions. Long-term demand forecasting usually covers from 1 to 10 years ahead in yearly values, and is explicitly intended for long-term capital investment studies.

The concept of medium forecasting is also used for maintenance and fuel scheduling for several years on a monthly basis.

The short-term forecast requires knowledge of the demand from 1 h up to a few weeks. Information derived from the short-term demand forecasts are vital to the system as operations in terms of short-term unit maintenance work, weekly, daily, and hourly demand scheduling of generating units, and economic and secure operation of power systems.

In this text, we will focus on the short-term demand forecasting with mathematical methods. First we introduce some basic foundations.

Production Theory for electric power systems is found by economic dispatch and unit commitment. Inventory Theory for electric power systems is based on fuel and hydro scheduling. Fuel scheduling is achieved by an optimal search of the various raw resources available under existing contracts or open markets. These policies were found in previous parts of this text. These policies are dependent on the forecast of the energy and ancillary services. Dependent upon some forecast of production, the optimal operating schedules are found. The optimal fuel schedules are found based on the production schedules. A competitive environment requires price forecasting as well as demand forecasting as the revenue received from sale has to exceed the costs of production.

Forecasting is an essential component of all operational functions, including transmission maintenance scheduling, unit outage scheduling for maintenance, nuclear fuel recycling, etc. However, forecasting need not be associated with scheduling and control. Other examples where forecasting is the key driver include marketing, financial planning, and labor policies.

Forecasts are obtained using qualitative or quantitative techniques. Qualitative techniques perform a forecast based on the result of an expression of one or more experts' personal judgment or opinion. Such methods are useful when historical data are not available or scarce. These methods include subjective curve fitting, Delphi method, and technological comparisons. This technique is often used for long-term forecasts. It is especially appropriate for events that are not periodic in nature. As an example, a major star is married on the tonight show, such as Tiny Tim. The experts in social responses would call in the demand experts to obtain their judgment on what to expect as the historical TV data did not contain such an event. This event was critical to the operating process. This forecast should have arrived at by consensus after prolonged discussion by the economic experts. This event almost caused a blackout since the impact of the event was not understood at that time. Many dispatchers understand the economic impacts sufficiently to out predict the quantitative techniques presented in this work.

The same process is used in leading consulting firms to determine the workload of each person for the expected projects over the task assignment horizon. Such techniques are not covered in this work.

www.mepcafe.com

Two distinct quantitative techniques are used in forecasting. They are both conventional statistical techniques. They are time series analysis and econometric analysis. Time series analysis is simply the reduction of a series of numerical values that a random variable takes on over a period of time. A random variable example is to consider the demand for energy each hour of every day. Another random variable example is to consider the clearing price for the exchange of electric energy for each hour of every day. Other common examples include the daily market closing prices of a particular fuel commodity over the period of a year. These are all time series. Time series analysis exploits techniques that utilize the patterns in the data for forecasting the values that the variable of interest will take on in future periods. It is noted that many time series are discrete in nature.

Most of the decision or control functions presented in this work require knowledge of future conditions, most especially demand. Table 12.1 provides a summary of existing and potential uses of short-term demand forecasting.

Automatic generation control (AGC) algorithms require knowledge of the inertia, governor, and frequency response for the next period of operation. The response capabilities of each unit have to be known to provide sufficient response capability as demand changes. The unit participation factors should be based on the economic dispatch to optimally follow demand changes. Some recent AGC packages do directly use demand forecast, as that knowledge of demand trends in the next few minutes can provide more optimal AGC control strategies. This is especially true when renewable energy, wind and solar, is included as such generation does not provide inertia, governor, or frequency response. AGC algorithms when augmented with knowledge of demand trends in the next few minutes permits the use of feed-forward or tracking controls which take into account rate of generation increase limits, valve-point loading, prohibited zone operation, loss of renewable generation, etc.

TABLE 12.1 Demand Forecast Applications

Function	Forecast Horizon	Forecast Intervals
Automatic generation control (AGC)	Next 15 min	5 sec
Economic dispatch (ED)	Next hour	30 sec
Power flow (PF)	Next 2 days	5 min
Optimal power flow (OPF)		
Contingency analysis (CA)	Next 2 days	10 min
Situational awareness (SA)	Next hour	120 samples per second
Voltage stability (VS)	Next hour	120 samples per second
Unit commitment (UC)	Next 14 days	Hourly
Transaction evaluation and management (TEM)	Next 14 days	Hourly
Wind forecasting	Next 5–60 min	30 sec
Hydro forecasting	Next 14 days	Hourly
Fuel scheduling	Next 1–6 months	Weekly

www.mepcafe.com

Economic dispatch presently does utilize forecasted demand to minimize crossing of prohibited zones. Valve point loading is generally not implemented as of this writing. Both algorithms are and have been ripe areas of forecasting research.

Operator or optimal power flow (PF/OPF) requires bus demand forecasts for power for the next 10–60 min for short-term planning, especially maintenance scheduling (MS) and transaction evaluation (TA). Bus demand forecasts are the future expected demands at each demand node in the power system to be studied. Operational planning can extend up to the next 1 or 6 months for maintenance scheduling of transmission equipment.

Security analysis (SA) requires bus demand forecasts the next hour, up to 24 h, to facilitate in predicting trouble periods. SA requires up to 6 months of bus load forecasts to include planned outages for maintenance scheduling.

Unit commitment traditionally required hourly system demands for the next 168–336 h. Some unit commitment programs even require knowledge of future demands for the next month, that is, 744 h (31 days). Unit commitment should include all of the ancillary services when searching for the optimal unit schedule.

Extensions to UC for Transaction Evaluation Management (TEM) require a 24–168 h forecast coupled to a good demand spike model will help in controlling demand peaks and valleys in spot pricing strategies.

12.2 ANALYTIC METHODS

They are both conventional statistical techniques: time series analysis and econometric analysis. Quantitative methods include regression analysis, decomposition methods, filtering such as exponential smoothing, and the Box-Jenkins methodology.

Time series analysis is simply the reduction of a series of numerical values that a random variable takes on over a period of time. Consider the demand for energy each hour of every day. Consider the clearing price for the exchange of electric energy for each hour of every day. One common example includes the daily market closing prices of a futures contract at the California Oregon Border (COB) over the period of a year. These are all time series. Time series analysis exploits techniques that identify the patterns in the data for forecasting the future values that the variable of interest will attain. Time series analysis is often broken into two components: deterministic and stochastic.

The time series analysis approach is to analyze the data using some type of analysis such as regression analysis (curve fitting optimization) problem or a time series (frequency fitting optimization) problem. The model form used as the basis for forecasting demand is based on system identification. System identification determines if the form is additive, multiplicative, or both. The first step is to remove demand trends, then slow frequency oscillations, and then faster frequency oscillations until all demand components outside the time horizon of interest are eliminated or modeled. Short-term demand forecasting requires that all components outside the next week have been explained and eliminated. They are added after the resulting time series has been identified and parameters estimated.

www.mepcafe.com

Econometric analysis identifies the economic drivers that lead to demand. The electric demand is often a function of the business index, the weather index (i.e., temperature), and other sociological indices. Econometric analysis requires that all demand is based on a cause-to-effect explanation. Econometric analysis is often based on least square curve fitting techniques classified as regression analysis.

Econometric analysis expresses the variable to be forecast (the dependent variable) as a mathematical function of another (independent) variables. Consider forecasting the total burger sales consumed each lunch period as functionally related to the unemployment rate during the same period in a given location. Data on demand in retail stores and over the Internet mail order sales as a function of total sales over previous periods may be used to forecast total sales in a future period given the Internet order sales for that period. The variable to be forecast (the dependent variable) is expressed as a mathematical function of other (independent) variables in regression analysis. For example, forecasting the total sales of burgers in a given period may be functionally related to the adjacent factory production during the same period.

Econometric techniques are usually additive. Each customer type is identified. Then the appliances or energy needs, for each customer is identified. Once all, or at least the major energy needs are modeled, the total demand is simply the sum of the various components for each customer.

Consider forecasting the total sales of room heaters purchased in a given period as functionally related to the increased demand for electricity for heat pumps during the same period. Data on heat pump demand in retail stores and over the Internet mail order sales as a function of total sales over previous periods may be used to forecast total sales in a future period given the Internet order sales for that period. Sale of heat pumps signals the conversion of homeowners from natural gas furnaces to electric demand. There is a legal case in Chicago where the gas company sued a homeowner for stealing gas after converting to a heat pump. The local gas company did not identify the conversion by inspection but based the complaint on decreased sales to that house.

Sales of microturbines signal the conversion of homes from all electric to combined heat and power (CHP) using natural gas or renewable (bio-fuel) gas. Indeed, the planned production of fuel cells to produce hydrogen, as an energy storage device, with or without carbon capture to make methane gas, provide a signal of homeowners to shift from electric energy on the grid to natural gas or to local renewable resources. Such shifts are encouraged by governments when tax incentives are given to reduce the amount of carbon emissions or other pollutants.

The economic analysis is a physical categorization to determine the demand equipment at each customer site and then aggregate the equipment demands based on number and type of customers. This model is dependent on the ability to identify which customers have which equipment, how many customers are of this type, and then aggregate the demand to the distribution feeder level for all distribution feeders. An example is to identify the appliances in a high income house, such as refrigerator, freezer, central air-conditioning, central heating by heat pump with energy storage, home theater, heated swimming pool, electric vehicle, etc. Then the number of high-income households is counted, normally through public records and past

www.mepcafe.com

electric invoices. This procedure is then repeated for each type of household, of commercial enterprise, or each type of industrial enterprise, etc. Note that each type of customer, including government enterprises (police, hospital, fire, etc.), has to be identified and inventoried.

Econometric analysis is expected to be far less costly once smart meters are implemented with feedback data on appliance usage. Such an approach based on sampling technology by Electric Power Research Institute (EPRI) and others has been combined with economic analysis and has been researched extensively in the past. This was a major project by the U.S. Department of Energy after the 1974 Oil Embargo. Such economic modeling is beyond the scope of this work.

The various types of forecasts are often used in conjunction. Indeed, the dispatcher "judgmental" technique is often used to modify an appropriate time series analysis when unusual events occur.

The art of forecasting is valuable across all industries and is worthy of a separate volume. There are several references that deal with forecasting for specific domains beyond the domain of this text.

The State Estimation Chapter provides the general background for computing the statistics of state and parameter estimates for certain cases where the weighted least squares (WLS) estimation approach is used. In the case of the state space mode, alternative techniques are available, such as Kalman filtering.

An alternative technique is to use an artificial life technique such as Artificial Neural Networks (ANNs). ANN is an approximation to the neurons of the brain. ANN is an alternative technique to the Box-Jenkins approach and is well-suited to replace curve fitting techniques. The EPRI has a package that is in use at several utilities. The ANN approach still requires the preprocessing required of time series analysis.

12.3 DEMAND MODELS

Power system demand is assumed to be time dependent, evolving according to a probabilistic law. It is common practice to employ a white noise sequence as input to a linear filter whose output is the power system demand. This is an adequate model for predicting the demand time series. The noise input is assumed normally distributed with zero mean and some variance σ^2. A number of classes of models exist for characterizing the linear filter.

The model form used as the basis for forecasting demand is based on system identification. System identification determines if the form is additive, multiplicative, or both.

Time series techniques may be additive or multiplicative. This work will use additive to simplify development. The first step is to remove base demand, then demand trends, then slow frequency oscillations, then faster frequency oscillations, until only an unexplained noise component remains. Note that all demand components outside the time horizon of interest are eliminated or modeled. Short-term demand forecasting requires that all components outside the short term have been explained from monitored data.

www.mepcafe.com

There are two steps to modeling demand: deterministic and stochastic analysis. The deterministic analysis is to find the slowest changing components, such as annual demand growth, seasonal variations, weather variations, etc. The deterministic is to analyze the data using some type of analysis such as a curve fitting optimization (regression) analysis or frequency identification (spectrum) analysis. Spectrum analysis is often accomplished using Fourier or Hartley transform analysis similar to the analysis found in systems and signals textbooks. The stochastic component is found using a time series analysis approach, as pioneered by Box and Jenkins.

12.4 COMMODITY PRICE FORECASTING

Price forecasting is one degree of complexity above demand forecasting. Price forecasting requires knowledge of the economic drivers such as fuel prices, unit outages (forced and planned), transmission congestion, fuel network delivery congestion (trains and barges), etc. Even regulatory policies are a driver for electric price forecasting when renewable energy resources are considered.

Price forecasting is first applied to fuel resources when procurement contracts are considered instead of future and spot markets. Price forecasting is applied to electric future and spot markets in a competitive environment.

Electric spot market pricing is critical for competitive market trading and production bidding. Both are forms of contracting as discussed in the business environments chapter.

Electric prices are subject to changes based on fuel price changes which are normally slowly occurring events over weeks or months. The oil embargo of 1974 is an example of an abrupt price change that was immediately captured by the markets. Electric prices are also a function of distribution network congestion. Floods curtail barge and train traffic resulting in shortages at some geographic locations. Such distribution congestion then leads to a different transmission flow pattern and potentially a congested transmission network due to unexpected flow patterns. The transmission network can also be congested due to the generation available due to market bidding or, most especially, to forced outages of generation or of transmission equipment. Indeed, several periods of high system stress has been caused by excessive scheduling of transmission equipment for maintenance at the same time. When such congestion events occur, the electric prices can spike to very high levels.

The nuclear unit transformer outages of 1994 led to the spike of electric prices in the Midwest, which led to the near bankruptcy of a major Midwest company. The importance of transformer maintenance was highlighted by this event. This event also highlighted the need for data mining to determine the forces that cause a certain pattern of bidding behavior. When a company offer bids for purchase and for sale, then commits to the purchases but not the sales, should alarm traders that there is an extreme shortage of resources within a geographic region.

Electric prices are also altered by the installation of gas turbines. Natural gas turbines are quickly installed where there are close connections of the natural gas and of the electric networks. This was a policy of ENRON to connect the natural gas

www.mepcafe.com

network and the electric transmission network whenever an arbitrage opportunity arose due to the near geographic locations of each network.

