

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.

1.2 COURSE SCOPE

Topics to be addressed include:

1. Power generation characteristics.
2. Economic dispatch and the general economic dispatch problem.
3. Thermal unit economic dispatch and methods of solution.
4. Optimization with constraints.
5. Using dynamic programming for solving economic dispatch and other optimization problems.

6. Transmission system effects:
 - a. power flow equations and solutions,
 - b. transmission losses,
 - c. effects on scheduling.
7. The unit commitment problem and solution methods:
 - a. dynamic programming,
 - b. the Lagrange relaxation method.
8. Generation scheduling in systems with limited energy supplies.
9. The hydrothermal coordination problem and examples of solution techniques.
10. Production cost models:
 - a. probabilistic models,
 - b. generation system reliability concepts.
11. Automatic generation control.
12. Interchange of power and energy:
 - a. interchange pricing,
 - b. centrally dispatched power pools,
 - c. transmission effects and wheeling,
 - d. transactions involving nonutility parties.
13. Power system security techniques.
14. An introduction to least-squares techniques for power system state estimation.
15. Optimal power flow techniques and illustrative applications.

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 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 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: $10^7 \text{ kW} \times 8760 \text{ h/yr} \times 0.60 = 5.256 \times 10^{10} \text{ kWh}$

Annual fuel consumption: $10,500 \text{ Btu/kWh} \times 5.256 \times 10^{10} \text{ kWh}$
 $= 55.188 \times 10^{13} \text{ Btu}$

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

To put this cost in perspective, it represents a direct requirement for revenues from the average customer of this system of 3.15 cents per 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. Recent fuel costs always seem to be ancient history and entirely inappropriate to current conditions. To avoid leaving false impressions about the actual value of the methods to be discussed, all the examples and problems that are in the text are expressed in a nameless, fictional monetary unit to be designated as an "R."

1.4 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 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 that "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 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 above, 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 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.

FURTHER READING

The books below are suggested as sources of information for the general area covered by this text. The first four are "classics;" the next seven are specialized or else are collections of articles or chapters on various topics involved in generation operation and control. Reference 12 has proven particularly helpful in reviewing various thermal cycles. The last two may be useful supplements in a classroom environment.

1. Steinberg, M. J., Smith, T. H., *Economy Loading of Power Plants and Electric Systems*, Wiley, New York, 1943.
2. Kirchmayer, L. K., *Economic Operation of Power Systems*, Wiley, New York, 1958.
3. Kirchmayer, L. K., *Economic Control of Interconnected Systems*, Wiley, New York, 1959.
4. Cohn, N., *Control of Generation and Power Flow on Interconnected Systems*, Wiley, New York, 1961.
5. Hano, I., *Operating Characteristics of Electric Power Systems*, Denki Shoin, Tokyo, 1967.
6. Handschin, E. (ed.), *Real-Time Control of Electric Power Systems*, Elsevier, Amsterdam, 1972.
7. Savulescu, S. C. (ed.), *Computerized Operation of Power Systems*, Elsevier, Amsterdam, 1976.
8. Sterling, M. J. H., *Power System Control*, Peregrinus, London, 1978.
9. El-Hawary, M. E., Christensen, G. S., *Optimal Economic Operation of Electric Power Systems*, Academic, New York, 1979.
10. Cochran, R. G., Tsoufanidis, N. M. I., *The Nuclear Fuel Cycle: Analysis and Management*, American Nuclear Society, La Grange Park, IL, 1990.
11. Stoll, H. G. (ed.), *Least-Cost Electric Utility Planning*, Wiley, New York, 1989.
12. El-Wakil, M. M., *Power Plant Technology*, McGraw-Hill, New York, 1984.

13. Debs, A. S., *Modern Power Systems Control and Operation*, Kluwer, Norwell, MA, 1988.
14. Strang, G., *An Introduction to Applied Mathematics*, Wellesley-Cambridge Press, Wellesley, MA, 1986.
15. Miller, R. H., Malinowski, J. H., *Power System Operation*, Third Edition, McGraw-Hill, New York, 1994.
16. Handschin, E., Petroianu, A., *Energy Management Systems*, Springer-Verlag, Berlin, 1991.

2 Characteristics of Power Generation Units

2.1 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 2.1. 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.

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

H = Btu per hour heat input to the unit (or MBtu/h)

F = Fuel cost times H is the R per hour (R/h) input to the unit for fuel

Occasionally the R per hour 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 2.2 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

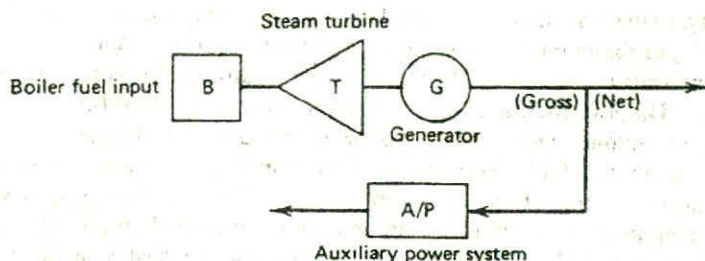


FIG. 2.1 Boiler-turbine-generator unit.

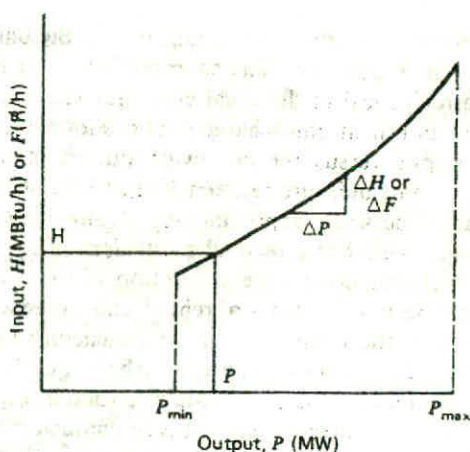


FIG. 2.2 Input-output curve of a steam turbine generator.

total cost per hour (R per hour). 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

capability, so that 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 2.3. 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 per kilowatt hour (or R per kilowatt hour) 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 R per 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 2.4. This characteristic is H/P versus P . It is proportional to the reciprocal of the usual efficiency characteristic developed for 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 that their unit heat rates range between approximately 11,400 Btu/kWh and 9800 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 2.2. The data obtained from heat rate tests or from the plant design engineers may be fitted by a polynomial curve. In many cases, quadratic

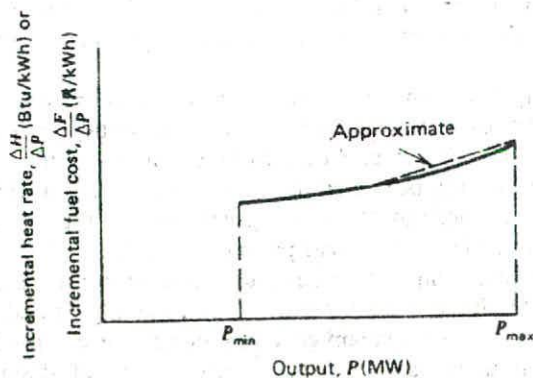


FIG. 2.3 Incremental heat (cost) rate characteristic.

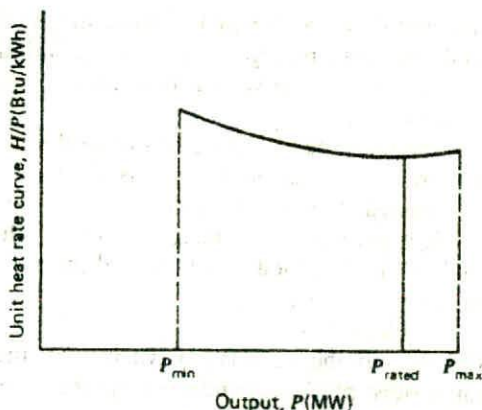


FIG. 2.4 Net heat rate characteristic of a steam turbine generator unit.

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 2.5 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 2.5 show a stepped incremental characteristic at 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

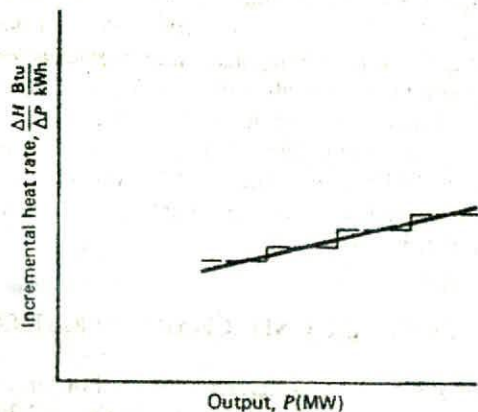


FIG. 2.5 Approximate representations of the incremental heat rate curve.

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/dP^2 equals zero if dF/dP is constant.)

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.

1. 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.
2. 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 in 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: Chapter 1, reference 12.)

A uniform standard must be adopted so that everyone uses the same heating value standard. In the USA, 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!

2.2 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 2.2 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 2.6 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 2.6. It is possible to use this 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 2.7 is a sketch of a rather complex

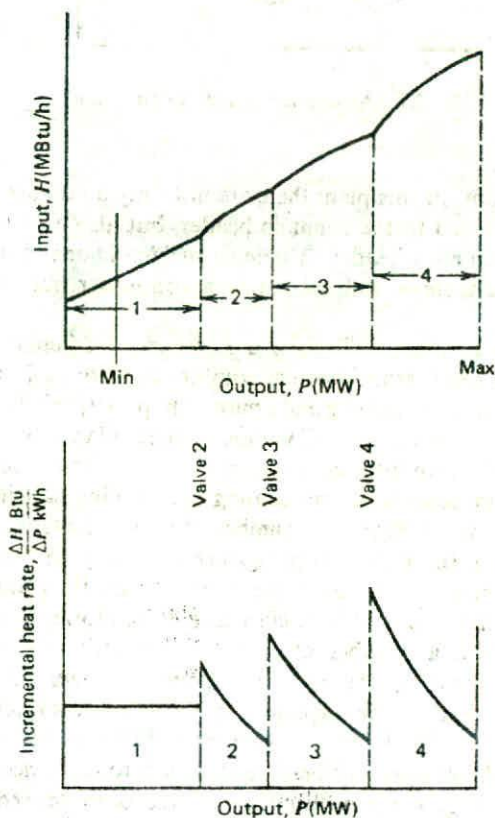


FIG. 2.6 Characteristics of a steam turbine generator with four steam admission valves.

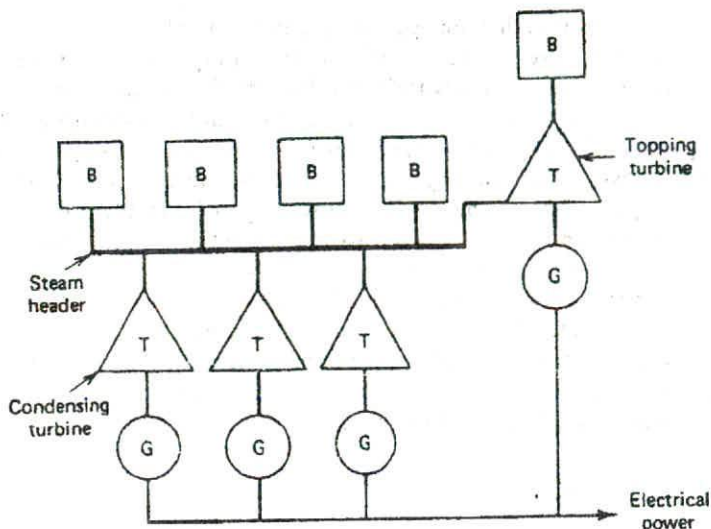


FIG. 2.7 A common-header steam plant.

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 (Chapter 1, reference 1) treat this type of plant quite extensively. Common-header plants were constructed originally not only to provide a large 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 2.7. 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 to 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 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 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. Net full-load heat rates were typically 13,600 Btu/kWh (HHV). In the 1980s and 1990s, new, large, simple cycle units with much improved heat rates were used for power generation. Figure 2.8 shows the approximate, reported range of heat rates

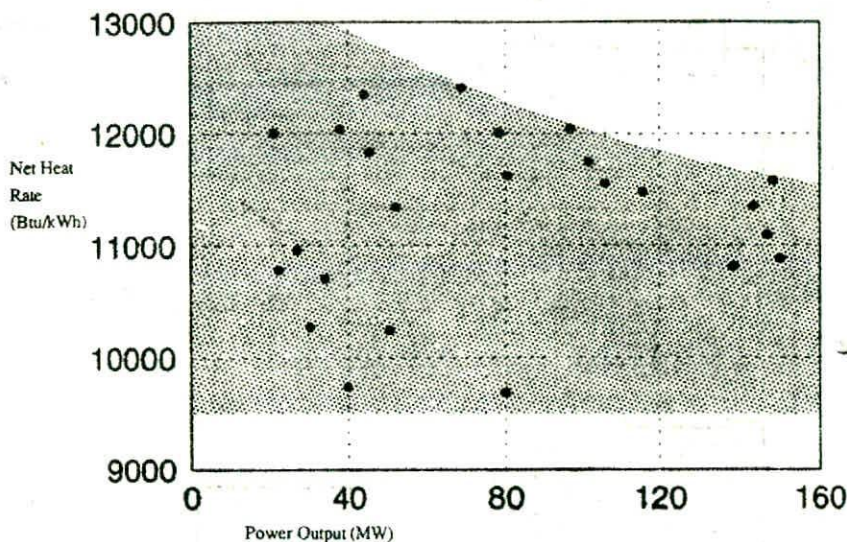


FIG. 2.8 Approximate net heat rates for a range of simple cycle gas turbine units. Units are fired by natural gas and represent performance at standard conditions of an ambient temperature of 15°C at sea level. (Heat rate data from reference 1 were adjusted by 13% to represent HHVs and auxiliary power needs.)

16 CHARACTERISTICS OF POWER GENERATION UNITS

for simple cycle units. These data were taken from a 1990 publication (reference 1) 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.

Combined cycle plants use the high-temperature exhaust gases from one or more gas turbines to generate steam in heat-recovery steam generators (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 2). A 50% efficiency would correspond to a net heat rate of 6825 Btu/kWh. Performance data vary with specific 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

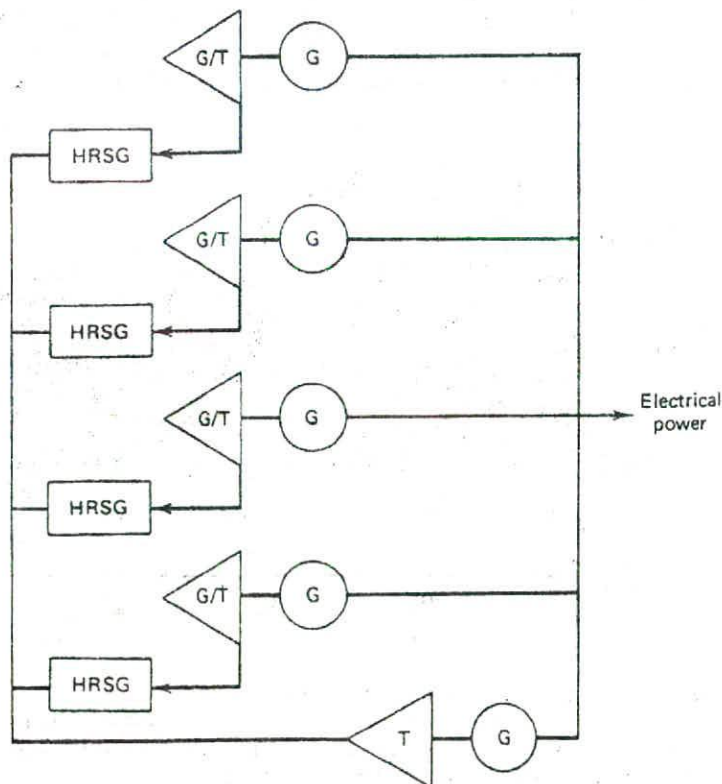


FIG. 2.9 A combined cycle plant with four gas turbines and a steam turbine generator.

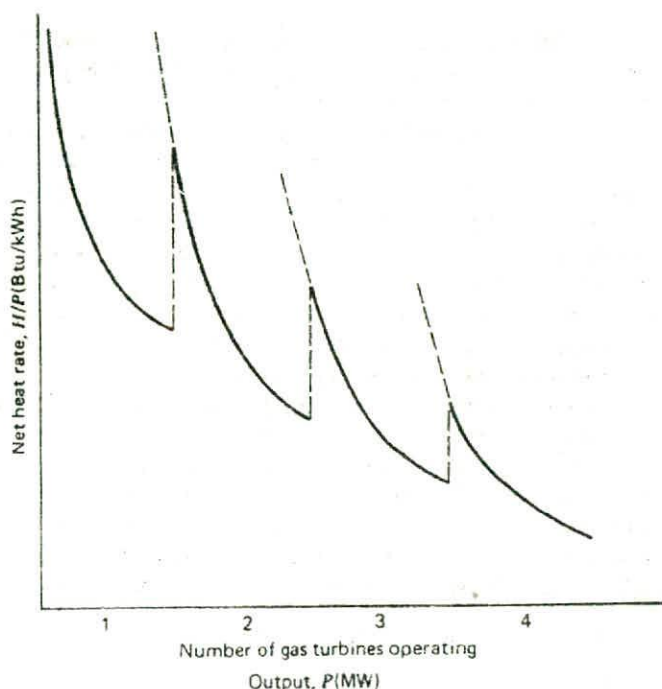


FIG. 2.10 Combined cycle plant heat rate characteristic.

from available information. Figure 2.9 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 2.10 illustrates this. Incremental heat rate characteristics tend to be flatter than those normally seen for steam turbine units.

2.3 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 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. 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 an excess electric energy to the utility. Figure 2.11 shows the input-output characteristics for a 50-MW single extraction unit. The data show the heat

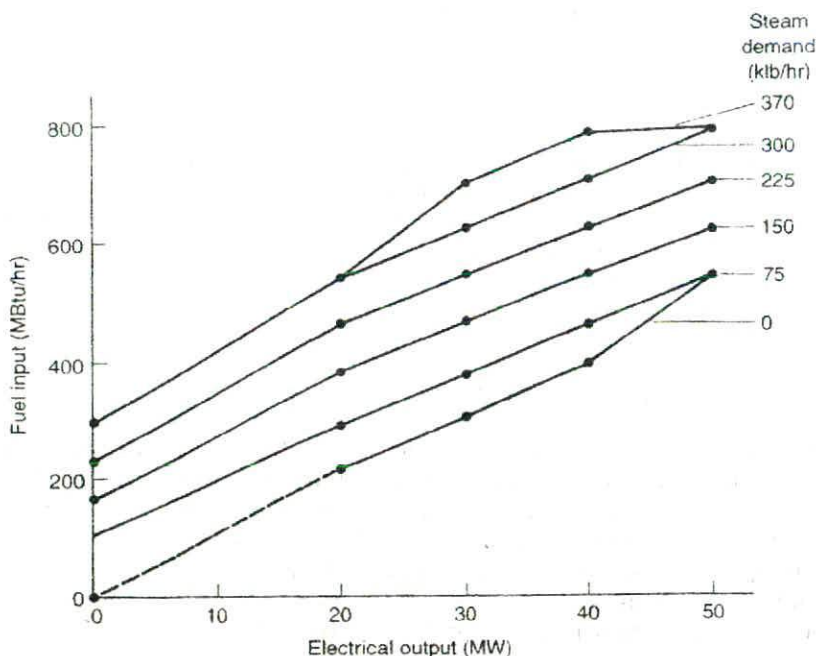


FIG. 2.11 Fuel input required for steam demand and electrical output for a single extraction steam turbine generator.

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 lbs/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 nonutility-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.

2.4 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.

In the remainder of this text, 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 nuclear engineering, as well as detailed understanding of the fuel cycle and its economic aspects (see Chapter 1, reference 10).

2.5 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 2.12 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 2.13. 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/sec) for the water rate.

Figure 2.14 shows the input-output characteristics of a hydroelectric plant

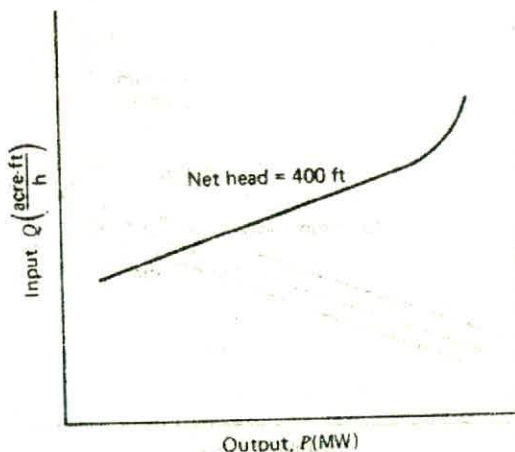


FIG. 2.12 Hydroelectric unit input-output curve.

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 2.14, the volume of water required for a given power output decreases as the head increases. (That is, $\partial Q/\partial \text{head}$ or $\partial Q/\partial \text{volume}$ are negative for a fixed power.) In a later section, methods are discussed that have been proposed

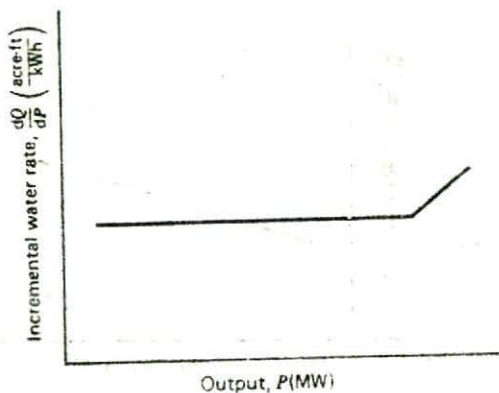


FIG. 2.13 Incremental water rate curve for hydroelectric plant.

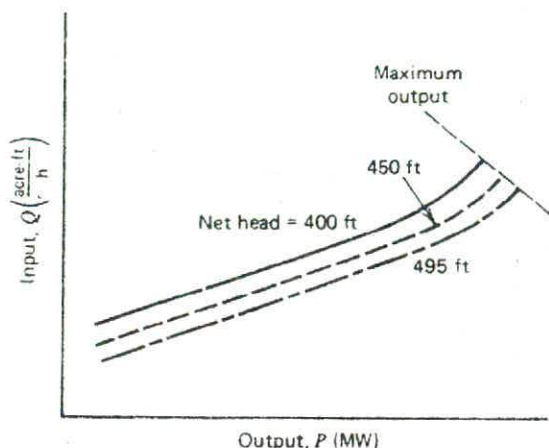


FIG. 2.14 Input-output curves for hydroelectric plant with a variable head.

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

Figure 2.15 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,

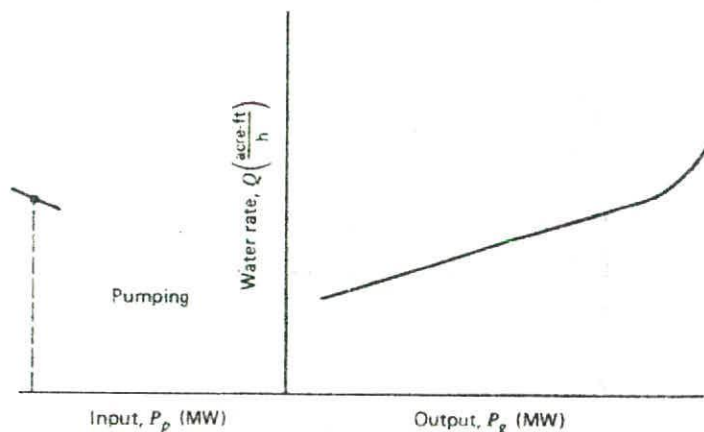


FIG. 2.15 Input-output characteristics for a pumped storage hydroplant with a fixed net hydraulic head.

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 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 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.

APPENDIX

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

TABLE 2.1 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	11000	11088	11429	12166	13409 ^a
Steam--oil	50	11500	11592	11949	12719	14019 ^a
Steam--gas	50	11700	11794	12156	12940	14262 ^a
Steam--coal	200	9500	9576	9871	10507	11581 ^a
Steam--oil	200	9900	9979	10286	10949	12068 ^a
Steam--gas	200	10050	10130	10442	11115	12251 ^a
Steam--coal	400	9000	9045	9252	9783	10674 ^a
Steam--oil	400	9400	9447	9663	10218	11148 ^a
Steam--gas	400	9500	9548	9766	10327	11267 ^a
Steam--coil	600	8900	8989	9265	9843	10814 ^a
Steam--oil	600	9300	9393	9681	10286	11300 ^a
Steam--gas	600	9400	9494	9785	10396	11421 ^a
Steam--coal	800-1200	8750	8803	9048	9625 ^a	
Steam--oil	800-1200	9100	9155	9409	10010 ^a	
Steam--gas	800-1200	9200	9255	9513	10120 ^a	

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

TABLE 2.2 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

technology encouraged the installation of reheat units having steam temperatures of 1000°F and pressures in the range of 1450 to 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 2.1. 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 curves 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 2.1 is summarized in Table 2.2 for the various types and sizes of fossil units.

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

Output (%)	Net Heat Rate (Btu/kWh)
100	10400
75	10442
50	10951

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 above. 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.

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 which 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 equipments. 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

TABLE 2.3 Typical Maintenance and Forced Outage Data

Unit Type	Size Range (MW)	Scheduled Maintenance Requirement (days/yr)	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.

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

1. Analytical (statistical and probability) methods and models,
2. Performance measures and acceptable standards,
3. 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 2.3 contains typical generating unit data on scheduled maintenance requirements, the "equivalent forced outage rate" and the "availability factor" that were taken from a 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:

$$\text{Equivalent forced outage rate} = \frac{\text{forced outage hours} + \text{equivalent forced derated hours}}{\text{forced outage hours} + \text{hours in service} + \text{equivalent forced derated hours during reserve shutdown}}$$

$$\text{Availability factor (AF)} = \frac{\text{available hours}}{\text{period hours}}$$

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

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 above, 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 above 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:

$$\text{EFOR} = \frac{\text{forced outage hours}}{\text{forced outage hours} + \text{hours in service}}$$

or

$$\text{EFOR}' = \frac{\text{forced outage hours}}{\text{forced outage hours} + \text{available hours}}$$

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:

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

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.

REFERENCES

1. 1990 Performance Specs, *Gas Turbine World*, Oct. 1990, Vol. 11, Pequot Publications, Inc., Fairfield, CT.
2. Foster-Pegg, R. W., *Cogeneration—Interactions of Gas Turbine, Boiler and Steam Turbine*, ASMS paper 84-JPGC-GT-12, 1984 Joint Power Generation Conference.

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* 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,

$$\begin{aligned} F_T &= F_1 + F_2 + F_3 + \cdots + F_N \\ &= \sum_{i=1}^N F_i(P_i) \end{aligned} \quad (3.1)$$

$$\phi = 0 = P_{load} - \sum_{i=1}^N P_i \quad (3.2)$$

* Generating units consume fuel at a specific rate (e.g., MBtu/h), which as noted in Chapter 2 can be converted to R/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 R/h .

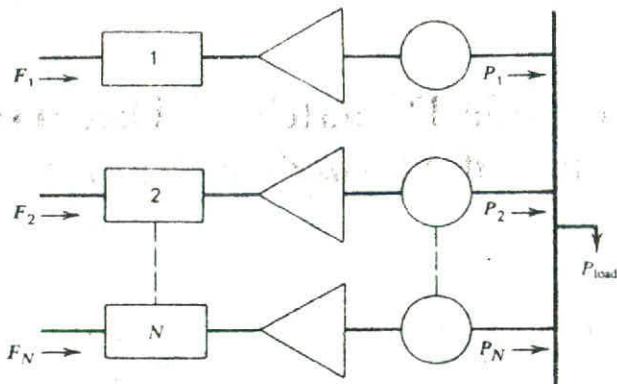


FIG. 3.1 N thermal units committed to serve a load of P_{load} .

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 Eq. 3.3.

$$\mathcal{L} = F_T + \lambda \phi \quad (3.3)$$

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 zero. 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 Eq. 3.4.

$$\frac{\partial \mathcal{L}}{\partial P_i} = \frac{dF_i(P_i)}{dP_i} - \lambda = 0$$

or

$$0 = \frac{dF_i}{dP_i} - \lambda \quad (3.4)$$

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 Eq. 3.5.

$$\begin{array}{ll}
 \frac{dF_i}{dP_i} = \lambda & N \text{ equations} \\
 P_{i,\min} \leq P_i \leq P_{i,\max} & 2N \text{ inequalities} \\
 \sum_{i=1}^N P_i = P_{\text{load}} & 1 \text{ constraint}
 \end{array} \quad (3.5)$$

When we recognize the inequality constraints, then the necessary conditions may be expanded slightly as shown in the set of equations making up Eq. 3.6.

$$\begin{array}{ll}
 \frac{dF_i}{dP_i} = \lambda & \text{for } P_{i,\min} < P_i < P_{i,\max} \\
 \frac{dF_i}{dP_i} \leq \lambda & \text{for } P_i = P_{i,\max} \\
 \frac{dF_i}{dP_i} \geq \lambda & \text{for } P_i = P_{i,\min}
 \end{array} \quad (3.6)$$

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

Input-output curve:

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

Unit 3: Oil-fired steam unit: Max output = 200 MW
Min output = 50 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.

