12 An Introduction to State Estimation in Power Systems

12.1 INTRODUCTION

State estimation is the process of assigning a value to an unknown system state variable based on measurements from that system according to some criteria. Usually, the process involves imperfect measurements that are redundant and the process of estimating the system states is based on a statistical criterion that estimates the true value of the state variables to minimize or maximize the selected criterion. A commonly used and familiar criterion is that of minimizing the sum of the squares of the differences between the estimated and "true" (i.e., measured) values of a function.

The ideas of least-squares estimation have been known and used since the early part of the nineteenth century. The major developments in this area have taken place in the twentieth century in applications in the aerospace field. In these developments, the basic problems have involved the location of an aerospace vehicle (i.e., missile, airplane, or space vehicle) and the estimation of its trajectory given redundant and imperfect measurements of its position and velocity vector. In many applications, these measurements are based on optical observations and/or radar signals that may be contaminated with noise and may contain system measurement errors. State estimators may be both static and dynamic. Both types of estimators have been developed for power systems. This chapter will introduce the basic development of a static-state estimator.

In a power system, the state variables are the voltage magnitudes and relative phase angles at the system nodes. Measurements are required in order to estimate the system performance in real time for both system security control and constraints on economic dispatch. The inputs to an estimator are imperfect power system measurements of voltage magnitudes and power, VAR, or ampere-flow quantities. The estimator is designed to produce the "best estimate" of the system voltage and phase angles, recognizing that there are errors in the measured quantities and that there may be redundant measurements. The output data are then used in system control centers in the implementation of the security-constrained dispatch and control of the system as discussed in Chapters 11 and 13.

12.2 POWER SYSTEM STATE ESTIMATION

As introduced in Chapter 11, the problem of monitoring the power flows and voltages on a transmission system is very important in maintaining system security. By simply checking each measured value against its limit, the power system operators can tell where problems exist in the transmission system—and, it is hoped, they can take corrective actions to relieve overloaded lines or out-of-limit voltages.

Many problems are encountered in monitoring a transmission system. These problems come primarily from the nature of the measurement transducers and from communications problems in transmitting the measured values back to the operations control center.

Transducers from power system measurements, like any measurement device, will be subject to errors. If the errors are small, they may go undetected and can cause misinterpretation by those reading the measured values. In addition, transducers may have gross measurement errors that render their output useless. An example of such a gross error might involve having the transducer connected up backward; thus, giving the negative of the value being measured. Finally, the telemetry equipment often experiences periods when communications channels are completely out; thus, depriving the system operator of any information about some part of the power system network.

It is for these reasons that power system state estimation techniques have been developed. A state estimator, as we will see shortly, can "smooth out" small random errors in meter readings, detect and identify gross measurement errors, and "fill in" meter readings that have failed due to communications failures.

To begin, we will use a simple DC load flow example to illustrate the principles of state estimation. Suppose the three-bus DC load flow of Example 4B were operating with the load and generation shown in Figure 12.1. The only information we have about this system is provided by three MW power flow meters located as shown in Figure 12.2.

Only two of these meter readings are required to calculate the bus phase angles and all load and generation values fully. Suppose we use M_{13} and M_{32} and further suppose that M_{13} and M_{32} give us perfect readings of the flows on their respective transmission lines.

$$M_{13} = 5 \text{ MW} = 0.05 \text{ pu}$$

$$M_{32} = 40 \text{ MW} = 0.40 \text{ pu}$$

Then, the flows on lines 1-3 and 3-2 can be set equal to these meter readings.