Several cities, such as Chicago, Illinois, and Sydney, New South Wales, have performed large conversions of buildings CHP facilities. Natural gas microturbines provide electricity and heat for heating and for cooling on a building by building basis. The city of Sydney embarked on a Trigeneration Master Plan as part of Sustainable Sydney 2030, a policy of not relying on coal-fired generation after 2030. Such programs severely alter the reliance on the transmission and distribution networks as well as the demand for other generation. Price spikes should be expected as such systems are installed as market penetration increase. Included as distributed generation, such systems highlight the impact of wind and solar generation as local resources when integrated with demand centers.

12.5 FORECASTING ERRORS

Unfortunately, all forecasting situations involve some degree of uncertainty which makes the errors unavoidable.

The forecast error for a particular forecast $\hat{X}_t$ with respect to actual value X_t is

$$e_t = X_t - \hat{X}_t \tag{12.1}$$

To avoid the offset of positive with negative errors, we need to use the absolute deviations.

$$|e_t| = |X_t - \hat{X}_t| \tag{12.2}$$

Hence, we can define a measure known as the mean absolute deviation (MAD) as follows:

$$\text{MAD} = \frac{\sum_{t=1}^{n} |e_t|}{n} = \frac{\sum_{t=1}^{n} |X_t - \hat{X}_t|}{n} \tag{12.3}$$

Another method is to use the mean-squared error (MSE) defined as follows:

$$\text{MSE} = \frac{\sum_{t=1}^{n} e_t^2}{n} = \frac{\sum_{t=1}^{n} (X_t - \hat{X}_t)^2}{n} \tag{12.4}$$

12.6 SYSTEM IDENTIFICATION

System identification determines the formula form for all components to minimize the forecast error. The major dimension to be identified is the number of components. There are two basic model forms: additive and multiplicative.

www.mepcafe.com

One normally removes the longer-term trends in demand growth (or decline) which occur over weeks, months, and years into the future. In general, short-term demand behavior is influenced by typical factors such as

a. Hour of the day
b. Day of the week
c. Weather conditions
d. Strikes or other political events
e. Distributed generation
f. Demand response strategies
g. Pricing strategies
h. Other economic conditions

When modeling the demand, one relates power consumption to each one of these factors to measurable variables associated with that factor. Past historical analysis of the data will help in estimating the parameters of the corresponding model. These models, together with the most recent information, are used for the forecasting process. Note that part of this analysis is econometric while the rest is time series analysis.

12.7 ECONOMETRIC MODELS

The application of econometric models to the demand forecasting problems is beyond the scope of this work. However, Environmental Control appliances have a major impact on the forecast. The modeling of such devices as a group is often included in the time series analysis. Thus, these models are an introduction to the impact of these appliances and least squares regression.

12.7.1 Linear Environmental Model

The major electrical demand is often air-conditioning or heating. A typical weather-sensitive study determines the dependence of demand variations on weather changes, especially average temperature. More detailed studies may isolate sensitivities to humidity, solar radiation, wind direction, and wind speed. Figure 12.1 illustrates a simple first-order temperature sensitive model. In that figure one is interested in the slopes of the summer and winter lines. One example model for the summer is

$$P_S(k) = A_S(\Delta\theta(k) - T_s) \tag{12.5}$$

$$P_W(k) = (T_W - A_W \Delta\theta(k)) \tag{12.6}$$

where T_s is a lower temperature limit for air-conditioning use, and $\theta(k)$ is the average temperature at time k. If one makes the measurements at periods when the base

www.mepcafe.com

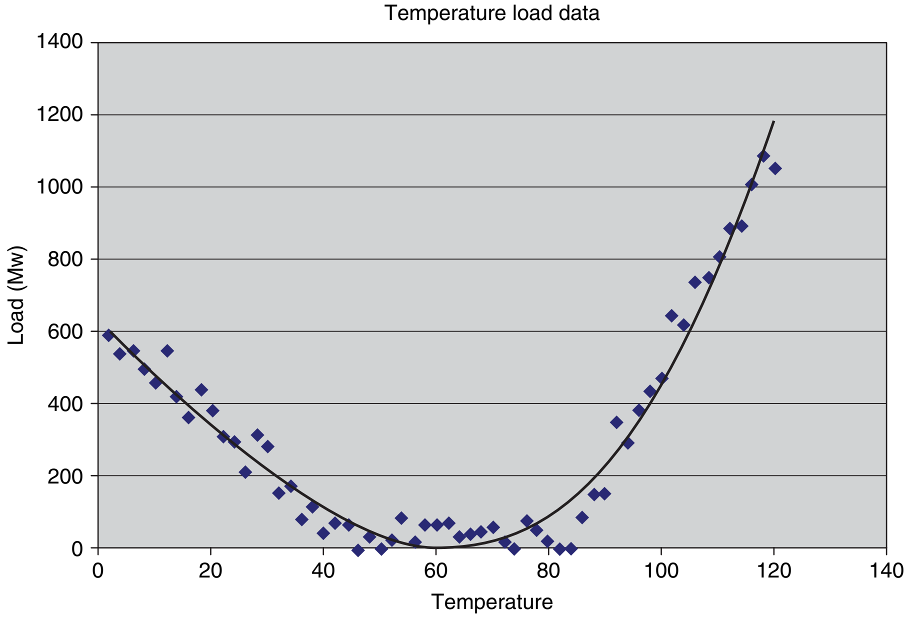


FIGURE 12.1 Example first-order weather model dependency for seasonal demands.

demand is approximately constant and where there are no special events, (e.g., strikes or games), then one can express the demand using the aforementioned relation as

$$P_D(k) = P_B + P_W(k) + \eta(k) \tag{12.7}$$

where the subscript S is for summer demands and W is for winter demands. T_W is the turn on temperature for heating use. Note that k refers here to a measurement sample rather than a specific time.

Note that these models assume that the periods for heating and for cooling are separated. The aforementioned straight line approximations are often used. Polynomial, quadratic, or learning curve fitting represents the bath tub curve function more directly. Weighted least squares curve fitting or least absolute value curve fitting are common approaches to find these models.

Example 12A: Demands $P_D(k)$ and temperatures $\Delta\Theta(k)$ are provided in a weather sensitivity demand study in Table 12.2. Assuming that $T = 40°F$ and that the noise is zero mean uncorrelated errors with constant variances; determine the best estimates of P_B and A, as per the model expressed in Equations 12.6 and 12.7. The following determines the change to the forecast due to the weather-dependent demand:

Solution follows the method of state estimation. The first step is to define

$$\mathbf{P}_D = \begin{bmatrix} P_{D1} \\ \vdots \\ P_{D6} \end{bmatrix} \tag{12.8}$$

www.mepcafe.com

TABLE 12.2 Data for Example

Hour	P_D (MW)	T (°F)
1	1059.4	16
2	1058.2	18
3	1057.0	20
4	1055.8	22
5	1055.8	22
6	1054.0	25

$$\mathbf{x} = \begin{bmatrix} P_B \\ A_W \end{bmatrix} \tag{12.9}$$

$$\mathbf{v} = \begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(6) \end{bmatrix} \tag{12.10}$$

$$\mathbf{H} = \begin{bmatrix} 1 & \Delta\theta(1) \\ 1 & \Delta\theta(2) \\ \vdots & \vdots \\ 1 & \Delta\theta(6) \end{bmatrix} \tag{12.11}$$

Then the equation can be expressed in vector form as

$$\boldsymbol{P}_D = \mathbf{Hx} + \mathbf{v} \tag{12.12}$$

Using the techniques of WLS estimation, the best estimate of x is given by

$$\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\boldsymbol{P}_D \tag{12.13}$$

since all variances of $v(k)$ are the same. As a result, we conclude

$$P_B = 1005.38\ \Delta MW, \quad A_W = -0.58\ \Delta MW/\Delta°F, \quad R^2 = 0.9945.$$

12.7.2 Weather-Sensitive Models

Another approach is to us a period differenced polynomial for summer demands and for winter demands. Note that k refers here to a measurement sample rather than a specific time. These more elaborate weather-sensitive models enable additional environmental modeling. Two important factors may be accounted for: time lag and

www.mepcafe.com

TABLE 12.3 Data for Example

k	$\Delta P = W(k)$ (MW)	ΔW (°F)
1	206	24
2	208	30
3	205	35
4	203	38
5	197	43
6	191	50
7	188	54
8	186	58
9	184	61

saturation effects. In a time lag model the weather-sensitive part of the demand may be expressed as

$$W_k = A_0\Delta\theta_k + A_1\Delta\theta_k + \cdots + A_m\Delta\theta_k \tag{12.14}$$

where m is an integer to be determined from actual data. The time lag is due to the fact that buildings have thermal mass, resulting in a lagged response of air-conditioning (or heating) equipment to temperature variations. Time lag may also appear when wind speeds are higher, thus increasing heating or cooling demands.

Saturation effects develop at extreme weather conditions like a heat wave when almost all of installed air-conditioning equipment will be operating continuously. A simple saturation model can be expressed as where

$$W_k = \min\left\{W_k^o, \overline{W}\right\} \tag{12.15}$$

$$W_k^o = A_o\Delta\theta_k + A_1\Delta\theta_k + \cdots + A_m\Delta\theta_k \tag{12.16}$$

and $\overline{W}$ is the saturation limit. In this case, one has to identify $\overline{W}$ in addition to the other parameters.

Example 12B: In a weather sensitivity modeling study, data for consecutive hours were collected in which the weather-sensitive component was extracted to be as shown in Table 12.3. Assuming the presence of a first-order lag ($m = 1$) and saturation effects, compute the best estimates of the coefficients.

$$W_k = \min\left\{w_k^o, w\right\} \tag{12.17}$$

A cursory look at the data indicates that data points for $k = 1,\ldots,4$, correspond to saturation since temperature fluctuations near the peak are not changing the demand very much. As a result, W is computed as the average of those readings: $W = 205.69$.

www.mepcafe.com

The remaining five data points are now used for estimating A_0 and A_1. The governing vector equation is given by the best estimates of A_0 and A_1 obtained by the WLS approach. Since those numbers are the same as the previous example, the results are the same.

Other weather effects can be included directly or indirectly. The direct approach will require extra terms associated with the weather variable. In the indirect approach, the temperature variable is modified to account for those factors. As a result, $\Delta\Theta(k)$ becomes an equivalent temperature to be used in the weather model.

12.8 TIME SERIES

Time series can be defined as a sequential set of data measured over time, such as the hourly, daily, or weekly peak demand. The basic idea of forecasting is to first build a pattern matching available data as accurately as possible, then find the forecasted value with respect to time using the established model.

Generally, series are often described as having the following components:

$$X_t = T_t + S_t + R_t \quad t = \cdots -1, 0, 1, 2, \ldots \tag{12.18}$$

Here, T_t is the trend term, S_t the seasonal term, and R_t is the irregular or random component. At this moment, we do not consider the cyclic terms since these fluctuations can have a duration from 2 to 10 years or even longer, which is not applicable to short-term demand forecasting.

It is often assumed, to make things a little easier for the moment, that the trend is a constant level.

If the trends and mean have been removed, then the seasonal effect has period s; that is, it repeats after s time periods. Or the sum of the seasonal components over a complete cycle or period is 0.

$$\sum_{j=1}^{s} S(t+j) = 0 \tag{12.19}$$

12.8.1 Time Series Models Seasonal Component

A simple example is to consider the data on fuel consumption given in the following table.

Example 12C: We can average the seasonal values over the series, and use these, minus the overall mean, as seasonal estimates shown as follows (overall mean is 761.65):

$$S(1) = 888.2 - 761.65 = 126.55 \qquad S(2) = 709.2 - 761.65 = -52.4$$
$$S(3) = 616.4 - 761.65 = -145.25 \qquad S(4) = 832.8 - 761.65 = 71.15$$

After subtraction of these values, the original series removes seasonal effects. It should be noted that this technique works well on series having linear trends with small slopes.

www.mepcafe.com

In addition, we can look at the averages for each complete seasonal cycle (the period) since the seasonal effect over an entire period is 0. To avoid losing too much data, a method called Moving Average (MA) is used here, which is simply the series of averages:

$$\frac{1}{s}\sum_{j=0}^{s-1}X_{t+j},\quad \frac{1}{s}\sum_{j=1}^{s}X_{t+j},\quad \frac{1}{s}\sum_{j=2}^{s+1}X_{t+j},\ldots \tag{12.20}$$

A problem is presented here if the period is even, since the adjusted series values do not correspond to the original ones at time points. To overcome this problem, the Centered Moving Average (CMA) is utilized to bring us back to the correct time points. It is shown in Tables 12.4 and 12.5.

TABLE 12.4 Primary Energy Consumption (Fuel Equivalent)

Year/Season	1	2	3	4
1965	874	679	616	816
1966	866	700	603	814
1967	843	719	594	819
1968	906	703	634	844
1969	952	745	635	871
Means	888.2	709.2	616.4	832.8

TABLE 12.5 Calculation of the Moving Average

Quarter	X_t	MA	CMA(Order-2)	Difference
1	874			
2	679			
3	616	746.25	745.25	−129.25
4	816	744.25	746.875	69.125
5	866	749.5	747.875	118.125
6	700	746.25	746	−46
7	603	745.75	741.75	−138.75
8	814	737.75	740.125	73.875
9	834	742.5	741.375	92.625
10	719	740.25	740.875	−21.875
11	594	741.5	750.5	−156.5
12	819	759.5	757.5	61.5
13	906	755.5	760.5	145.5
14	703	765.5	768.625	−65.625
15	634	771.75	777.5	−143.5
16	844	783.25	788.5	55.5
17	952	793.75	793.875	158.125
18	745	794	797.375	−52.375
19	635	800.75		
20	871			

www.mepcafe.com

TABLE 12.6 Prediction of Energy Consumption

Period	Trend	Seasonal	Predicted	Actual
t	T	S	P	X
21	789.963	127.340	917.303	981
22	793.610	−47.741	747.123	759
23	797.257	−143.254	654.003	674
24	800.904	63.746	864.650	900

Through looking at the differences between CMA and the original series, we can estimate the kth seasonal effect simply by average the kth quarter differences:

$$S(1) = 128.954 \qquad S(2) = -46.487 \qquad S(3) = -142 \qquad S(4) = 65$$

But the sum of these four values is 5.107. Recall that we assume the seasonal terms to sum to 0, we need to add a correction factor of $-5.107/4 = -1.254$ to give

$$S(1) = 127.34 \qquad S(2) = -47.741 \qquad S(3) = -143.254 \qquad S(4) = 63.746$$

Now the irregular component can be easily calculated by subtracting both the CMA and the seasonal effects.