Unit 1: fuel cost = 1.1 R/MBtu

Unit 2: fuel cost = 1.0 R/MBtu

Unit 3: fuel cost = 1.0 R/MBtu

Then

$$F_1(P_1) = H_1(P_1) \times 1.1 = 561 + 7.92P_1 + 0.001562P_1^2 \text{ R/h}$$

$$F_2(P_2) = H_2(P_2) \times 1.0 = 310 + 7.85P_2 + 0.00194P_2^2 \text{ R/h}$$

$$F_3(P_3) = H_3(P_3) \times 1.0 = 78 + 7.97P_3 + 0.00482P_3^2 \text{ R/h}$$

Using Eq. 3.5, the conditions for an optimum dispatch are

$$\frac{dF_1}{dP_1} = 7.92 + 0.003124P_1 = \lambda$$

$$\frac{dF_2}{dP_2} = 7.85 + 0.00388P_2 = \lambda$$

$$\frac{dF_3}{dP_3} = 7.97 + 0.00964P_3 = \lambda$$

and

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

Solving for λ , one obtains

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

and then solving for P_1 , P_2 , and P_3 ,

$$P_1 = 393.2 \text{ MW}$$

$$P_2 = 334.6 \text{ MW}$$

$$P_3 = 122.2 \text{ MW}$$

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 R/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

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

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

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

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

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

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

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.

$$P_1 = 600 \text{ MW}$$

$$\left. \frac{dF_2}{dP_2} \right|_{P_2} = 7.85 + 0.00388P_2 = \lambda$$

$$\left. \frac{dF_3}{dP_3} \right|_{P_3} = 7.97 + 0.00964P_3 = \lambda$$

$$P_2 + P_3 = 850 - P_1 = 250 \text{ MW}$$

which results in

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

and

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

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

Note that this dispatch meets the conditions of Eq. 3.6 since

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

which is less than λ , while

$$\left. \frac{dF_2}{dP_2} \right|_{P_2} \text{ and } \left. \frac{dF_3}{dP_3} \right|_{P_3}$$

both equal λ .

3.2 THERMAL SYSTEM DISPATCHING WITH NETWORK LOSSES CONSIDERED

Figure 3.2 shows symbolically an all-thermal power generation system connected to an equivalent load bus through a transmission network. The economic-dispatching 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 Eq. 3.1. However, the constraint equation previously shown in Eq. 3.2 must now be expanded to the one shown in Eq. 3.7.

$$P_{\text{load}} + P_{\text{loss}} - \sum_{i=1}^N P_i = \phi = 0 \quad (3.7)$$

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 Eq. 3.8. 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 values of P_i results in Eq. 3.9. There are N equations of this type to be satisfied along with the constraint equation shown in Eq. 3.7. This collection, Eq. 3.9 plus Eq. 3.7, is known collectively as the *coordination equations*.

$$\mathcal{L} = F_T + \lambda \phi \quad (3.8)$$

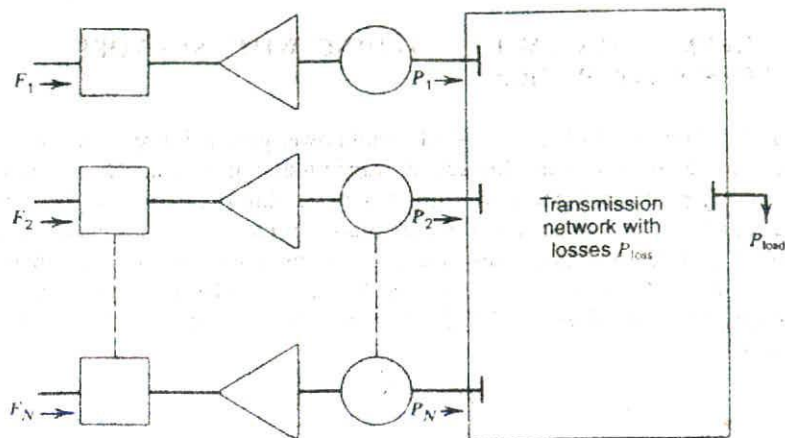
$$\frac{\partial \mathcal{L}}{\partial P_i} = \frac{dF_i}{dP_i} - \lambda \left(1 - \frac{\partial P_{\text{loss}}}{\partial P_i} \right) = 0 \quad (3.9)$$

or

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

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

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, reference 2). The other

FIG. 3.2 N thermal units serving load through transmission network.

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 3C

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 Eqs. 3.8 and 3.9,

$$\frac{dF_1}{dP_1} = \lambda \left(1 - \frac{\partial P_{\text{loss}}}{\partial P_1} \right)$$

becomes

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

Similarly for P_2 and P_3 ,

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

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

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:

$$7.92 + 0.003124P_1 = \lambda(1 - 0.0240) = \lambda(0.9760)$$

$$7.85 + 0.00388P_2 = \lambda(1 - 0.0540) = \lambda(0.9460)$$

$$7.97 + 0.00964P_3 = \lambda(1 - 0.0360) = \lambda(0.9640)$$

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 are solved for

$$7.92 + 0.003124P_1 = \lambda(1 - 0.0264) = \lambda(0.9736)$$

$$7.85 + 0.00388P_2 = \lambda(1 - 0.0538) = \lambda(0.9462)$$

$$7.97 + 0.00964P_3 = \lambda(1 - 0.0301) = \lambda(0.9699)$$

$$P_1 + P_2 + P_3 - 850 - 15.78 = P_1 + P_2 + P_3 - 865.78 = 0$$

resulting in $\lambda = 9.5275 \text{ ₹/MWh}$ and

$$P_1 = 433.94 \text{ MW}$$

$$P_2 = 300.11 \text{ MW}$$

$$P_3 = 131.74 \text{ MW}$$

⋮

Table 3.1 summarizes the iterative process used to solve this problem.

TABLE 3.1 Iterative Process Used to Solve Example 3

Iteration	P_1 (MW)	P_2 (MW)	P_3 (MW)	Losses (MW)	λ (R/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	434.13	299.99	130.71	15.83	9.5284

3.3 THE LAMBDA-ITERATION METHOD

Figure 3.3 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

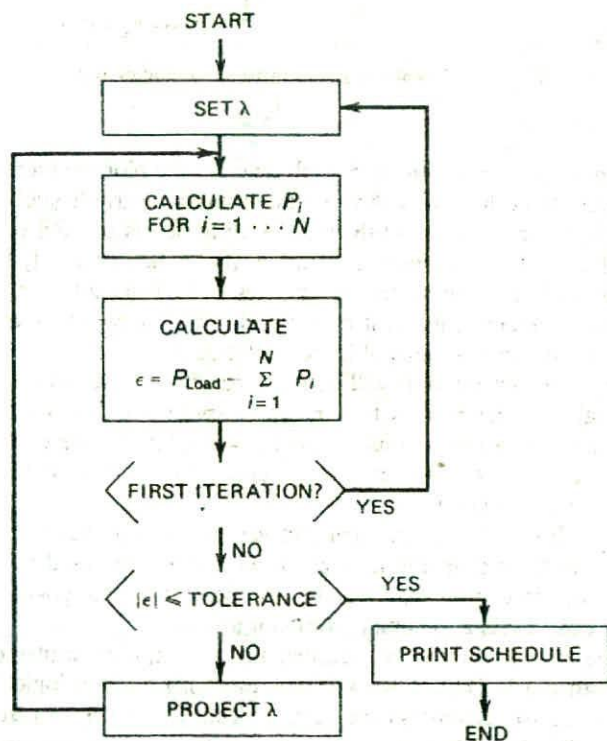


FIG. 3.3 Economic dispatch by the lambda-iteration method.

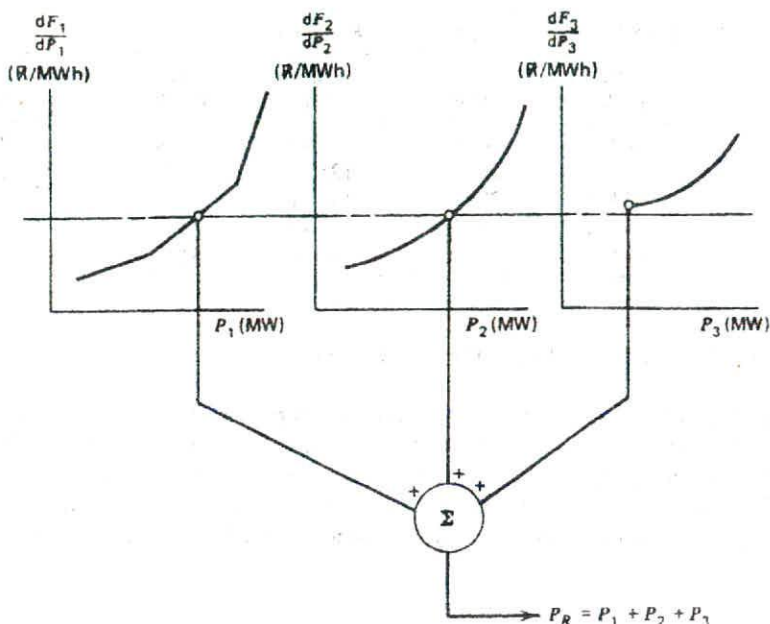


FIG. 3.4 Graphical solution to economic dispatch.

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 3.4. 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 3.5).

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 3.3. 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

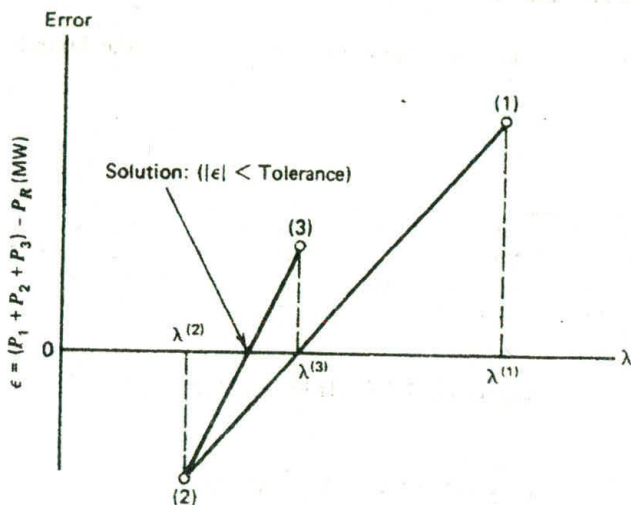


FIG. 3.5 Lambda projections.

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 3.3 and is essentially a rule based on finding the proper operating point within a specified tolerance. The other, not shown in Figure 3.3, 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 3.3, 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 zero.

EXAMPLE 3D

Assume that one wishes to use cubic functions to represent the input-output characteristics of generating plants as follows.

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

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 R/MBtu for each unit and unit limits as follows.

$$320 \text{ MW} \leq P_1 \leq 800 \text{ MW}$$

$$300 \text{ MW} \leq P_2 \leq 1200 \text{ MW}$$

$$275 \text{ MW} \leq P_3 \leq 1100 \text{ MW}$$

Two sample calculations are shown, both using the flowchart in Figure 3.3. 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 3.5.

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

$$P_{\text{load}} = 2500 \text{ MW}$$

$$\lambda_{\text{start}} = 8.0 \text{ R/MWh}$$

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

$$P_{\text{load}} = 2500 \text{ MW}$$

$$\lambda_{\text{start}} = 10.0 \text{ R/MWh}$$

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

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

3.4 GRADIENT METHODS OF ECONOMIC DISPATCH

Note that the lambda search technique always requires that one be able to find the power output of a generator, given an incremental cost for that generator. In the case of a quadratic function for the cost function, or in the case where the incremental cost function is represented by a piecewise linear function, this is possible. However, it is often the case that the cost function is much more complex, such as the one below:

$$F(P) = A + BP + CP^2 + D \exp\left[\frac{(P - E)}{F}\right]$$

In this case, we shall propose that a more basic method of solution for the optimum be found.

3.4.1 Gradient Search

This method works on the principle that the minimum of a function, $f(x)$, can be found by a series of steps that always take us in a downward direction. From any starting point, x^0 , we may find the direction of "steepest descent" by noting that the gradient of f , i.e.,

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} \quad (3.10)$$

always points in the direction of maximum ascent. Therefore, if we want to move in the direction of maximum descent, we negate the gradient. Then we should go from x^0 to x^1 using:

$$x^1 = x^0 - \nabla f \alpha \quad (3.11)$$

Where α is a scalar to allow us to guarantee that the process converges. The best value of α must be determined by experiment.

3.4.2 Economic Dispatch by Gradient Search

In the case of power system economic dispatch this becomes:

$$f = \sum_{i=1}^N F_i(P_i) \quad (3.12)$$

and the object is to drive the function to its minimum. However, we have to be concerned with the constraint function:

$$\Phi = \left(P_{\text{load}} - \sum_{i=1}^N P_i \right) \quad (3.13)$$

To solve the economic dispatch problem which involves minimizing the objective function and keeping the equality constraint, we must apply the gradient technique directly to the Lagrange function itself.

The Lagrange function is:

$$\mathcal{L} = \sum_{i=1}^N F_i(P_i) + \lambda \left(P_{\text{load}} - \sum_{i=1}^N P_i \right) \quad (3.14)$$

and the gradient of this function is:

$$\nabla \mathcal{L} = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial P_1} \\ \frac{\partial \mathcal{L}}{\partial P_2} \\ \frac{\partial \mathcal{L}}{\partial P_3} \\ \frac{\partial \mathcal{L}}{\partial \lambda} \end{bmatrix} = \begin{bmatrix} \frac{d}{dP_1} F_1(P_1) - \lambda \\ \frac{d}{dP_2} F_2(P_2) - \lambda \\ \frac{d}{dP_3} F_3(P_3) - \lambda \\ P_{\text{load}} - \sum_{i=1}^N P_i \end{bmatrix} \quad (3.15)$$

The problem with this formulation is the lack of a guarantee that the new points generated each step will lie on the surface Φ . We shall see that this can be overcome by a simple variation of the gradient method.

The economic dispatch algorithm requires a starting λ value and starting values for P_1 , P_2 , and P_3 . The gradient for \mathcal{L} is calculated as above and the new values of λ , P_1 , P_2 , and P_3 , etc., are found from:

$$\mathbf{x}^1 = \mathbf{x}^0 - (\nabla \mathcal{L})\alpha \quad (3.16)$$

where the vector \mathbf{x} is:

$$\mathbf{x} = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ \vdots \\ \lambda \end{bmatrix} \quad (3.17)$$

EXAMPLE 3E

Given the generator cost functions found in Example 3A, solve for the economic dispatch of generation with a total load of 800 MW.

Using $\alpha = 100$ and starting from $P_1^0 = 300$ MW, $P_2^0 = 200$ MW, and $P_3^0 = 300$ MW, we set the initial value of λ equal to the average of the incremental costs of the generators at their starting generation values. That is:

$$\lambda^0 = \frac{1}{3} \sum_{i=1}^3 \left[\frac{d}{dP_i} F_i(P_i^0) \right]$$

This value is 9.4484.

The progress of the gradient search is shown in Table 3.2. The table shows that the iterations have led to no solution at all. Attempts to use this formulation

TABLE 3.2 Economic Dispatch by Gradient Method

Iteration	P_1	P_2	P_3	P_{total}	λ	Cost
1	300	200	300	800	9.4484	7938.0
2	300.59	200.82	298.59	800	9.4484	7935
3	301.18	201.64	297.19	800.0086	9.4484	7932
4	301.76	202.45	295.8	800.025	9.4570	7929.3
5	302.36	203.28	294.43	800.077	9.4826	7926.9
⋮						
10	309.16	211.19	291.65	811.99	16.36	8025.6

will result in difficulty as the gradient cannot guarantee that the adjustment to the generators will result in a schedule that meets the correct total load of 800 MW.

A simple variation of this technique is to realize that one of the generators is always a dependent variable and remove it from the problem. In this case, we pick P_3 and use the following:

$$P_3 = 800 - P_1 - P_2$$

Then the total cost, which is to be minimized, is:

$$\text{Cost} = F_1(P_1) + F_2(P_2) + F_3(P_3) = F_1(P_1) + F_2(P_2) + F_3(800 - P_1 - P_2)$$

Note that this function stands by itself as a function of two variables with no load-generation balance constraint (and no λ). The cost can be minimized by a gradient method and in this case the gradient is:

$$\nabla \text{Cost} = \begin{bmatrix} \frac{d}{dP_1} \text{Cost} \\ \frac{d}{dP_2} \text{Cost} \end{bmatrix} = \begin{bmatrix} \frac{dF_1}{dP_1} - \frac{dF_3}{dP_1} \\ \frac{dF_2}{dP_2} - \frac{dF_3}{dP_2} \end{bmatrix}$$

Note that this gradient goes to the zero vector when the incremental cost at generator 3 is equal to that at generators 1 and 2. The gradient steps are performed in the same manner as previously, where:

$$x^1 = x^0 - \nabla \text{Cost} \times x$$

and

$$x = \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}$$

Each time a gradient step is made, the generation at generator 3 is set to 800 minus the sum of the generation at generators 1 and 2. This method is often called the "reduced gradient" because of the smaller number of variables.

EXAMPLE 3F

Reworking example 3E with the reduced gradient we obtain the results shown in Table 3.3. This solution is much more stable and is converging on the optimum solution.

TABLE 3.3 Reduced Gradient Results ($\alpha = 10$)

Iteration	P_1	P_2	P_3	P_{total}	Cost
1	300	200	300	800	7938.0
2	320.04	222.36	257.59	800	7858.1
3	335.38	239.76	224.85	800	7810.4
4	347.08	253.33	199.58	800	7781.9
5	355.97	263.94	180.07	800	7764.9
...					
10	380.00	304.43	115.56	800	7739.2

3.5 NEWTON'S METHOD

We may wish to go a further step beyond the simple gradient method and try to solve the economic dispatch by observing that the aim is to always drive

$$\nabla \mathcal{L}_x = 0 \quad (3.18)$$

Since this is a vector function, we can formulate the problem as one of finding the correction that exactly drives the gradient to zero (i.e., to a vector, all of whose elements are zero). We know how to find this, however, since we can use Newton's method. Newton's method for a function of more than one variable is developed as follows.

Suppose we wish to drive the function $g(x)$ to zero. The function g is a vector and the unknowns, x , are also vectors. Then, to use Newton's method, we observe:

$$g(x + \Delta x) = g(x) + [g'(x)]\Delta x = 0 \quad (3.19)$$

If we let the function be defined as:

$$g(x) = \begin{bmatrix} g_1(x_1, x_2, x_3) \\ g_2(x_1, x_2, x_3) \\ g_3(x_1, x_2, x_3) \end{bmatrix} \quad (3.20)$$

then

$$g'(x) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \frac{\partial g_1}{\partial x_3} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \frac{\partial g_2}{\partial x_3} \\ \frac{\partial g_3}{\partial x_1} & \frac{\partial g_3}{\partial x_2} & \frac{\partial g_3}{\partial x_3} \end{bmatrix} \quad (3.21)$$

which is the familiar Jacobian matrix. The adjustment at each step is then:

$$\Delta x = -[g'(x)]^{-1}g(x) \quad (3.22)$$

Now, if we let the g function be the gradient vector $\nabla \mathcal{L}_x$ we get:

$$\Delta \mathbf{x} = - \left[\frac{\partial}{\partial \mathbf{x}} \nabla \mathcal{L}_x \right]^{-1} \Delta \mathcal{L} \quad (3.23)$$

For our economic dispatch problem this takes the form:

$$\mathcal{L} = \sum_{i=1}^N F_i(P_i) + \lambda \left(P_{\text{load}} - \sum_{i=1}^N P_i \right) \quad (3.24)$$

and $\nabla \mathcal{L}$ is as it was defined before. The Jacobian matrix now becomes one made up of second derivatives and is called the Hessian matrix:

$$\left[\frac{\partial}{\partial \mathbf{x}} \nabla \mathcal{L}_x \right] = \begin{bmatrix} \frac{d^2 \mathcal{L}}{dx_1^2} & \frac{d^2 \mathcal{L}}{dx_1 dx_2} & \dots \\ \frac{d^2 \mathcal{L}}{dx_2 dx_1} & \dots & \dots \\ \vdots & \vdots & \vdots \\ \frac{d^2 \mathcal{L}}{d\lambda dx_1} & \dots & \dots \end{bmatrix} \quad (3.25)$$

Generally, Newton's method will solve for the correction that is much closer to the minimum generation cost in one step than would the gradient method.

EXAMPLE 3G

In this example we shall use Newton's method to solve the same economic dispatch as used in Examples 3E and 3F.

The gradient is the same as in Example 3E, the Hessian matrix is:

$$[H] = \begin{bmatrix} \frac{d^2 F_1}{dP_1^2} & 0 & 0 & -1 \\ 0 & \frac{d^2 F_2}{dP_2^2} & 0 & -1 \\ 0 & 0 & \frac{d^2 F_3}{dP_3^2} & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}$$

In this example, we shall simply set the initial λ equal to 0, and the initial generation values will be the same as in Example 3E as well. The gradient of the Lagrange function is:

$$\nabla \mathcal{L} = \begin{bmatrix} 8.8572 \\ 8.6260 \\ 10.8620 \\ 0 \end{bmatrix}$$

The Hessian matrix is:

$$[H] = \begin{bmatrix} 0.0031 & 0 & 0 & -1 \\ 0 & 0.0039 & 0 & -1 \\ 0 & 0 & 0.0096 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}$$

Solving for the correction to the x vector and making the correction, we obtain

$$x = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ \lambda \end{bmatrix} = \begin{bmatrix} 369.6871 \\ 315.6965 \\ 114.6164 \\ 9.0749 \end{bmatrix}$$

and a total generation cost of 7738.8. Note that no further steps are necessary as the Newton's method has solved in one step. When the system of equations making up the generation cost functions are quadratic, and no generation limits are reached, the Newton's method will solve in one step.

We have introduced the gradient, reduced gradient and Newton's method here mainly as a way to show the variations of solution of the generation economic dispatch problem. For many applications, the lambda search technique is the preferred choice. However, in later chapters, when we introduce the optimal power flow, the gradient and Newton formulations become necessary.

3.6 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 3.6 are representative of such functions. Note that were we 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

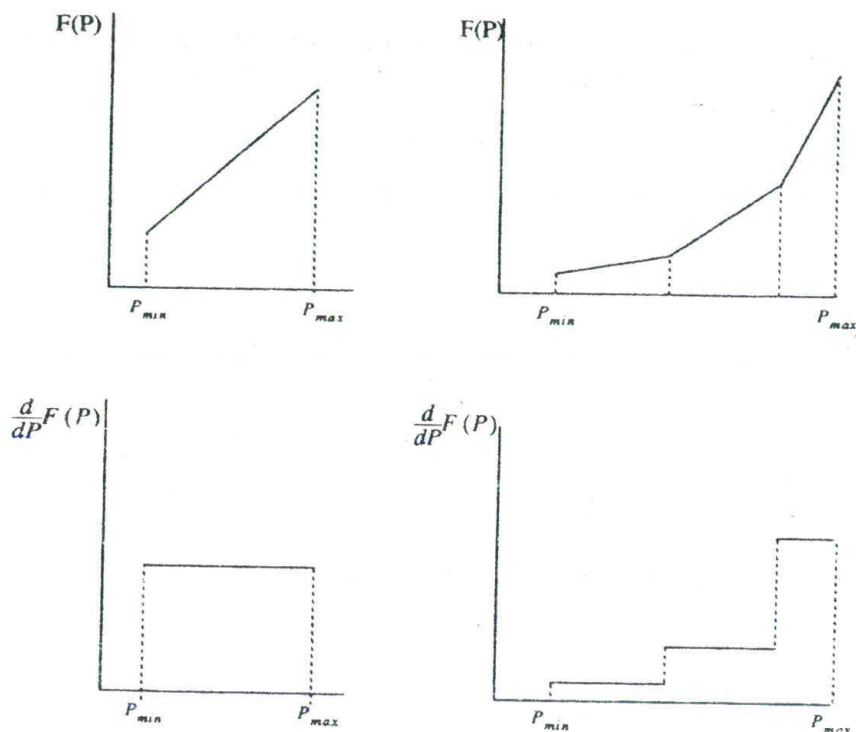


FIG. 3.6 Piecewise linear cost functions.

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 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 which is partially loaded for one segment. Note, that if there are two units with exactly the same incremental cost, we simply load them equally.

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.

3.7 ECONOMIC DISPATCH USING DYNAMIC PROGRAMMING

As we saw in Chapter 2 when we considered the valve points in the input-output curve (for example, Figure 2.6), 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 the 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 which uses dynamic programming (DP). If the reader has no background in DP, Appendix 3B of this chapter should be read at this time.

The dynamic programming solution to economic dispatch is done as an allocation problem, as given in Appendix 3B. 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 3H

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) $P_1 = P_2 = P_3$	Costs (₹/hour)		
	F_1	F_2	F_3
0	∞	∞	∞
50	810	750	806
75	1355	1155	1108.5
100	1460	1360	1411
125	1772.5	1655	11704.5
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

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

52 ECONOMIC DISPATCH OF THERMAL UNITS

over the allowable range of P_2 and for $100 \leq D \leq 350$ MW. The search data are given in the table below. 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.

D (MW)	$P_2 = 0$ $F_2(P_2) = \infty$ $F_1(D)$ (R/h)							f_2 (R/h)	P_2^* (MW)
	0	50	75	100	125	150 (MW)	1950 (R/h)		
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 $50 \leq P_3 \leq 175$ 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.

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

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

One problem that is common to economic dispatch with dynamic programming is the poor control performance of the generators. We shall deal with the control of generators in Chapter 9 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 dynamic programming and the cost curves are nonconvex, 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 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 $t = 1 \dots t_{\max}$, with load levels of P_{load}^t , and N generators on-line to supply the load:

$$\sum_{i=1}^N P_i^t = P_{\text{load}}^t \quad (3.26)$$

Each unit must obey a rate limit such that:

$$P_i^{t+1} = P_i^t + \Delta P_i \quad (3.27)$$

and

$$-\Delta P_i^{\max} \leq \Delta P_i \leq \Delta P_i^{\max} \quad (3.28)$$

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

$$F^{\text{total}} = \sum_{t=1}^{t_{\max}} \sum_{i=1}^N F_i(P_i^t) \quad (3.29)$$

subject to:

$$\sum_{i=1}^N P_i^t = P_{\text{load}}^t \quad \text{for } t = 1 \dots t_{\max} \quad (3.30)$$

and

$$P_i^{t+1} = P_i^t + \Delta P_i \quad (3.31)$$

with

$$-\Delta P_i^{\max} \leq \Delta P_i \leq \Delta P_i^{\max} \quad (3.32)$$

This optimization problem can be solved with dynamic programming and the "control performance" of the dispatch will be considerably better than that using dynamic programming and no ramp limit constraints (see Chapter 9, reference 19).

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 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 i^{th} unit is given in Figure 3.7. 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,

$$\Delta\lambda_i = \Delta\lambda \cong F_i''(\lambda^0)\Delta P_i \quad (3.33)$$

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,

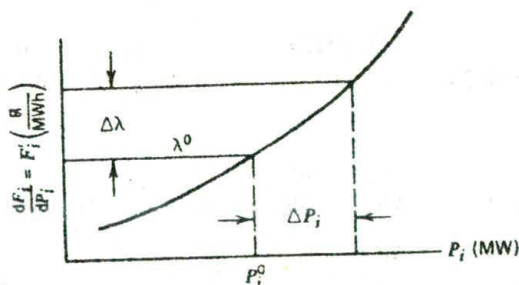


FIG. 3.7 Relationship of $\Delta\lambda$ and ΔP_i .

the sum of the individual unit changes. Let P_D be the total demand on the generators (where $P_D = P_{\text{load}} + P_{\text{loss}}$), then

$$\begin{aligned}\Delta P_D &= \Delta P_1 + \Delta P_2 + \cdots + \Delta P_N \\ &= \Delta \lambda \sum_i \left(\frac{1}{F_i''} \right)\end{aligned}\quad (3.34)$$

The earlier equation, 3.33, can be used to find the *participation factor* for each unit as follows.

$$\left(\frac{\Delta P_i}{\Delta P_D} \right) = \frac{(1/F_i'')}{\sum_i \left(\frac{1}{F_i''} \right)} \quad (3.35)$$

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 break points on piecewise linear incremental cost functions, or have nonconvex cost curves.

EXAMPLE 3I

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 Eq. 3.24,

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

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

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

Then for each unit

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

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

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

3.9 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 *assumes* that there are N units already connected to the system. The purpose of the economic dispatch problem is to find the optimum operating policy for these N units. This is the problem that we have been investigating so far in this text.

On the other hand, the unit commitment problem is more complex. We may assume that we have N 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 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?)

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 zero. 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.

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

This is shown in Figure 3.8 for various values of f .