$$f_{13} = \frac{1}{x_{13}}(\theta_1 - \theta_3) = M_{13} = 0.05 \text{ pu}$$

$$f_{32} = \frac{1}{x_{23}}(\theta_3 - \theta_2) = M_{32} = 0.40 \text{ pu}$$

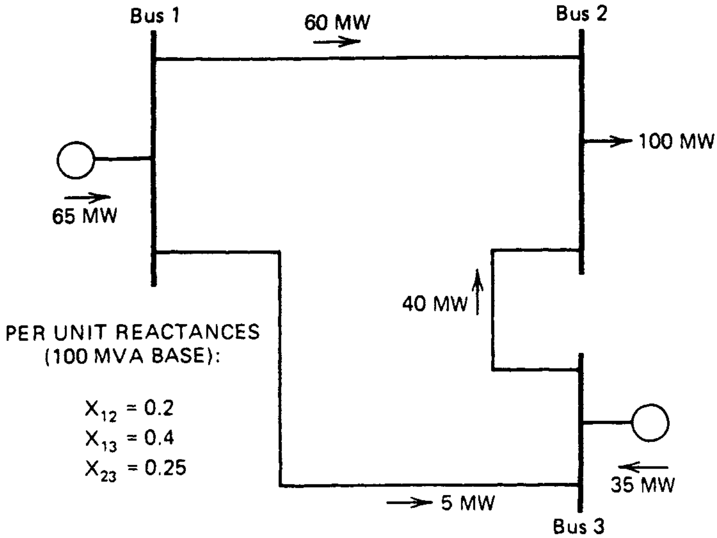

FIG. 12.1 Three-bus system from Example 4B.

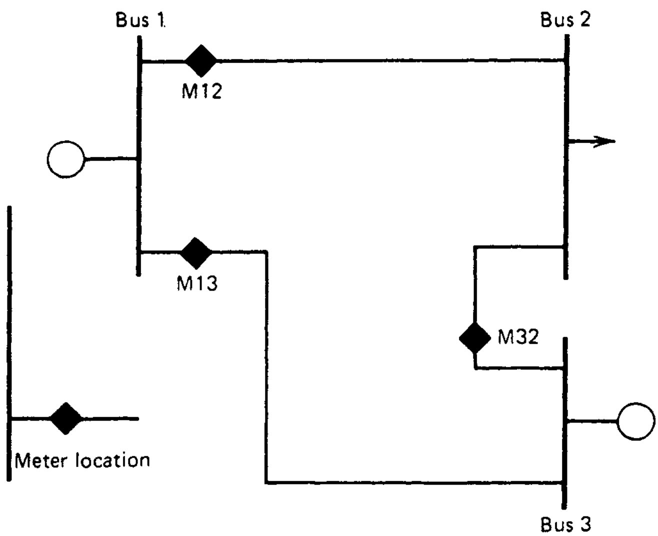

FIG. 12.2 Meter placement.

Since we know that $\theta_3 = 0$ rad, we can solve the f_{13} equation for θ_1, and the f_{32} equation for θ_2, resulting in

$$\theta_1 = 0.02 \text{ rad}$$

$$\theta_2 = -0.10 \text{ rad}$$

We will now investigate the case where all three meter readings have slight errors. Suppose the readings obtained are

$$M_{12} = 62 \text{ MW} = 0.62 \text{ pu}$$
$$M_{13} = 6 \text{ MW} = 0.06 \text{ pu}$$
$$M_{32} = 37 \text{ MW} = 0.37 \text{ pu}$$

If we use only the M_{13} and M_{32} readings, as before, we will calculate the phase angles as follows:

$$\theta_1 = 0.024 \text{ rad}$$
$$\theta_2 = -0.0925 \text{ rad}$$
$$\theta_3 = 0 \text{ rad (still assumed to equal zero)}$$

This results in the system flows as shown in Figure 12.3. Note that the predicted flows match at M_{13} and M_{32}, but the flow on line 1-2 does not match the reading of 62 MW from M_{12}. If we were to ignore the reading on M_{13} and use M_{12} and M_{32}, we could obtain the flows shown in Figure 12.4.

All we have accomplished is to match M_{12}, but at the expense of no longer matching M_{13}. What we need is a procedure that uses the information available from all three meters to produce the best estimate of the actual angles, line flows, and bus load and generations.