If we suppose the model is appropriate, then we can use it to make predictions. To simplify, we omit the random data, and so all we need to do is to predict the trend, say a linear trend:

$$T(t) = a + bt \tag{12.21}$$

With the application to the CMA, we have

$$\hat{T}(t) = 713.376 + 3.647t \tag{12.22}$$

Hence, a prediction is shown in Table 12.6.

12.8.2 Auto-Regressive (AR)

In this model, the current value X_t of the time series is expressed linearly in terms of its previous values $X_{t-1}, X_{t-2}, \ldots$ and a white noise series $\{\varepsilon_t\}$ with zero mean and variance σ^2.

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_p X_{t-p} + \varepsilon_t \tag{12.23}$$

By introducing the backshift operator B that defines $X_{t-1} = BX_t$, and consequently $X_{t-m} = B^m X_t$, Equation 12.15 can be written in the form:

$$\varphi(B) X_t = \varepsilon_t \tag{12.24}$$

www.mepcafe.com

where

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p \tag{12.25}$$

Note that this model has a similar form to the multiple linear regression model. The difference is that in regression the variable of interest is regressed onto a linear function of other known (explanatory) variables, whereas here X_t is expressed as a linear function of its own past values—thus the description "auto-regressive." As the values of X_t at p previous times are involved in the model, it is said to be an AR (p) model.

Now we need to calculate the parameters φ_i for prediction. There are two such methods: least squares estimation and maximum likelihood estimation (MLE).

To calculate the least squares estimators, we need to minimize the following expression (here we let $p = 2$):

$$\sum_{t=1}^{N} (X_t - \varphi_1 X_{t-1} - \varphi_2 X_{t-2})^2 \tag{12.26}$$

with respect to φ_1 and φ_2. But since we do not have the information for $t = 1$ or $t = 2$, an assumption is made here that X_1 and X_2 are fixed, and excluding the first two terms from the sum of squares. That is, to minimize

$$\sum_{t=3}^{N} (X_t - \varphi_1 X_{t-1} - \varphi_2 X_{t-2})^2 \tag{12.27}$$

Then, a similar approach to linear regression is used to obtain the parameters.

12.8.3 Moving Average (MA)

In the MA process, the current value of the time series X_t is expressed linearly in terms of current and previous values of a white noise series (ε_t). This noise series is constructed from the forecast errors or residuals when demand observations become available. The order of this process depends on the oldest noise value at which X_t is regressed on. For an MA of order q (i.e., MA(q)), this model can be written as

$$X_t = \eta_t - \theta_1 \eta_{t-1} - \theta_2 \eta_{t-2} - \cdots - \theta_q \eta_{t-q} \tag{12.28}$$

A similar application of the backshift operator on the white noise series would allow the equation to be written as

$$X_t = \theta(B)\eta_t \tag{12.29}$$

where

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \tag{12.30}$$

www.mepcafe.com

Moving Average models can be converted to AR models as can AR models be converted to MA models. The end result is an infinite series that is harder to manage.

Example 12D: Consider the MA model:

$$X(k) = \eta(k) - 0.5\eta(k-1) \tag{12.31}$$

How does one obtain the equivalent infinite AR model?
A useful mathematical tool for solving this problem is the so-called backward shift operator B (Box–Jenkins), which is defined as

$$BX(k) = X(k-1) \tag{12.32}$$

Based on that one defines B'' to be the second difference:

$$X(k) = \eta(k) - 0.5\eta(k-1) = \eta(k) - 0.5B\eta(k) = (1-0.5B)\eta(k) \tag{12.33}$$

Consequently,

$$(1-0.5B)^{-1} X(k) = \eta(k) \tag{12.34}$$

where

$$(1-0.5B)^{-1} = 1 + 0.5B + (0.5B)^2 + (0.5B)^3 + \cdots = \sum_{i=0}^{\infty}(0.5B)^i \tag{12.35}$$

This is an infinite AR model:

$$X(k) = \eta(k) - 0.5\eta(k-1) - 0.25\eta(k-2) - 0.5\eta(k-n) - \cdots \tag{12.36}$$

12.8.4 Auto-Regressive Moving Average (ARMA): Box-Jenkins

If we combine the MA and AR models together, we can present a broader class of model, that is, ARMA model as follows:

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_p X_{t-p} + \eta_t + \theta_1 \eta_{t-1} + \theta_2 \eta_{t-2} + \cdots + \theta_q \eta_{t-q} \tag{12.37}$$

where φ_i and θ_j are called the ARMA parameters, respectively. It is an ARMA(p,q) model.

A methodology for ARMA models was developed largely by Box and Jenkins (1976), and so the models are often called Box-Jenkins models.

Recall that an MA process is equivalent to an infinite AR process. Similarly, one can show that an ARMA process can be expressed as an infinite AR process or an infinite MA process.

www.mepcafe.com

Example 12E: Consider converting the following ARMA(1,1) process into an infinite AR process:

$$X(k) = -0.7X(k-1) + \eta(k) - 0.5\eta(k-1) \tag{12.38}$$

By rearranging terms and using the backward shift operator B, the earlier process can be written as

$$(1+0.7B)X(k) = (1-0.5B)\eta(k) \tag{12.39}$$

Multiplying both sides by the inverse of the right hand side shift:

$$(1-0.5B)^{-1}(1+0.7B)X(k) = \eta(k) \tag{12.40}$$

Expanding the denominator:

$$[1+0.5B+(0.5B)^2+\cdots](1+0.7B)X(k) = \eta(k) \tag{12.41}$$

Expanding the expression:

$$(1+.58+(.58)^2+\cdots)(1+.7B)X(k) = \eta(k) \tag{12.42}$$

Expanding this expression, one obtains

$$X(k)+1.2X(k-1)+0.6X(k-2)+0.518X(k-3)+\cdots = \eta(k) \tag{12.43}$$

A well behaved AR equivalent of an ARMA process requires the coefficients to converge in a geometric fashion to zero. Hence, one constructs a finite approximation of order l:

$$X(k) \approx \sum_{i=1}^{p} c_i X(k-i) + \eta(k) \tag{12.44}$$

where the c_i's are the coefficients of the equivalent AR model. Given a sufficiently large "p," WLS estimation can obtain the best estimates of the coefficients. These estimates can be used to evaluate the following estimates of $\hat{\eta}(k)$:

$$\hat{\eta}(k) = X(k) - \sum_{i=1}^{p} c_i X(k-i) \tag{12.45}$$

Next, the estimates $\hat{\eta}(k)$ are used in the original ARMA(n, m) model:

$$X(k) \approx \sum_{i=1}^{n} a_i X(k-i) + \eta(k) + \sum_{j=1}^{m} b_j \hat{\eta}(k-j) \tag{12.46}$$

www.mepcafe.com

Since the $\hat{\eta}(k-j)$ terms in the aforementioned expression are given inputs, WLS estimation techniques can be used to evaluate the best estimates of a_i and b_j for all n and m.

12.8.5 Auto-Regressive Integrated Moving-Average (ARIMA): Box-Jenkins

The time series defined previously as an AR, MA, or as an ARMA process is called a stationary process. This means that the mean of the series of any of these processes and the covariances among its observations do not change with time. Unfortunately, this is not often true in power demand. But previous knowledge is definitely useful in that the nonstationary series can be transformed into a stationary one with some tricks. This can be achieved, for the time series that are nonstationary in the mean, by a differencing process. By introducing the ∇ operator, a differenced time series of order 1 can be written as

$$\nabla X_t = X_t - X_{t-1} = (1-B)X_t \tag{12.47}$$

Consequently, an order d differenced time series is written as

$$\nabla^d X_t = (1-B)^d X_t \tag{12.48}$$

The differenced stationary series can be modeled as an AR, MA, or ARMA to yield an ARI, IMA, or ARIMA time series processes. For a series that needs to be differenced d times and has orders p and q for the AR and MA components (i.e., ARIMA(p,d,q)), the model is written as

$$\varphi(B)\nabla^d X_t = \theta(B)\eta_t \tag{12.49}$$

However, as a result of daily, weekly, yearly, or other periodicities, many time series exhibit periodic behaviors in response to one or more of these periodicities. Therefore, a seasonal ARIMA model is appropriate. It has been shown that the general multiplicative model $(p,d,q) \times (P,D,Q)_s$ for a time series model can be written in the form:

$$\varphi(B)\Phi(B^S)\nabla^d \nabla_S^D X_t = \theta(B)\Theta\left(B^S\right)\eta_t \tag{12.50}$$

where definitions for $\Phi(B^S), \nabla_S^D, \Theta(B^S)$ are given in the following:

$$\nabla_S^D = (X_t - X_{t-S})^D = (1-B^S)^D X_t \tag{12.51}$$

$$\Phi(B^S) = 1 - \Phi_1 B^S - \Phi_2 B^{2S} - \cdots - \Phi_p B^{pS} \tag{12.52}$$

$$\Theta(B^S) = 1 - \Theta_1 B^S - \Theta_2 B^{2S} - \cdots - \Theta_q B^{qS} \tag{12.53}$$

www.mepcafe.com

The model presented earlier can obviously be extended to the case where two seasonalities are accounted. An example demonstrating the seasonal time series modeling is the model for an hourly demand data with daily cycle as is very common for short-term demand forecasting. Such a model can be expressed using the model with $S = 24$.

12.8.6 Others (ARMAX, ARIMAX, SARMAX, NARMA)

Auto-Regressive Moving Average and ARIMA use only the time and demand as input parameters. Since demand generally depends on the weather and the time of day, exogenous variables (X) sometimes can be included to give an ARMAX and ARIMAX model. There are other versions of ARIMA, such as Piecewise ARMA (PARMA), Seasonal ARMA (SARMAX), etc. The concept is to identify the time series model as a single expression instead of preprocessing and post-processing.

12.9 TIME SERIES MODEL DEVELOPMENT

Total real demand (for the system, or a specific demand bus), may be expressed as either an addition or multiplicative model. An addition model is often used of the following form with deterministic components and a random component:

$$P_D(t) = B(t) + W(t) + S(t) + \eta(t) \tag{12.54}$$

where

$$\begin{aligned} P_D(t) &= \text{total demand at time } t \\ W(t) &= \text{weather} - \text{sensitive demand component at time } t \\ B(t) &= \text{base (normal) demand at time } t \\ S(t) &= \text{demand incrememnt due to special events} \\ \eta(t) &= \text{random demand component.} \end{aligned}$$

The time t is measured periodically, for example, every hour, minute, 10 min, and so on, depending on the type of application considered. Note that the P with a subscript has been dropped for simplicity.

Defining Δt to be the specified time increment, then the index k is used to specify the kth sample of information as follows:

$$t_k = t_0 + k\Delta t, \quad k = 1, 2, \ldots, \tag{12.55}$$

where t_0 is the specified initial time. For all subsequent discussions t will be replaced by t_k or simply k, that is,

$$P_D(k) = B(k) + W(k) + S(k) + \eta(k) \tag{12.56}$$

www.mepcafe.com

This decomposition of the demand into separate components is dependent on an overall modeling process based on data collected to find the coefficients for the model. The critical modeling step is referred to as model identification. Forecasting is performed during the process of demand prediction using the identified model and the estimated parameters.

12.9.1 Base Demand Models

Given the historical record for the past month or so, one can subtract the weather-sensitive component from the total demand, smooth out demand excursions due to special events, then the remaining component is the base demand.

In the absence of a super demand model that can account for everything, the most convenient way to deal with the base demand is to treat it as a stochastic time series. By so doing, one can make use of well-established techniques.

As an illustrative step, let $B(k)$ be the average over the set of sample points provided (e.g., k may correspond to 3–4 p.m. every Tuesday for the past 4 weeks). Defining $y(k)$ to be

$$y(k) = B(k) - \overline{B}(k) \tag{12.57}$$

one obtains a time series for $y(k)$ that can now be modeled. The immediate discussion will focus on three closely related time series models: the ARMA, and ARMA models.

12.9.2 Trend Models

Trends are modeled are linear, quadratic, exponential, or polynomial functions of very low frequency. The first process is to determine the spectrum of the time series using Fourier or Hartley transform to determine the predominant frequencies present. These long-term components are primarily of interest to power system planning and are not included within this work. While the practice is to remove them, such long-term components are often lumped into the base demand component. This work assumes that these trends are removed and added later during the forecasting process. Recall that Fourier or Hartley series analysis is essentially a curve fitting process as demonstrated in the following section.

12.9.3 Linear Regression Method

This method supposes that the demand is affected by some factors such as high and low temperatures, weather condition, and economic growth, etc. This relation can be expressed as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \eta \tag{12.58}$$

where y is the demand, x_i are the affecting factors, β_i are the regression parameters with respect to x_i, and η is an error term.

www.mepcafe.com

For this model, we always assume that the error term η has a mean value equal to 0 and constant variance.

Since parameters β_i are unknown, they should be estimated from observations of y and x_i. Let b_i ($i = 0,1,2,\ldots, k$) be the estimates in terms of β_i ($i = 0,1,2,\ldots,k$). Recall that the error term has a 50% chance of being positive and negative respectively, we omit this term in calculating parameters, which means:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k \tag{12.59}$$

Then, we use the least square estimates method which minimizes the sum of squared residuals or sum of squared errors (SSE) to obtain the parameters b_i:

$$\underline{B} = [b_0 \quad b_1 \quad b_2 \quad \cdots \quad b_k]^{\mathrm{T}} = (\underline{X}^{\mathrm{T}}\underline{X})^{-1}\underline{X}^{\mathrm{T}}\underline{Y} \tag{12.60}$$

where Y and X are the following column vector and matrix:

$$\underline{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \text{and} \quad \underline{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \tag{12.61}$$

After the parameters are calculated, this model can be used for prediction. It will be accurate in predicting y values if the standard error s is small.

$$s = \sqrt{\frac{\text{SSE}}{n-(k+1)}}, \quad \text{SSE} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \quad y_i : \text{observed}, \quad \hat{y}_i : \text{estimated} \tag{12.62}$$

There are also other ways to check the validity of a regression model, such as the R_h ratio.