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

$$\begin{array}{ll} \text{Minimize:} & f(x_1, x_2) = 0.25x_1^2 + x_2^2 \\ \text{Subject to the constraint:} & \omega(x_1, x_2) = 0 \\ \text{Where:} & \omega(x_1, x_2) = 5 - x_1 - x_2 \end{array} \quad (3A.2)$$

This optimization problem can be pictured as in Figure 3.9.

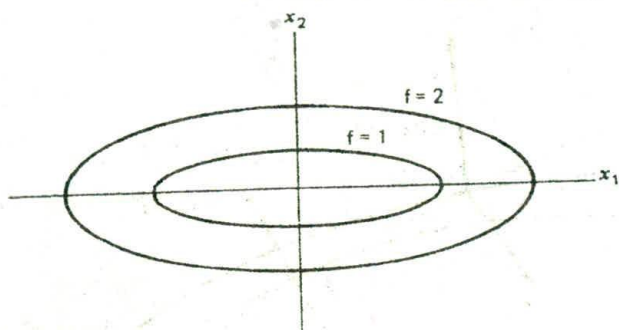


FIG. 3.8 Elliptical objective function.

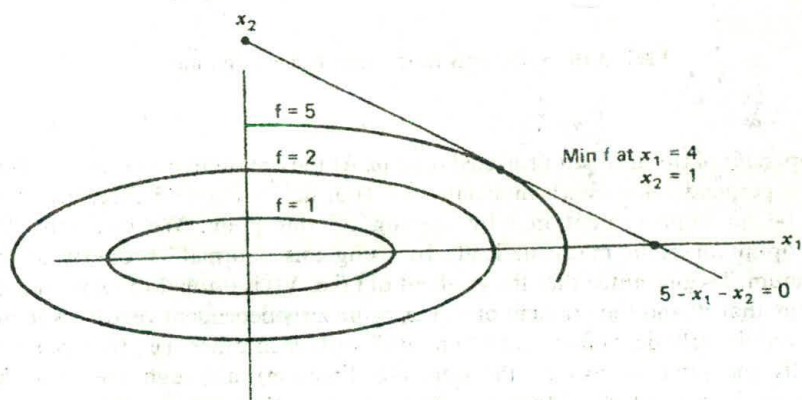


FIG. 3.9 Elliptical objective function with equality constraint.

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 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 3.10 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

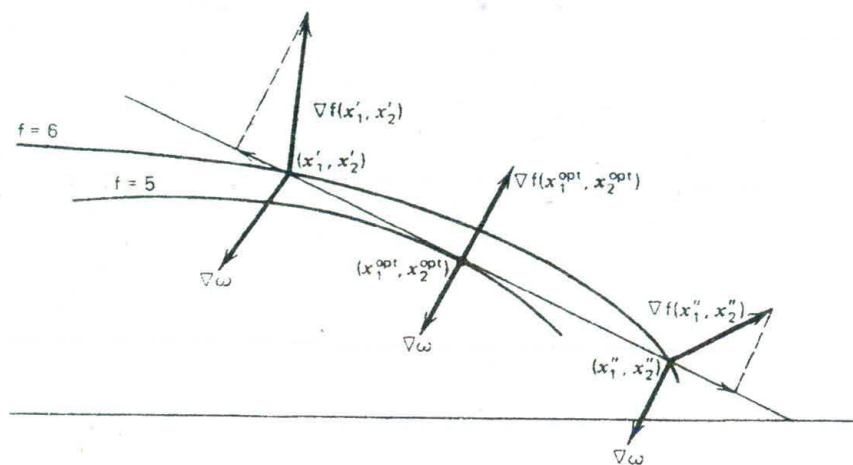


FIG. 3.10 Gradients near a constrained optimum.

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 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.

$$\nabla f + \lambda \nabla\omega = 0 \quad (3A.3)$$

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 Eq. 3A.3, we will restate them as

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

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 Eq. 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

zero. That is,

At the optimum:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial x_1} &= 0 \\ \frac{\partial \mathcal{L}}{\partial x_2} &= 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= 0\end{aligned}\quad (3A.5)$$

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

$$\begin{aligned}\mathcal{L}(x_1, x_2, \lambda) &= 0.25x_1^2 + x_2^2 + \lambda(5 - x_1 - x_2) \\ \frac{\partial \mathcal{L}}{\partial x_1} &= 0.5x_1 - \lambda = 0 \\ \frac{\partial \mathcal{L}}{\partial x_2} &= 2x_2 - \lambda = 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= 5 - x_1 - x_2 = 0\end{aligned}\quad (3A.6)$$

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

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

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.

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

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,

$$\nabla f + \lambda_1 \nabla \omega_1 + \lambda_2 \nabla \omega_2 + \lambda_3 \nabla \omega_3 = 0 \quad (3A.9)$$

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

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

whose optimum occurs at

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

Up until now, we have assumed that all the constraints in the problem were equality constraints; that is, $\omega(x_1, x_2, \dots) = 0$. In general, however, optimization problems involve inequality constraints; that is, $g(x_1, x_2, \dots) \leq 0$, 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 $g(x_1, x_2, \dots) = 0$ at the optimum.

The fundamental rule that tells when the optimum has been reached is presented in a famous paper by Kuhn and Tucker (reference 3). *The Kuhn-Tucker conditions*, as they are called, are presented here.

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

Then, forming the Lagrange function,

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

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

1. $\frac{\partial \mathcal{L}}{\partial x_i}(\mathbf{x}^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0) = 0$ for $i = 1 \dots N$
2. $\omega_i(\mathbf{x}^0) = 0$ for $i = 1 \dots N\omega$

$$\begin{array}{ll}
 3. & g_i(\mathbf{x}^0) \leq 0 \quad \text{for } i = 1 \dots Ng \\
 4. & \left. \begin{array}{l} \mu_i^0 g_i(\mathbf{x}^0) = 0 \\ \mu_i^0 \geq 0 \end{array} \right\} \quad \text{for } i = 1 \dots Ng
 \end{array}$$

The first condition is simply the familiar set of partial derivatives of the Lagrange function that must equal zero 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 zero, either μ_i^0 is equal to zero or $g_i(\mathbf{x}^0)$ is equal to zero, or both are equal to zero. If μ_i^0 is equal to zero, $g_i(\mathbf{x}^0)$ is free to be nonbinding; if μ_i^0 is positive, then $g_i(\mathbf{x}^0)$ must be zero. 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.

$$\begin{array}{ll}
 \text{Minimize:} & f(x_1, x_2) = 0.25x_1^2 + x_2^2 \\
 \text{Subject to:} & \omega(x_1, x_2) = 5 - x_1 - x_2 = 0 \\
 & g(x_1, x_2) = x_1 + 0.2x_2 - 3 \leq 0
 \end{array}$$

which can be illustrated as in Figure 3.11.

First, set up the Lagrange equation for the problem.

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

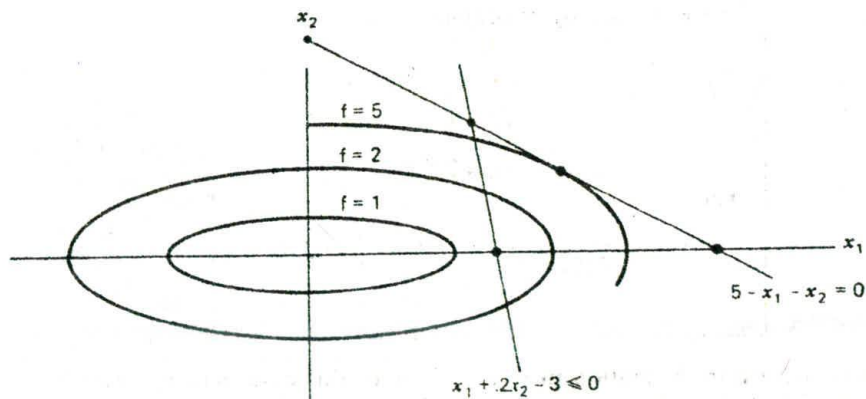


FIG. 3.11 Elliptical objective function with equality and inequality constraints.

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

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

The fourth condition gives

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

$$\mu \geq 0$$

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 zero. 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

$$g(x_1, x_2) \Big|_{\substack{x_1=4 \\ x_2=1}} = 4 + 0.2(1) - 3 = 1.2 \neq 0$$

Now we will try a solution in which $\mu > 0$. In this case, $g(x_1, x_2)$ must be exactly zero 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

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

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 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,

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

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

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

$$C'_1(x_1) - \lambda + \mu_1 - \mu_2 = 0$$

$$C'_2(x_2) - \lambda + \mu_3 - \mu_4 = 0$$

Condition 2 gives

$$L - x_1 - x_2 = 0$$

Condition 3 gives

$$x_1 - x_1^+ \leq 0$$

$$x_1^- - x_1 \leq 0$$

$$x_2 - x_2^+ \leq 0$$

$$x_2^- - x_2 \leq 0$$

Condition 4 gives

$$\mu_1(x_1 - x_1^+) = 0 \quad \mu_1 \geq 0$$

$$\mu_2(x_1^- - x_1) = 0 \quad \mu_2 \geq 0$$

$$\mu_3(x_2 - x_2^+) = 0 \quad \mu_3 \geq 0$$

$$\mu_4(x_2^- - x_2) = 0 \quad \mu_4 \geq 0$$

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 zero 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 zero. 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 zero. 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

$$x_1 - x_1^+ = 0$$

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:

Minimize: $C_1(x_1) + C_2(x_2) + \dots + C_N(x_N)$

Subject to: $L - x_1 - x_2 - \dots - x_N = 0$

$$\text{And: } \left. \begin{array}{l} x_i - x_i^+ \leq 0 \\ x_i^- - x_i \leq 0 \end{array} \right\} \text{ for } i = 1 \dots N$$

Let the optimum lie at $x_i = x_i^{\text{opt}}$ $i = 1 \dots N$ and assume that at least one x_i is not at limit. Then,

$$\text{If } x_i^{\text{opt}} < x_i^+ \text{ and } x_i^{\text{opt}} > x_i^-, \text{ then } C_i(x_i^{\text{opt}}) = \lambda$$

$$\text{If } x_i^{\text{opt}} = x_i^+ \quad C_i'(x_i^{\text{opt}}) \leq \lambda$$

$$\text{If } x_i^{\text{opt}} = x_i^- \quad C_i'(x_i^{\text{opt}}) \geq \lambda$$

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.

$$\text{If: } g(x_1) = x_1 - x_1^+ \leq 0$$

$$\text{Then: } g(x_1, S_1) = x_1 - x_1^+ + S_1^2 = 0$$

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.

$$\text{Minimize: } f(x_1, x_2) = C_1(x_1) + C_2(x_2)$$

$$\text{Subject to: } \omega(x_1, x_2) = L - x_1 - x_2 = 0$$

$$\text{And: } g_1(x_1) = x_1 - x_1^+ \leq 0 \quad \text{or} \quad g_1(x_1, S_1) = x_1 - x_1^+ + S_1^2 = 0$$

$$g_2(x_1) = x_1^- - x_1 \leq 0 \quad g_2(x_1, S_2) = x_1^- - x_1 + S_2^2 = 0$$

$$g_3(x_2) = x_2 - x_2^+ \leq 0 \quad g_3(x_2, S_3) = x_2 - x_2^+ + S_3^2 = 0$$

$$g_4(x_2) = x_2^- - x_2 \leq 0 \quad g_4(x_2, S_4) = x_2^- - x_2 + S_4^2 = 0$$

The resulting Lagrange function is

$$\mathcal{L} = f(x_1, x_2) + \lambda_0 \omega(x_1, x_2) + \lambda_1 g_1(x_1, S_1) \\ + \lambda_2 g_2(x_1, S_2) + \lambda_3 g_3(x_2, S_3) + \lambda_4 g_4(x_2, S_4)$$

Note that all constraints are now equality constraints, so we have used only λ values as Lagrange multipliers.

Condition 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: $L - x_1 - x_2 = 0$

$$(x_1 - x_1^+ + S_1^2) = 0$$

$$(x_1^- - x_1 + S_2^2) = 0$$

$$(x_2 - x_2^+ + S_3^2) = 0$$

$$(x_2^- - x_2 + S_4^2) = 0$$

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 zero). Since there are as many problem variables whether one uses the slack variable form or the inequality constraint form, there is little advantage to either, other than perhaps a conceptual advantage to the student.

Dual Variables

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 just solved 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 Eq. 3A.2 as:

Minimize: $f(x_1, x_2) = 0.25x_1^2 + x_2^2$

Subject to: $w(x_1, x_2) = 5 - x_1 - x_2$

and its Lagrangian function is:

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

If we define a dual function, $q(\lambda)$, as:

$$q(\lambda) = \min_{x_1, x_2} \mathcal{L}(x_1, x_2, \lambda) \quad (3A.12)$$

Then the "dual problem" is to find

$$q^*(\lambda) = \max_{\lambda \geq 0} q(\lambda) \quad (3A.13)$$

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 which maximizes $q(\lambda)$. We then take this value of λ and, holding it constant, we find values of x_1 and x_2 which 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 Eq. 3A.2, we can simplify the procedure above 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; namely, we first optimize on λ and then on the problem variables, and then go back and update λ , etc. Since the dual problem requires that we find

$$q^*(\lambda) = \max_{\lambda \geq 0} q(\lambda)$$

and we do not have an explicit function in λ (as we did above), 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

$$\lambda^1 = \lambda^0 + \left[\frac{d}{d\lambda} q(\lambda) \right] \alpha$$

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:

$$\alpha = 0.5 \quad \text{when} \quad \frac{d}{d\lambda} q(\lambda) \text{ is positive}$$

and

$$\alpha = 0.1 \quad \text{when} \quad \frac{d}{d\lambda} q(\lambda) \text{ is negative}$$

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 Eq. 3A.2 and the optimal value will be called J^* and it is defined as:

$$J^* = \min \mathcal{L} \quad (3A.14)$$

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:

$$\frac{J^* - q^*}{q^*} \quad (3A.15)$$

TABLE 3.4 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

For a convex problem with continuous variables, the duality gap will become zero at the final solution. When we again take up the dual optimization method in Chapter 5, we will be dealing with nonconvex problems with noncontinuous variables and the duality gap will never actually go to zero.

Using the dual optimization approach on the problem given in Eq. 3A.2 and starting at $\lambda = 0$, we obtain the results shown in Table 3.4. 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.

APPENDIX 3B Dynamic-Programming Applications

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 dynamic programming. 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 willingness to express the particular optimization problem in terms appropriate for a dynamic-programming (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 text will touch on all three areas.

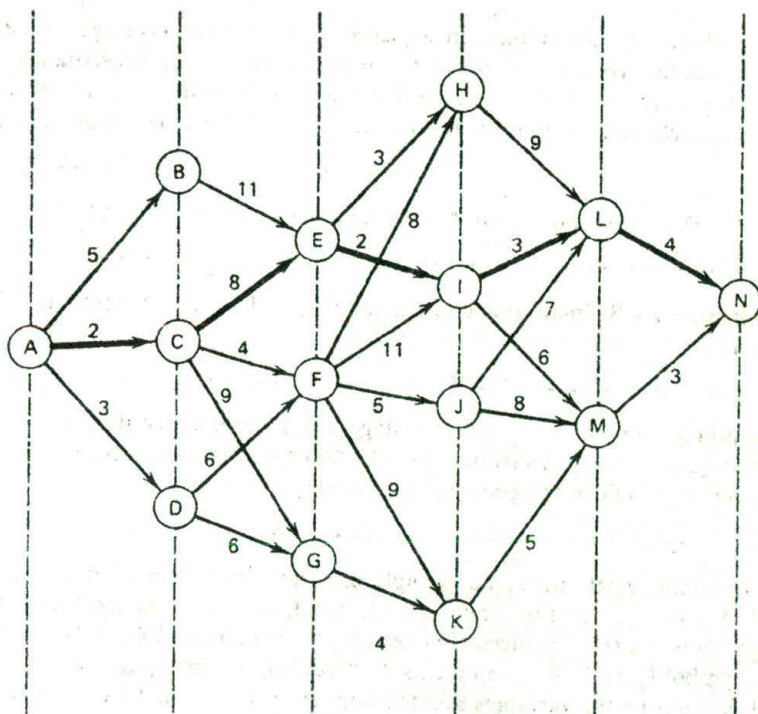


FIG. 3.12 Dynamic-programming example.

First, however, it will be as well to introduce some of the notions of DP by means of some one-dimensional examples. Figure 3.12 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 dynamic programming. The first two examples are from reference 18 and are used by permission.

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.

In reference 20, Bellman and Dreyfus 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 3.13 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 [$\{X_3\} = \{H, I, J, K\}$]. The symbol $V_a(X_i, X_i + 1)$ represents the "cost" of traversing stage $a (= I, \dots, V)$ 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 3.13 represent this minimum cost.

$$\begin{array}{lll} \{X_0\}: A & \{X_2\}: E, F, G & \{X_4\}: L, M \\ \{X_1\}: B, C, D & \{X_3\}: H, I, J, K & \{X_5\}: N \end{array}$$

$f_1(X_i)$: Minimum cost for the first stage is obvious:

$$f_1(B) = V_1(A, B) = 5$$

$$f_1(C) = V_1(A, C) = 2$$

$$f_1(D) = V_1(A, D) = 3$$

$f_{II}(X_2)$: Minimum cost for stages I and II as a function of X_2 :

$$f_{II}(E) = \min_{\{X_1\}} [f_1(X_1) + V_{II}(X_1, E)]$$

$$= \min[5 + 11, 2 + 8, 3 + \infty] = 10$$

$$X_1 = B \quad = C \quad = D \quad X_2 = C$$

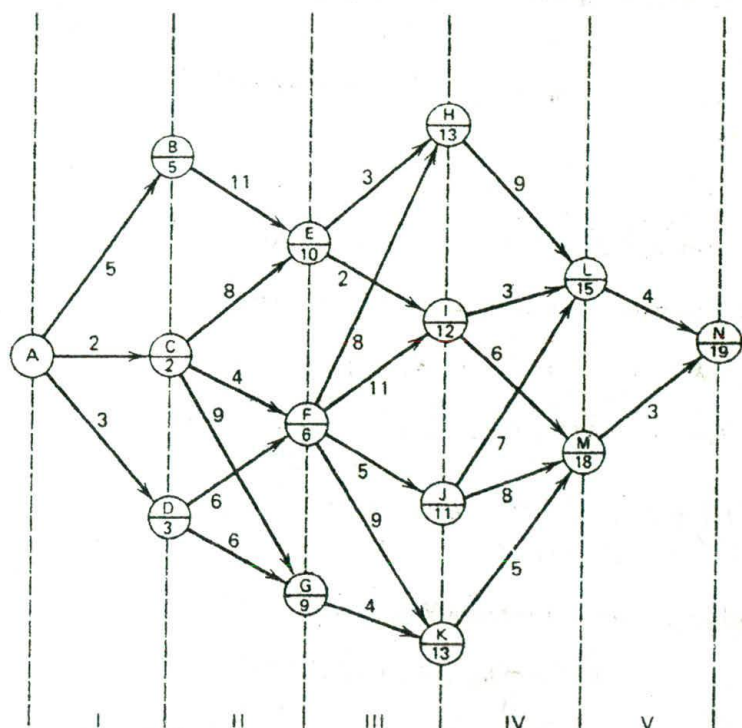


FIG. 3.13 Dynamic-programming example showing minimum cost at each node.

The cost is infinite for node D since there is no path from D to E:

$$f_{II}(F) = \min_{(X_1)} [f_I(X_1) + V_{II}(X_1, F)] = \min[\infty, 6, 9] = 6, X_1 = C$$

$$f_{II}(G) = \min_{(X_1)} [f_I(X_1) + V_{II}(X_1, G)] = \min[\infty, 11, 9] = 9, X_1 = D$$

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.

(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 :

$$f_{III}(H) = \min_{(X_2)} [f_{II}(X_2) + V_{III}(X_2, H)] = \min[13, 14, \infty] = 13 \quad \text{with } X_2 = E$$

In general,

$$f_{III}(X_3) = \min_{(X_2)} [f_{II}(X_2) + V_{III}(X_2, X_3)]$$

Giving,

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 :

$$f_{IV}(X_4) = \min_{(X_3)} [f_{III}(X_3) + V_{IV}(X_3, X_4)]$$

$$f_{IV}(L) = \min[13 + 9, 12 + 3, 11 + 7, 13 + \infty] = 15, \quad X_3 = I$$

$$X_3 = H = I = J = K$$

$$f_{IV}(M) = [13 + \infty, 12 + 6, 11 + 8, 13 + 5] = 18 \quad X_3 = I \text{ or } K$$

$$X_3 = H = I = J = K$$

f_V : Minimum cost of I through V as a function of X_5 :

$$f_V(N) = \min_{(X_4)} [f_{IV}(X_4) + V_V(X_4, X_5)]$$

$$= \min[15 + 4, 18 + 3] = 19 \quad X_4 = L$$

$$X_4 = L = M$$

Tracing back, the path of minimum cost is found as follows:

Stage 1	$\{X_i\}$	f_i
1	B, C, D	5, 2, 3
2	E, F, G	10, 6, 9
3	H, I, J, K	13, 12, 11, 13
4	L, M	15, 18
5	N	19

It would be possible to carry out this procedure in the opposite direction just as easily.

An Allocation Problem

Table 3.5 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

TABLE 3.5 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

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

X_1, X_2, X_3, X_4 be investments in I through IV

$V(X_1), V(X_2), V(X_3), V(X_4)$ be profits

$X_1 + X_2 + X_3 + X_4 = 10$ is the constraint; that is,
10 units must be invested

To transform this into a multistage problem, let the stages be

$$X_1, U_1, U_2, A$$

where

$$U_1 = X_1 + X_2 \quad U_1 \leq A \quad U_2 \leq A$$

$$U_2 = U_1 + X_3 \quad \{A\} = 0, 1, 2, 3, \dots, 10$$

$$A = U_2 + X_4$$

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

$$f_2(U_1) = \max_{X_1=0, 1, \dots, U_1} [V_1(X_1) + V_2(U_1 - X_1)]$$

$X_1, X_2, \text{ or } U_1$	$V_1(X_1)$	$V_2(X_2)$	$f_2(U_1)$	Optimal Subpolicies for I & 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,

$$f_3(U_2) = \max_{U_1=0, 1, 2, \dots, U_2} [f_2(U_1) + V_3(U_2 - U_1)]$$

$U_1, U_2, \text{ or } X_3$	$f_2(U_1)$	$V_3(X_3)$	$f_3(U_2)$	Optimal Subpolicies	
				For I & II	For I, II, & 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

$$f_4(A) = \max_{(U_2)} [f_3(U_2) + V_4(A - U_2)]$$

U_2, A or X_4	$f_3(U_2)$	$V_4(X_4)$	$f_4(A)$	Optimal Subpolicy for I, II, & 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:

1. It was not necessary to enumerate all possible solutions. Instead, we used an orderly, stagewise search, the form of which was the same at each stage.
2. The solution was obtained not only for $A = 10$, but for the complete set of A values $\{A\} = 0, 1, 2, \dots, 10$.
3. The optimal policy contains only optimal subpolicies. For instance, $A = 10$, (4, 3, 1, 2) is the optimal policy. For stages I, II, III, and $U_2 = 8$, (4, 3, 1) is the optimal subpolicy. For stages I and II, and $U_1 = 7$, (4, 3) is the optimal subpolicy. For stage I only, $X_1 = 4$ fixes the policy.
4. 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.

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

- 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.
- Calculate the net heat rate in Btu per kilowatt-hour, and plot against output in megawatts.
- Assume that the cost of fuel is 1.5 R/MBtu. Calculate the incremental production cost in R/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 curves (IHR) specified as three connected line segments (four points as shown in Figure 3.14).

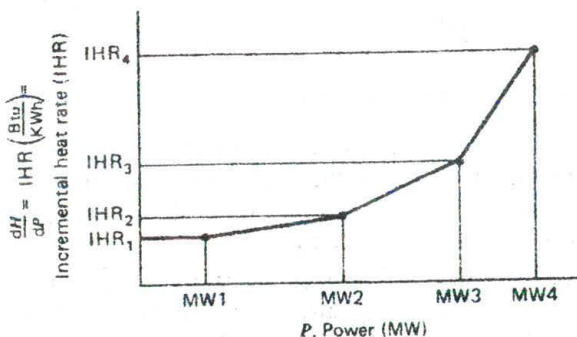


FIG. 3.14 Piecewise linear incremental heat rate curve for Problem 3.2.

Unit 1:

Point	MW	IHR (Btu/kWh)
1	100	7000
2	200	8200
3	300	8900
4	400	11000

Fuel cost for unit 1 = 1.60 R/MBtu

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 R/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 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 5.
- 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.

- 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?
 - What is the economic schedule for the period of time from 0600 to 1800 (load = 150 MW)?
 - What is the economic schedule for the period of time from 1800 to 0600 (load = 50 MW)?
- 3.7 Assume that all three of the thermal units described below 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

- Use the lambda-iteration method to find the economic dispatch for a total demand of 450 MW.
- 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.
- 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_i = heat input in MBtu/h; P_i = electric power output in MW; $i = 1, 2, 3$.)

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 incremental heat rate 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	13512.5
	300	9900.0
	350	9918.0
2	45	22764.5
	200	11465.0
	300	11060.0
3	350	11117.9
	47.5	16039.8
	200	10000.0
	300	9583.3
	450	9513.9

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.

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.

$$H_1 = 400 + 5P_1 + 0.01P_1^2, \quad \text{MBtu/h, } 20 \leq P_1 \leq 200 \text{ MW}$$

$$H_2 = 600 + 4P_2 + 0.015P_2^2, \quad \text{MBtu/h, } 20 \leq P_2 \leq 200 \text{ MW}$$

The units each burn coal with a heat content of 11,500 Btu/lb that costs 13.50 R per ton (2000 lb). The combustion process in each unit results in 11.75% of the coal by weight going up the stack as fly ash.

- Calculate the net heat rates of both units at 200 MW.
 - Calculate the incremental heat rates; schedule each unit for optimum economy to serve a total load of 250 MW with both units on-line.
 - Calculate the cost of supplying that load.
 - 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.
 - 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.
 - Calculate the rate of emission of ash and the cost for this schedule to serve the 250 MW load. What is the cost penalty?
 - 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 3.15 represent relative cost in hundreds of millions R ($R \cdot 10^8$). 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

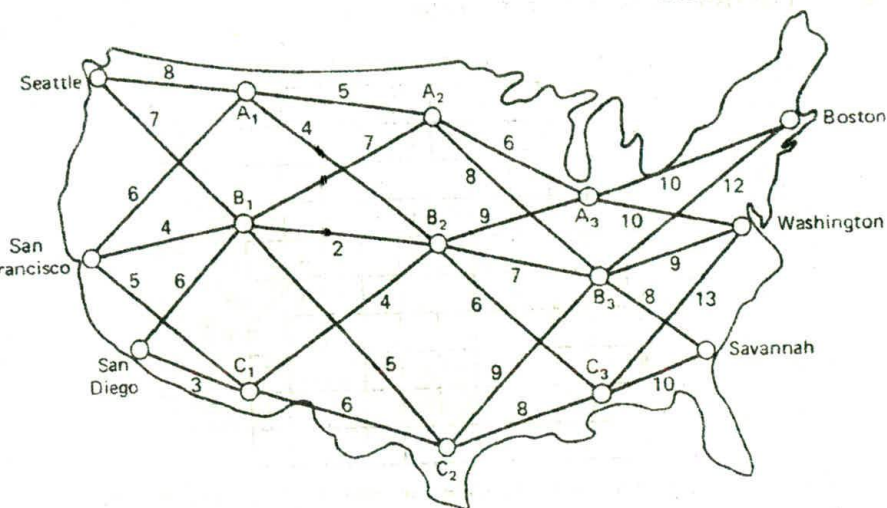


FIG. 3.15 Possible oil pipeline routes for Problem 3.11.

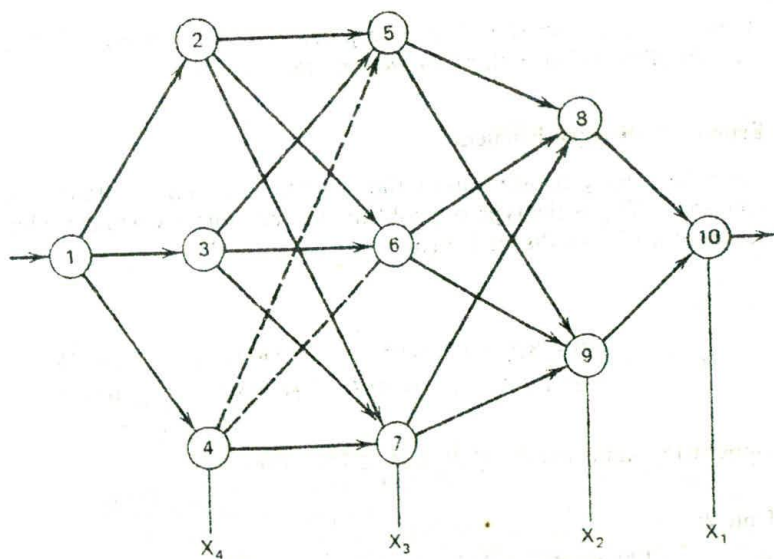


FIG. 3.16 Possible stagecoach routes for Problem 3.12.

are shown in Figure 3.16. 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

$i \backslash j$	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										

FIG. 3.17 Cost to go from state i to state j in Problem 3.12. Costs not shown are infinite.

stagecoach run from state i to state j , denoted as C_{ij} , is given in Figure 3.17. Find the safest path(s) for the salesman to take.