Before proceeding, let's discuss what we have been doing. Since the only thing we know about the power system comes to us from the measurements,

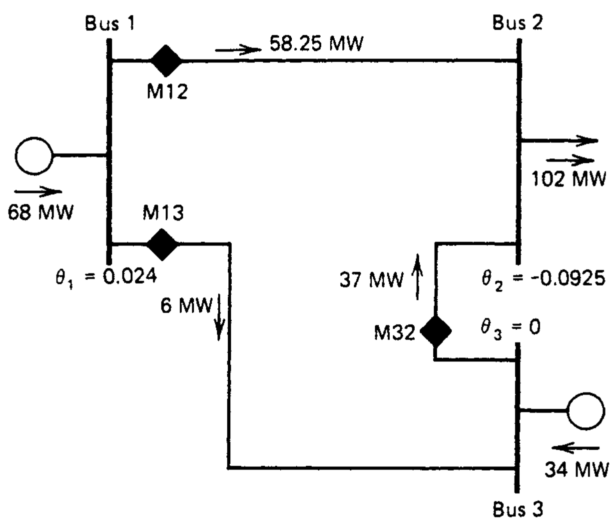

FIG. 12.3 Flows resulting from use of meters M_{13} and M_{32}.

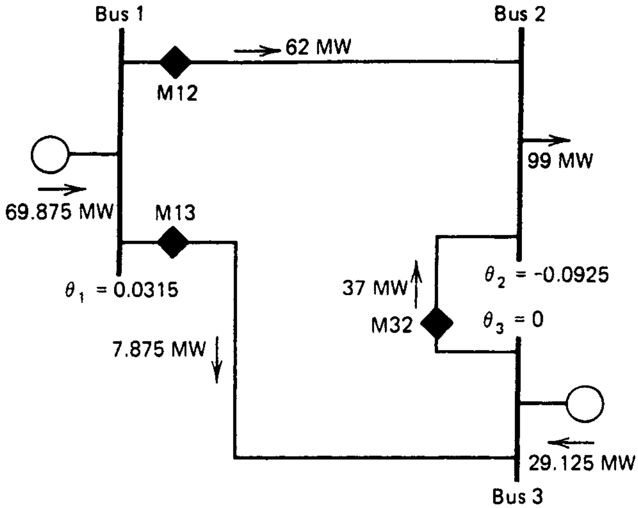

FIG. 12.4 Flows resulting from use of meters M_{12} and M_{32}.

we must use the measurements to estimate system conditions. Recall that in each instance the measurements were used to calculate the bus phase angles at bus 1 and 2. Once these phase angles were known, all unmeasured power flows, loads, and generations could be determined. We call θ_1 and θ_2 the *state variables* for the three-bus system since knowing them allows all other quantities to be calculated. In general, the state variables for a power system consist of the bus voltage magnitude at all buses and the phase angles at all but one bus. The swing or reference bus phase angle is usually assumed to be zero radians. Note that we could use real and imaginary components of bus voltage if desired. If we can use measurements to estimate the "states" (i.e., voltage magnitudes and phase angles) of the power system, then we can go on to calculate any power flows, generation, loads, and so forth that we desire. This presumes that the network configuration (i.e., breaker and disconnect switch statuses) is known and that the impedances in the network are also known. Automatic load tap changing autotransformers or phase angle regulators are often included in a network, and their tap positions may be telemetered to the control as a measured quantity. Strictly speaking, the transformer taps and phase angle regulator positions should also be considered as states since they must be known in order to calculate the flows through the transformers and regulators.

To return to the three-bus DC power flow model, we have three meters providing us with a set of redundant readings with which to estimate the two states θ_1 and θ_2. We say that the readings are redundant since, as we saw earlier, only two readings are necessary to calculate θ_1 and θ_2, the other reading is always "extra." However, the "extra" reading does carry useful information and ought not to be discarded summarily.

This simple example serves to introduce the subject of *static-state estimation*, which is the art of estimating the exact system state given a set of imperfect measurements made on the power system. We will digress at this point to develop the theoretical background for static-state estimation. We will return to our three-bus system in Section 12.3.4.

12.3 MAXIMUM LIKELIHOOD WEIGHTED LEAST-SQUARES ESTIMATION

12.3.1 Introduction

Statistical estimation refers to a procedure where one uses samples to calculate the value of one or more unknown parameters in a system. Since the samples (or measurements) are inexact, the estimate obtained for the unknown parameter is also inexact. This leads to the problem of how to formulate a "best"estimate of the unknown parameters given the available measurements.

The development of the notions of state estimation may proceed along several lines, depending on the statistical criterion selected. Of the many criteria that have been examined and used in various applications, the following three are perhaps the most commonly encountered.