Example 12F: The best way to understand a new process is to examine each algorithm individually. This example shows the suggested procedure for this chapter. When implementing an algorithm, it is best to give the algorithm input for which the output is known. Consider curve fitting a Fourier series with the regression algorithm. The process is to generate a time series from a known Fourier equation without noise being added. If the procedure cannot find the correct output, there is an error.
Consider curve fitting the series:

$$x = \sin(2 \times pi \times a \times t) + \sin(2 \times pi \times b \times t) + \eta; \quad 0 \le t \le 10, \quad \Delta t = \frac{1}{100} \tag{12.63}$$

where

$$a = 10, \quad b = 25, \quad \eta = 0$$

Then the solution found is the original numbers. As noise is added, the values for the coefficients will drift depending on the amount of noise added.

www.mepcafe.com

Recall that the Fourier series:

$$y = a_0 + \sum_{i=1}^{n} (a_i \cos(nwx) + b_i \sin(nwx)) \tag{12.64}$$

For length N input sequence x, the discrete Fourier transform (DFT) is a length N vector, the fast Fourier transform implement the relationships

$$\begin{aligned} X(k) &= \sum_{n=1}^{N} \left(x(n) e^{-j2\prod(k-1)\left(\frac{n-1}{N}\right)} \right), \quad 1 \le k \le N \\ x(k) &= \sum_{k=1}^{N} \left(X(k) e^{-j2\prod(k-1)\left(\frac{n-1}{N}\right)} \right), \quad 1 \le n \le N \end{aligned} \tag{12.65}$$

Can be written as the discrete Fourier series:

$$x(n) = \frac{1}{N} \sum_{k=1}^{N} \left[a(k) \cos\left(\frac{2\prod(k-1)(n-1)}{N} \right) + b(k) \sin\left(\frac{2\prod(k-1)(n-1)}{N} \right) \right] \tag{12.66}$$

$$\begin{aligned} a(k) &= \text{real}((X(k)) \\ b(k) &= -\text{imaginary}(X(k)), \quad 1 \le n \le N \end{aligned} \tag{12.67}$$

The application of the Fast Fourier Transform or the Fast Hartley Transform is beyond this work. Such transform are often used as part of a time series analysis package.

12.9.4 Seasonal Models

The demand varies from season to season primarily as a function of weather but also of changes in social behavior. Summer time is more for vacations. Winter is more for manufacturing. Seasonal models are also used to remove any periodic information, such as daily or weekly repetition. Seasonal adjustments are used in a very genetic sense in these cases.

12.9.5 Stationarity

The use of the AR, MA, and ARMA models is predicated upon the assumption that the driving noise process is white. There are standard techniques for checking if a process is stationary. For example, $y(k)$ should have a constant or zero mean regardless of the number of samples chosen, its autocovariance should be a function of time differences, etc. Should the process prove to be nonstationary, then some form of data preprocessing is needed. Box and Jenkins have shown that factors associated with periodic behavior, (time of day, or day of the week) may create nonstationary behavior. As a result, they make the suggestion to construct a new time series $x(k)$ from $y(k)$ such that $x(k)$ is stationary.

www.mepcafe.com

Examples of such constructions are

$$
\begin{aligned}
x_a(k) &= y(k) - y(k-1) = (1-B)y(k) \\
x_b(k) &= y(k) - y(k-24) = (1-B^{24})y(k) \\
x_c(k) &= y(k) - y(k-168) = (1-B^{168})y(k)
\end{aligned}
\tag{12.68}
$$

$x_a(k)$ series is obtained by taking the differences of $y(k)$ from 1 h to the next. This will eliminate persistent trend effects that destroy stationarity.

$x_b(k)$ series accounts for time-of-day effects, and $x_c(k)$ series for the hour of the week. A composite of these is also possible, for example,

$$x_d(k) = (1-B)(1-B^{24})y(k) \tag{12.69}$$

whereby both hourly trends and hour-of-the-day effects are eliminated.

Obviously, a certain amount of experimentation is needed to identify the proper composite series. But once one obtains a resulting stationary series, then the earlier techniques of model identification will be very powerful. The use of an ARMA $x(k)$ series and its integration later on to a $y(k)$ series is called an integrated ARIMA process.

If w_k behaves like white Gaussian noise with zero mean and constant variance, then the aforementioned process is stationary. A simple stationarity test can be employed against the time series $y(k)$, $k = 1,\ldots, N$, by showing that the shifted mean

$$\bar{y}_i = \frac{1}{T}\sum_{k=i+1}^{i+T} y(k) \tag{12.70}$$

and the shifted variance

$$\sigma_i = \frac{1}{T}\sum_{K=i+1}^{i+T} (y(k) - \bar{y}(k))^2 \tag{12.71}$$

for $i = 0, 1, 2,\ldots$ are independent of i.

Another test is to show that the autocovariance

$$\gamma_i = \operatorname{cov}\left[y(k), y(k+i)\right] = E\left[(y(k)-\bar{y})(y(k+i)-\bar{y})\right] \tag{12.72}$$

is independent of k and dependent only on the value of i. If the time series $y(k)$, $k = 1,\ldots, N$, proves to be stationary and is of the AR type, then there are simple means of obtaining the best estimates of these parameters, such as the Durban Levinson presented later. However, if the series is nonstationary, then differencing to remove seasonality, or removing the mean and trend are to be employed to convert it to a stationary one.

www.mepcafe.com

12.9.6 WLS Estimation Process

Assuming that the order n of an AR model is known, WLS estimation can be used to estimate the coefficients in the same fashion as the state estimation problem is solved.

Express the AR model in vector form:

$$\begin{bmatrix} y(n+1) \\ y(n+2) \\ \vdots \\ y(N) \end{bmatrix} = \begin{bmatrix} y(n) & y(n-1) & \cdots & y(1) \\ y(n+1) & y(n) & \cdots & y(2) \\ \vdots & \vdots & \ddots & \vdots \\ y(N-1) & y(N-2) & \cdots & y(N-n) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} w(n+1) \\ w(n+2) \\ \vdots \\ w(N) \end{bmatrix} \tag{12.73}$$

This can be re-expressed in matrix form:

$$\mathbf{y} = \mathbf{Hn} + \mathbf{w} \tag{12.74}$$

The best estimate of a is given by the same solution as used by state estimation:

$$\hat{\mathbf{a}} = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{y} \tag{12.75}$$

Example 12G: The AR process

$$y(k) = a_1 y(k-1) + a_2 y(k-2) + a_3 y(k-3) + w(k), \tag{12.76}$$

corresponds to a third-order difference equation. Hence, the following state variables can be defined:

$$\begin{aligned} x_1(k) &= y(k-2) \\ x_2(k) &= y(k-1) \\ x_3(k) &= y(k). \end{aligned} \tag{12.77}$$

Based on these definitions:

$$\begin{aligned} x_1(k+1) &= x_2(k) \\ x_2(k+1) &= x_3(k) \\ x_3(k+1) &= a_1 x_3(k) + a_2 x_2(k) + a_3 x_1(k) + w(k) \end{aligned} \tag{12.78}$$

or more compactly

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a_3 & a_2 & a_1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} w(k) \tag{12.79}$$

www.mepcafe.com

$$y(k) = x_3(k)$$

As is well-known from linear systems theory one can construct an *n*th order state space model of an ARMA(*p*, *q*) provided that $p > q$ (Franklin and Powell).

Example 12H: Construct an appropriate state space model for the ARMA(2,1) process:

$$y(k) = -2y(k-1) - 3y(k-2) + w(k) - w(k-1) \tag{12.80}$$

Solution is started by defining the state equations:

$$\begin{aligned} x_1(k+1) &= x_2(k) \\ x_2(k+1) &= -3x_1(k) - 2x_2(k) + w(k) \\ y(k) &= -3x_1(k) - 3x_2(k) + w(k) \end{aligned} \tag{12.81}$$

One can show that this state space representation satisfies the ARMA(2,1) model of the example. In general, the following state space representation is valid:

$$\begin{aligned} &x_1(k+1) = x_2(k) \\ &x_2(k+1) = x_3(k) \\ &\vdots \\ &x_n(k+1) = a_n x_1(k) + \cdots + a_1 x_n(k) + w(k) \\ &y(k) = a_n x_1(k) + \cdots + a_1 x_n(k) + w(k) + b_1 x_n(k) + b_2 x_{n-2}(k) + \cdots + b_m x_{n-m}(k) \end{aligned} \tag{12.82}$$

As a result, the methods of extended Kalman filtering or maximum likelihood identification can be used to estimate all of the known parameters (Franklin and Powell).

State Estimation provides the general background for computing the statistics of state and parameter estimates for certain cases where the WLS estimation approach is used. In the case of the state space mode, alternative techniques are available. The state space model is convenient for many applications where the estimation of the unknown parameters is possible by means of an extended Kalman filter or the maximum likelihood method. However, the WLS approach is more often used in practice.

12.9.7 Order and Variance Estimation

Estimation of the orders n and m of the ARMA process requires considerable care. The accepted wisdom of forecasting specialists is that the orders should be as small as possible without jeopardizing the accuracy of the models. This is called parsimonious. This trade-off can be measured by a performance index, to be minimized, to obtain the optimal order of the postulated model.

www.mepcafe.com

Two such indices are:

$$J_1 = s^2 \exp\left(\frac{2l}{N}\right), \tag{12.83}$$

and

$$J_2 = s^2 N^{\frac{l}{N}}, \tag{12.84}$$

where N is the number of sample points, and l is the number of known coefficients (e.g., $f = n + m$ for the ARMA process), and

$$s^2 = \frac{1}{N}\sum_{k=1}^{N}[y(k) - \hat{y}(k)]^2 \tag{12.85}$$

$y(k)$ is the estimated value of $y(k)$ from the postulated model. As f increases, s^2 will normally decrease while J_1 or J_2 will increase. At the minimum of either index, one attains an acceptable trade-off between complexity (larger i) and accuracy (smaller s^2).

The variance data is attainable from the particular estimators used. The estimates of variances and cross-correlations of the estimated coefficients are more critical.

12.9.8 Yule-Walker Equations

The Yule-Walker equations arise naturally in the problem of linear prediction of any zero-mean weakly stationary process x_t based on a finite number of sequential observations. Consider the case where x_t is an AR(p) process:

$$\mathrm{AR}(p) = \sum_{k=0}^{p} a_k x_{t-k} = \eta_t \tag{12.86}$$

where $a_0 = 1$ and $\operatorname{var} \eta_t = \sigma_p^2$. If c_k are the autocovariances, then

$$E(x_t \eta_t) = E\left(x_t \sum_{k=0}^{p} a_k x_{t-k}\right) = \sum_{k=0}^{p} a_t c_k \tag{12.87}$$

On the other hand,

$$E(x_t \eta_t) = E\left[\left(\eta_t - \sum_{k=0}^{p} a_k x_{t-k}\right)\eta_t\right] = E\left[\eta_t^2\right] = \sigma_p^2 \tag{12.88}$$

www.mepcafe.com

Therefore,

$$\sum_{k=0}^{p} a_k c_k = \sigma_p^2 \tag{12.89}$$

For $a_i \geq 0$, we have

$$0 = E(x_t \eta_{t+l}) = E\left[x_t\left(\sum_{k=0}^{p} a_k x_{t+l-k}\right)\eta_t\right] = \sum_{k=0}^{p} a_k c_{l-k} \tag{12.90}$$

Writing the aforesaid for $a_i = 1,\ldots,p$ in matrix form, we have the Yule-Walker equations:

$$\begin{bmatrix} c_0 & c_1 & \cdots & c_p \\ c_1 & c_0 & \cdots & c_{p-1} \\ \cdots & \cdots & \cdots & \cdots \\ c_p & c_{p-1} & \cdots & c_0 \end{bmatrix}\begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} \sigma_p^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{12.91}$$

Let Σ_p denote the $(p1) \times (p1)$ Toeplitz covariance matrix. Note that Σ_p is nonnegative definite, since for any p 1-dimensional vector $b = [b_0, \ldots, b_p]^{\mathrm{T}}$ we have

$$b' \sum_p b = \sum_{j=0}^{p}\sum_{k=0}^{p} b_j c_{j-k} b_k = \operatorname{var}\left[\sum_{j=0}^{p} b_j x_{t-j}\right] \geq 0 \tag{12.92}$$

Note that we did not need to assume that x_t was AR(p) to get this last result.

A Toeplitz matrix is a diagonal-constant matrix, named after Otto Toeplitz, in which each descending diagonal from left to right is constant.

Suppose now that x_t is any weakly stationary zero mean process with autocovariance sequence $\{c_r\}$, and one wants the best linear predictor of x_t based on $x_{t-1}, \ldots, x_{t-p}$. Writing the (one step) predictor as

$$\hat{x}_t = -\sum_{k=1}^{p} b_k x_{t-k} \tag{12.93}$$

the mean-squared prediction error (with $b = 1$) follows:

$$\begin{aligned} E[x_t - \hat{x}_t]^2 &= E\left[\sum_{k=0}^{p} b_k x_{t-k}\right]^2 = E\left[\sum_{j=0}^{p} b_j x_{t-j} \sum_{k=0}^{p} b_k x_{t-k}\right]^2 \\ &= \sum_{j=0}^{p}\sum_{k=0}^{p} b_j c_{j-k} b_k = b' \sum b_p \end{aligned} \tag{12.94}$$

www.mepcafe.com

We will now show that $b'\Sigma_p b$ is minimized (subject to the constraint that $b_0 = 1$) by taking $b \geq a$, where $a = (1, a_1,\ldots,a_p)$'is the solution to the Yule-Walker equations, and that the resulting minimum attainable mean-squared prediction error is σ_p^2. Thus, by solving the Yule-Walker equations a_i, and σ_p^2, we obtain the coefficients of the best linear predictor of x_i based on $x_{t-1}, \ldots, x_{t-p}$ and the corresponding minimum mean-squared error of prediction, even if x_i itself is not AR (p).