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:

$$H_1(P_1) = \begin{cases} 80 + 8P_1 + 0.024P_1^2 & 20 \text{ MW} \leq P_1 \leq 60 \text{ MW} \\ 196.4 + 3P_1 + 0.075P_1^2 & 60 \text{ MW} \leq P_1 \leq 100 \text{ MW} \end{cases}$$

Generation limits are $20 \text{ MW} \leq P_1 \leq 100 \text{ MW}$.

Unit 2:

$$H_2(P_2) = \begin{cases} 120 + 6P_2 + 0.04P_2^2 & 20 \text{ MW} \leq P_2 \leq 40 \text{ MW} \\ 157.335 + 3.3333P_2 + 0.08333P_2^2 & 40 \text{ MW} \leq P_2 \leq 100 \text{ MW} \end{cases}$$

Generation limits are $20 \text{ MW} \leq P_2 \leq 100 \text{ MW}$.

Unit 3:

$$H_3(P_3) = \begin{cases} 100 + 4.6666P_3 + 0.13333P_3^2 & 20 \text{ MW} \leq P_3 \leq 50 \text{ MW} \\ 316.66 + 2P_3 + 0.1P_3^2 & 50 \text{ MW} \leq P_3 \leq 100 \text{ MW} \end{cases}$$

Generation limits are $20 \text{ MW} \leq P_3 \leq 100 \text{ MW}$. Fuel costs = 1.5 R/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:

$$P_D = 100 \text{ MW}$$

$$P_D = 140 \text{ MW}$$

$$P_D = 180 \text{ MW}$$

$$P_D = 220 \text{ MW}$$

$$P_D = 260 \text{ MW}$$

where

$$P_D = P_1 + P_2 + P_3$$

Solve using dynamic programming 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 dynamic programming? If you think you know how, try solving for $P_D = 100 \text{ MW}$.

3.14 Given: the two generating units below with piecewise linear cost functions $F(P)$ as shown.

Unit 1: $P_1^{\min} = 25 \text{ MW}$ and $P_1^{\max} = 200 \text{ MW}$

$P_1(\text{MW})$	$F_1(P_1)(\text{R h})$
25	289.0
100	971.5
150	1436.5
200	1906.5

Unit 2: $P_2^{\min} = 50 \text{ MW}$ and $P_2^{\max} = 400 \text{ MW}$

$P_2(\text{MW})$	$F_2(P_2)(\text{R h})$
50	3800
100	4230
200	5120
400	6960

Find the optimum generation schedule for a total power delivery of 350 MW (assume both generators are on-line).

- 3.15 Given: two generator units with piecewise linear incremental cost functions as shown.

Unit 1: $P_1^{\min} = 100 \text{ MW}$ and $P_1^{\max} = 400 \text{ MW}$

$P_1(\text{MW})$	$\frac{d}{dP_1} F_1(P_1)(\text{R/MWh})$
100	6.5
200	7.0
300	8.0
400	11.0

Unit 2: $P_2^{\min} = 120 \text{ MW}$ and $P_2^{\max} = 300 \text{ MW}$

$P_2(\text{MW})$	$\frac{d}{dP_2} F_2(P_2)(\text{R/MWh})$
120	8.0
150	8.3
200	9.0
300	12.5

- Find the optimum schedule for a total power delivery of 500 MW.
- Now assume that there are transmission losses in the system and the incremental losses for the generators are:

$$\frac{dP_{\text{loss}}}{dP_1} = -0.05263$$

and

$$\frac{dP_{\text{loss}}}{dP_2} = 0.04762$$

Find the optimum schedule for a total power delivery of 650 MW; that is, 650 equals the load plus the losses.

FURTHER READING

Since this chapter introduces several optimization concepts, it would be useful to refer to some of the general works on optimization such as references 1 and 2. The importance of the Kuhn-Tucker theorem is given in their paper (reference 3). A very thorough discussion of the Kuhn-Tucker theorem is found in Chapter 1 of reference 4.

For an overview of recent power system optimization practices see references 5 and 6. Several other applications of optimization have been presented. Reference 7 discusses the allocation of regulating margin while dispatching generator units. References 8-11 discuss how to formulate the dispatch problem as one that minimizes air pollution from power plants.

Reference 12 explains how dynamic economic dispatch is developed. Reference 13 is a good review of recent work in economic dispatch. References 14 and 15 show how special problems can be incorporated into economic dispatch, while references 16 and 17 show how altogether different, nonconventional algorithms can be applied to economic dispatch. References 18-21 are an overview of dynamic programming, which is introduced in one of the appendices of this chapter.

1. *Application of Optimization Methods in Power System Engineering*, IEEE Tutorial Course Text 76CH1107-2-PWR, IEEE, New York, 1976.
2. Wilde, P. J., Beightler, C. S., *Foundations of Organization*, Prentice-Hall, Englewood Cliffs, NJ, 1967.
3. Kuhn, H. W., Tucker, A. W., "Nonlinear Programming," in *Second Berkeley Symposium on Mathematical Programming Statistics and Probability*, 1950, University of California Press, Berkeley, 1951.
4. Wismer, D. A., *Optimization Methods for Large-Scale Systems with Applications*, McGraw-Hill, New York, 1971.
5. IEEE Committee Report, "Present Practices in the Economic Operation of Power Systems," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-90, July/August 1971, pp. 1768-1775.
6. Najaf-Zadeh, K., Nikolas, J. T., Anderson, S. W., "Optimal Power System Operation Analysis Techniques," *Proceedings of the American Power Conference*, 1977.
7. Stadlin, W. O., "Economic Allocation of Regulating Margin," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-90, July/August 1971, pp. 1777-1781.
8. Gent, M. R., Lamont, J. W., "Minimum Emission Dispatch," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-90, November/December 1971, pp. 2650-2660.
9. Sullivan, R. L., "Minimum Pollution Dispatching," IEEE Summer Power Meeting, Paper C-72-468, 1972.
10. Friedman, P. G., "Power Dispatch Strategies for Emission and Environmental Control," *Proceedings of the Instrument Society of America*, Vol. 16, 1973.
11. Gjengedal, T., Johansen, S., Hansen, O., "A Qualitative Approach to Economic-Environment Dispatch—Treatment of Multiple Pollutants," *IEEE Transactions on Energy Conversion*, Vol. 7, No. 3, September 1992, pp. 367-373.
12. Ross, D. W., "Dynamic Economic Dispatch," *IEEE Transactions on Power Apparatus and Systems*, November/December 1980, pp. 2060-2068.
13. Chowdhury, B. H., Rahman, S., "A Review of Recent Advances in Economic Dispatch," *IEEE Transactions on Power Systems*, Vol. 5, No. 4, November 1990, pp. 1248-1259.
14. Bobo, D. R., Mauzy, D. M., "Economic Generation Dispatch with Responsive Spinning Reserve Constraints," 1993 *IEEE Power Industry Computer Applications Conference*, 1993, pp. 299-303.
15. Lee, F. N., Breipohl, A. M., "Reserve Constrained Economic Dispatch with

- Prohibited Operating Zones," *IEEE Transactions on Power Systems*, Vol. 8, No. 1, February 1993, pp. 246-254.
16. Walters, D. C., Sheble, G. B., "Genetic Algorithm Solution of Economic Dispatch with Valve Point Loading," *IEEE Transactions on Power Systems*, Vol. 8, No. 3, August 1993, pp. 1325-1332.
 17. Wong, K. P., Fung, C. C., "Simulated Annealing Bbased Economic Dispatch Algorithm," *IEE Proceedings, Part C: Generation, Transmission and Distribution*, Vol. 140, No. 6, November 1993, pp. 509-515.
 18. Kaufmann, A., *Graphs. Dynamic Programming and Finite Games*, Academic Press, New York, 1967.
 19. Howard, R. A., *Dynamic Programming and Markov Processes*, Wiley and Technology Press, New York, 1960.
 20. Bellman, R. E., Dreyfus, S. E., *Applied Dynamic Programming*, Princeton University Press, Princeton, NJ, 1962.
 21. Larson, R. E., Costi, J. L., *Principles of Dynamic Programming*, M. Decker, New York, 1978-1982.

4 Transmission System Effects

As we saw in the previous chapter, the transmission network's incremental power losses may cause a bias in the optimal economic scheduling of the generators. The *coordination equations* include the effects of the incremental transmission losses and complicate the development of the proper schedule. The network elements lead to two other, important effects:

1. The total real power loss in the network increases the total generation demand, and
2. The generation schedule may have to be adjusted by shifting generation to reduce flows on transmission circuits because they would otherwise become overloaded.

It is the last effect that is the most difficult to include in optimum dispatching. In order to include constraints on flows through the network elements, the flows must be evaluated as an integral part of the scheduling effort. This means we must solve the power flow equations along with the generation scheduling equations. (Note that earlier texts, papers, and even the first edition of this book referred to these equations as the "load flow" equations.)

If the constraints on flows in the networks are ignored, then it is feasible to use what are known as *loss formulae* that relate the total and incremental, real power losses in the network to the power generation magnitudes. Development of loss formulae is an art that requires knowledge of the power flows in the network under numerous "typical" conditions. Thus, there is no escaping the need to understand the methods involved in formulating and solving the power flow equations for an AC transmission system.

When the complete transmission system model is included in the development of generation schedules, the process is usually imbedded in a set of computer algorithms known as the *optimal power flow* (or OPF). The complete OPF is capable of establishing schedules for many controllable quantities in the bulk power system (i.e., the generation and transmission systems), such as transformer tap positions, VAR generation schedules, etc. We shall defer a detailed examination of the OPF until Chapter 13.

Another useful set of data that are obtainable when the transmission network is incorporated in the scheduling process is the incremental cost of power at various points in the network. With no transmission effects considered (that is, ignoring all incremental losses and any constraints on power flows), the network

is assumed to be a single node and the incremental cost of power is equal to λ everywhere. That is,

$$\frac{dF_i}{dP_i} = \lambda$$

Including the effect of incremental losses will cause the incremental cost of real power to vary throughout the network. Consider the arrangement in Figure 3.2 and assume that the coordination equations have been solved so the values of dF_i/dP_i and λ are known. Let the "penalty factor" of bus i be defined as

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

so that the relationship between the incremental costs at any two buses, i and j , is

$$P_{f_i} F'_i = P_{f_j} F'_j$$

where $F'_k = dF_k/dP_k$ is the bus incremental cost. There is no requirement that bus i is a generator bus. If the network effects are included using a network model or a loss formula, bus i might be a load bus or a point where power is delivered to an interconnected system. The incremental cost (or "value") of power at bus i is then,

$$\text{Incremental cost at } i = F'_i = (P_{f_j}/P_{f_i})F'_j$$

where j is any real generator bus where the incremental cost of production is known. So if we can develop a network model to be used in optimum generation scheduling that includes all of the buses, or at least those that are of importance, and if the incremental losses ($\partial P_L/\partial P_k$) can be evaluated, the coordination equations can be used to compute the incremental cost of power at any point of delivery.

When the schedule is determined using a complete power flow model by using an OPF, the flow constraints can be included and they may affect the value of the incremental cost of power in parts of the network. Rather than attempt a mathematical demonstration, consider a system in which most of the low cost generation is in the north, most of the load is in the south along with some higher cost generating units, and the northern and southern areas are interconnected by a relatively low capacity transmission network. The network north-to-south transfer capability limits the power that can be delivered from the northern area to satisfy the higher load demands. Under a schedule that is constrained by this transmission flow limitation, the southern area's generation would need to be increased above an unconstrained, optimal level in order to satisfy some of the load in that region. The constrained economic schedule

would split the system into two regions with a higher incremental cost in the southern area. In most actual cases where transmission does constrain the economic schedule, the effect of the constraints is much more significant than the effects of incremental transmission losses.

This chapter develops the power flow equations and outlines methods of solution. Operations control centers frequently use a version of the power flow equations known as the "decoupled power flow." The power flow equations form the basis for the development of loss formulae. Scheduling methods frequently use penalty factors to incorporate the effect of incremental real power losses in dispatch. These can be developed from the loss formulae or directly from the power flow relationships.

Power flow is the name given to a network solution that shows currents, voltages, and real and reactive power flows at every bus in the system. It is normally assumed that the system is balanced and the common use of the term power flow implies a positive sequence solution only. Full three-phase power-flow solution techniques are available for special-purpose calculations. As used here, we are only interested in balanced solutions. Power flow is not a single calculation such as $E = IR$ or $E = [Z]I$ involving linear circuit analysis. Such circuit analysis problems start with a given set of currents or voltages, and one must solve for the linearly dependent unknowns. In the power-flow problem we are given a nonlinear relationship between voltage and current at each bus and we must solve for all voltages and currents such that these nonlinear relationships are met. The nonlinear relationships involve, for example, the real and reactive power consumption at a bus, or the generated real power and scheduled voltage magnitude at a generator bus. As such, the power flow gives us the electrical response of the transmission system to a particular set of loads and generator unit outputs. Power flows are an important part of power system design procedures (system planning). Modern digital computer power-flow programs are routinely run for systems with up to 5000 or more buses and also are used widely in power system control centers to study unique operating problems and to provide accurate calculations of bus penalty factors. Present, state-of-the-art system control centers use the power flow as a key, central element in the scheduling of generation, monitoring of the system, and development of interchange transactions. OPF programs are used to develop optimal economic schedules and control settings that will result in flows that are within the capabilities of the elements of the system, including the transmission network, and bus voltage magnitudes that are within acceptable tolerances.

4.1 THE POWER FLOW PROBLEM AND ITS SOLUTION

The power flow problem consists of a given transmission network where all lines are represented by a Pi-equivalent circuit and transformers by an ideal voltage transformer in series with an impedance. Generators and loads represent

the boundary conditions of the solution. Generator or load real and reactive power involves products of voltage and current. Mathematically, the power flow requires a solution of a system of simultaneous nonlinear equations.

4.1.1 The Power Flow Problem on a Direct Current Network

The problems involved in solving a power flow can be illustrated by the use of direct current (DC) circuit examples. The circuit shown in Figure 4.1 has a resistance of 0.25Ω tied to a constant voltage of 1.0 V (called the *reference voltage*). We wish to find the voltage at bus 2 that results in a net inflow of 1.2 W . Buses are electrical nodes. Power is said to be "injected" into a network; therefore, loads are simply negative injections.

The current from bus 2 to bus 1 is

$$I_{21} = (E_2 - 1.0) \times 4 \quad (4.1)$$

Power P_2 is

$$P_2 = 1.2 = E_2 I_{21} = E_2 (E_2 - 1) \times 4 \quad (4.2)$$

or

$$4E_2^2 - 4E_2 - 1.2 = 0 \quad (4.3)$$

The solutions to this quadratic equation are $E_2 = 1.24162 \text{ V}$ and $E_2 = -0.24162 \text{ V}$. Note that 1.2 W enter bus 2, producing a current of 0.96648 A ($E_2 = 1.24162$), which means that 0.96648 W enter the reference bus and 0.23352 W are consumed in the $0.25\text{-}\Omega$ resistor.

Let us complicate the problem by adding a third bus and two more lines (see Figure 4.2). The problem is more complicated because we cannot simply write out the solutions using a quadratic formula. The admittance equations are

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} 14 & -4 & -10 \\ -4 & 9 & -5 \\ -10 & -5 & 15 \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix} \quad (4.4)$$

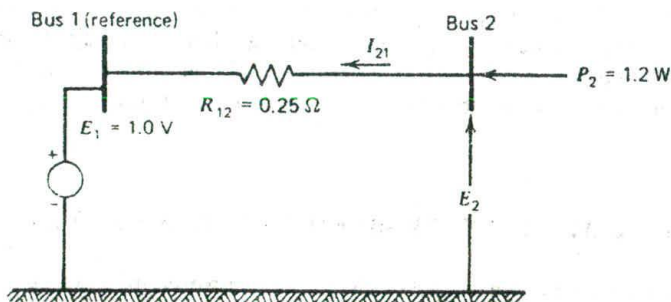


FIG. 4.1 Two-bus DC network.

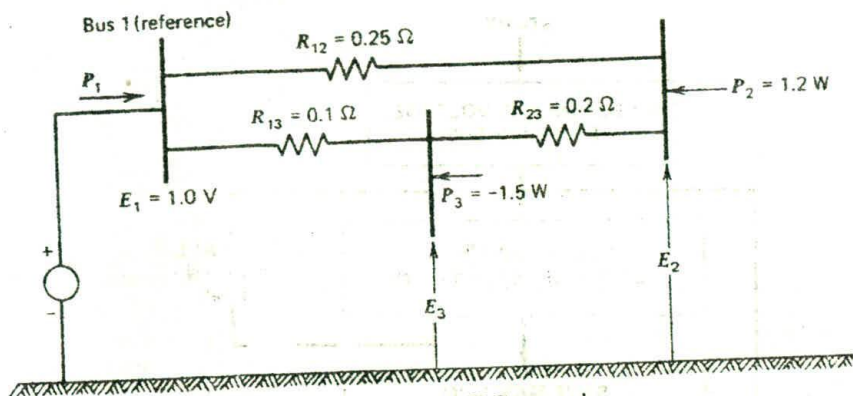


FIG. 4.2 Three-bus DC network.

In this case, we know the power injected at buses 2 and 3 and we know the voltage at bus 1. To solve for the unknowns (E_2 , E_3 and P_1), we write Eqs. 4.5, 4.6, and 4.7. The solution procedure is known as the *Gauss-Seidel procedure*, wherein a calculation for a new voltage at each bus is made, based on the most recently calculated voltages at all neighbouring buses.

$$\text{Bus 2: } I_2 = \frac{P_2}{E_2} = -4(1.0) + 9E_2 - 5E_3$$

$$E_2^{\text{new}} = \frac{1}{9} \left(\frac{1.2}{E_2^{\text{old}}} + 4 + 5E_3^{\text{old}} \right) \quad (4.5)$$

where E_2^{old} and E_3^{old} are the initial values for E_2 and E_3 , respectively.

$$\text{Bus 3: } I_3 = \frac{P_3}{E_3} = -10(1.0) - 5E_2^{\text{new}} + 15E_3$$

$$E_3^{\text{new}} = \frac{1}{15} \left[\frac{-1.5}{E_3^{\text{old}}} + 10 + 5E_2^{\text{new}} \right] \quad (4.6)$$

where E_2^{new} is the voltage found in solving Eq. 4.5, and E_3^{old} is the initial value of E_3 .

$$\text{Bus 1: } P_1 = E_1 I_1^{\text{new}} = 1.0 I_1^{\text{new}} = 14 - 4E_2^{\text{new}} - 10E_3^{\text{new}} \quad (4.7)$$

The Gauss-Seidel method first assumes a set of voltages at buses 2 and 3 and then uses Eqs. 4.5 and 4.6 to solve for new voltages. The new voltages are compared to the voltage's most recent values, and the process continues until

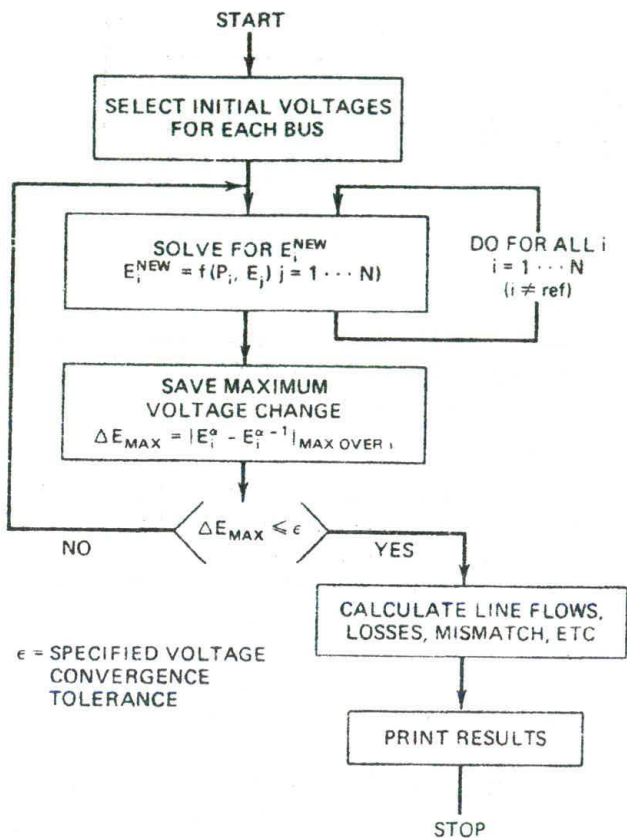


FIG. 4.3 Gauss-Seidel power-flow solution.

the change in voltage is very small. This is illustrated in the flowchart in Figure 4.3 and in Eqs. 4.8 and 4.9.

First iteration:

$$E_2^{(0)} = E_3^{(0)} = 1.0$$

$$E_2^{(1)} = \frac{1}{9} \left(\frac{1.2}{1.0} + 4 + 5 \right) = 1.133$$

$$E_3^{(1)} = \frac{1}{15} \left[\frac{-1.5}{1.0} + 10 + 5(1.133) \right] = 0.944$$

(4.8)

$$\Delta E_{\max} = 0.133 \text{ too large}$$

Note: In calculating $E_3^{(1)}$ we used the new value of E_2 found in the first correction.

$$\begin{aligned} \text{Second iteration: } E_2^{(2)} &= \frac{1}{9} \left[\frac{1.2}{1.133} + 4 + 5(0.944) \right] = 1.087 \\ E_3^{(2)} &= \frac{1}{15} \left[\frac{-1.5}{0.944} + 10 + 5(1.087) \right] = 0.923 \\ \Delta E_{\max} &= 0.046 \end{aligned} \quad (4.9)$$

And so forth until $\Delta E_{\max} < \epsilon$.

4.1.2 The Formulation of the AC Power Flow

AC power flows involve several types of bus specifications, as shown in Figure 4.4. Note that $[P, \theta]$, $[Q, |E|]$, and $[Q, \theta]$ combinations are generally not used.

The transmission network consists of complex impedances between buses and from the buses to ground. An example is given in Figure 4.5. The equations are written in matrix form as

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} = \begin{bmatrix} y_{12} & -y_{12} & 0 & 0 \\ -y_{12} & (y_{12} + y_{2g} + y_{23}) & -y_{23} & 0 \\ 0 & -y_{23} & (y_{23} + y_{3g} + y_{34}) & -y_{34} \\ 0 & 0 & -y_{34} & (y_{34} + y_{4g}) \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ E_3 \\ E_4 \end{bmatrix} \quad (4.10)$$

(All I^s, E^s, y^s complex)

This matrix is called the *network Y matrix*, which is written as

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} & Y_{13} & Y_{14} \\ Y_{21} & Y_{22} & Y_{23} & Y_{24} \\ Y_{31} & Y_{32} & Y_{33} & Y_{34} \\ Y_{41} & Y_{42} & Y_{43} & Y_{44} \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ E_3 \\ E_4 \end{bmatrix} \quad (4.11)$$

The rules for forming a Y matrix are

If a line exists from i to j

$$Y_{ij} = -y_{ij}$$

and

$$Y_{ii} = \sum_j y_{ij} + y_{ig}$$

j over all lines connected to i .

Bus Type	P	Q	$ E $	θ	Comments
Load	✓	✓			Usual load representation
Voltage Controlled	✓		✓		Assume $ E $ is held constant no matter what Q is
Generator or Synchronous Condenser	✓		✓ when $Q^- < Q_g < Q^+$		Generator or synchronous condenser ($P = 0$) has VAR limits
	✓	✓ when $Q_g < Q^-$ $Q_g > Q^+$			Q^- minimum VAR limit Q^+ maximum VAR limit $ E $ is held as long as Q_g is within limit
Fixed Z to Ground					Only Z is given
Reference			✓	✓	"Swing bus" must adjust net power to hold voltage constant (essential for solution)

FIG. 4.4 Power-flow bus specifications (quantities checked are the bus boundary conditions).

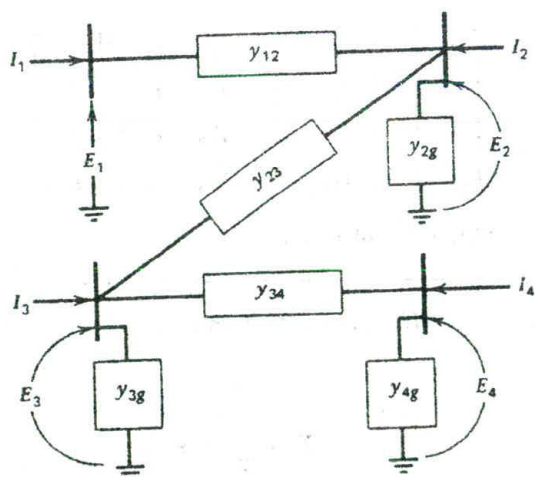


FIG. 4.5. Four-bus AC network.

The equation of net power injection at a bus is usually written as

$$\frac{P_k - jQ_k}{E_k^*} = \sum_{\substack{j=1 \\ j \neq k}}^n Y_{kj} E_j + Y_{kk} E_k \quad (4.12)$$

4.1.2.1 The Gauss-Seidel Method

The voltages at each bus can be solved for by using the Gauss-Seidel method. The equation in this case is

$$E_k^{(\alpha)} = \frac{1}{Y_{kk}} \frac{(P_k - jQ_k)}{E_k^{(\alpha-1)*}} - \frac{1}{Y_{kk}} \left[\sum_{j < k} Y_{kj} E_j^{(\alpha)} + \sum_{j > k} Y_{kj} E_j^{(\alpha-1)} \right] \quad (4.13)$$

Voltage at
iteration α

The Gauss-Seidel method was the first AC power-flow method to be developed for solution on digital computers. This method is characteristically long in solving due to its slow convergence and often difficulty is experienced with unusual network conditions such as negative reactance branches. The solution procedure is the same as shown in Figure 4.3.

4.1.2.2 The Newton-Raphson Method

One of the disadvantages of the Gauss-Seidel method lies in the fact that each bus is treated independently. Each correction to one bus requires subsequent correction to all the buses to which it is connected. The Newton-Raphson method is based on the idea of calculating the corrections while taking account of all the interactions.

Newton's method involves the idea of an error in a function $f(x)$ being driven to zero by making adjustments Δx to the independent variable associated with the function. Suppose we wish to solve

$$f(x) = K \quad (4.14)$$

In Newton's method, we pick a starting value of x and call it x^0 . The error is the difference between K and $f(x^0)$. Call the error ε . This is shown in Figure 4.6 and given in Eq. 4.15.

$$f(x^0) + \varepsilon = K \quad (4.15)$$

To drive the error to zero, we use a Taylor expansion of the function about x^0 .

$$f(x^0) + \frac{df(x^0)}{dx} \Delta x + \varepsilon = K \quad (4.16)$$

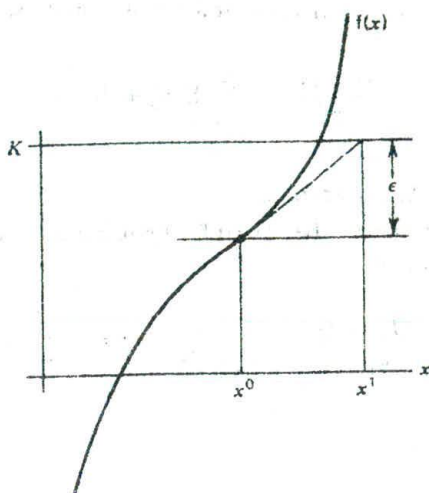


FIG. 4.6 Newton's method.