1. *The maximum likelihood criterion*, where the objective is to maximize the probability that the estimate of the state variable, $\hat{\mathbf{x}}$, is the true value of the state variable vector, $\mathbf{x}$ (i.e., maximize $P(\hat{\mathbf{x}}) = \mathbf{x}$).
2. *The weighted least-squares criterion*, where the objective is to minimize the sum of the squares of the weighted deviations of the estimated measurements, $\hat{\mathbf{z}}$, from the actual measurements, $\mathbf{z}$.
3. *The minimum variance criterion*, where the object is to minimize the expected value of the sum of the squares of the deviations of the estimated components of the state variable vector from the corresponding components of the true state variable vector.

When normally distributed, unbiased meter error distributions are assumed, each of these approaches results in identical estimators. This chapter will utilize the maximum likelihood approach because the method introduces the measurement error weighting matrix $[R]$ in a straightforward manner.

The maximum likelihood procedure asks the following question: "What is the probability (or likelihood) that I will get the measurements I have obtained?" This probability depends on the random error in the measuring device (transducer) as well as the unknown parameters to be estimated. Therefore, a reasonable procedure would be one that simply chose the estimate as the value that maximizes this probability. As we will see shortly, the maximum likelihood estimator assumes that we know the probability density function (PDF) of the random errors in the measurement. Other estimation schemes could also be used. The "least-squares" estimator does not require that we know the probability density function for the sample or measurement errors. However, if we assume that the probability density function of sample or measurement error is a normal (Gaussian) distribution, we will end up with the same estimation formula. We will proceed to develop our estimation formula using the maximum likelihood criterion assuming normal distributions for measurement errors. The result will be a "least-squares" or more precisely a "weighted least-squares" estimation formula, even though we will develop the formulation using the maximum likelihood criteria. We will illustrate this method with a simple electrical circuit and show how the maximum likelihood estimate can be made.

First, we introduce the concept of *random measurement error*. Note that we have dropped the term "sample" since the concept of a measurement is much more appropriate to our discussion. The measurements are assumed to be in error: that is, the value obtained from the measurement device is close to the true value of the parameter being measured but differs by an unknown error. Mathematically, this can be modeled as follows.

Let z^{meas} be the value of a measurement as received from a measurement device. Let z^{true} be the true value of the quantity being measured. Finally, let η be the random measurement error. We can then represent our measured value as

$$z^{\text{meas}} = z^{\text{true}} + \eta \tag{12.1}$$

The random number, η, serves to model the uncertainty in the measurements. If the measurement error is unbiased, the probability density function of η is usually chosen as a normal distribution with zero mean. Note that other measurement probability density functions will also work in the maximum likelihood method as well. The probability density function of η is

$$\text{PDF}(\eta) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\eta^2/2\sigma^2) \tag{12.2}$$

where σ is called the standard deviation and σ^2 is called the variance of the random number. PDF(η) describes the behavior of η. A plot of PDF(η) is shown in Figure 12.5. Note that σ, the standard deviation, provides a way to model the seriousness of the random measurement error. If σ is large, the measurement is relatively inaccurate (i.e., a poor-quality measurement device), whereas a small value of σ denotes a small error spread (i.e., a higher-quality measurement device). The normal distribution is commonly used for modeling measurement errors since it is the distribution that will result when many factors contribute to the overall error.

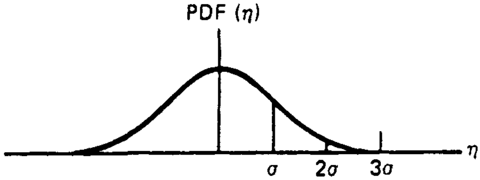

FIG. 12.5 The normal distribution.