Example 12I: Find the Yule–Walker equations and the autocorrelation equation for the AR(2) process:

$$X_t = .33X_{t-1} + .22X_{t-2} + \eta_t \tag{12.95}$$

The Yule-Walker equations are:

$$\rho_k = -.33\rho_{k-1} + .22\rho_{k-2} = 0, \quad k \geq 2 \tag{12.96}$$

Let

$$\rho_k = A\lambda_k \tag{12.97}$$

Then find the roots:

$$\begin{aligned} &\lambda^2 - .33\lambda - .22 = 0 \\ &\lambda_1 = .66, \lambda_2 = -.33 \end{aligned} \tag{12.98}$$

Substitute into the autocorrelation function (ACF):

$$\begin{aligned} \rho_k &= A(.66)^{|k|} + B(-.33)^{|k|} \\ \rho_0 &= A + B = 1 \\ \rho_1 &= .33 + .22\rho_1 = .43 \end{aligned} \tag{12.99}$$

Then solve for A and B:

$$\begin{aligned} &.66A - .33B = .43 \\ &A = .76 \\ &B = .24 \end{aligned} \tag{12.100}$$

So the ACF follows:

$$\rho_k = .76(.66)^{|k|} + .24(.33)^{|k|}, \quad k \in Z \tag{12.101}$$

www.mepcafe.com

12.9.9 Durbin-Levinson Algorithm

Consider the problem of estimating the parameters of an AR(p) model. Start from the Yule-Walker equations. The Durbin-Levinson algorithm provides an alternative that avoids the matrix inversion in the Yule-Walker equations. It is actually a prediction algorithm. It also can be used for parameter estimation for the AR(p) model. Nice side effect of the DL-algorithm is that the partial autocorrelations are found automatically and mean-squared errors associated with our predictions!

The DL-algorithm is a recursive prediction algorithm. Suppose we predict:

$$X_{n+1} = f(X_1, X_2, \ldots, X_n) \tag{12.102}$$

Assume that time goes by and we observe X_{n+1}, but now we want to predict:

$$X_{n+2} = f(X_1, X_2, \ldots, X_{n+1}) \tag{12.103}$$

We could start from "scratch" with the Yule Walker equations, or we use what was learned from predicting X_{n+1} and update that function.

The setup for the DL-algorithm is a mean zero, stationary process $\{X_t\}$ with covariance function $X(h)$. If this is not valid, then subtract the mean, predict, and add the mean back into the forecast.

The best linear predictor of X_{n+1} given the past n values:

$$\hat{X}_{n+1} = b_{n,n}X_1 + b_{n,n-1}X_2 + \cdots + b_{n,1}X_n = \sum_{i=1}^{n}[b_{n,i}X_{n-i+1}] \tag{12.104}$$

The mean-squared prediction error is

$$v_n = E\left[\left(\mathrm{X}_{n+1} - \hat{X}_{n+1}\right)^2\right] \tag{12.105}$$

We want to recursively compute the "best" b's, and, at the same time, compute the v's.

The Durbin-Levinson Algorithm in summary:

Step Zero Set the initial values:

$$\begin{aligned} b_{00} &= 0 \\ v_0 &= \gamma_X(0) \\ n &= 1 \end{aligned} \tag{12.106}$$

www.mepcafe.com

Step One Compute:

$$b_{n,n} = \left[\gamma_X(n) - \sum_{i=1}^{n-1} b_{n-1,i} \gamma_X(n-i) \right] v_{n-1}^{-1} \tag{12.107}$$

Step Two, for $n \geq 2$, compute:

$$\begin{pmatrix} b_{n,1} \\ \vdots \\ b_{n,n-1} \end{pmatrix} = \begin{pmatrix} b_{n-1,1} \\ \vdots \\ b_{n-1,n-1} \end{pmatrix} - b_{n,n} \begin{pmatrix} b_{n-1,n-1} \\ \vdots \\ b_{n-1,1} \end{pmatrix} \tag{12.108}$$

Step Three Compute:

$$v_n = v_{n-1}\left(1 - b_{n,n}^2\right) \tag{12.109}$$

Set $n = n + 1$ and return to Step One.

The DL-algorithm requires:

$$\begin{aligned} &\gamma_X(0) = 0 \\ &\gamma_X(n) \to \infty \quad \text{as } n \to \infty \end{aligned} \tag{12.110}$$

Example 12J: Consider the AR(2) process:

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \eta_t \\ \eta_t &= WN\left(0, \sigma_\eta^2\right) \\ \phi_1 &= 0.5 \\ \phi_2 &= 0.3 \end{aligned} \tag{12.111}$$

This is a prediction exercise, not just an estimation of parameters. Note that the process is causal. Let us recursively predict:

$$\hat{X}_{n+1} \quad \text{for} \quad n = 1, 2, \ldots, N \tag{12.112}$$

based on previous values, and given the MSE of the predictions.
Data needed includes:

$$\gamma_X(0), \gamma_X(1), \gamma_X(2), \ldots \tag{12.113}$$

but we can solve for them as needed. We set up the standard equations by multiplying the AR equation by X_{t-k} and taking expectations:

$$\begin{aligned} \gamma_X(0) - \varphi_1 \gamma_X(1) - \varphi_2 \gamma_X(2) &= \sigma_\eta^2 \\ \gamma_X(1) - \varphi_1 \gamma_X(0) - \varphi_2 \gamma_X(1) &= 0 \\ \gamma_X(2) - \varphi_1 \gamma_X(1) - \varphi_2 \gamma_X(0) &= 0 \end{aligned} \tag{12.114}$$

www.mepcafe.com

which yields:

$$\begin{aligned}\gamma_X(0) &= \frac{1-\varphi_2}{(1+\varphi_2)(1-(\varphi_1+\varphi_2)(\varphi_1+1-\varphi_2))}\sigma_\eta^2 \\ \gamma_X(1) &= \frac{\varphi_1}{(1+\varphi_2)(1-(\varphi_1+\varphi_2)(\varphi_1+1-\varphi_2))}\sigma_\eta^2 \\ \gamma_X(2) &= \frac{\varphi_1^2+\varphi_2-\varphi_2^2}{(1+\varphi_2)(1-(\varphi_1+\varphi_2)(\varphi_1+1-\varphi_2))}\sigma_\eta^2\end{aligned} \tag{12.115}$$

Notice that all three have the same denominator!
Since we have the following, we can easily get additional terms as needed:

$$\gamma_X(k) = \varphi_1\gamma_X(k-1)+\varphi_2\gamma_X(k-2), \quad \text{for} \quad k \geq 2 \tag{12.116}$$

Example 12K: The DL-algorithm is applied to the AR(2) process defined earlier:

$$\begin{aligned} n &= 1 \\ b_{00} &= 0 \\ v_0 &= \gamma_X(0) \\ b_{1,1} &= \rho_X(1) = \frac{\varphi_1}{1-\varphi_2} \\ v_1 &= v_0\left(1-b_{1,1}^2\right)\end{aligned} \tag{12.117}$$

So, the best linear predictor of X_2 based on X_1 is

$$\hat{X}_2 = b_{1,1}X_1 = \frac{\varphi_1}{1-\varphi_2}X_1 \tag{12.118}$$

and the MSE of this predictor is v_1.
The second iteration includes one more term:

$$\begin{aligned} n &= 2 \\ b_{2,2} &= [\gamma_X(2)-b_{1,1}\gamma_X(1)]v_1^{-1} \\ &= \frac{\rho_X(2)-\rho_X^2(1)}{1-\rho_X^2(1)} = \varphi_2 \\ b_{2,1} &= b_{1,1}-b_{2,2}b_{1,1} = \frac{\varphi_1}{1-\varphi_2}(1-\varphi_2) = \varphi_1 \\ v_2 &= v_1\left(1-b_{2,2}^2\right) = \gamma_X(0)\left[1-\frac{\varphi_1^2}{(1-\varphi_2)^2}\right]\left(1-\varphi_2^2\right)\end{aligned} \tag{12.119}$$

www.mepcafe.com

So, the best linear predictor of X_3 based on X_1 and X_2 is

$$\hat{X}_3 = b_{2,2}X_1 + b_{2,1}X_2 = \varphi_2 X_1 + \varphi_1 X_2 \tag{12.120}$$

and the MSE associated with this prediction is v_2.
The next iteration yields:

$$n = 3$$
$$b_{3,3} = \frac{\rho_X(3) - \varphi_1\rho_X(2) - \varphi_2\rho_X(1)}{1 - \varphi_1\rho_X(1) - \varphi_2\rho_X(2)} = \varphi_3 \equiv 0 \tag{12.121}$$

Now the reason for the transformation to this representation becomes apparent! That numerator is 0 for this AR(2) model! This shows agreement with the model specified. Continue for the next term.

$$\begin{pmatrix} b_{3,1} \\ \\ b_{3,2} \end{pmatrix} = \begin{pmatrix} b_{2,1} \\ \\ b_{2,2} \end{pmatrix} - b_{3,3}\begin{pmatrix} b_{2,2} \\ \\ b_{2,1} \end{pmatrix} = \begin{pmatrix} b_{2,1} \\ \\ b_{2,2} \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \\ \phi_1 \end{pmatrix} \tag{12.122}$$

So, the best linear predictor of this series given the three previous readings is

$$\hat{X}_4 = b_{3,3}X_1 + b_{3,2}X_2 + b_{3,1}X_3 = \varphi_2 X_2 + \varphi_1 X_3 \tag{12.123}$$

In fact, for all future n, we will find that the get b values are 0 and

$$\hat{X}_n = \varphi_1 X_n + \varphi_2 X_{n-1} \tag{12.124}$$

Note that we also now know the Partial-autocorrelation function (PACF) for this AR(2) model:

$$\begin{aligned} \mathrm{PACF}_X(1) &= b_{1,1} = \frac{\varphi_1}{1-\varphi_2} \\ \mathrm{PACF}_X(2) &= b_{2,2} = \varphi_2 \\ \mathrm{PACF}_X(n) &= b_{n,n} = 0 \quad \text{for } n \geq 2 \end{aligned} \tag{12.125}$$

12.9.10 Innovations Estimation for MA and ARMA Processes

Innovation algorithm for preliminary estimation of coefficients of MA(q) and ARMA(p; q) processes is the preferred method to initially find the "q" coefficients. To understand the name of the algorithm and how it works, we should make an excursion into forecasting.

Assume that we observe values X_i ($i = 1,n$) coming from MA (q) process:

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots + \theta_q Z_{t-q} \tag{12.126}$$

www.mepcafe.com

Assume that we would like to predict (forecast) future value X_{n+1}, given the first n observations. Take conditional expectation:

$$\begin{aligned}\hat{X}_{n+1} &= E(Z_{n+1} + \theta_1 Z_n + \theta_2 Z_{n-1} + \cdots + \theta_q Z_{n+1-q} \mid X_1, \ldots, X_n) \\ &\equiv 0 + \theta_1 Z_n + \theta_2 Z_{n-1} + \cdots + \theta_q Z_{n+1-q}\end{aligned} \quad (12.127)$$

Note that

$$X_{n+1} - \hat{X}_{n+1} = Z_{n+1} \quad (12.128)$$

Then, in general,

$$X_t - \hat{X}_t = Z_t \quad (12.129)$$

Substitute these differences instead of shocks (Z_t) which are not observable:

$$\hat{X}_{n+1} = \theta_1 \left(X_n - \hat{X}_{n+1} \right) + \cdots + \theta_q \left(X_{n+1-q} - \hat{X}_{n+1-q} \right) \quad (12.130)$$

The differences $(X_{n+1} - \hat{X}_{n+1})$ are called innovations. The forecast is expressed via innovations.

The Innovation algorithm is applicable to all (also *nonstationary*) time series and gives the coefficients of innovations in the expansion:

$$\hat{X}_{n+1} = \sum_{j=1}^{n} \theta_{n,j} \left(X_{n+1-j} - \hat{X}_{n+1-j} \right), \quad n \geq 1 \quad (12.131)$$

Note when $j = n$, the $n + 1$ estimate is the expected value of X_1 which is 0.

Assume

$$E(X_t) = 0, E(X_t X_s) = \kappa(t,s), E \mid X_t \mid^2 < \infty \quad (12.132)$$

Denote

$$v_n = E\left(X_{n+1} - \hat{X}_{n+1} \right)^2 \quad (12.133)$$

This is the innovation variance or mean square prediction error. Then, compute the parameters $\theta_{n1}, \ldots, \theta_{nn}$ recursively from the equations:

$$\begin{aligned}v_0 &= \kappa(1,1) \\ \theta_{n,n-k} &= v_k^{-1}(\kappa(n+1, k+1) - \sum_{j=0}^{k-1} (\theta_{k,k-j} \theta_{n,n-j} v_j), \quad 0 \leq k < n, \\ v_n &= \kappa(n+1, n+1) - \sum_{j=0}^{n-1} \left(\theta_{n,n-j}^2 v_j \right)\end{aligned} \quad (12.134)$$

www.mepcafe.com

Solve iteratively v_0, $\theta_{1,1}$, v_1, $\theta_{2,2}$, $\theta_{2,1}$, v_2, $\theta_{3,3}$, $\theta_{3,2}$, $\theta_{3,1}$, v_3, $\theta_{1,1}$, etc. Note that $\theta_{0,1}$, = 0 for all j. Note the algorithm also gives mean square error:

$$v_n = E\left(X_{n+1} - \hat{X}_{n+1}\right)^2 \tag{12.135}$$

For sample $(x_1,\ldots, x_n)$ we estimate the parameters $(\gamma'(0), \gamma'(1),\ldots,\gamma'(m))$ for $m < n$. Provided that $\gamma'(0) > 0$, and that we *have assumed to fit pure MA model*:

$$\begin{aligned} X_t &= Z_t + \hat{\theta}_{m,1} Z_{t-1} + \cdots + \hat{\theta}_{m,m} Z_{t-m} \\ Z_{t-m} &\sim WN(0, \hat{v}_m) \end{aligned} \tag{12.136}$$

As before, the coefficients are also coefficients in the forecast:

$$\hat{X}_{m+1} = \hat{\theta}_{m,1}(X_m - \hat{X}_m) + \cdots + \hat{\theta}_{m,m}(X_{m+1-m} - \hat{X}_{m+1-m}) \tag{12.137}$$

Innovation algorithm then gives estimates of the parameters $\hat{\theta}_{m,1},\ldots,\hat{\theta}_{m,m},\hat{v}_m$. We use to specify that the theoretical autocovariance function is replaced by a sample auto-covariance function. Normality holds, as in the case of Yule-Walker estimates, and corresponding ratios *Rh*'s are given.