Setting the error to zero, we calculate

$$\Delta x = \left(\frac{df(x^0)}{dx} \right)^{-1} [K - f(x^0)] \quad (4.17)$$

When we wish to solve a load flow, we extend Newton's method to the multivariable case (the multivariable case is called the *Newton-Raphson method*). An equation is written for each bus "i."

$$P_i + jQ_i = E_i I_i^* \quad (4.18)$$

where

$$I_i = \sum_{k=1}^N Y_{ik} E_k$$

then

$$\begin{aligned} P_i + jQ_i &= E_i \left(\sum_{k=1}^N Y_{ik} E_k \right)^* \\ &= |E_i|^2 Y_{ii}^* + \sum_{\substack{k=1 \\ k \neq i}}^N Y_{ik}^* E_i E_k^* \end{aligned}$$

As in the Gauss-Seidel method, a set of starting voltages is used to get things going. The $P + jQ$ calculated is subtracted from the scheduled $P + jQ$ at the bus, and the resulting errors are stored in a vector. As shown in the following, we will assume that the voltages are in polar coordinates and that we are going to adjust each voltage's magnitude and phase angle as separate independent

variables. Note that at this point, two equations are written for each bus: one for real power and one for reactive power. For each bus,

$$\begin{aligned}\Delta P_i &= \sum_{k=1}^N \frac{\partial P_i}{\partial \theta_k} \Delta \theta_k + \sum_{k=1}^N \frac{\partial P_i}{\partial |E_k|} \Delta |E_k| \\ \Delta Q_i &= \sum_{k=1}^N \frac{\partial Q_i}{\partial \theta_k} \Delta \theta_k + \sum_{k=1}^N \frac{\partial Q_i}{\partial |E_k|} \Delta |E_k|\end{aligned}\quad (4.19)$$

All the terms are arranged in a matrix (the Jacobian matrix) as follows.

$$\begin{bmatrix} \Delta P_1 \\ \Delta Q_1 \\ \Delta P_2 \\ \Delta Q_2 \\ \vdots \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{\partial P_1}{\partial \theta_1} & \frac{\partial P_1}{\partial |E_1|} & \cdots \\ \frac{\partial Q_1}{\partial \theta_1} & \frac{\partial Q_1}{\partial |E_1|} & \cdots \\ \vdots & \vdots & \vdots \end{bmatrix}}_{\text{Jacobian matrix}} \begin{bmatrix} \Delta \theta_1 \\ \Delta |E_1| \\ \vdots \end{bmatrix}\quad (4.20)$$

The Jacobian matrix in Eq. 4.20 starts with the equation for the real and reactive power at each bus. This equation, Eq. 4.18, is repeated below:

$$P_i + jQ_i = E_i \sum_{k=1}^N Y_{ik}^* E_k^*$$

This can be expanded as:

$$\begin{aligned}P_i + jQ_i &= \sum_{k=1}^N |E_i| |E_k| (G_{ik} - jB_{ik}) e^{j(\theta_i - \theta_k)} \\ &= \sum_{k=1}^N \{ |E_i| |E_k| [G_{ik} \cos(\theta_i - \theta_k) + B_{ik} \sin(\theta_i - \theta_k)] \\ &\quad + j[|E_i| |E_k| [G_{ik} \sin(\theta_i - \theta_k) - B_{ik} \cos(\theta_i - \theta_k)]] \}\end{aligned}\quad (4.21)$$

where

θ_i, θ_k = the phase angles at buses i and k , respectively;

$|E_i|, |E_k|$ = the bus voltage magnitudes, respectively

$G_{ik} + jB_{ik} = Y_{ik}$ is the ik term in the Y matrix of the power system.

The general practice in solving power flows by Newton's method has been to use

$$\frac{\Delta|E_i|}{|E_i|}$$

instead of simply $\Delta|E_i|$; this simplifies the equations. The derivatives are:

$$\begin{aligned} \frac{\partial P_i}{\partial \theta_k} &= |E_i||E_k|[G_{ik} \sin(\theta_i - \theta_k) - B_{ik} \cos(\theta_i - \theta_k)] \\ \frac{\partial P_i}{\left(\frac{\partial|E_k|}{|E_k|}\right)} &= |E_i||E_k|[G_{ik} \cos(\theta_i - \theta_k) + B_{ik} \sin(\theta_i - \theta_k)] \\ \frac{\partial Q_i}{\partial \theta_k} &= -|E_i||E_k|[G_{ik} \cos(\theta_i - \theta_k) + B_{ik} \sin(\theta_i - \theta_k)] \\ \frac{\partial Q_i}{\left(\frac{\partial|E_k|}{|E_k|}\right)} &= |E_i||E_k|[G_{ik} \sin(\theta_i - \theta_k) - B_{ik} \cos(\theta_i - \theta_k)] \end{aligned} \quad (4.22)$$

For $i = k$:

$$\begin{aligned} \frac{\partial P_i}{\partial \theta_i} &= -Q_i - B_{ii}E_i^2 \\ \frac{\partial P_i}{\left(\frac{\partial|E_i|}{|E_i|}\right)} &= P_i + G_{ii}E_i^2 \\ \frac{\partial Q_i}{\partial \theta_i} &= P_i - G_{ii}E_i^2 \\ \frac{\partial Q_i}{\left(\frac{\partial|E_i|}{|E_i|}\right)} &= Q_i - B_{ii}E_i^2 \end{aligned}$$

Equation 4.20 now becomes

$$\begin{bmatrix} \Delta P_1 \\ \Delta Q_1 \\ \Delta P_2 \\ \Delta Q_2 \\ \vdots \end{bmatrix} = [J] \begin{bmatrix} \Delta \theta_1 \\ \frac{\Delta|E_1|}{|E_1|} \\ \Delta \theta_2 \\ \frac{\Delta|E_2|}{|E_2|} \\ \vdots \end{bmatrix} \quad (4.23)$$

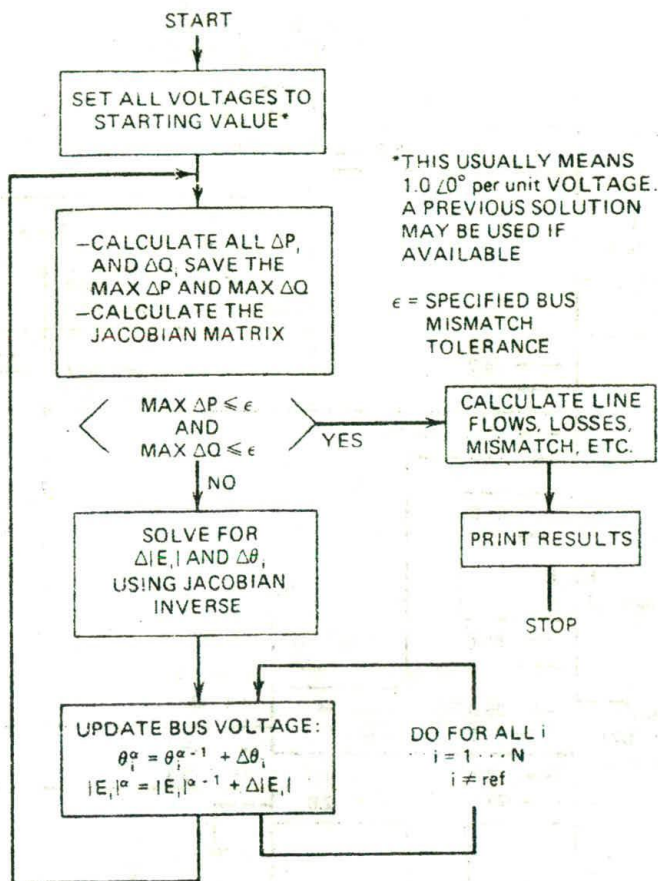


FIG. 4.7 Newton-Raphson power-flow solution.

The solution to the Newton-Raphson power flow runs according to the flowchart in Figure 4.7. Note that solving for $\Delta \theta$ and $\Delta |E|$ requires the solution of a set of linear equations whose coefficients make up the Jacobian matrix. The Jacobian matrix generally has only a few percent of its entries that are nonzero. Programs that solve an AC power flow using the Newton-Raphson method are successful because they take advantage of the Jacobian's "sparsity." The solution procedure uses Gaussian elimination on the Jacobian matrix and does not calculate J^{-1} explicitly. (See reference 3 for introduction to "sparsity" techniques.)

EXAMPLE 4A

The six-bus network shown in Figure 4.8 will be used to demonstrate several aspects of load flows and transmission loss factors. The voltages and flows

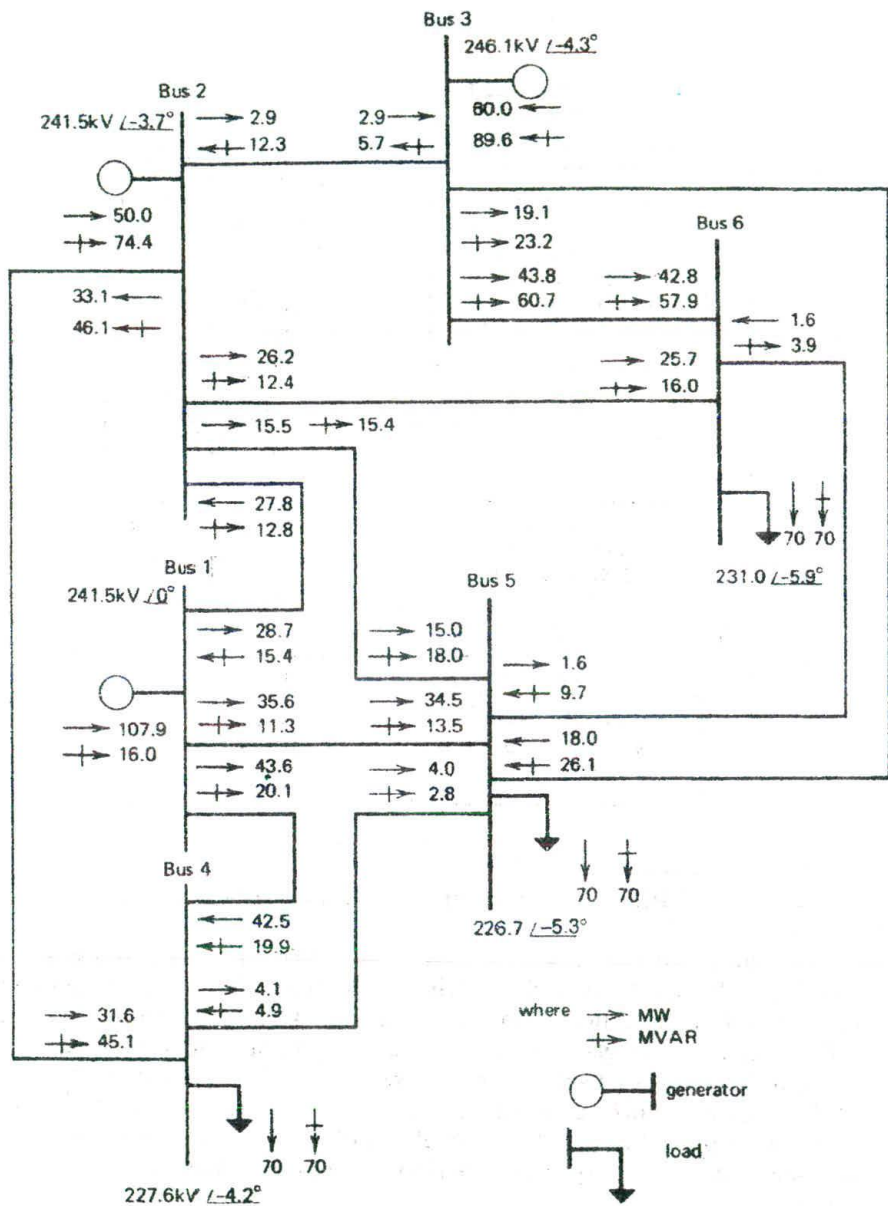


FIG. 4.8 Six-bus network base case AC power flow.

shown are for the "base case" of 210 MW total load. The impedance values and other data for this system may be found in the appendix of this chapter.

4.1.3 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 Eq. 4.23 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. Reference 11 shows the development of 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).

Starting with the terms in the Jacobian matrix (see Eq. 4.22), the following simplifications are made:

- Neglect and interaction between P_i and any $|E_k|$ (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

$$\frac{\partial P_i}{\left(\frac{\partial |E_k|}{|E_k|}\right)}$$

will be considered to be zero.

- Neglect any interaction between Q_i and θ_k (see the note above—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 zero.

- 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

$$Q_i \ll B_{ii} |E_i|^2$$

This leaves the derivatives as:

$$\frac{\partial P_i}{\partial \theta_k} = -|E_i||E_k|B_{ik} \quad (4.24)$$

$$\frac{\partial Q_i}{\left(\frac{\partial |E_k|}{|E_k|}\right)} = -|E_i||E_k|B_{ik} \quad (4.25)$$

If we now write the power flow adjustment equations as:

$$\Delta P_i = \left(\frac{\partial P_i}{\partial \theta_k}\right) \Delta \theta_k \quad (4.26)$$

$$\Delta Q_i = \left[\frac{\partial Q_i}{\left(\frac{\partial |E_k|}{|E_k|}\right)} \right] \frac{\Delta |E_k|}{|E_k|} \quad (4.27)$$

then, substituting Eq. 4.24 into Eq. 4.26, and Eq. 4.25 into Eq. 4.27, we obtain:

$$\Delta P_i = -|E_i||E_k|B_{ik} \Delta \theta_k \quad (4.28)$$

$$\Delta Q_i = -|E_i||E_k|B_{ik} \frac{\Delta |E_k|}{|E_k|} \quad (4.29)$$

Further simplification can then be made:

- Divide Eqs. 4.28 and 4.29 by $|E_i|$.
- Assume $|E_k| \cong 1$ in Eq. 4.28.

which results in:

$$\frac{\Delta P_i}{|E_i|} = -B_{ik} \Delta \theta_k \quad (4.30)$$

$$\frac{\Delta Q_i}{|E_i|} = -B_{ik} \Delta |E_k| \quad (4.31)$$

We now build Eqs. 4.30 and 4.31 into two matrix equations:

$$\begin{bmatrix} \frac{\Delta P_1}{|E_1|} \\ \frac{\Delta P_2}{|E_2|} \\ \vdots \end{bmatrix} = \begin{bmatrix} -B_{11} & -B_{12} & \cdots \\ -B_{21} & -B_{22} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \Delta \theta_1 \\ \Delta \theta_2 \\ \vdots \end{bmatrix} \quad (4.32)$$

$$\begin{bmatrix} \frac{\Delta Q_1}{|E_1|} \\ \frac{\Delta Q_2}{|E_2|} \\ \vdots \end{bmatrix} = \begin{bmatrix} -B_{11} & -B_{12} & \dots \\ -B_{21} & -B_{22} & \dots \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \Delta|E_1| \\ \Delta|E_2| \\ \vdots \end{bmatrix} \quad (4.33)$$

Note that both Eqs. 4.32 and 4.33 use the same matrix. Further simplification, however, will make them different.

Simplifying the $\Delta P - \Delta\theta$ relationship of Eq. 4.32:

- 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 which arise from autotransformers.

Simplifying the $\Delta Q - \Delta|E|$ relationship of Eq. 4.33:

- Omit all effects from phase shift transformers.

The resulting equations are:

$$\begin{bmatrix} \frac{\Delta P_1}{|E_1|} \\ \frac{\Delta P_2}{|E_2|} \\ \vdots \end{bmatrix} = [B'] \begin{bmatrix} \Delta\theta_1 \\ \Delta\theta_2 \\ \vdots \end{bmatrix} \quad (4.34)$$

$$\begin{bmatrix} \frac{\Delta Q_1}{|E_1|} \\ \frac{\Delta Q_2}{|E_2|} \\ \vdots \end{bmatrix} = [B''] \begin{bmatrix} \Delta|E_1| \\ \Delta|E_2| \\ \vdots \end{bmatrix} \quad (4.35)$$

where the terms in the matrices are:

$$B'_{ik} = -\frac{1}{x_{ik}}, \text{ assuming a branch from } i \text{ to } k \text{ (zero otherwise)}$$

$$B'_{ii} = \sum_{k=1}^N \frac{1}{x_{ik}}$$

$$B''_{ik} = -B_{ik} = -\frac{x_{ik}}{r_{ik}^2 + x_{ik}^2}$$

$$B''_{ii} = \sum_{k=1}^N -B_{ik}$$

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 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 Eqs. 4.34 and 4.35.

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.

Note that Eq. 4.34 is often referred to as the P - θ Eq. and Eq. 4.35 as the Q - E (or Q - V) equation.

A flowchart of the algorithm is shown in Figure 4.9. A comparison of the convergence of the Gauss-Seidel, the full Newton and the decoupled power flow algorithms is shown in Figure 4.10.

4.1.4 The "DC" Power Flow

A further simplification of the power flow algorithm involves simply dropping the Q - V equation (Eq. 4.35) altogether. This results in a completely linear, noniterative, power flow algorithm. To carry this out, we simply assume that all $|E_i| = 1.0$ per unit. Then Eq. 4.34 becomes:

$$\begin{bmatrix} \Delta P_1 \\ \Delta P_2 \\ \vdots \end{bmatrix} = [B'] \begin{bmatrix} \Delta \theta_1 \\ \Delta \theta_2 \\ \vdots \end{bmatrix} \quad (4.36)$$

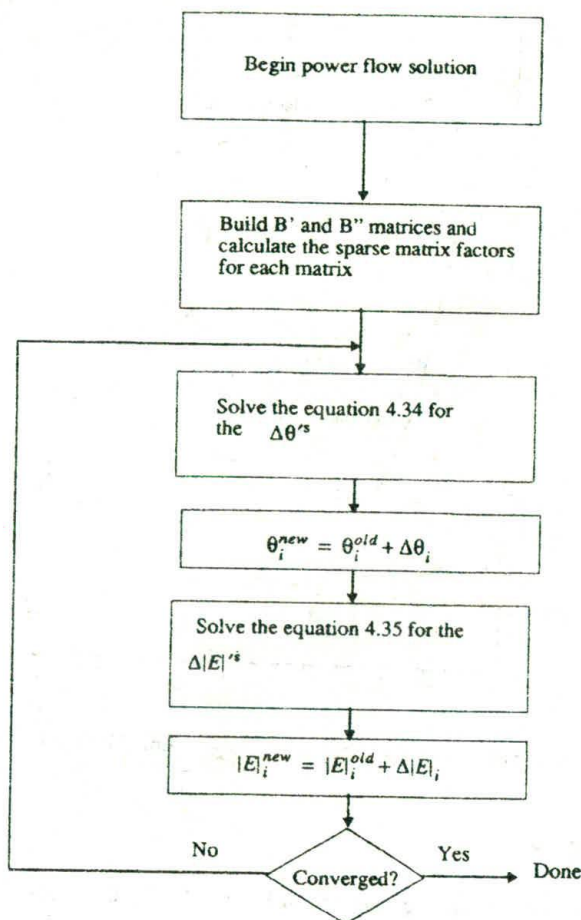


FIG. 4.9 Decoupled power flow algorithm.

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:

$$P_{ik} = \frac{1}{X_{ik}} (\theta_i - \theta_k) \quad (4.37)$$

and

$$P_i = \sum_{\substack{k = \text{buses} \\ \text{connected to } i}}^N P_k$$

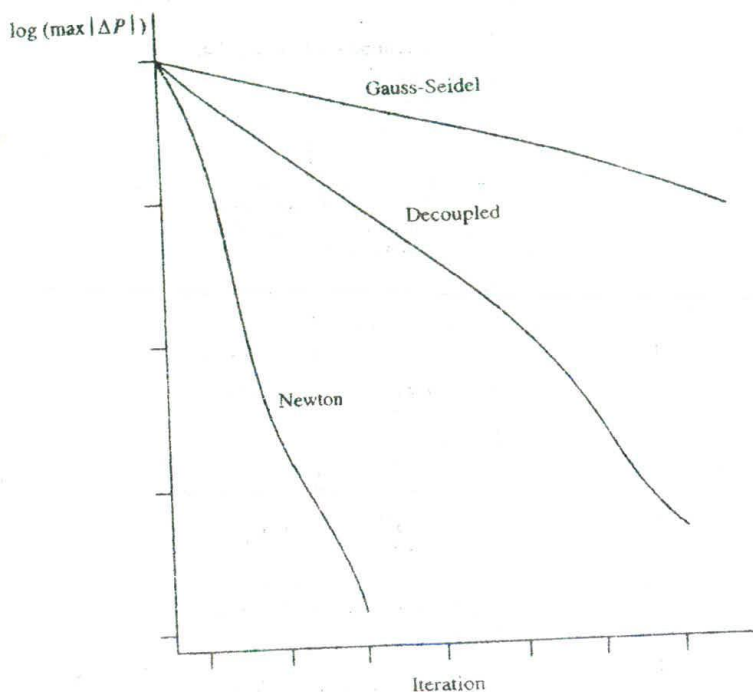


FIG. 4.10 Comparison of three power flow algorithm convergence characteristics.

EXAMPLE 4B

The megawatt flows on the network in Figure 4.11 will be solved using the DC power flow. The B' matrix equation is:

$$\begin{bmatrix} 7.5 & -5.0 \\ -5.0 & 9.0 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}$$

$$\theta_3 = 0$$

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:

$$\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 0.2118 & 0.1177 \\ 0.1177 & 0.1765 \end{bmatrix} \begin{bmatrix} 0.65 \\ -1.00 \end{bmatrix} = \begin{bmatrix} 0.02 \\ -0.1 \end{bmatrix}$$

The resulting flows are shown in Figure 4.12 and calculated using Eq. 4.37. Note that all flows in Figure 4.12 were converted to actual megawatt values.

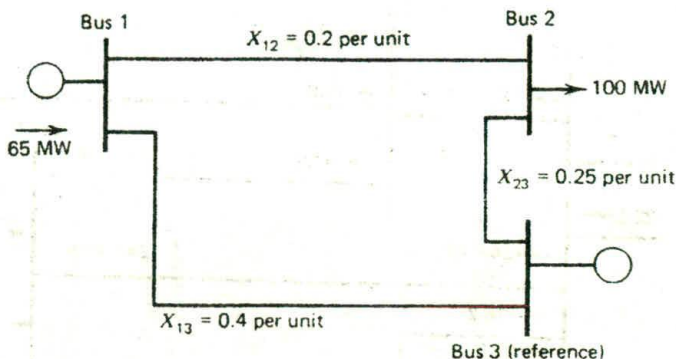


FIG. 4.11 Three-bus network.

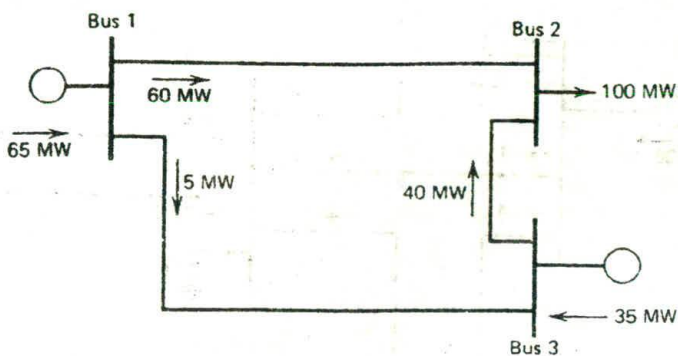


FIG. 4.12 Three-bus network showing flows calculated by DC power flow.

EXAMPLE 4C

The network of Example 4A was solved using the DC power flow with resulting power flows as shown in Figure 4.13. 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.

4.2 TRANSMISSION LOSSES

4.2.1 A Two-Generator System

We are given the power system in Figure 4.14. 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 4.15.

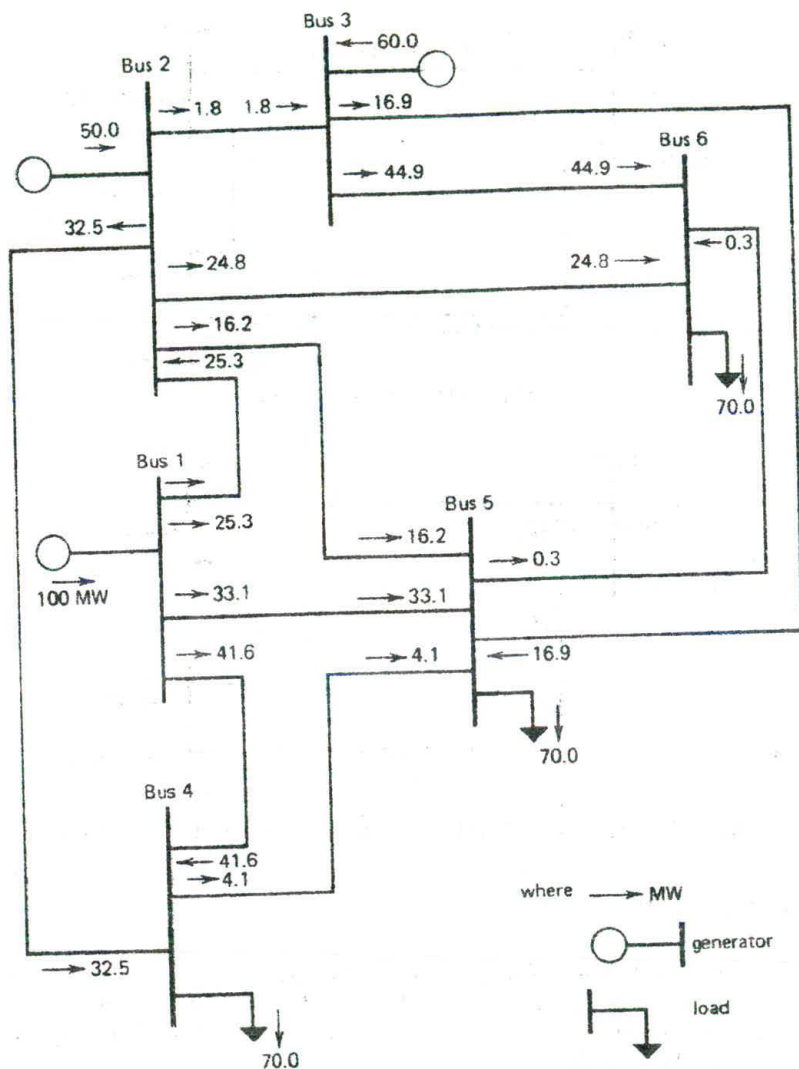


FIG. 4.13 Six-bus network base case DC power flow for Example 4C.

Where should the extra 12.5 MW be generated? Solve the Lagrange equation that was given in Chapter 3.

$$\mathcal{L} = F_1(P_1) + F_2(P_2) + \lambda(500 + P_{\text{loss}} - P_1 - P_2) \quad (4.38)$$

where

$$P_{\text{loss}} = 0.0002P_1^2$$

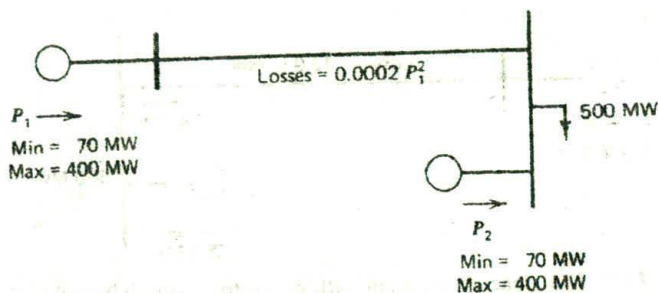


FIG. 4.14 Two-generator system.

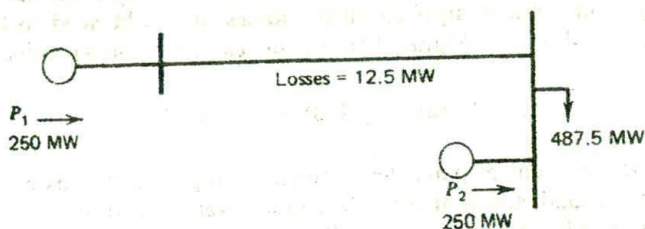


FIG. 4.15 Two-generator system with both generators at 250 MW output.

then

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial P_1} &= \frac{dF_1(P_1)}{dP_1} - \lambda \left(1 - \frac{\partial P_{\text{loss}}}{\partial P_1} \right) = 0 \\ \frac{\partial \mathcal{L}}{\partial P_2} &= \frac{dF_2(P_2)}{dP_2} - \lambda \left(1 - \frac{\partial P_{\text{loss}}}{\partial P_2} \right) = 0 \\ P_1 + P_2 - 500 - P_{\text{loss}} &= 0\end{aligned}\quad (4.39)$$

Substituting into Eq. 4.39,

$$\begin{aligned}7.0 + 0.004P_1 - \lambda(1 - 0.0004P_1) &= 0 \\ 7.0 + 0.004P_2 - \lambda &= 0 \\ P_1 + P_2 - 500 - 0.0002P_1^2 &= 0\end{aligned}$$

Solution:

$$P_1 = 178.882$$

$$P_2 = 327.496$$

Production cost:

$$F_1(P_1) + F_2(P_2) = 4623.15 \text{R/h}$$

Losses:

$$6.378 \text{ MW}$$

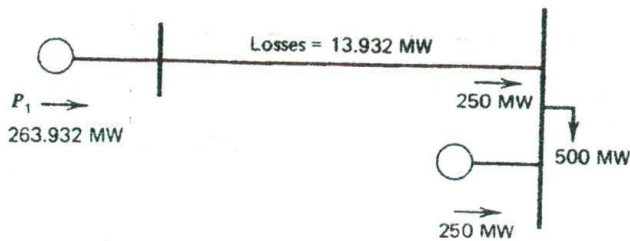


FIG. 4.16 Two-generator system with generator 1 supplying all losses.

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 4.16. In this case, the total production cost would be

$$F_1(263.932) + F_2(250) = 4661.84 \text{ R/h}$$

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

$$F_1(102.084) + F_2(400) = 4655.43 \text{ R/h}$$

$$\text{Min losses} = 2.084 \text{ MW}$$

4.2.2 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.