12.3.2 Maximum Likelihood Concepts

The principle of maximum likelihood estimation is illustrated by using a simple DC circuit example as shown in Figure 12.6. In this example, we wish to estimate the value of the voltage source, x^{true}, using an ammeter with an error having a known standard deviation. The ammeter gives a reading of z_1^{meas}, which is equal to the sum of z_1^{true} (the true current flowing in our circuit) and η_1 (the error present in the ammeter). Then we can write

$$z_1^{meas} = z_1^{true} + \eta_1 \tag{12.3}$$

Since the mean value of η_1 is zero, we then know that the mean value of z_1^{meas} is equal to z_1^{true}. This allows us to write a probability density function for z_1^{meas} as

$$\text{PDF}(z_1^{meas}) = \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left[\frac{-(z_1^{meas} - z_1^{true})^2}{2\sigma_1^2}\right] \tag{12.4}$$

where σ_1 is the standard deviation for the random error η_1. If we assume that the value of the resistance, r_1, in our circuit is known, then we can write

$$\text{PDF}(z_1^{meas}) = \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left[\frac{-\left(z_1^{meas} - \dfrac{1}{r_1}x\right)^2}{2\sigma_1^2}\right] \tag{12.5}$$

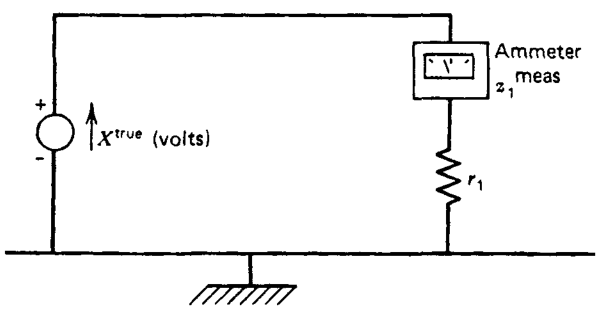

FIG. 12.6 Simple DC circuit with current measurement.

Coming back to our definition of a maximum likelihood estimator, we now wish to find an estimate of x (called x^{est}) that maximizes the probability that the observed measurement z_1^{meas} would occur. Since we have the probability density function of z_1^{meas}, we can write

$$\begin{aligned}\text{prob}(z_1^{\text{meas}}) &= \int_{z_1^{\text{meas}}}^{z_1^{\text{meas}}+\mathrm{d}z_1^{\text{meas}}} \text{PDF}(z_1^{\text{meas}})\,\mathrm{d}z_1^{\text{meas}} \quad \text{as } \mathrm{d}z_1^{\text{meas}} \to 0 \\ &= \text{PDF}(z_1^{\text{meas}})\,\mathrm{d}z_1^{\text{meas}} \end{aligned} \tag{12.6}$$

The maximum likelihood procedure then requires that we maximize the value of $\text{prob}(z_1^{\text{meas}})$, which is a function of x. That is,

$$\max_x \text{prob}(z_1^{\text{meas}}) = \max_x \text{PDF}(z_1^{\text{meas}})\,\mathrm{d}z_1^{\text{meas}} \tag{12.7}$$

One convenient transformation that can be used at this point is to maximize the natural logarithm of $\text{PDF}(z_1^{\text{meas}})$ since maximizing the Ln of $\text{PDF}(z_1^{\text{meas}})$ will also maximize $\text{PDF}(z_1^{\text{meas}})$. Then we wish to find

$$\max_x \text{Ln}[\text{PDF}(z_1^{\text{meas}})]$$

or

$$\max_x \left[-\text{Ln}(\sigma_1\sqrt{2\pi}) - \frac{\left(z_1^{\text{meas}} - \dfrac{1}{r_1}x\right)^2}{2\sigma_1^2} \right]$$

Since the first term is constant, it can be ignored. We can maximize the function in brackets by minimizing the second term since it has a negative coefficient, that is,

$$\max_x \left[-\text{Ln}(\sigma_1\sqrt{2\pi}) - \frac{\left(z_1^{\text{meas}} - \dfrac{1}{r_1}x\right)^2}{2\sigma_1^2} \right]$$

is the same as

$$\min_x \left[\frac{\left(z_1^{\text{meas}} - \dfrac{1}{r_1}x\right)^2}{2\sigma_1^2} \right] \tag{12.8}$$

The value of x that minimizes the right-hand term is found by simply taking the first derivative and setting the result to zero:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\left(z_1^{\text{meas}} - \dfrac{1}{r_1}x\right)^2}{2\sigma_1^2} \right] = \frac{-\left(z_1^{\text{meas}} - \dfrac{1}{r_1}x\right)}{r_1\sigma_1^2} = 0 \tag{12.9}$$