The coefficients' estimates $\hat{\theta}_q = \left(\hat{\theta}_{q,1},\ldots,\hat{\theta}_{q,q}\right)^t$, regrettably, do not converge to $\theta_q = (\theta_q, \ldots, \theta_q)^t$ as $n \to \infty$ for q fixed. Convergence holds when the order of the MA(m) increases as a function of a sample size n: $m = m(n)$ (the order on MA m depends on the sample size n) so that $m \to \infty$ and $n^{-1/3}m(n) \to 0$ as $n \to \infty$.

One would practically run innovation algorithm to estimate coefficients $(\hat{\theta}_{m,1},\ldots,\hat{\theta}_{m,m},\hat{v}_m)$ for increasing m until the part of the vector we are interested in, $\hat{\theta}_{m,1},\ldots,\hat{\theta}_{m,q},\hat{v}_q$ stabilizes comparing with its standard deviation or the ratios R_h's:

$$R_h = \frac{\hat{\theta}_{m,h}}{1.96 \times \text{estimated standard deviation of } \hat{\theta}_{m,h}} \tag{12.138}$$

The largest lag h for which $R_h > 1$ is the order of MA to be fitted. This is the signature to correct MA order q, as necessary.

12.9.11 ARIMA Overall Process

The Box-Jenkins model building process is deceptively simple when applied to demand models of time series. The Box-Jenkins model building process consists of the following steps:

1. Stationary identification
   a. Difference ("Integrate") until it is covariant stationary
   b. Consider alternative extensions to the ARIMA modeling

www.mepcafe.com

   c. De-meaning the process is often the first step
   d. De-trending the process is the second step for short-term demand forecasting
2. Model identification
   a. Estimate ACF
   b. Estimate PACF
3. Model dimension
   a. Make an initial guess at small values of p and q for an ARMA(p, q) model that might describe the (transformed) series.
   b. Estimate the parameters for alternative models: ARMA($p-1$, $q-1$), ARMA($p-1$, q), ARMA (p, $q-1$), $p+1$, q), ARMA(p, $q+1$), and ARMA ($p+1$, $q+1$).
4. Parameter estimation
   a. Objective is to minimize the sum of squares of errors.
   b. Perform diagnostic analysis so the model is consistent with the observed features of the data.
   c. Consider the alternative terms using the goal of a "Parsimonious" solution that uses the fewest terms.
5. Model validation
   a. Diagnostics are used to check the validity of the model.
   b. Is the model adequate for the applications needing this date?
6. Model forecasting
   a. Perform the forecasting of the future expected value of the series as needed.
   b. Repeat the modeling process as needed when future conditions show modeling errors.

The ARIMA Model is written as ARIMA(p, d, q) where p is the number of AR terms, d is the order of differencing, and q is the number of moving average terms.

Integrated at the order d (e.g., the dth difference) is often required to eliminate seasonality. Seasonality is the repetition of a component on some cyclic order such as a daily cycle, weekly cycle, especially weather season cycle (winter, spring, summer, and fall).

ARIMA(1,1,0) is a first-order AR model with one order of differencing. All texts use the "regular" AR, MA, or ARMA model using the same notation. ARIMA(1,0,0) is an AR(1) process. ARIMA(0,0,1) is an MA(1) process. An ARIMA(1,0,1) is a combined ARMA(1,1) process.

Steps 1, 2, and 3 require information on the correlation structure. Auto-regressive (AR) and moving average (MA) are simple to calculate, so, what is the big difference? The AR model includes lagged terms of the time series itself. The MA model includes lagged terms on the noise or residuals. How do we decide which to use? The answer is to examine the ACF and PACF values as the number of lags increases.

www.mepcafe.com

If large values of ACF terms persist, suspect nonstationarity. Take the difference. Usually, application of differencing is one or at most two times. The first available stationary difference should be used. Beware of over differencing as indicated by: increasing variance after differencing, increasing order of MA after differencing, noninvertible MA (unit root). Alternatively, one can examine the frequency spectrum and identify oscillations in the data that are at a longer period than the period to be forecast.

Autocorrelation functions (ACFs) and PACFs are the key signatures throughout this process. Recall that the ACF is a set of correlation coefficients between the series and lags of itself over time. The PACF is the partial correlation coefficients between the series and lags of itself over time. The amount of correlation between a variable and a lag of itself that is not explained by correlations at all lower-order lags. ACF at lag 1 "propagates" to lag 2 and presumably to higher-order lags. PACF at lag 2 is difference between the actual correlation at lag 2 and expected correlation due to propagation of correlation at lag 1.

If the time series is white noise, then all of the ACF and PACF terms are 0.

Autoregressive models have a definitive signature. An AR model of order "p" (AR(p)) indicates that the current value of X_t can be found from past values, plus a random shock (noise). This is similar to a multiple regression model, but X_t is regressed on past values of X_t.

The AR(1) model has a simple signature. A simple way to model dependence over time is with the "auto-regressive model of order 1" (AR(1)). This is a WLS model of X_t regressed on lagged X_{t-1}. What does the model say for the $t + 1$ observation? The AR(1) model expresses what we do not know in terms of what we do know at time t.

If a_1 is 0, X_t depends purely on the random component (η), and there is no temporal dependence. If a_1 is greater than zero (0), previous values of X_t influence the value of X_t as the ACF components show direct exponential decay. If a_1 is less than 0, then the ACF components show oscillating decay. The PACF component will show the first value as nonzero and all higher terms 0 for any value of a_1.

If our model successfully captures the dependence structure in the data, then the residuals should look random. There should be no dependence in the residuals! We then check the AR(1) model; check the residuals from the regression for any "leftover" dependence.

Identifying an AR process of order "p" has a signature of PACF spikes through all lags "p" with all higher-order components are 0. As the PACF displays a sharp cutoff, the ACF decays toward 0 more slowly (i.e., has significant spikes at higher lags), we say that the series displays an "AR signature. The lag at which the PACF cuts off is the indicated number of AR terms.

Moving-average models have a defined signature complementarity. An MA model of order "q" MA(q) has a defined signature. MA implies that the current value of X_t can be found from past shocks/error (eta), plus a new shock/error (etan). The time series is regarded as an MA (unevenly weighted, because of different coefficients) of a random shock series eta.

The MA(1) model has the following signature. If b_1 is greater than 0, X_t ACF shows a positive spike for the first component. Then the PACF shows oscillating

www.mepcafe.com

decay. If b_1 is less than 0, there is a negative spike for the first ACF component. The ACF first component is nonzero and all higher-order terms are 0. Then the PACF shows geometric decay.

Identifying an MA process is complementary to the AR process signature. If the ACF of the differenced series displays a sharp cutoff and/or the lag − 1, autocorrelation is negative then consider adding an MA term to the model. The lag at which the ACF cuts off is the indicated number of MA terms.

Reiterate the signatures one more time. The diagnostic patterns of ACF and PACF for an AR(1) model are

ACF: declines in geometric progression from its highest value at lag 1
PACF: cuts off abruptly after lag 1

The opposite types of patterns apply to an MA(1) process:

ACF: cuts off abruptly after lag 1
PACF: declines in geometric progression from its highest value atlag 1

Mixed ARMA models are again an identification of the signature. The ARMA(1,1) signature is that both the ACF and the PACF exponentially decrease. If a1 is greater than 0, then ACF is nonoscillating decay. The PACF shows oscillating decay beginning at lag 1. If a1 is less than 0, then the ACF shows oscillating decay beginning at lag 1. The PACF shows exponential decay beginning at lag 1.

An ARMA process of the order (p, q) shows ACF decay (either direct or oscillatory) beginning at lag q. The PACF shows decay (either direct or oscillatory) beginning at lag p.

Much of fitting ARMA models is guess work and trial-and-error! Sometimes you can use lower-order models by combining MA and AR terms, such as ARMA(1,1) versus AR(3,0). Lower-order models are better! Parsimonious is a goal to be achieved.

How is the best model selected? In most cases, the best model turns out a model that uses either only AR terms or only MA terms. It is possible for an AR term and an MA term to cancel each other's effects, even though both may appear significant in the model. If a mixed ARMA model seems to fit the data, also try a model with one fewer AR term and one fewer MA term. As with all WLS methods, simpler models are better! Parsimonious!

12.10 ARTIFICIAL NEURAL NETWORKS

The field of ANNs has experienced a resurgence of interest in the last 10 years. This is due in part to the use of neural networks for a variety of new applications, including forecasting. Neural networks are ideally suited to this application when the forecast is to be made based on the complex interactions of several variables, where the exact functional relationships between the variables are not known.

www.mepcafe.com

Artificial Neural Networks have been used to select the Fourier or Hartley Waveforms based on the type of day, day of week, weather variables (temperature, humidity, wind speed, etc.), and day of month. This approach is based on finding the spectrum of each day in the past and identifying those days that are being repeated.

Artificial Neural Networks have been used to find a function based on the type of day, day of week, weather variables (temperature, humidity, wind speed, etc.), and day of month.

Artificial Neural Networks have been used to model the econometric model for heating or cooling demand as an adjustment to the ANN that has found the base demand for that type of day, day of week, etc.

There are several alternative approaches to the application of ANNs to the short-term demand forecasting problem. A review of the complete range of designs is beyond the scope of this work.

The most common form of neural network training—the determining of weights for interconnections between neurons—is known as the back-propagation algorithm. This algorithm is suited to networks having smooth, differentiable transfer functions. Unfortunately, back-propagation converges to only a local minimum in some instances and fails to converge at all. The development of new training algorithms for ANNs is an active area of research today.

12.10.1 Introduction to Artificial Neural Networks

The human brain contains on the order of 100 billion neurons, simple processing elements that give off brief electrical pulses when they are sufficiently excited. Each neuron is electrochemically connected to about 10,000 other neurons by way of branches called axons and chemical-filled gaps known as synapses. The pulses from a neuron travel through these connections to either excite or inhibit other neurons. It is this massively interconnected network that allows us as humans to accomplish a myriad of tasks including controlling our speech and movement.

The field of ANNs was based on a simple mathematical model for the chemical reactions occurring within biological neural networks.

The brain is also the site of human memory. When a child recognizes his or her mother's voice, or a senior citizen remembers his or her high school graduation, the mind is actually retrieving patterns that are stored away in the brain. But how does the brain store and retrieve memories? It is not done in the way that one of today's computers does so, whereby a discrete memory location is allocated for a bit of data, that location is assigned a value of high or low corresponding to the piece of information, and when asked to recall the data the value at the location is read and used wherever the program desires. Instead, the brain stores information in a distributed fashion—in its structure. It is the network of neurons, synapses, and the specific strengths of the electrical interconnections than contain human memory.

An extremely crude abstraction of human memory would be a matrix whose elements represent the strength of the electrical interconnections between different

www.mepcafe.com

neurons. Just looking at this matrix, of course, we could not decipher any of the information held within it, but we could know which neurons were better connected electrically to other neurons and vice versa. We could perform mathematical operations on parts of the matrix, though, and possibly we could interpret these results. If the matrix were for a living being, its elements would not be static, but instead would change as new memories are added and other memories are lost (though, hopefully, not at the same rate). Looking at this process on a microscale, D.O. Hebb explained it as "when an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased." This is how changes in memory take place. As a person is exposed to events, his sensory nerves cause neurons in the brain to be excited and pulse. If this pulsing is frequent enough, the interconnection strength between neurons change, and new information is added to the memory.

It was this theory that prompted researchers to begin studying ANNs—abstractions of human neural structure that can be "trained" to recognize patterns, classify items, and make forecasts, among other applications. These networks rarely contain more than 1000 neurons and, thus, are not able to generalize on the scale that the human brain works. The neural network is usually represented mathematically as a set of weights representing the interconnection strengths, along with some model or models of how the neurons behave when excited. Neural network research has included varying the interconnection schemes between neurons, altering neuron transfer functions, experimenting with mathematical techniques to determine the interconnection strengths which optimize the system, and determining where the networks are best suited for application.

12.10.2 Artificial Neurons

A detailed model for an artificial neuron used in ANNs is shown in the following Figure 12.2. The inputs are applied to the neuron and are scaled and summed. A bias term is added to the sum which is then passed through the transfer function of the neuron to yield the output. The transfer function may include a threshold trigger to give an output only if that threshold is reached. The standard symbol for a neuron is simply a circle with the interconnections shown going into and out of the neuron.

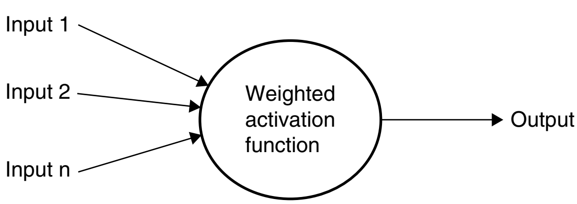


FIGURE 12.2 A model of an artificial neuron.

www.mepcafe.com

12.10.3 Neural network applications

Artificial neural networks have a wide range of application. One of the first applications of the networks was the pattern recognition problem. In this type of problem, items are taken from a set and classified into distinct groups based upon their characteristics. In some instances, ANNs can outperform the classical techniques for this application. This is one method to apply ANN to forecasting.