Minimize:

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

Where:

$$F_T = \sum_{i=1}^N F_i(P_i)$$

$$\phi = P_{\text{load}} + P_{\text{loss}}(P_1, P_2, \dots, P_N) - \sum_{i=1}^N P_i$$

Solution:

$$\frac{\partial \mathcal{L}}{\partial P_i} = 0 \quad \text{for all } P_{i\text{min}} \leq P_i \leq P_{i\text{max}}$$

Then

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

The equations are rearranged

$$\left(\frac{1}{1 - \frac{\partial P_{\text{loss}}}{\partial P_i}} \right) \frac{dF_i(P_i)}{dP_i} = \lambda \quad (4.40)$$

where

$$\frac{\partial P_{\text{loss}}}{\partial P_i}$$

is called the *incremental loss* for bus i , and

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

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 each unit the same. We can still use this concept by observing that the penalty factor, Pf_i , will have the following effect. For $Pf_i > 1$ (positive increase in P_i results in increase in losses)

$$Pf_i \frac{dF_i(P_i)}{dP_i}$$

acts as if

$$\frac{dF_i(P_i)}{dP_i}$$

had been slightly increased (moved up). For $Pf_i < 1$ (positive increase in P_i results in decrease in losses)

$$Pf_i \frac{dF_i(P_i)}{dP_i}$$

acts as if

$$\frac{dF_i(P_i)}{dP_i}$$

had been slightly decreased (moved down). The resulting set of equations look like

$$P_{f_i} \frac{dF_i(P_i)}{dP_i} = \lambda \quad \text{for all } P_{i\min} \leq P_i \leq P_{i\max} \quad (4.41)$$

and are called *coordination equations*. The P_i values that result when penalty factors are used will be somewhat different from the dispatch which ignores the losses (depending on the P_{f_i} and $dF_i(P_i)/dP_i$ values). This is illustrated in Figure 4.17.

4.2.3 The B Matrix Loss Formula

The B matrix loss formula was originally introduced in the early 1950s as a practical method for loss and incremental loss calculations. At the time, automatic dispatching was performed by analog computers and the loss formula was "stored" in the analog computers by setting precision potentiometers. The equation for the B matrix loss formula is as follows.

$$P_{\text{loss}} = \mathbf{P}^T[\mathbf{B}]\mathbf{P} + B_0^T\mathbf{P} + B_{00} \quad (4.42)$$

where

\mathbf{P} = vector of all generator bus net MW

$[\mathbf{B}]$ = square matrix of the same dimension as \mathbf{P}

B_0 = vector of the same length as \mathbf{P}

B_{00} = constant

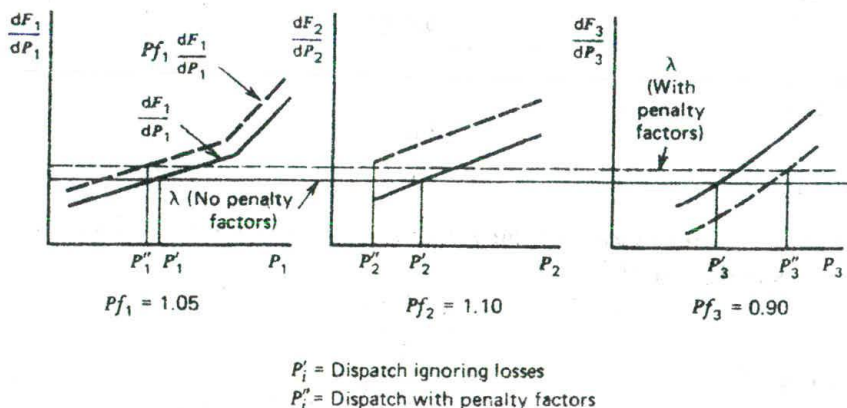


FIG. 4.17 Economic dispatch, with and without penalty factors.

This can be written:

$$P_{\text{loss}} = \sum_i \sum_j P_i B_{ij} P_j + \sum_i B_{i0} P_i + B_{00} \quad (4.43)$$

Before we discuss the calculation of the B coefficients, we will discuss how the coefficients are used in an economic dispatch calculation. Substitute Eq. 4.43 into Eqs. 3.7, 3.8, and 3.9.

$$\phi = - \sum_{i=1}^N P_i + P_{\text{load}} + \left(\sum_i \sum_j P_i B_{ij} P_j + \sum_i B_{i0} P_i + B_{00} \right) \quad (4.44)$$

Then

$$\frac{\partial \phi}{\partial P_i} = \frac{dF_i(P_i)}{dP_i} - \lambda \left(1 - 2 \sum_j B_{ij} P_j - B_{i0} \right) \quad (4.45)$$

Note that the presence of the incremental losses has coupled the coordination equations; this makes solution somewhat more difficult. A method of solution that is often used is shown in Figure 4.18.

EXAMPLE 4D

The B matrix loss formula for the network in Example 4A is given here. (Note that all P_i values must be per unit on 100 MVA base, which results in P_{loss} in per unit on 100 MVA base.)

$$P_{\text{loss}} = [P_1 \quad P_2 \quad P_3] \begin{bmatrix} 0.0676 & 0.00953 & -0.00507 \\ 0.00953 & 0.0521 & 0.00901 \\ -0.00507 & -0.00901 & 0.0294 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} \\ + [-0.0766 \quad -0.00342 \quad 0.0189] \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} + 0.040357$$

From the base case power flow we have

$$P_1 = 107.9 \text{ MW}$$

$$P_2 = 50.0 \text{ MW}$$

$$P_3 = 60.0 \text{ MW}$$

$$P_{\text{loss}} = 7.9 \text{ MW (as calculated by the power flow)}$$

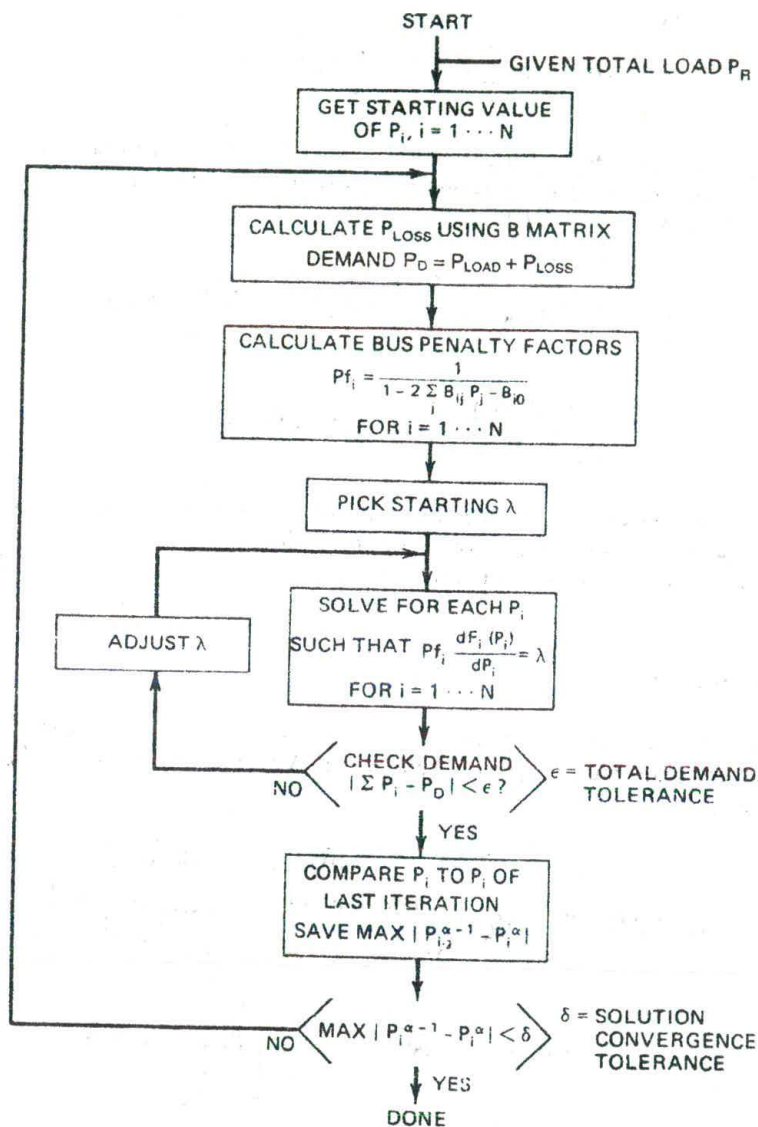


FIG. 4.18 Economic dispatch with updated penalty factors.

With these generation values placed in the B matrix, we see a very close agreement with the power flow calculation.

$$\begin{aligned}
 P_{\text{loss}} &= [1.079 \quad 0.50 \quad 0.60] \begin{bmatrix} 0.0676 & 0.00953 & -0.00507 \\ 0.00953 & 0.0521 & 0.00901 \\ -0.00507 & 0.00901 & 0.0294 \end{bmatrix} \begin{bmatrix} 1.079 \\ 0.50 \\ 0.60 \end{bmatrix} \\
 &+ [-0.0766 \quad -0.00342 \quad 0.0189] \begin{bmatrix} 1.079 \\ 0.50 \\ 0.60 \end{bmatrix} + 0.040357 \\
 &= 0.07877 \text{ pu (or 7.877 MW) loss}
 \end{aligned}$$

EXAMPLE 4E

Let the fuel cost curves for the three units in the six-bus network of Example 4A be given as

$$F_1(P_1) = 213.1 + 11.669P_1 + 0.00533P_1^2 \text{ R/h}$$

$$F_2(P_2) = 200.0 + 10.333P_2 + 0.00889P_2^2 \text{ R/h}$$

$$F_3(P_3) = 240.0 + 10.833P_3 + 0.00741P_3^2 \text{ R/h}$$

with unit dispatch limits

$$50.0 \text{ MW} \leq P_1 \leq 200 \text{ MW}$$

$$37.5 \text{ MW} \leq P_2 \leq 150 \text{ MW}$$

$$45.0 \text{ MW} \leq P_3 \leq 180 \text{ MW}$$

A computer program using the method of Figure 4.17 was run using:

$$P_{\text{load}} \text{ (total load to be supplied)} = 210 \text{ MW}$$

The resulting iterations (Table 4.1) show how the program must redispatch again and again to account for the changes in losses and penalty factors.

Note that the flowchart of Figure 4.18 shows a "two-loop" procedure. The "inner" loop adjusts λ until total demand is met; then the outer loop recalculates the penalty factors. (Under some circumstances the penalty factors are quite sensitive to changes in dispatch. If the incremental costs are relatively "flat," this procedure may be unstable and special precautions may need to be employed to insure convergence.)

TABLE 4.1 Iterations for Example 4E

Iteration	λ	P_{loss}	P_D	P_1	P_2	P_3
1	12.8019	17.8	227.8	50.00	85.34	92.49
2	12.7929	11.4	221.4	74.59	71.15	75.69
3	12.8098	9.0	219.0	73.47	70.14	75.39
4	12.8156	8.8	218.8	73.67	69.98	75.18
5	12.8189	8.8	218.8	73.65	69.98	75.18
6	12.8206	8.8	218.8	73.65	69.98	75.18

4.2.4 Exact Methods of Calculating Penalty Factors

4.2.4.1 A Discussion of Reference Bus Versus Load Center Penalty Factors

The B matrix assumes that all load currents conform to an equivalent total load current and that the equivalent load current is the negative of the sum of all generator currents. When incremental losses are calculated, something is implied.

$$\text{Total loss} = \mathbf{P}^T[B]\mathbf{P} + \mathbf{B}_0^T\mathbf{P} + B_{00}$$

$$\text{Incremental loss at generator bus } i = \frac{\partial P_{\text{loss}}}{\partial P_i}$$

The incremental loss is the change in losses when an increment is made in generation output. As just derived, the incremental loss for bus i assumed that all the other generators remained fixed. By the original assumption, however, the load currents all conform to each other and always balance with the generation; then the implication in using a B matrix is that an incremental increase in generator output is matched by an equivalent increment in load.

An alternative approach to economic dispatch is to use a reference bus that always moves when an increment in generation is made. Figure 4.19 shows a

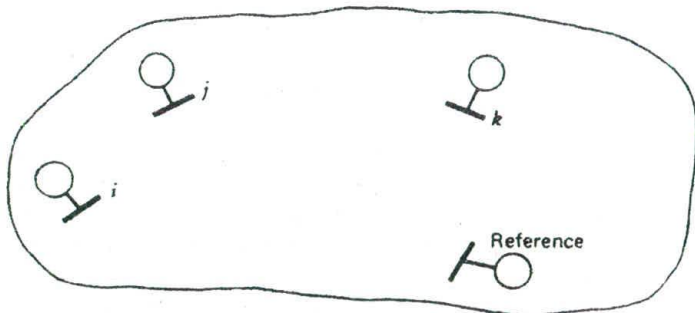


FIG. 4.19 Power system with reference generator.

power system with several generator buses and a reference-generator bus. Suppose we change the generation on bus i by ΔP_i ,

$$P_i^{\text{new}} = P_i^{\text{old}} + \Delta P_i \quad (4.46)$$

Furthermore, we will assume that *load stays constant* and that to compensate for the increase in ΔP_i , the reference bus just drops off by ΔP_{ref} .

$$P_{\text{ref}}^{\text{new}} = P_{\text{ref}}^{\text{old}} + \Delta P_{\text{ref}} \quad (4.47)$$

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,

$$\Delta P_{\text{ref}} = -\Delta P_i + \Delta P_{\text{loss}} \quad (4.48)$$

Next, we can define β_i as the ratio of the negative change in the reference-bus power to the change ΔP_i .

$$\beta_i = \frac{-\Delta P_{\text{ref}}}{\Delta P_i} = \frac{(\Delta P_i - \Delta P_{\text{loss}})}{\Delta P_i} \quad (4.49)$$

or

$$\beta_i = 1 - \frac{\partial P_{\text{loss}}}{\partial P_i} \quad (4.50)$$

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

$$\text{Total production cost} = \sum F_i(P_i)$$

then the change in production cost with a shift ΔP_i from plant i is

$$\Delta \text{Production cost} = \frac{dF_i(P_i)}{dP_i} \Delta P_i + \frac{dF_{\text{ref}}(P_{\text{ref}})}{dP_{\text{ref}}} \Delta P_{\text{ref}} \quad (4.51)$$

but

$$\Delta P_{\text{ref}} = -\beta_i \Delta P_i$$

then

$$\Delta \text{Production cost} = \frac{dF_i(P_i)}{dP_i} \Delta P_i - \beta_i \frac{dF_{\text{ref}}(P_{\text{ref}})}{dP_{\text{ref}}} \Delta P_i \quad (4.52)$$

To satisfy the economic conditions,

$$\Delta \text{Production cost} = 0$$

or

$$\frac{dF_i(P_i)}{dP_i} = \beta_i \frac{dF_{\text{ref}}(P_{\text{ref}})}{dP_{\text{ref}}} \quad (4.53)$$

which could be written as

$$\frac{1}{\beta_i} \frac{dF_i(P_i)}{dP_i} = \frac{dF_{\text{ref}}(P_{\text{ref}})}{dP_{\text{ref}}} \quad (4.54)$$

This is very similar to Eq. 4.40. To obtain an economic dispatch solution, pick a value of generation on the reference bus and then set all other generators according to Eq. 4.54, and check for total demand and readjust reference generation as needed until a solution is reached.

Note further that this method is exactly the first-order gradient method with losses.

$$\Delta F_T = \sum_{i \neq \text{ref}} \left[\frac{dF_i}{dP_i} - \beta_i \frac{dF_{\text{ref}}}{dP_{\text{ref}}} \right] \Delta P_i \quad (4.55)$$

4.2.4.2 Reference-Bus Penalty Factors Direct from the AC Power Flow

The reference-bus 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.

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

$$\begin{aligned} \Delta P_{\text{ref}} &= \sum_i \frac{\partial P_{\text{ref}}}{\partial \theta_i} \Delta \theta_i + \sum_i \frac{\partial P_{\text{ref}}}{\partial |E_i|} \Delta |E_i| \\ &= \sum_i \frac{\partial P_{\text{ref}}}{\partial \theta_i} \frac{\partial \theta_i}{\partial P_i} \Delta P_i + \sum_i \frac{\partial P_{\text{ref}}}{\partial |E_i|} \frac{\partial |E_i|}{\partial P_i} \Delta P_i \end{aligned} \quad (4.56)$$

To carry out the matrix manipulations, we will also need the following.

$$\begin{aligned} \Delta P_{\text{ref}} &= \sum_i \frac{\partial P_{\text{ref}}}{\partial \theta_i} \Delta \theta_i + \sum_i \frac{\partial P_{\text{ref}}}{\partial |E_i|} \Delta |E_i| \\ &= \sum_i \frac{\partial P_{\text{ref}}}{\partial \theta_i} \frac{\partial \theta_i}{\partial Q_i} \Delta Q_i + \sum_i \frac{\partial P_{\text{ref}}}{\partial |E_i|} \frac{\partial |E_i|}{\partial Q_i} \Delta Q_i \end{aligned} \quad (4.57)$$

The terms $\partial P_{\text{ref}}/\partial \theta_i$ and $\partial P_{\text{ref}}/|E_i|$ are derived by differentiating Eq. 4.18 for the reference bus. The terms $\partial \theta_i/\partial P_i$ and $\partial |E_i|/\partial P_i$ are from the inverse Jacobian matrix (see Eq. 4.20). We can write Eqs. 4.56 and 4.57 for every bus i in the network. The resulting equation is

$$\begin{aligned} & \begin{bmatrix} \frac{\partial P_{\text{ref}}}{\partial P_1} & \frac{\partial P_{\text{ref}}}{\partial Q_1} & \frac{\partial P_{\text{ref}}}{\partial P_2} & \frac{\partial P_{\text{ref}}}{\partial Q_2} & \cdots & \frac{\partial P_{\text{ref}}}{\partial P_N} & \frac{\partial P_{\text{ref}}}{\partial Q_N} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial P_{\text{ref}}}{\partial \theta_1} & \frac{\partial P_{\text{ref}}}{\partial |E_1|} & \frac{\partial P_{\text{ref}}}{\partial \theta_2} & \frac{\partial P_{\text{ref}}}{\partial |E_2|} & \cdots & \frac{\partial P_{\text{ref}}}{\partial \theta_N} & \frac{\partial P_{\text{ref}}}{\partial |E_N|} \end{bmatrix} [J^{-1}] \end{aligned} \quad (4.58)$$

By transposing we get

$$\begin{bmatrix} \frac{\partial P_{\text{ref}}}{\partial P_1} \\ \frac{\partial P_{\text{ref}}}{\partial Q_1} \\ \frac{\partial P_{\text{ref}}}{\partial P_2} \\ \frac{\partial P_{\text{ref}}}{\partial Q_2} \\ \vdots \\ \frac{\partial P_{\text{ref}}}{\partial P_N} \\ \frac{\partial P_{\text{ref}}}{\partial Q_N} \end{bmatrix} = [J^T]^{-1} \begin{bmatrix} \frac{\partial P_{\text{ref}}}{\partial \theta_1} \\ \frac{\partial P_{\text{ref}}}{\partial |E_1|} \\ \frac{\partial P_{\text{ref}}}{\partial \theta_2} \\ \frac{\partial P_{\text{ref}}}{\partial |E_2|} \\ \vdots \\ \frac{\partial P_{\text{ref}}}{\partial \theta_N} \\ \frac{\partial P_{\text{ref}}}{\partial |E_N|} \end{bmatrix} \quad (4.59)$$

In practice, instead of calculating J^T explicitly, we use Gaussian elimination on J^T in the same way we operate on J in the Newton power flow solution.

APPENDIX

Power Flow Input Data for Six-Bus System

Figure 4.20 lists the input data for the six-bus sample system used in the examples in Chapter 4. The impedances are per unit on a base of 100 MVA. The generation cost functions are contained in Example 4E.

Line Data

From bus	To bus	R(pu)	X(pu)	BCAP ^a (pu)
1	2	0.10	0.20	0.02
1	4	0.05	0.20	0.02
1	5	0.08	0.30	0.03
2	3	0.05	0.25	0.03
2	4	0.05	0.10	0.01
2	5	0.10	0.30	0.02
2	6	0.07	0.20	0.025
3	5	0.12	0.26	0.025
3	6	0.02	0.10	0.01
4	5	0.20	0.40	0.04
5	6	0.10	0.30	0.03

^a BCAP = half total line charging susceptance.

Bus Data

Bus number	Bus type	Voltage schedule (pu V)	P_{gen} (pu MW)	P_{load} (pu MW)	Q_{load} (pu MVAR)
1	Swing	1.05			
2	Gen.	1.05	0.50	0.0	0.0
3	Gen.	1.07	0.60	0.0	0.0
4	Load		0.0	0.7	0.7
5	Load		0.0	0.7	0.7
6	Load		0.0	0.7	0.7

FIG. 4.20 Input data for six-bus sample power system.

PROBLEMS

- 4.1 The circuit elements in the 138 kV circuit in Figure 4.21 are in per unit on a 100 MVA base with the nominal 138 kV voltage as base. The $P + jQ$ load is scheduled to be 170 MW and 50 MVAR.

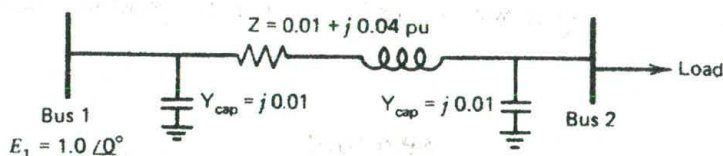


FIG. 4.21 Two-bus AC system for Problem 4.1.

- Write the Y matrix for this two-bus system.
- Assume bus 1 as the reference bus and set up the Gauss-Seidel correction equation for bus 2. (Use $1.0 \angle 0^\circ$ as the initial voltage on

- bus 2.) Carry out two or three iterations and show that you are converging.
- c. Apply the "DC" load flow conventions to this circuit and solve for the phase angle at bus 2 for the same load real power of 1.7 per unit.

4.2 Given the network in Figure 4.22 (base = 100 MVA):

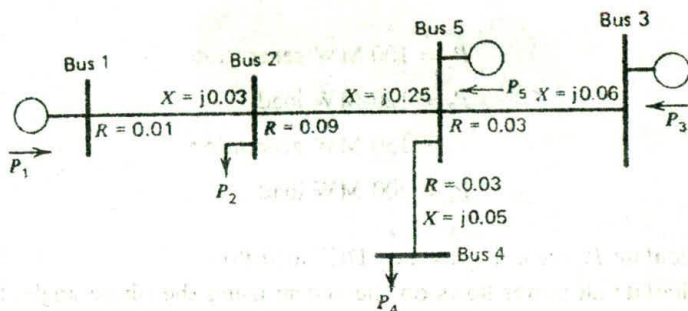


FIG. 4.22 Five-bus network for Problem 4.2.

- a. Develop the $[B']$ matrix for this system.

$$\begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \end{bmatrix} = [B'] \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \\ \theta_5 \end{bmatrix} \quad \begin{array}{l} P \text{ in per unit MW} \\ \theta \text{ in radians (rad)} \end{array}$$

- b. Assume bus 5 as the reference bus. To carry out a "DC" load flow, we will set $\theta_5 = 0$ rad. Row 5 and column 5 will be zeroed.

$$\begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \end{bmatrix} = \begin{bmatrix} & & & & 0 \\ & & & & 0 \\ \text{Remainder} & & & & \\ \text{of } B' & & & & \\ & & & & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \\ \theta_5 \end{bmatrix}$$

Solve for the $[B']^{-1}$ matrix.

$$\begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \\ \theta_5 \end{bmatrix} = [B']^{-1} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \end{bmatrix}$$

c. Calculate the phase angles for the set of power injections.

$$P_1 = 100 \text{ MW generation}$$

$$P_2 = 120 \text{ MW load}$$

$$P_3 = 150 \text{ MW generation}$$

$$P_4 = 200 \text{ MW load}$$

d. Calculate P_5 according to the "DC" load flow.

e. Calculate all power flows on the system using the phase angles found in part c.

f. (Optional) Calculate the reference-bus penalty factors for buses 1, 2, 3, and 4. Assume all bus voltage magnitudes are 1.0 per unit.

4.3 Given the following loss formula (use P values in MW):

$$B_{ij} = \begin{matrix} & \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \begin{bmatrix} 1.36255 \times 10^{-4} & 1.753 \times 10^{-5} & 1.8394 \times 10^{-4} \\ 1.754 \times 10^{-5} & 1.5448 \times 10^{-4} & 2.82765 \times 10^{-4} \\ 1.8394 \times 10^{-4} & 2.82765 \times 10^{-4} & 1.6147 \times 10^{-3} \end{bmatrix} \end{matrix}$$

B_{i0} and B_{00} are neglected. Assume three units are on-line and have the following characteristics.

Unit 1: $H_1 = 312.5 + 8.25P_1 + 0.005P_1^2$, MBtu/h

$$50 \leq P_1 \leq 250 \text{ MW}$$

$$\text{Fuel cost} = 1.05 \text{ R/MBtu}$$

Unit 2: $H_2 = 112.5 + 8.25P_2 + 0.005P_2^2$, MBtu/h

$$5 \leq P_2 \leq 150 \text{ MW}$$

$$\text{Fuel cost} = 1.217 \text{ R/MBtu}$$

Unit 3: $H_3 = 50 + 8.25P_3 + 0.005P_3^2$, MBtu/h
 $15 \leq P_3 \leq 100$ MW
 Fuel cost = 1.1831 R/MBtu

a. No Losses Used in Scheduling

- Calculate the optimum dispatch and total cost neglecting losses for $P_D = 190$ MW.*
- Using this dispatch and the loss formula, calculate the system losses.

b. Losses Included in Scheduling

- Find the optimum dispatch for a total generation of $P_D = 190$ MW* using the coordination equations and the loss formula.
- Calculate the cost rate.
- Calculate the total losses using the loss formula.
- Calculate the resulting load supplied.

4.4 All parts refer to the three-bus system shown in Figure 4.23.

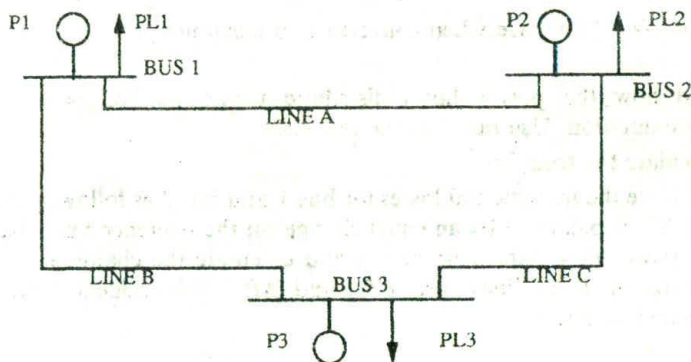


FIG. 4.23 Network for Problem 4.4.

Data for this problem is as follows:

Unit 1: $P_1 = 570$ MW

Unit 2: $P_2 = 330$ MW

Unit 3: $P_3 = 200$ MW

Loads:

$P_{L1} = 200$ MW

$P_{L2} = 400$ MW

$P_{L3} = 500$ MW

* $P_{\text{demand}} = P_1 + P_2 + P_3 = P_D$

P_{loss} = power loss

$P_{\text{load}} = P_D - P_{\text{loss}}$ = net load

Transmission line data:

$$P_{\text{loss}} \text{ in line A} = 0.02P_A^2 \text{ (where } P_A = P \text{ flow from bus 1 to bus 2)}$$

$$P_{\text{loss}} \text{ in line B} = 0.02P_B^2 \text{ (where } P_B = P \text{ flow from bus 1 to bus 3)}$$

$$P_{\text{loss}} \text{ in line C} = 0.02P_C^2 \text{ (where } P_C = P \text{ flow from bus 2 to bus 3)}$$

Note: the above data are for P_{loss} in per unit when power flows P_A or P_B or P_C are in per unit.

Line reactances:

$$X_A = 0.2 \text{ per unit}$$

$$X_B = 0.3333 \text{ per unit}$$

$$X_C = 0.05 \text{ per unit}$$

(assume 100-MVA base when converting to per unit).

- Find how the power flows distribute using the DC power flow approximation. Use bus 3 as the reference.
- Calculate the total losses.
- Calculate the incremental losses for bus 1 and bus 2 as follows: assume that ΔP_1 is balanced by an equal change on the reference bus. Use the DC power flow data from part a and calculate the change in power flow on all three lines ΔP_A , ΔP_B , and ΔP_C . Now calculate the line incremental loss as:

$$\Delta P_{\text{loss}_A} = \left(\frac{\partial}{\partial P_A} \right) \Delta P_A = (0.04P_A) \Delta P_A$$

Similarly, calculate for lines B and C.

- Find the bus penalty factors calculated from the line incremental losses found in part c.
- 4.5 The three-bus, two-generator power system shown in Figure 4.24 is to be dispatched to supply the 500-MW load. Each transmission line has losses

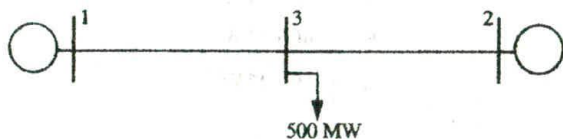


FIG. 4.24 Circuit for Problem 4.5

that are given by the equations below.

$$P_{\text{loss}_{1,3}} = 0.0001P_1^2$$

$$P_{\text{loss}_{2,3}} = 0.0002P_2^2$$

$$F_1(P_1) = 500 + 8P_1 + 0.002P_1^2$$

$$50 \text{ MW} < P_1 < 500 \text{ MW}$$

$$F_2(P_2) = 400 + 7.9P_2 + 0.0025P_2^2$$

$$50 \text{ MW} < P_2 < 500 \text{ MW}$$

You are to attempt to solve for both the economic dispatch of this system and the "power flow." The power flow should show what power enters and leaves each bus of the network. If you use an iterative solution, show at least two complete iterations. You may use the following initial conditions: $P_1 = 250 \text{ MW}$ and $P_2 = 250 \text{ MW}$.

FURTHER READING

The basic papers on solution of the power flow can be found in references 1-5. The development of the loss-matrix equations is based on the work of Kron (reference 6), who developed the reference-frame transformation theory. Other developments of the transmission-loss formula are seen in references 7 and 8. Meyer's paper (9) is representative of recent adaptation of sparsity programming methods to calculation of the loss matrix.

The development of the reference-bus penalty factor method can be seen in references 10 and 11. Reference 12 gives an excellent derivation of the reference-bus penalty factors derived from the Newton power-flow equations. Reference 12 provides an excellent summary of recent developments in power system dispatch.

1. Ward, J. B., Hale, H. W., "Digital Computer Solution of Power-Flow Problems," *AIEE Transactions, Part III Power Apparatus and Systems*, Vol. 75, June 1956, pp. 398-404.
2. VanNess, J. E., "Iteration Methods for Digital Load Flow Studies," *AIEE Transactions on Power Apparatus and Systems*, Vol. 78A, August 1959, pp. 583-588.
3. Tinney, W. F., Hart, C. E., "Power Flow Solution by Newton's Method," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-86, November 1967, pp. 1449-1460.
4. Stott, B., Alsac, O., "Fast Decoupled Load Flow," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-93, May/June 1974, pp. 859-869.
5. Stott, B., "Review of Load-Flow Calculation Methods," *Proceedings of the IEEE*, Vol. 62, No. 2, July 1974, pp. 916-929.
6. Kron, G., "Tensorial Analysis of Integrated Transmission Systems—Part I: The Six Basic Reference Frames," *AIEE Transactions*, Vol. 70, Part I, 1951, pp. 1239-1248.
7. Kirchmayer, L. K., Stagg, G. W., "Analysis of Total and Incremental Losses in Transmission Systems," *AIEE Transactions*, Vol. 70, Part I, 1951, pp. 1179-1205.

8. Early, E. D., Watson, R. E., "A New Method of Determining Constants for the General Transmission Loss Equation," *AIEE Transactions on Power Apparatus and Systems*, Vol. PAS-74, February 1956, pp. 1417-1423.
9. Meyer, W. S., "Efficient Computer Solution for Kron and Kron Early-Loss Formulas," *Proceedings of the 1973 PICA Conference*, IEEE 73 CHO 740-1, PWR, pp. 428-432.
10. Shipley, R. B., Hochdorf, M., "Exact Economic Dispatch—Digital Computer Solution," *AIEE Transactions on Power Apparatus and Systems*, Vol. PAS-75, November 1956, pp. 1147-1152.
11. Dommel, H. W., Tinney, W. F., "Optimal Power Flow Solutions," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-87, October 1968, pp. 1866-1876.
12. Happ, H. H., "Optimal Power Dispatch," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-93, May/June 1974, pp. 820-830.

5 Unit Commitment

5.1 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 5A, 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.

EXAMPLE 5A

Suppose one had the three units given here:

Unit 1:	Min = 150 MW
	Max = 600 MW
	$H_1 = 510.0 + 7.2P_1 + 0.00142P_1^2$ MBtu/h
Unit 2:	Min = 100 MW
	Max = 400 MW
	$H_2 = 310.0 + 7.85P_2 + 0.00194P_2^2$ MBtu/h
Unit 3:	Min = 50 MW
	Max = 200 MW
	$H_3 = 78.0 + 7.97P_3 + 0.00482P_3^2$ MBtu/h

with fuel costs:

$$\text{Fuel cost}_1 = 1.1 \text{ R/MBtu}$$

$$\text{Fuel cost}_2 = 1.0 \text{ R/MBtu}$$

$$\text{Fuel cost}_3 = 1.2 \text{ R/MBtu}$$

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 5.1.

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 5.1a. 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 5B, where a simple priority list scheme is developed.

TABLE 5.1 Unit Combinations and Dispatch for 550-MW Load of Example 5A

Unit 1	Unit 2	Unit 3	Max Generation	Min Generation	P_1	P_2	P_3	F_1	F_2	F_3	Total Generation 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

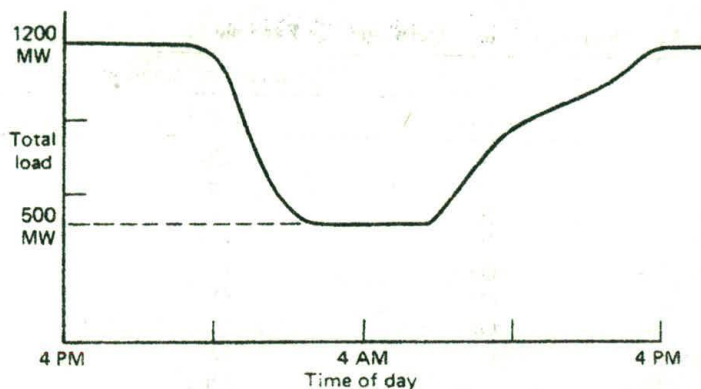


FIG. 5.1a Simple "peak-valley" load pattern.

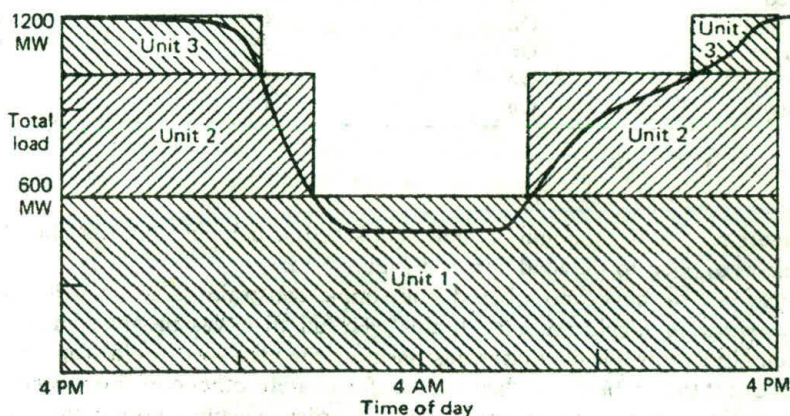


FIG. 5.1b Unit commitment schedule using shut-down rule.

EXAMPLE 5B

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 5A, with the load varying from a peak of 1200 MW to a valley of 500 MW. To obtain a "shut-down rule," simply use a brute-force technique wherein all combinations of units will be tried (as in Example 5A) 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 5.2. Our shut-down rule is quite simple.

When load is above 1000 MW, run all three units; between 1000 MW and 600 MW, run units 1 and 2; below 600 MW, run only unit 1.

TABLE 5.2 "Shut-down Rule" Derivation for Example 5B

Load	Optimum Combination		
	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 5.1b shows the unit commitment schedule derived from this shut-down rule as applied to the load curve of Figure 5.1a.

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.

5.1.1 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.

5.1.2 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 9). 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 generation-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 9) 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 5C

Suppose a power system consisted of two isolated regions: a western region and an eastern region. Five units, as shown in Figure 5.2, 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 5.2. What can we say about the allocation of spinning reserve in this system?

The data for the system in Figure 5.2 are given in Table 5.3. 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 reserves. However, the tie capacity of only 550 MW limits the transfer. Therefore, the loss of unit 4 cannot

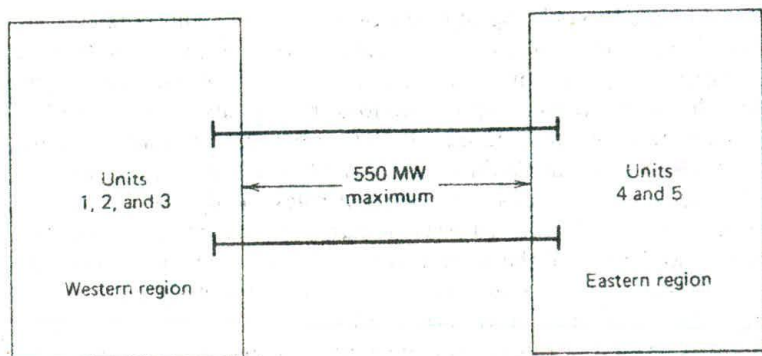


FIG. 5.2 Two-region system.

TABLE 5.3 Data for the System in Figure 5.2

Region	Unit	Unit Capacity (MW)	Unit Output (MW)	Regional Generation (MW)	Spinning Reserve	Regional Load (MW)	Interchange (MW)
Western	1	1000	900	1740	100	1900	160 in
	2	800	420		380		
	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	

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.

5.1.3 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 up time:** once the unit is running, it should not be turned off immediately.
- **Minimum down time:** 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.

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.

$$\text{Start-up cost when cooling} = C_c(1 - e^{-t/\alpha}) \times F + C_f$$

where

C_c = cold-start cost (MBtu)

F = fuel cost

C_f = fixed cost (includes crew expense, maintenance expenses) (in $\text{\$}$)

α = thermal time constant for the unit

t = time (h) the unit was cooled

$$\text{Start-up cost when banking} = C_i \times t \times F + C_f$$

where

C_i = 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 5.3.

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.

5.1.4 Other Constraints

5.1.4.1 Hydro-Constraints

Unit commitment cannot be completely separated from the scheduling of hydro-units. In this text, 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.

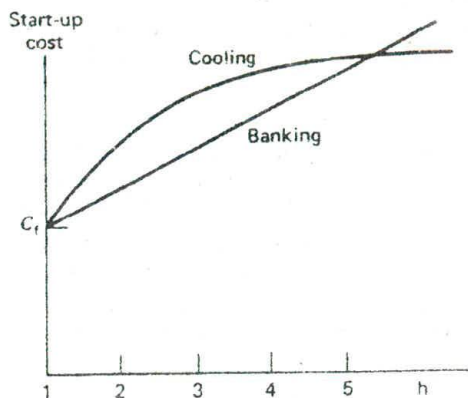


FIG. 5.3 Time-dependent start-up costs.

5.1.4.2 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.

5.1.4.3 Fuel Constraints

We will treat the "fuel scheduling" problem briefly in Chapter 6. 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.

5.2 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 units to commit and dispatch.
- The M load levels and operating limits on the N 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,

$$C(N, 1) + C(N, 2) + \dots + C(N, N - 1) + C(N, N) = 2^N - 1$$

where $C(N, j)$ is the combination of N items taken j at a time. That is,

$$C(N, j) = \left[\frac{N!}{(N-j)!j!} \right]$$

$$j! = 1 \times 2 \times 3 \times \dots \times j$$

For the total period of M intervals, the maximum number of possible combinations is $(2^N - 1)^M$, which can become a horrid number to think about.

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 $(2^N - 1)^{24}$ becomes the following.

N	$(2^N - 1)^{24}$
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 relation (LR).

5.2.1 Priority-List Methods

The simplest unit commitment solution method consists of creating a priority list of units. As we saw in Example 5B, a simple shut-down 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 5B 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 5D

Construct a priority list for the units of Example 5A. (Use the same fuel costs as in Example 5A.) First, the full-load average production cost will be calculated:

Unit	Full Load Average Production Cost (R/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:

Unit	R/MWh	Min MW	Max MW
2	9.48	100	400
1	9.79	150	600
3	11.188	50	200

and the commitment scheme would (ignoring min up/down time, start-up costs, etc.) simply use only the following combinations.

Combination	Min MW from Combination	Max MW from Combination
2 + 1 + 3	300	1200
2 + 1	250	1000
2	100	400

Note that such a scheme would not completely parallel the shut-down sequence described in Example 5B, 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 shut-down 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 shut-down time for the unit, keep commitment as is and go to last step; if not, go to 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 dynamic-programming methods usually create the same type of priority list for use in the DP search.

5.2.2 Dynamic-Programming Solution

5.2.2.1 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

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 zero.
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 dynamic-programming 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.

5.2.2.2 Forward DP Approach

One could set up a dynamic-programming 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 generator 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 dynamic-program 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 dynamic-programming algorithm is shown by the flowchart in Figure 5.4.

The recursive algorithm to compute the minimum cost in hour K with combination I is,

$$F_{\text{cost}}(K, I) = \min_{(L)} [P_{\text{cost}}(K, I) + S_{\text{cost}}(K-1, L; K, I) + F_{\text{cost}}(K-1, L)] \quad (5.1)$$

where

$F_{\text{cost}}(K, I)$ = least total cost to arrive at state (K, I)

$P_{\text{cost}}(K, I)$ = production cost for state (K, I)

$S_{\text{cost}}(K-1, L; K, I)$ = transition cost from state $(K-1, L)$ to state (K, I)

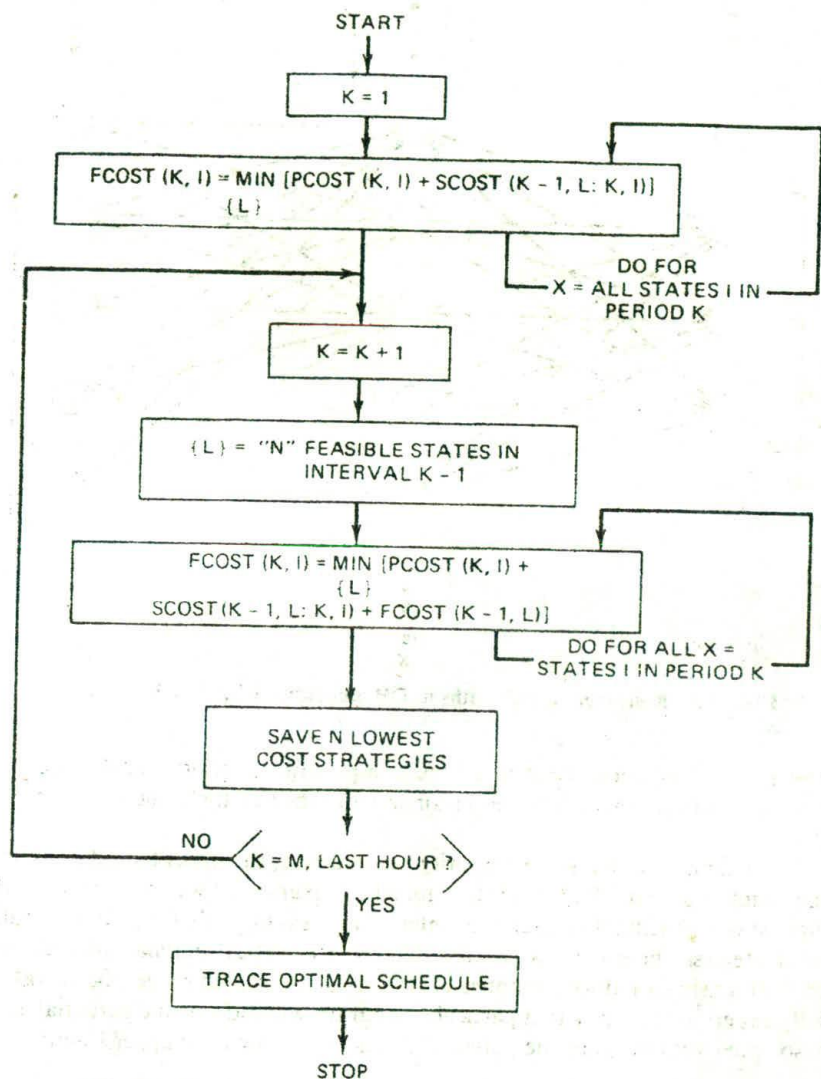


FIG. 5.4 Unit commitment via forward dynamic programming.

State (K, I) is the I^{th} combination in hour K . For the forward dynamic-programming 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 5.4.

X = number of states to search each period

N = number of strategies, or paths, to save at each step

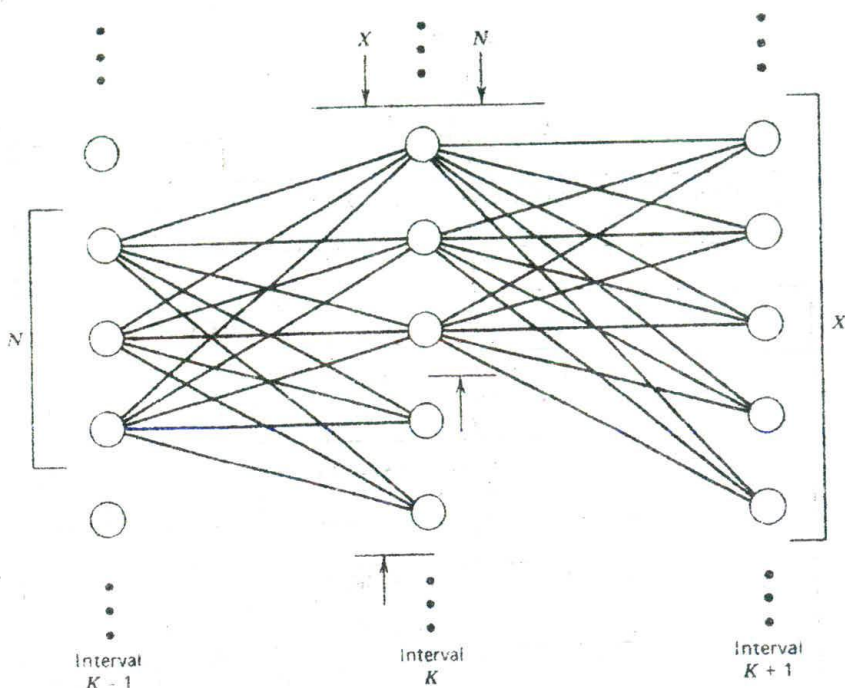


FIG. 5.5 Restricted search paths in DP algorithm with $N = 3$ and $X = 5$.

These variables allow control of the computational effort (see Figure 5.5). For complete enumeration, the maximum number of the value of X or N is $2^n - 1$.

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 5E

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 down times. The third case includes the hot-start costs, as well as the minimum up and down times. 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 5.4.

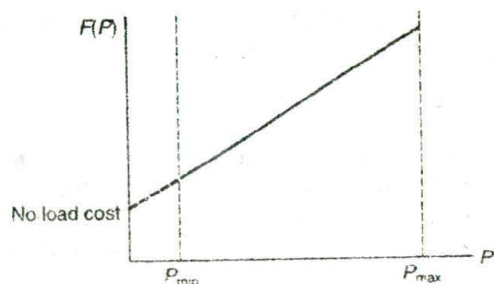
TABLE 5.4 Unit Characteristics, Load Pattern, and Initial Status for the Cases in Example 5E

Unit	Max (MW)	Min (MW)	Incremental Heat Rate (Btu/kWh)	No-Load Cost (R/h)	Full-Load Ave. Cost (R/mWh)	Minimum Times (h)	
						Up	Down
1	80	25	10440	213.00	23.54	4	2
2	250	60	9000	585.62	20.34	5	3
3	300	75	8730	684.74	19.74	5	4
4	60	20	11900	252.00	28.00	1	1

Unit	Initial Conditions		Start-Up Costs		
	Hours Off-Line (-) or On-Line (+)		Hot (R)	Cold (R)	Cold Start (h)
1	-5		150	350	4
2	8		170	400	5
3	8		500	1100	5
4	-6		0	0.02	0

Load Pattern	
Hour	Load (MW)
1	450
2	530
3	600
4	540
5	400
6	280
7	290
8	500

In order to make the required computations more efficiently, a simplified model of the unit characteristics is used. In practical applications, two- or three-section stepped incremental curves might be used, as shown in Figure 5.6. For our example, only a single step between minimum and the maximum power points is used. The units in this example have linear $F(P)$ functions:



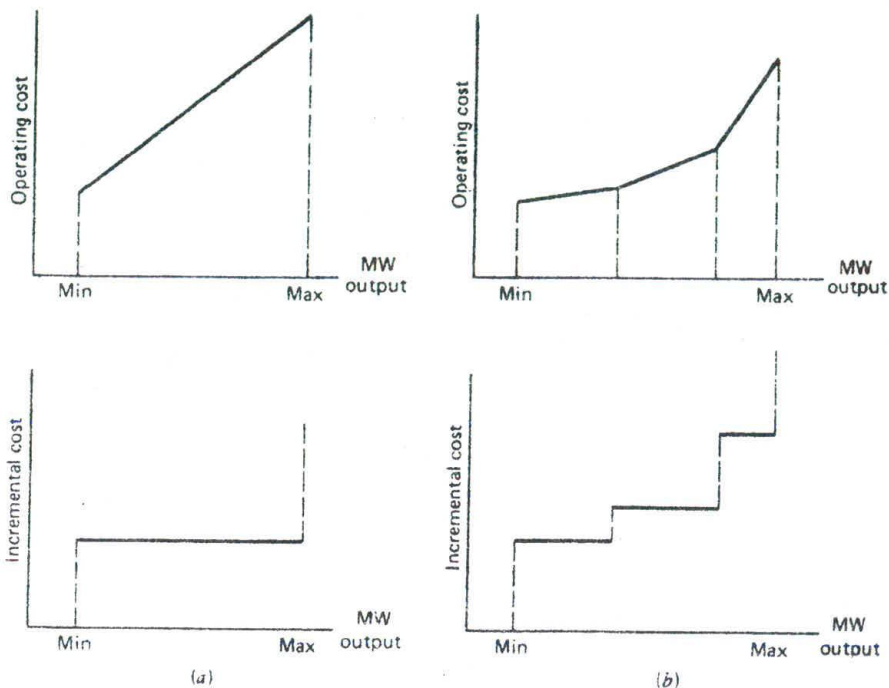


FIG. 5.6 (a) Single-step incremental cost curve and (b) multiple-step incremental cost curve.

The $F(P)$ function is:

$$F(P) = \text{No-load cost} + \text{Inc cost} \times P$$

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: unit 3, unit 2, unit 1, unit 4. For the first two cases, the minimum up and down times 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 5.5, 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.

TABLE 5.5 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).

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

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, and so on. 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

$$F_{\text{cost}}(J, K) = \min_{\{L\}} [P_{\text{cost}}(J, K) + S_{\text{cost}}(J-1, L: J, K) + F_{\text{cost}}(J-1, L)]$$

Allowable states are

$$\{ \} = \{0010, 0110, 1110, 1111\} = \{5, 12, 14, 15\}$$

In hour 0 $\{L\} = \{12\}$, initial condition.

J = 1: 1st hour

$$\begin{array}{l} \frac{K}{15} \\ 15 \end{array} \quad F_{\text{cost}}(1, 15) = P_{\text{cost}}(1, 15) + S_{\text{cost}}(0, 12: 1, 15) \\ \quad \quad \quad = 9861 + 350 = 10211$$

$$14 \quad F_{\text{cost}}(1, 14) = 9493 + 350 = 9843$$

$$12 \quad F_{\text{cost}}(1, 12) = 9208 + 0 = 9208$$

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 $\{L\} = \{12, 14\}$,

$$\begin{array}{l} \frac{K}{15} \\ 15 \end{array} \quad F_{\text{cost}}(2, 15) = \min_{\{12, 14\}} [P_{\text{cost}}(2, 15) + S_{\text{cost}}(1, L: 2, 15) + F_{\text{cost}}(1, L)] \\ \quad \quad \quad = 11301 + \min \left[\begin{array}{l} (350 + 9208) \\ (0 + 9843) \end{array} \right] = 20859$$

and so on.

Case 2

In Case 2, complete enumeration is tried with a limit of $(2^4 - 1) = 15$ dispatches each of the eight hours, so that there is a theoretical maximum of $15^8 = 2.56 \cdot 10^9$ possibilities. Fortunately, most of these are not feasible because they do not supply sufficient capacity, and can be discarded with little analysis required.

Figure 5.7 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.

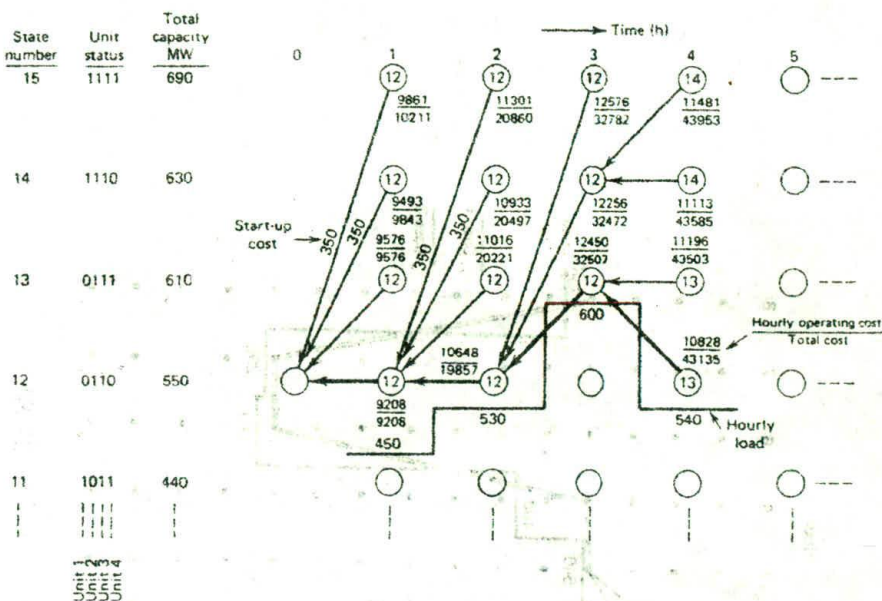


FIG. 5.7 Example 5E, Cases 1 and 2 (first 4 h).

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 R165, or R0.104/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 R2.00 MBtu. A savings of R165 every 3 h is equivalent to R481,800/yr.

The total 8-h trajectories for Cases 1 and 2 are shown in Figure 5.8. The neglecting of start-up and shut-down 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.

Case 3

In case 3, the original unit data are used so that the minimum shut-down and operating times are observed. The forward dynamic-programming 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 5.9. However, when only four strategies were saved, the procedure flounders (i.e., fails to find a feasible path) in

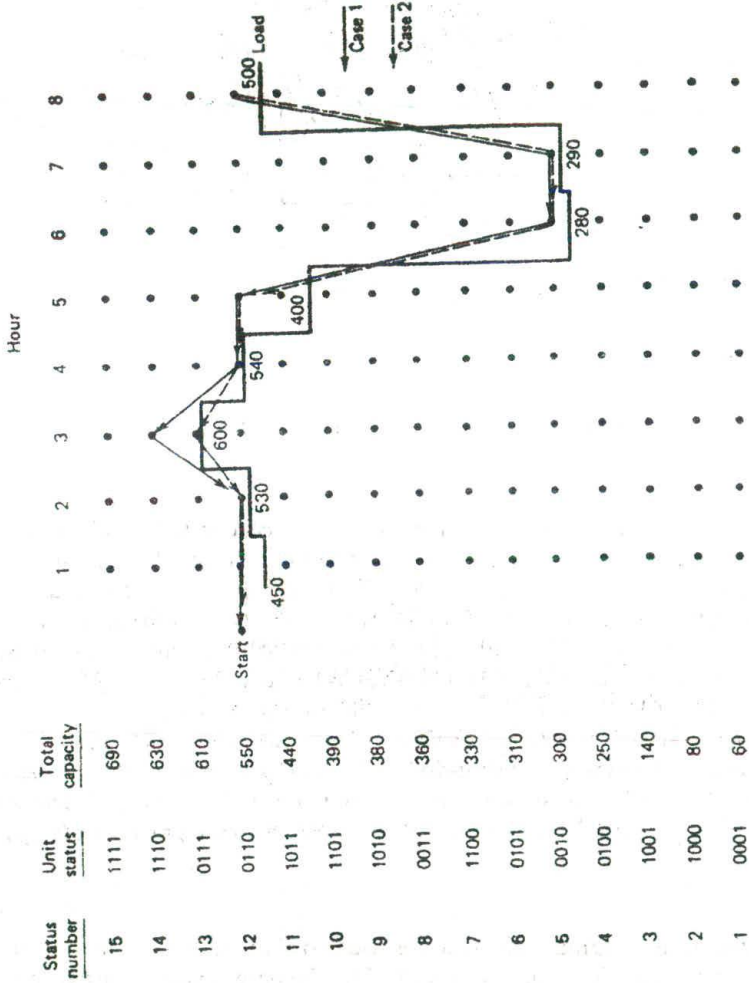


FIG. 5.8 Example 5E, Cases 1 and 2 (complete solutions).