Neural networks can also serve as a model for human memory. The Hopfield neural network—discussed in a later section—is one such model that is capable of recalling specific binary or bipolar patterns with which it has been trained. This network has also shown promise as a tool for optimization because of a special energy function that may be associated with the network's state.

One of the most common applications for ANNs is forecasting future trends in many industries. The best types of problems to apply ANNs to are ones where the variables that affect the output are known or can be determined, while the exact functional relationships between these variables cannot be determined. Neural networks are presently being used to forecast such things as stock market prices, and the risk associated with different loan applications.

Our example of this condition is the power system short-term demand forecasting problem. The inability to determine these functional relationships makes the use of traditional techniques such as regression analysis impractical for these forecasts. ANNs are particularly valuable for problems that are extremely nonlinear.

A given ANN is not valuable for the applications listed earlier until certain network parameters are determined which cause the network to behave in a meaningful manner. This process is known as "training" the ANN. Specifically, neural network training involves the determination of weights for interconnections between neurons. The goal of the training is usually to minimize the sum of a function of output errors for a given set of inputs, or to ensure that the stable points of a network correspond to a set of training patterns. For example, training data may consist of a set of input vectors and a set of output vectors, where each input pattern is associated with a single output pattern. The goal of the training is to have the actual output of the network be as close, in some performance measure, to the associated output pattern when a given input pattern is applied. The back-propagation algorithm and the interior point linear programming algorithms are two such approaches.

12.10.4 Hopfield Neural Networks

The Hopfield network has a single layer of neurons, each of which is fed back to the inputs of each of the other neurons. The neurons are characterized by a bipolar or binary output and a threshold transfer function. The output of each neuron is fed back to the input of each other neuron via weighted interconnections. Conditions are often imposed on these feedback connections, including symmetric weights and no self-feedback. Absolute stability is guaranteed for networks having these particular restrictions on the weights. The asynchronous network operates as follows. First, an initial output vector is selected and network output is held at this

www.mepcafe.com

state. Then according to some probability distribution, a neuron is selected for update. If neuron i is selected during iteration k, its output is updated according to a given rule. The rule is based on the neuron's state (x), the weight for the interconnection to neuron i from neuron j ($W_{i,j}$), and the threshold for each neuron (t_i). The output of the remaining neurons does not change at this stage. The network is then checked for stability; that is, to see if any neuron from the network would have its output change if it were selected for update. If the network is stable, stop. Otherwise, select another neuron for update and repeat the process. This is very similar to the Gauss Power Flow algorithm.

The usual goal of training a Hopfield network is to increase the likelihood that the network will stabilize to one of the training patterns during operation. Hopfield networks may also stabilize at spurious solutions, enter a limit cycle, or wander chaotically during operation, depending on the restrictions placed on the weights. The most common method for training Hopfield networks is known as the sum of outer products algorithm. Another method uses linear programming to maximize the regions of convergence about the training patterns.

12.10.5 Feed-Forward Networks

Feed-forward networks are the most common type of ANNs. The inputs are applied to the first layer of neurons, which usually have a linear transfer function. This input layer is connected to the hidden layer via weighted interconnections. Each hidden layer neuron has an associated transfer function, and the output of the hidden layer neurons serves as input to the next layer of neurons. Usually, only one hidden layer is present so this layer is directly connected to the output layer by weighted interconnections. Feed-forward ANNs may have more than one hidden layer, but this topology is rarely used. The output layer neurons may have either a linear or nonlinear transfer function, depending on the design of the network. There are no limits to the number of neurons in a given layer. A simple feed-forward network is shown in Figure 12.4. This network has four inputs, two hidden layer neurons, and one output. The network is fully connected, meaning that each neuron is connected to all the neurons in the successive layer, and a bias is connected to the hidden and output layer neurons.

The most common nonlinear activation function for a neuron is the sigmoid. This shape is shown in Figure 12.3.

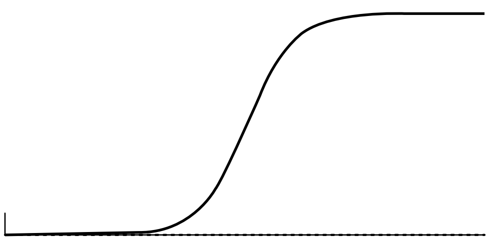


FIGURE 12.3 Sigmoid function.

www.mepcafe.com

The sigmoid typically starts at 0 and increases to 1.0 according to the function:

$$S(t) = \frac{1}{1+e^{-t}} \tag{12.139}$$

Other examples of activation functions include: the ***Gompertz curve***, the ***Ogee curve***, the ***logistic***, and ***hyperbolic tangent*** functions. Sigmoid curves based on statistical cumulative distributions include: the integrals of the ***logistic distribution***, the ***normal distribution***, and ***Student's t probability density functions***. The common neuron transfer functions which appear often in the research literature include: learning curve, Gaussian, and step functions.

An alternative is the piecewise-linear equivalent function. Functionally this is expressed as

$$F(k) = \{a \text{ if } x > \max, d \text{ if } x < \min, b + cx \text{ otherwise}\} \tag{12.140}$$

where a, b, c, and d are constants selected to mimic the sigmoid function. Note that as x tends to infinity the sigmoid approximates a threshold function. Linear programming (LP) is widely used when this transfer function is assumed.

A feed-forward network is normally three layers: input, hidden, and output layer as shown in Figure 12.4. The number of neurons per layer is one of the parameters to identify when fitting the data. This example network was used to add terms as needed from historical patterns as an additive process.

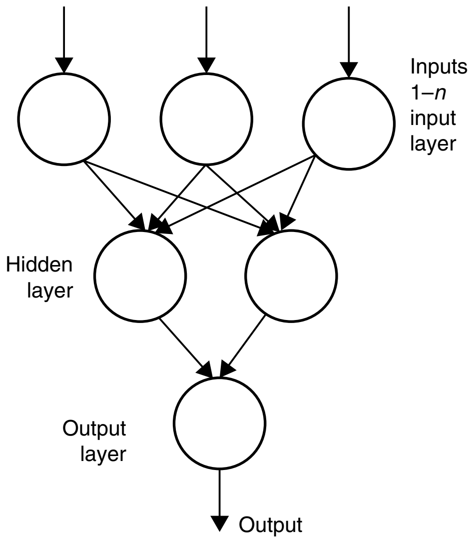


FIGURE 12.4 A feed-forward artificial neural network.

www.mepcafe.com

The training set for a feed-forward network is a set of input–output vectors. When a given training input pattern is applied to the input layer of the ANN, the error is some measure of the difference between the actual output of the network and the desired output vector. The interconnection weights are determined in the training process so that the output error for the training cases is minimized. The most common training algorithm for feed-forward networks is a gradient technique known as back-propagation. This iterative method is a blame-assigning algorithm that changes a weight based on that interconnection's contribution to the error over the individual patterns. Each training pattern is presented several—possibly several thousand—times in the algorithm. Detailed discussions of the back-propagation algorithm may be found in any text on neural networks.

The load profiles tend to change with the time of the year. This is due to factors such as seasonal loads (i.e., heating and cooling) and growth. ANNs can include these seasonal variables in two ways. First, a given model may be retrained regularly, sometimes daily, to reflect the load trend. The training data in this case is selected from a time-limited set, assuring that the training data is always recent enough to be unaffected by seasonal variations. This is known as the "moving window" technique for training data selection. The other method for dealing with seasonal variables is the use of separate models for different times of the year. Each model includes variables that are pertinent only to that particular season of the year.

The ANN method models the multivariate forecast problem without making complex dependency assumptions about the input variables, but instead relies on the selection of appropriate training and input data. The ANN performs what is essentially a pattern recognition function, based upon the historical data that is used to train the network. With its ability to be retrained using recent historical data, the ANN is inherently updatable and eliminates the need for huge databases; thus, the amount of data needed to train the network is at a minimum. The short-term ANN methods reviewed do not have the ability to capture the seasonal and annual load growth trends that other techniques can capture. A proper comparison of techniques would require the ANNs to include all models incorporated within the classical techniques.

The use of ANNs for STLF may be reduced to several tasks. The process begins by determining the length and type of the forecast to be made. Forecasts are usually daily or weekly, and may estimate the hour-by-hour load or the peak load for the period, among other possibilities. Next, the appropriate type and quantity of training data is selected. The specific inputs required for a forecast vary by geography, demographics, and time of year. In general, a neural network that is used to forecast load for one region is not an optimal design for another region. The amount of training data required to train the network also vanes. With the moving window approach, it is useful to select data from several weeks prior to the forecast period. This width of this window is normally on the order of four weeks to prevent negative effects from seasonal load changes. Historical data from the same period in previous years may also be used, but caution must be exercised with regard to annual load growth which could tend to hurt results. Next, the structure of the ANN is determined. The number of input neurons is determined by the number of input variables, and the number of output neurons is

www.mepcafe.com

determined in a similar fashion. No method exists for determining the required number of hidden layer neurons without experimenting. One popular method consists of adding neurons until no additional benefit is seen during the training period. Another approach "prunes" neurons that have low values on incoming and outgoing interconnections, because these neurons increase training time but do not play a large role in the forecast. Once all this has been completed, the neural network is trained with the algorithm of choice, a forecast is made, and the process begins again.

12.10.6 Back-Propagation Algorithm

The back-propagation algorithm (McClelland and Rumelhart, 1988, see reference 1) is used in layered feed-forward ANNs. This means that the artificial neurons are organized in layers, and send their signals "forward," and then the errors are propagated backward. The network receives inputs by neurons in the input layer, and the output of the network is given by the neurons on an output layer. There may be one or more intermediate hidden layers.

The back-propagation algorithm uses supervised learning, which means that we provide the algorithm with examples of the inputs and outputs we want the network to compute, and then the error (difference between actual and expected results) is calculated. The idea of the back-propagation algorithm is to reduce this error, until the ANN learns the training data. The training begins with random weights, and the goal is to adjust them so that the error will be minimal.

The activation function of the artificial neurons in ANNs implementing the back-propagation algorithm is a weighted sum (the sum of the inputs × multiplied by their respective weights *w*:

$$A_j(\overline{x},\overline{w}) = \sum_{j=0}^{n} x_j w_{j,i} \tag{12.141}$$

We can see that the activation depends only on the inputs and the weights.

If the output function would be the identity (output = activation), then the neuron would be called linear. But these have severe limitations.

The most common output function is the sigmoidal function as presented earlier. The sigmoidal function is very close to one for large positive numbers, 0.5 at 0, and very close to 0 for large negative numbers. This allows a smooth transition between the low and high output of the neuron (close to 0 or close to 1). One can see that the output depends only in the activation, which in turn depends on the values of the inputs and their respective weights.

Now, the goal of the training process is to obtain a desired output when certain inputs are given. Since the error is the difference between the actual and the desired output, the error depends on the weights, and we need to adjust the weights in order to minimize the error. We can define the error function for the output of each neuron:

$$E_j(\overline{x},\overline{w},d) = (O_j(\overline{x},\overline{w}) - d_j)^2 \tag{12.142}$$

www.mepcafe.com

We take the square of the difference between the output and the desired target because it will be always positive, and because it will be greater if the difference is big, and lesser if the difference is small. The error of the network will simply be the sum of the errors of all the neurons in the output layer:

$$E(\overline{x},\overline{w},\overline{d}) = \sum_{j=1}^{n}(O_j(\overline{x},\overline{w}) - d_j)^2 \tag{12.143}$$

The back-propagation algorithm next calculates how the error depends on the output, inputs, and weights. After we find this, we can adjust the weights using the method of gradient descendent:

$$\Delta w_{j,i} = -\eta \frac{\partial E}{\partial w_{j,i}} \tag{12.144}$$

This formula can be interpreted in the following way: the adjustment of each weight $w_{j,i}$ will be the negative of a constant eta multiplied by the dependence of the *i*th previous weight on the error of the network, which is the derivative of E with respect to w. The size of the adjustment will depend on O, and on the contribution of the weight to the error of the function. That is, if the weight contributes a lot to the error, the adjustment will be greater than if it contributes in a smaller amount. The weight is adjusted until we find appropriate weights (the error is minimal). The mathematical proof of the back-propagation algorithm is out of the scope of this material.

One only needs to find the derivative of E in respect to the weights (w). This is the goal of the back-propagation algorithm, since we need to achieve this backward. First, we need to calculate how much the error depends on the output, which is the derivative of E in respect to the output:

$$\frac{\partial E}{\partial O_j} = 2(O_j - d_j) \tag{12.145}$$

And then, how much the output depends on the activation, which in turn depends on the weights:

$$\frac{\partial O_j}{\partial w_{j,i}} = \frac{\partial O_j}{\partial A_j}\frac{\partial A_j}{\partial w_{j,i}} \tag{12.146}$$

And one can use the chain rule:

$$\frac{\partial E_j}{\partial w_{j,i}} = \frac{\partial E_j}{\partial O_j}\frac{\partial O_j}{\partial w_{j,i}} \tag{12.147}$$

www.mepcafe.com

And so, the adjustment to each weight will be:

$$\Delta w_{j,i} = -2\eta(O_j - d_j)O_j(1-O_j)x_t \tag{12.148}$$

We can use this adjustment as it is for training an ANN with two layers. Now, for training the network with one more layer we need to make some observations. If we want to adjust the weights (let us call them v) of a previous layer, we need first to calculate how the error depends not on the weight, but in the input from the previous layer. This is easy as we would just need to change x with w in the aforementioned equations. But we also need to see how the error of the network depends on the adjustment of v. So:

$$\Delta v_{l,k} = -\eta \frac{\partial E_j}{\partial v_{j,i}} = -\eta \frac{\partial E_j}{\partial x_l}\frac{\partial x_l}{\partial v_{j,i}} \tag{12.149}$$

where

$$\frac{\partial E}{\partial w_j} = 2(O_j - d_j)O_j(1-O_j)w_{j,i} \tag{12.150}$$

And, assuming that there are inputs u into the neuron with v:

$$\frac{\partial x_j}{\partial v_{l,k}} = x_l(1-x_j)v_{l,k} \tag{12.151}$$

If we want to add yet another layer, we can do the same, calculating how the error depends on the inputs and weights of the first layer. We should just be careful with the indexes, since each layer can have a different number of neurons, and we should not confuse them.

For practical reasons, ANNs implementing the back-propagation algorithm do not have too many layers, since the time for training the networks grows exponentially. Also, there are refinements to the back-propagation algorithm which allow a faster learning.

Example 12L: Artificial Neural Network example of a curve fit to a nonlinear equation.

Assume that our model has three inputs, a, b, and c, and generates an output y. Assume that the actual process is given by the following equation:

$$y = 3a + bc + 8c;$$

Use this model to generate data. In practice, we do not have the mathematical model but obtain the data by observing the real system.

Let us make a very simple neural network for this case: 1 input and 1 output layer. Assume that the input layer size is 5 neurons. Since we have three inputs (a, b, c), our input layer will take the three values and transform it to a vector of size 5. Our output

www.mepcafe.com

layer takes this 5 element vector as input and transforms it to a vector of size 1 because there is only one output.

First, we will make a matrix of the training vector (R) which is size 3×2. The first column shows the minimum of all three inputs and the second will show the maximum of three inputs. In our case, all three inputs range from 0 to 1. So:

$$R = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \tag{12.152}$$

Next, make a size matrix which has the neural nodes for all the layers:

$$S = [51] \tag{12.153}$$

The weight matrix of the trained model follows:

$$W(1) = \begin{bmatrix} -0.368 & 0.031 & -0.540 \\ 0.464 & 0.234 & 0.588 \\ 1.957 & -1.689 & 1.540 \\ 1.114 & 1.084 & 0.244 \end{bmatrix} \tag{12.154}$$

$$W(2) = [-11.20, 9.459, -1.000, -0.914] \tag{12.155}$$

Now test it again on some other data. What about $a = 1$, $b = 1$, and $c = 1$?
So input matrix is

$$S = [1 \quad 1 \quad 1] \tag{12.156}$$

The result returned is 12 which is the actual output (3 × 1 + 1 × 1 + 8 × 1);

12.10.7 Interior Point Linear Programming Algorithms

At about the same time that ANNs became popular again, the field of linear programming experienced a major advance with the introduction of Karmarkar's interior point algorithm. This algorithm guaranteed polynomial-time solutions to large linear programs, where traditional methods such as the simplex technique could not. The interior point algorithms operate by shooting through the interior of the feasible region rather than moving from vertex to vertex of this region as in the simplex approach. Subsequent refinements to the Karmarkar's algorithm have reduced the computational complexity while at the same time speeding convergence to the solution. One of these improved algorithms is the affine scaling algorithm. The affine scaling algorithm is guaranteed to converge to the optimal solution for a linear program where all the variables are constrained to be nonnegative. The original

www.mepcafe.com

version of this algorithm operates by rescaling the variables so they are equidistant from the nonnegativity constraint boundaries, and then the solution moves in the direction of a projected gradient.

The IPLP algorithm is well-suited to the linearized ANN function that is a piecewise linear function, such as the following:

$$O_j = \left[\min\left(1.0, \max\left(0.0, \sum_{i=1}^{n} W_i I_i\right)\right)\right] \tag{12.157}$$

The material in optimal power flow chapter details this algorithm.

12.11 MODEL INTEGRATION

To recapitulate, historical data for the immediate past (normally, last five weeks of hourly demand and weather data) are used to identify an overall model of the form:

$$P_{\mathrm{D}}(k) = B(k) + W(k) + S(k) + v(k) \tag{12.158}$$

As a first step, unexplained demands due to special events are smoothed out since they do not represent normal demand behavior. Next, the parameters of the weather-sensitive part are identified and $W(k)$ (or its estimate) is subtracted from $P_{\mathrm{D}}(k)$ to obtain estimates of the base demand $B(k)$. This is then treated as an ARIMA process whose parameters are estimated using any of a variety of available techniques. The statistics of all estimates are also computed since they impact on the prediction step. This procedure is summarized in the following process:

1. Collect historical data
2. Remove special events
3. Smooth to remove higher-order frequencies
4. Identify parameters of weather model(s)
5. Identify base demand model
6. Identify order of ARMA model
7. Estimate parameters of ARMA model
8. Validate models

12.12 DEMAND PREDICTION

Depending on the type of application considered the models thus identified can be used to predict both the demand and its statistics. On the whole we shall consider three types of demand prediction: (a) Hourly system demand forecast, (b) One-Step ahead forecast, and (c) Hourly bus demand forecasts.

www.mepcafe.com

12.12.1 Hourly System Demand Forecasts

Hourly system forecasts for the next 72–168 h are needed for unit commitment, hydro scheduling, short-term maintenance scheduling, and production costing. The prediction process may be conducted once or a few times during the day. It consists of two steps. The first is an updating step of all model parameters based on a fixed set of the past *N* hours. This produces all the estimates of model parameters. The second is a prediction step based on forecasted weather and special events. With the previously estimated parameters, the weather component of the demand and its variance can be computed. Demands due to special events are estimated by the system operator on the basis of his or her experienced judgment.

The base demand component is obtained by means of the relations:

$$B(k+l) = \overline{B}(k+l) + \hat{y}(k+l) \tag{12.159}$$

$$\hat{y}(k+1) = \sum_{i=1}^{n} a_i \hat{y}(k+l-i) + \sum_{j=1}^{m} b_j \hat{w}(k+l-j) \tag{12.160}$$

where $B(k + 1)$ is the level of $B(k + 1)$ based on past averaging of base demand for the specified hour of the day, and $y(k + 1)$ is the best estimate of $y(k + 1)$, i hours from the present time k. Furthermore,

$$\hat{y}(k+l-i) = \begin{cases} y(k+l-i), & i \geq l \\ \hat{y}(k+l-i), & \text{otherwise} \end{cases} \tag{12.161}$$

$$\hat{w}(k+l-i) = \begin{cases} \hat{w}(k+l-i), & i \geq l \\ 0, & \text{otherwise.} \end{cases} \tag{12.162}$$

What these relations indicate is that y is to be replaced by its estimate whenever one considers a future hour. Similarly, for past data estimates of w are available, but for future data w is replaced by its mean, which is 0. In estimating the variance of the predicted demand one may use the following approximation:

$$\sigma^2_{y(k+l)} = \sigma^2\left(1 + d_1^2 + d_2^2 + \cdots + d_{l-1}^2\right) \tag{12.163}$$

where σ^2 is the variance of $w(k)$ and d_1 are the coefficients of the equivalent MA process associated with the basic ARMA process.

12.12.2 One-Step Ahead Forecasts

In those applications where a one-step-ahead forecast is all that is needed (e.g., AGC, security assessment, or the optimal power flow), then a simple ARMA (or ARIMA) model can be used. The coefficients of the models used are updated periodically every time a new PD_k is introduced such that only the last *N* records are

www.mepcafe.com

retained. In such cases, the sequential updating algorithm (matrix inversion lemma) can be very useful.

There are several filtering techniques that yield the next step forecast. Such methods are appropriate for control algorithms, such as AGC.

12.12.3 Hourly Bus Demand Forecasts

Bus demand forecasts for the next 24 h or so are essential for line outage maintenance, security assessment, detailed unit scheduling, and other applications.

There are two basic approaches to this application. In the first approach, each bus is assigned a participation factor $p(k)$, which is the fraction of the demand served by the h bus at hour k. Given the overall system demand forecast one can easily compute the ith bus demand by the relation:

$$P_{Di}(k) = \rho_i(k) P_D(k) \tag{12.164}$$

In a second approach, on-line state estimator data, together with weather data are used to develop an independent forecast for each bus using a combined base demand and weather-sensitive demand model. The forecasts are then adjusted to be consistent with the overall system demand forecast.

12.13 CONCLUSION

Short-term demand forecasting is required by existing and projected functions of power system control and operation. The targeted applications include AGC, security assessment, OPF, unit commitment, hydro scheduling, and others.

Discussion focused first on modeling of the demand and the identification of model parameters. For weather-sensitive demands models, the techniques of weighted least squares estimation can prove to be quite effective. Base demands can be modeled as ARMA processes whose parameters can be identified by a variety of methods. The identified models can be used in predicting the demand and its variances over the forecasting period.

Other useful methods implementing evolutionary programming (EP) and fuzzy logic (FL) into conventional time series models have been also proposed. We will not consider these methods in detail here.

A fact is that no one method can be applicable to all situations. So a method should be chosen considering many factors, such as the time frame, pattern of data, cost of forecasting, desired accuracy, availability of data, and easiness of operation and understanding. Therefore, more work still needs to be done.

Demands that result from special events are harder to predict. System operator judgment is perhaps the most reliable in these cases. One should consider the future development of an expert system or ANN to perform that function.

www.mepcafe.com

PROBLEMS

12.1 For the AR process:

$$y(k)+\frac{1}{3}y(k-1)+\frac{1}{3}y(k-2)=w(k)$$

where var[$w(k)$] = s^2, determine the variance of $y(k)$.

12.2 For the ARMA (1,1) process:

$$y(k)+0.4y(k-1)=w(k)-0.6w(k-1)$$

determine an equivalent MA process:

$$y(k)=w(k)+d_1w(k-1)+d_2w(k-2)+\cdots$$

Compute the coefficients d_1.

12.3 Find the ACF and the PACF for various assumed ARIMA(1,0,0), ARIMA(0,0,1), ARIMA(1,0,1), ARIMA(2,0,0), ARIMA(0,0,2), and ARIMA(2,0,2). Note that the patterns of exponential decay are present in the ACF, the PACF, or both.

12.4 Given the process:

$$y(k)+0.4y(k-1)=w(k)-0.6w(k-1)$$

generate the time series data and fit a Fourier (Hartley) series to this data using the WLS curve fitting process. How many samples are needed? What is the Frequency Spectrum? Compare to the fast transfer algorithms for each.

12.5 Prove the method of Yule-Walker in Section 12.12.

12.6 Find the ARMA parameters for the process of Problem 12.2. First, simulate the ARMA process to generate the data without adding noise. Then add noise to determine how quickly the estimated parameters change. Use the WLS algorithm.

12.7 Find the AR parameters for the process of Problem 12.1. First, simulate the AR process to generate the data without adding noise. Then add noise to determine how quickly the estimated parameters change.

12.8 Find the parameters of the following MA process using Innovations estimator. The MA process is of second order, with O1 = 0.5 and O2 = 0.2. Generate the data without noise, then find the parameters. Subsequently, add noise.

www.mepcafe.com

12.9 The patterns of the Problem 12.3 are used to identify the order of the ARIMA process. Develop an ANN to pick the best order given this information.

12.10 Compare the algorithms for curve fitting a Fourier (Hartley) series and the fast transforms.

12.11 Repeat Example 12F for the following series:

a. Two-term exponential model:

$$y = ae^{bx}$$
$$y = ae^{bx} + ce^{dx}$$

b. Gaussian model is used for fitting peaks, and is given by the equation

$$y = \sum_{i=1}^{n} \left(a_i e^{\left[-\left(\frac{x - b_i}{c_i} \right)^2 \right]} \right)$$

where a is the amplitude, b is the centroid (location), c is related to the peak width, n is the number of peaks to fit.

c. Polynomial models are given by

$$y = \sum_{i=1}^{n+1} \left(p_i x^{n+1-i} \right)$$

where $n + 1$ is the order of the polynomial and n is the degree of the polynomial. The order gives the number of coefficients to be fit, and the degree gives the highest power of the predictor variable.

d. One-term and a two-term power series model:

$$y = ax^b$$
$$y = a + bx^c$$

e. Rational models are defined as ratios of polynomials:

$$y = \frac{\sum_{i=1}^{n+1} \left(p_i x^{n+1-i} \right)}{x^m + \sum_{i=1}^{m} \left(q_i x^{m-i} \right)}$$

where n is the degree of the numerator polynomial and m is the degree of the denominator polynomial. Note that the coefficient associated with the

www.mepcafe.com

last power of m is always 1. This makes the numerator and denominator unique when the polynomial degrees are the same.

f. Sum of sines model is used for fitting periodic functions:

$$y = \sum_{i=1}^{n} \left(a_i \sin\left(b_i x + c_i \right) \right)$$

where a is the amplitude, b is the frequency, and c is the phase angle constant for each sine wave term. And n is the number of terms in the series. This equation is related to the Fourier series described previously as the sin and cosine functions are shifted by ninety degrees (90°). The main difference is that the sum of sines equation includes the phase constant, and does not include a DC offset term.

g. The Weibull distribution is widely used in reliability and life (failure rate) data analysis. The toolbox provides the two-parameter Weibull distribution

$$y = abx^{b-1} e^{-ax^b}$$

where a is the scale parameter and b is the shape parameter. Note that there is also a three-parameter Weibull distribution and a one-parameter Weibull distribution that are widely used.

12.12 Suppose $y(N + 1)$ is made available. Show that the parameters vector a $(N + 1)$ with the new data can be updated as follows:

$$\hat{\mathbf{a}}^{N+1} = \hat{\mathbf{a}} + \frac{\mathbf{P}_N}{1 + \mathbf{h}^{\mathrm{T}} \mathbf{P}_N h} (y(N+1) - \mathbf{h}^{\mathrm{T}} \mathbf{a})$$

where

$$\mathbf{P}_N = (\mathbf{H}^{\mathrm{T}} \mathbf{H})^{-1}$$

$$h = \begin{bmatrix} y(N) \\ y(N-1) \\ \vdots \\ y(N-n+1) \end{bmatrix}$$

REFERENCE

1. McClelland, J. L., Rumelhart, D. E., *Explorations in Parallel Distributed Processing – a Handbook of Models, Programs, and Exercises*, MIT Press, Cambridge, 1988.

www.mepcafe.com

INDEX

Power Generation, Operation, and Control, Third Edition. Allen J. Wood, Bruce F. Wollenberg, and Gerald B. Sheblé.

© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com

www.mepcafe.com