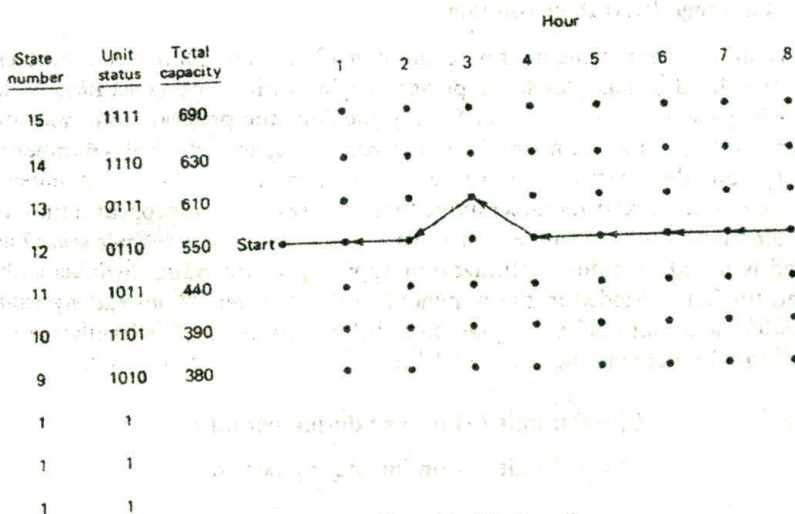


FIG. 5.9 Example 5E, Case 3.

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 5.4 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 5.6. These cases illustrate the forward dynamic-programming method and also point out the problems involved in the practical application of the method.

TABLE 5.6 Summary of Cases 1-3

Case	Conditions	Total Cost (R)
1	Priority order. Up and down times neglected	73439
2	Enumeration ($X \leq 15$) with 4 strategies (N) saved, Up and down times neglected	73274
3	$X \leq 15$. Up and down times observed	No solution
	$N = 4$ strategies	74110
	$B = 8$ strategies	74110
	$N = 10$ strategies	74110

5.2.3 Lagrange Relaxation Solution

The dynamic-programming 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 of forcing the dynamic-programming 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 3A and further expanded in the appendix 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:

$$U_i^t = 0 \text{ if unit } i \text{ is off-line during period } t$$

$$U_i^t = 1 \text{ if unit } i \text{ is on-line during period } t$$

We shall now define several constraints and the objective function of the unit commitment problem:

1. Loading constraints:

$$P_{\text{load}}^t - \sum_{i=1}^N P_i^t U_i^t = 0 \quad \text{for } t = 1 \dots T \quad (5.2)$$

2. Unit limits:

$$U_i^t P_i^{\text{min}} \leq P_i^t \leq U_i^t P_i^{\text{max}} \quad \text{for } i = 1 \dots N \quad \text{and } t = 1 \dots T \quad (5.3)$$

3. Unit minimum up- and down-time constraints. Note that other constraints can easily be formulated and added to the unit commitment problem. These include transmission security constraints (see Chapter 11), 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:

$$\sum_{t=1}^T \sum_{i=1}^N [F_i(P_i^t) + \text{Start up cost}_{i,t}] U_i^t = F(P^t, U^t) \quad (5.4)$$

We can then form the Lagrange function similar to the way we did in the economic dispatch problem:

$$\mathcal{L}(P, U, \lambda) = F(P^t, U^t) + \sum_{t=1}^T \lambda^t \left(P_{\text{load}}^t - \sum_{i=1}^N P_i^t U_i^t \right) \quad (5.5)$$

The unit commitment problem requires that we minimize the Lagrange function

above, subject to the local unit constraints 2 and 3, which can be applied to each unit separately. Note:

1. The cost function, $F(P_i^t, U_i^t)$, together with constraints 2 and 3 are each *separable over units*. That is, what is done with one unit does not affect the cost of running another unit, as far as the cost function and the unit limits (constraint 2) and the unit up- and down-time (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 the appendix 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:

$$q^*(\lambda) = \max_{\lambda^t} q(\lambda) \quad (5.6)$$

where

$$q(\lambda) = \min_{P_i^t, U_i^t} \mathcal{L}(P, U, \lambda) \quad (5.7)$$

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:

$$\mathcal{L} = \sum_{t=1}^T \sum_{i=1}^N [F_i(P_i^t) + \text{Start up cost}_{i,t}] U_i^t + \sum_{t=1}^T \lambda^t \left(P_{\text{load}}^t - \sum_{i=1}^N P_i^t U_i^t \right) \quad (5.8)$$

This is now rewritten as:

$$\mathcal{L} = \sum_{t=1}^T \sum_{i=1}^N [F_i(P_i^t) + \text{Start up cost}_{i,t}] U_i^t + \sum_{t=1}^T \lambda^t P_{\text{load}}^t - \sum_{t=1}^T \sum_{i=1}^N \lambda^t P_i^t U_i^t \quad (5.9)$$

The second term above is constant and can be dropped (since the λ^t are fixed). Finally, we write the Lagrange function as:

$$\mathcal{L} = \sum_{i=1}^N \left(\sum_{t=1}^T \{ [F_i(P_i^t) + \text{Start up cost}_{i,t}] U_i^t - \lambda^t P_i^t U_i^t \} \right) \quad (5.10)$$

Here, we have achieved our goal of separating the units from one another. The term inside the outer brackets; that is:

$$\sum_{t=1}^T \{ [F_i(P_i^t) + \text{Start up cost}_{i,t}] U_i^t - \lambda^t P_i^t U_i^t \}$$

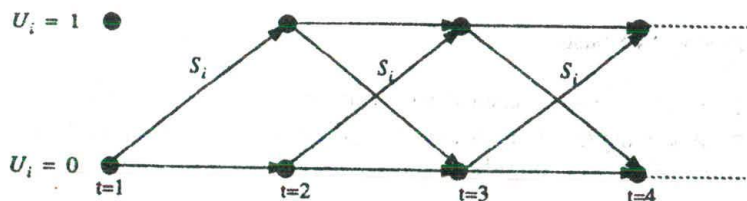
can be solved separately for each generating unit, without regard for what is happening on the other generating units. The minimum of the Lagrangian is found by solving for the minimum for each generating unit over all time periods; that is:

$$\min q(\lambda) = \sum_{i=1}^N \min_{t=1}^T \{ [F_i(P_i^t) + \text{Start up cost}_{i,t}] U_i^t - \lambda^t P_i^t U_i^t \} \quad (5.11)$$

Subject to

$$U_i^t P_i^{\min} \leq P_i^t \leq U_i^t P_i^{\max} \quad \text{for } t = 1 \dots T$$

and the up- and down-time constraints. This is easily solved as a dynamic-programming problem in one variable. This can be visualized in the figure below, 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 zero); 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):

$$\min [F_i(P_i^t) - \lambda^t P_i^t] \quad (5.12)$$

The minimum of this function is found by taking the first derivative:

$$\frac{d}{dP_i^t} [F_i(P_i^t) - \lambda^t P_i^t] = \frac{d}{dP_i^t} F_i(P_i^t) - \lambda^t = 0 \quad (5.13)$$

The solution to this equation is

$$\frac{d}{dP_i^t} F_i(P_i^{\text{opt}}) = \lambda^t \quad (5.14)$$

There are three cases to be concerned with depending on the relation of P_i^{opt} and the unit limits:

1. If $P_i^{\text{opt}} \leq P_i^{\text{min}}$ then:

$$\min [F_i(P_i) - \lambda^t P_i^t] = F_i(P_i^{\text{min}}) - \lambda^t P_i^{\text{min}} \quad (5.15a)$$

2. If $P_i^{\text{min}} \leq P_i^{\text{opt}} \leq P_i^{\text{max}}$, then:

$$\min [F_i(P_i) - \lambda^t P_i^t] = F_i(P_i^{\text{opt}}) - \lambda^t P_i^{\text{opt}} \quad (5.15b)$$

3. If $P_i^{\text{opt}} \geq P_i^{\text{max}}$, then:

$$\min [F_i(P_i) - \lambda^t P_i^t] = F_i(P_i^{\text{max}}) - \lambda^t P_i^{\text{max}} \quad (5.15c)$$

The solution of the two-state dynamic program for each unit proceeds in the normal manner as was done for the forward dynamic-programming solution of the unit commitment problem itself. Note that since we seek to minimize $[F_i(P_i) - \lambda^t P_i^t]$ at each stage and that when $U_i^t = 0$ this value goes to zero, then the only way to get a value lower is to have

$$[F_i(P_i) - \lambda^t P_i^t] < 0$$

The dynamic program should take into account all the start-up costs, S_i , for each unit, as well as the minimum up and down time for the generator. Since we are solving for each generator independently, however, we have avoided the dimensionality problems that affect the dynamic-programming solution.

5.2.3.1 Adjusting λ

So far, we have shown how to schedule generating units with fixed values of λ^t for each time period. As shown in the appendix to this chapter, the adjustment of λ^t must be done carefully so as to maximize $q(\lambda)$. Most references to work on the Lagrange relaxation procedure use a combination of gradient search and various heuristics to achieve a rapid solution. Note that unlike in the appendix, 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 5D, 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 5D, however, the λ adjustment factors are different:

$$\lambda^t = \lambda^t + \left[\frac{d}{d\lambda} q(\lambda) \right] \alpha \quad (5.16)$$

where

$$\alpha = 0.01 \quad \text{when} \quad \frac{d}{d\lambda} q(\lambda) \text{ is positive} \quad (5.17)$$

and

$$\alpha = 0.002 \quad \text{when} \quad \frac{d}{d\lambda} q(\lambda) \text{ is negative} \quad (5.18)$$

Each λ^i is treated separately. The reader should consult the references listed at the end of this chapter for more efficient methods of adjusting the λ values. The overall Lagrange relaxation unit commitment algorithm is shown in Figure 5.10.

Reference 15 introduces the use of what this text called the "relative duality gap" or $(J^* - q^*)/q^*$. The relative duality gap is used in Example 5D as a measure of the closeness to the solution. Reference 15 points out several useful things about dual optimization applied to the unit commitment problem.

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 the above are demonstrated in Example 5D. 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 dynamic-programming algorithm to get a good starting point, then run the dual optimization iterations, and finally, at the end, they use heuristic logic or restricted dynamic programming to get to a final solution. The result is a solution that is not limited to search windows, such as had to be done in strict application of dynamic programming.

EXAMPLE 5D

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

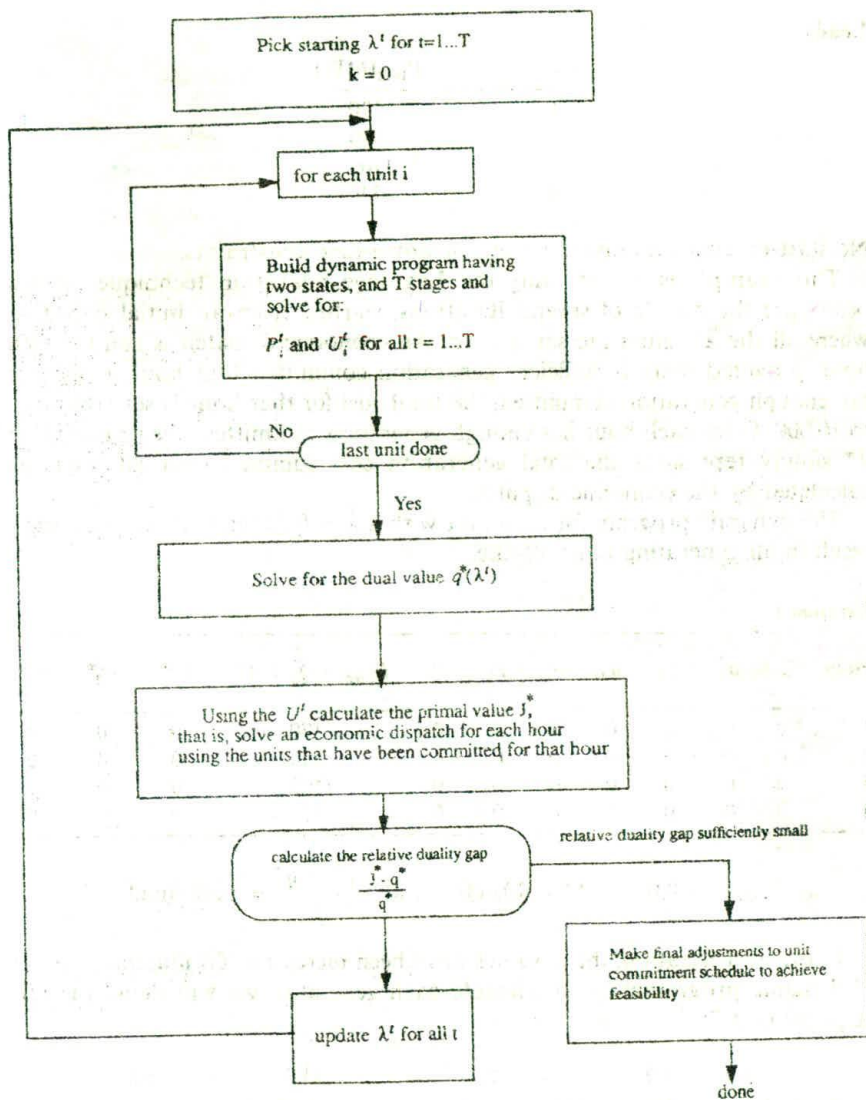


FIG. 5.10 Lagrange relaxation procedure for unit commitment.

units below:

$$F_1(P_1) = 500 + 10P_1 + 0.002P_1^2 \quad \text{and} \quad 100 < P_1 < 600$$

$$F_2(P_2) = 300 + 8P_2 + 0.0025P_2^2 \quad \text{and} \quad 100 < P_2 < 400$$

$$F_3(P_3) = 100 + 6P_3 + 0.005P_3^2 \quad \text{and} \quad 50 < P_3 < 200$$

Load:

t	P_{load}^t (MW)
1	170
2	520
3	1100
4	330

No start-up costs, no minimum up- or down-time constraints.

This example is solved using the Lagrange relaxation technique. Shown below are the results of several iterations, starting from an initial condition where all the λ^t values are set to zero. 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.

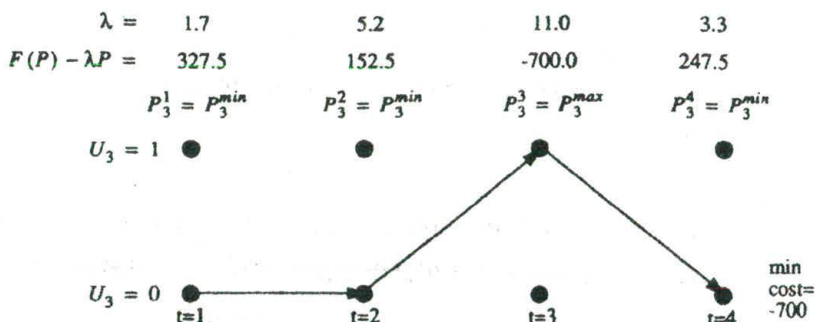
The dynamic program for each unit with a $\lambda^t = 0$ for each hour will always result in all generating units off-line.

Iteration 1

Hour	λ	u_1	u_2	u_3	P_1	P_2	P_3	$P_{load}^t - \sum_{i=1}^N P_i U_i^t$	P_1^{dc}	P_2^{dc}	P_3^{dc}
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 dynamic programming to schedule each generator, we will detail the DP steps for unit 3:



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	$P_{\text{load}}^i - \sum_{i=1}^N P_i^i U_i^i$	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

$$q(\lambda) = 14,982, \quad J^* = 40,000, \quad \text{and} \quad \frac{J^* - q^*}{q^*} = 1.67$$

Iteration 3

Hour	λ	u_1	u_2	u_3	P_1	P_2	P_3	$P_{\text{load}}^i - \sum_{i=1}^N P_i^i U_i^i$	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	$P_{\text{load}}^i - \sum_{i=1}^N P_i^i U_i^i$	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

$$q(\lambda) = 19,532, \quad J^* = 36,024, \quad \text{and} \quad \frac{J^* - q^*}{q^*} = 0.844$$

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 zero, 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).

APPENDIX

Dual Optimization on a Nonconvex Problem

We introduced the concept of dual optimization in Appendix 3A 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 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 text in Section 5.2.3. 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 below. It illustrates the ability of the dual optimization technique to solve the unit commitment problem. Given:

$$J(x_1, x_2, u_1, u_2) = (0.25x_1^2 + 15)u_1 + (0.255x_2^2 + 15)u_2 \quad (5A.1)$$

subject to:

$$\omega = 5 - x_1u_1 - x_2u_2 \quad (5A.2)$$

and

$$0 \leq x_1 \leq 10 \quad (5A.3)$$

$$0 \leq x_2 \leq 10 \quad (5A.4)$$

where x_1 and x_2 are continuous real numbers, and:

$$u_1 = 1 \quad \text{or} \quad 0$$

$$u_2 = 1 \quad \text{or} \quad 0$$

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 zero, 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:

$$\mathcal{L}(x_1, x_2, \lambda) = (0.25x_1^2 + 15) + (0.255x_2^2 + 15) + \lambda(5 - x_1 - x_2) \quad (5A.5)$$

The resulting optimum is at $x_1 = 2.5248$, $x_2 = 2.4752$, and $\lambda = 1.2642$, 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 one above, as follows. Define the Lagrange function as:

$$\mathcal{L}(x_1, x_2, u_1, u_2, \lambda) = (0.25x_1^2 + 15)u_1 + (0.255x_2^2 + 15)u_2 + \lambda(5 - x_1u_1 - x_2u_2) \quad (5A.6)$$

As shown in Appendix 3A, we define $q(\lambda)$ as:

$$q(\lambda) = \min_{x_1, x_2, u_1, u_2} \mathcal{L} \quad (5A.7)$$

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

$$q^*(\lambda) = \max_{\lambda \geq 0} q(\lambda) \quad (5A.8)$$

This is different from the dual optimization approach used in the Appendix 3A 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 the Appendix 3A.

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

$$(0.25x_1^2 + 15)u_1 + (0.255x_2^2 + 15)u_2 + \lambda^k(5 - x_1u_1 - x_2u_2)$$

where the value of λ^k is fixed.

We can then rearrange the equation above 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 above 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 zero (that is with $u = 0$) or it will be negative, with $u = 1$ and 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):

$$\frac{d}{dx_1} (0.25x_1^2 + 15 - x_1\lambda^k) = 0 \quad (5A.9)$$

If the value of x_1 which satisfies the above 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):

$$\frac{d}{dx_2} (0.255x_2^2 + 15 - x_2\lambda^k) = 0 \quad (5A.10)$$

and if the value of x_2 which satisfies the above value falls outside the 0 to 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 \mathcal{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

$$\nabla q = \frac{dq}{d\lambda} \quad (5A.11)$$

which for our problem is

$$\frac{dq}{d\lambda} = 5 - x_1u_1 - x_2u_2 \quad (5A.12)$$

we adjust λ according to

$$\lambda^{k+1} = \lambda^k + \frac{dq}{d\lambda} \alpha \quad (5A.13)$$

where α is a multiplier chosen to move λ only a short distance. (This is simply a gradient search method as was introduced in Chapter 3). Note also that if both u_1 and u_2 are zero, the gradient will be 5, indicating a positive value telling us to increase λ . Eventually, increasing λ will result in a negative value for

$$(0.25x_1^2 + 15 - x_1\lambda^k)$$

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 above, 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 λ^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

$$[(0.25x_1^2 + 15)u_1 + (0.255x_2^2 + 15)u_2 + \lambda(5 - x_1u_1 - x_2u_2)]$$

using the techniques in Appendix 3A (since the u_1 and u_2 variables are now known).

The solution to this minimum will be at $x_1 = \bar{x}_1$, $x_2 = \bar{x}_2$ and $\lambda = \bar{\lambda}$. For the case where u_1 and u_2 are both zero, we shall arbitrarily set this value to a large value (here we set it to 50). We shall call this minimum value

TABLE 5.7 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

$J^*(\bar{x}_1, \bar{x}_2, u_1, u_2)$ 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 zero, 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 5.7. 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 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 5.7.

PROBLEMS

- 5.1 Given the unit data in Tables 5.8 and 5.9, use forward dynamic-programming to find the optimum unit commitment schedules covering the 8-h period. Table 5.9 gives all the combinations you need, as well as the operating cost for each at the loads in the load data. A "x" indicates that a combination cannot supply the load. The starting conditions are: at the beginning of the first period units 1 and 2 are up, units 3 and 4 are down and have been down for 8 h.
- 5.2 Table 5.10 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.
 - 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 up-/min down-times 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 up/min down time rules.

TABLE 5.8 Unit Commitment Data for Problem 5.1

Unit	Max (MW)	Min (MW)	Incremental Heat Rate (Btu/kWh)	No-Load Energy Input (MBtu/h)	Start-Up Energy (MBtu)
1	500	70	9950	300	800
2	250	40	10200	210	380
3	150	30	11000	120	110
4	150	30	11000	120	110

Load data (all time periods = 2 h);

Time Period	Load (MW)
1	600
2	800
3	700
4	950

Start-up and shut-down rules:

Unit	Minimum Up Time (h)	Minimum Down Time (h)
1	2	2
2	2	2
3	2	4
4	2	4

Fuel cost = 1.00 R/MBtu.

TABLE 5.9 Unit Combinations and Operating Cost for Problem 5.1

Combination	Unit 1	Unit 2	Unit 3	Unit 4	Operating Cost (R/h)			
					Load 600 MW	Load 700 MW	Load 800 MW	Load 950 MW
A	1	1	0	0	6505	7525	×	×
B	1	1	1	0	6649	7669	8705	×
C	1	1	1	1	6793	7813	8833	10475

1 = up; 0 = down.

- 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 up-/min down-time rules.

When using a dynamic-programming method to solve a unit commitment problem with minimum up- and down-time rules, one must save an additional piece of information at each state, each hour. This information

TABLE 5.10 The Unit Characteristic and Load Pattern for Problem 5.2

Unit	Max (MW)	Net		Min (MW)	No-Load Cost (MBtu/h)	Start-Up Cost (MBtu)	Min Up/Down Time (h)
		Full-Load Heat Rate (Btu/kWh)	Incremental Heat Rate (Btu/kWh)				
1	200	11000	9900	50	220	400	8
2	60	11433	10100	15	80	150	8
3	50	12000	10800	15	60	105	4
4	40	12900	11900	5	40	0	4
5	25	13500	12140	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	2. Ignore losses, spinning reserve, etc. The only requirement is that the generation be able to supply the load. 3. Fuel costs for all units may be taken as 1.40 R/MBtu
5-6	110	
7-8	75	

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/shut-down rules should be given a value of infinity.

5.3 Lagrange Relaxation Problem

Given the three generating units below:

$$F_1(P_1) = 30 + 10P_1 + 0.002P_1^2 \quad \text{and} \quad 100 < P_1 < 600$$

$$F_2(P_2) = 20 + 8P_2 + 0.0025P_2^2 \quad \text{and} \quad 100 < P_2 < 400$$

$$F_3(P_3) = 10 + 6P_3 + 0.005P_3^2 \quad \text{and} \quad 50 < P_3 < 200$$

Load:

t	$P'_{\text{load}}(\text{MW})$
1	300
2	500
3	1100
4	400

No start-up costs, no minimum up- or down-time constraints.

- a. Solve for the unit commitment by conventional dynamic programming.
- b. Set up and carry out four iterations of the Lagrange relaxation method. Let the initial values of λ^t be zero for $t = 1 \dots 4$.
- c. Resolve with the added condition that the third generator has a minimum up time of 2 h.

FURTHER READING

Some good introductory references to the unit commitment problem are found in references 1–3. A survey of the state-of-the-art (as of 1975) of unit commitment solutions is found in reference 4. References 5 and 6 provide a good look at two commercial unit commitment programs in present use.

References 7–11 deal with unit commitment as an integer-programming problem. Much of the pioneering work in this area was done by Garver (reference 7), who also sounded a note of pessimism in a discussion of reference 8, written together with Happ in 1968. Further research (references 9–11) has refined the unit commitment solution by integer programming but has never really overcome the Garver–Happ limitations presented in the 1968 discussion, thus leaving dynamic programming and Lagrange relaxation as the only viable solution techniques to large-scale unit commitment problems.

The reader should see references 12 and 13 for a discussion of valve-point loading and for a thorough development of economic dispatch via dynamic programming.

Reference 14 provides the reader with a good overview of unit commitment scheduling. References 15, 16, and 17 are recommended for an understanding of the Lagrange relaxation method, while references 18–21 cover some of the special problems encountered in unit commitment scheduling.

1. Baldwin, C. J., Dale, K. M., Dittich, R. F., "A Study of Economic Shutdown of Generating Units in Daily Dispatch," *AIEE Transactions on Power Apparatus and Systems*, Vol. PAS-78, December 1959, pp. 1272–1284.
2. Burns, R. M., Gibson, C. A., "Optimization of Priority Lists for a Unit Commitment Program," IEEE Power Engineering Society Summer Meeting, Paper A-75-453-1, 1975.
3. Davidson, P. M., Kohbrman, F. J., Master, G. L., Schafer, G. R., Evans, J. R., Lovewell, K. M., Payne, T. B., "Unit Commitment Start-Stop Scheduling in the Pennsylvania–New Jersey–Maryland Interconnection," 1967 PICA Conference Proceedings, IEEE, 1967, pp. 127–132.
4. Gruhl, J., Schweppe, F., Ruane, M., "Unit Commitment Scheduling of Electric Power Systems," *Systems Engineering for Power: Status and Prospects*, Henniker, NH, U.S. Government Printing Office, Washington, DC, 1975.
5. Pang, C. K., Chen, H. C., "Optimal Short-Term Thermal Unit Commitment," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-95, July/August 1976, pp. 1336–1346.
6. Happ, H. H., Johnson, P. C., Wright, W. J., "Large Scale Hydro-Thermal Unit Commitment—Method and Results," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-90, May/June 1971, pp. 1373–1384.
7. Garver, L. L., "Power Generation Scheduling by Integer Programming—

- Development of Theory," *AIEE Transactions on Power Apparatus and Systems*, Vol. PAS-82, February 1963, pp. 730-735.
8. Muckstadt, J. A., Wilson, R. C., "An Application of Mixed-Integer Programming Duality to Scheduling Thermal Generating Systems," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-87, December 1968, pp. 1968-1978.
 9. Ohuchi, A., Kaji, I., "A Branch-and-Bound Algorithm for Start-up and Shut-down Problems of Thermal Generating Units," *Electrical Engineering in Japan*, Vol. 95, No. 5, 1975, pp. 54-61.
 10. Dillon, T. S., Egan, G. T., "Application of Combinational Methods to the Problems of Maintenance Scheduling and Unit Commitment in Large Power Systems," *Proceedings of IFAC Symposium on Large Scale Systems Theory and Applications*, Udine, Italy, 1976.
 11. Dillon, T. S., Edwin, K. W., Kochs, H. D., Taud, R. J., "Integer Programming Approach to the Problem of Optimal Unit Commitment with Probabilistic Reserve Determination," *IEEE Transactions on Power Apparatus and Systems*, Vol. 97, November/December 1978, pp. 2154-2166.
 12. Happ, H. H., Ille, W. B., Reisinger, R. M., "Economic System Operation Considering Valve Throttling Losses, I—Method of Computing Valve-Loop Heat Rates on Multivalve Turbines," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-82, February 1963, pp. 609-615.
 13. Ringlee, R. J., Williams, D. D., "Economic Dispatch Operation Considering Valve Throttling Losses, II—Distribution of System Loads by the Method of Dynamic Programming," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-82, February 1963, pp. 615-622.
 14. Cohen, A. I., Sherkat, V. R., "Optimization-Based Methods for Operations Scheduling," *Proceedings IEEE*, December 1987, pp. 1574-1591.
 15. 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.
 16. Merlin, A., Sandrin, P., "A New Method for Unit Commitment at Electricité de France," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-102, May 1983, pp. 1218-1225.
 17. Zhuang, F., Galiana, F. D., "Towards a More Rigorous and Practical Unit Commitment by Lagrangian Relaxation," *IEEE Transactions on Power Systems*, Vol. 3, No. 2, May 1988, pp. 763-773.
 18. Lee, F. N., Chen, Q., Breipohl, A., "Unit Commitment Risk with Sequential Rescheduling," *IEEE Transactions on Power Systems*, Vol. 6, No. 3, August 1991, pp. 1017-1023.
 19. Vemuri, S., Lemonidis, L., "Fuel Constrained Unit Commitment," *IEEE Transactions on Power Systems*, Vol. 7, No. 1, February 1992, pp. 410-415.
 20. Wang, C., Shahidepour, S. M., "Optimal Generation Scheduling with Ramping Costs," 1993 *IEEE Power Industry Computer Applications Conference*, pp. 11-17.
 21. Shaw, J. J., "A Direct Method for Security-Constrained Unit Commitment," *IEEE Transactions Paper 94 SM 591-8 PWRS*, presented at the IEEE Power Engineering Society Summer Meeting, San Francisco, CA, July 1994.