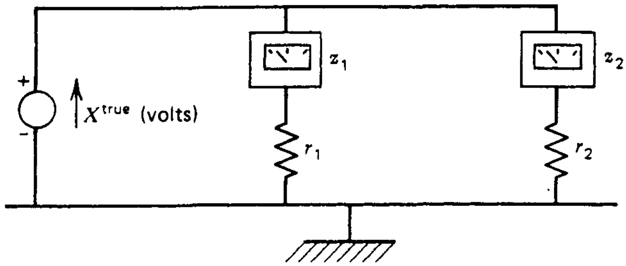

FIG. 12.7 DC circuit with two current measurements.

or

$$x^{\text{est}} = r_1 z_1^{\text{meas}}$$

To most readers this result was obvious from the beginning. All we have accomplished is to declare the maximum likelihood estimate of our voltage as simply the measured current times the known resistance. However, by adding a second measurement circuit, we have an entirely different situation in which the best estimate is not so obvious. Let us now add a second ammeter and resistance as shown in Figure 12.7.

Assume that both r_1 and r_2 are known. As before, model each meter reading as the sum of the true value and a random error:

$$\begin{aligned} z_1^{\text{meas}} &= z_1^{\text{true}} + \eta_1 \\ z_2^{\text{meas}} &= z_2^{\text{true}} + \eta_2 \end{aligned} \tag{12.10}$$

where the errors will be represented as independent zero mean, normally distributed random variables with probability density functions:

$$\begin{aligned} \text{PDF}(\eta_1) &= \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left(\frac{-(\eta_1)^2}{2\sigma_1^2}\right) \\ \text{PDF}(\eta_2) &= \frac{1}{\sigma_2\sqrt{2\pi}} \exp\left(\frac{-(\eta_2)^2}{2\sigma_2^2}\right) \end{aligned} \tag{12.11}$$

and as before we can write the probability density functions of z_1^{meas} and z_2^{meas} as

$$\begin{aligned} \text{PDF}(z_1^{\text{meas}}) &= \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left[\frac{-\left(z_1^{\text{meas}} - \dfrac{1}{r_1}x\right)^2}{2\sigma_1^2}\right] \\ \text{PDF}(z_2^{\text{meas}}) &= \frac{1}{\sigma_2\sqrt{2\pi}} \exp\left[\frac{-\left(z_2^{\text{meas}} - \dfrac{1}{r_2}x\right)^2}{2\sigma_2^2}\right] \end{aligned} \tag{12.12}$$

The likelihood function must be the probability of obtaining the measurements z_1^{meas} and z_2^{meas}. Since we are assuming that the random errors η_1 and η_2 are independent random variables, the probability of obtaining z_1^{meas} and z_2^{meas} is simply the product of the probability of obtaining z_1^{meas} and the probability of obtaining z_2^{meas}.

$$\begin{aligned}
\text{prob}(z_1^{\text{meas}} \text{ and } z_2^{\text{meas}}) &= \text{prob}(z_1^{\text{meas}}) \times (\text{prob}(z_2^{\text{meas}}) \\
&= \text{PDF}(z_1^{\text{meas}})\, \text{PDF}(z_2^{\text{meas}})\, \mathrm{d}z_1^{\text{meas}}\, \mathrm{d}z_2^{\text{meas}} \\
&= \left[\frac{1}{\sigma_1\sqrt{2\pi}} \exp\left(\frac{-\left(z_1^{\text{meas}} - \dfrac{1}{r_1}x\right)^2}{2\sigma_1^2}\right)\right] \\
&\quad \times \left[\frac{1}{\sigma_2\sqrt{2\pi}} \exp\left(\frac{-\left(z_2^{\text{meas}} - \dfrac{1}{r_2}x\right)^2}{2\sigma_2^2}\right)\right] \mathrm{d}z_1^{\text{meas}}\, \mathrm{d}z_2^{\text{meas}}
\end{aligned} \tag{12.13}$$

To maximize the function we will again take its natural logarithm:

$$\begin{aligned}
&\max_x \text{prob}(z_1^{\text{meas}} \text{ and } z_2^{\text{meas}}) \\
&= \max_x \left[-\text{Ln}(\sigma_1\sqrt{2\pi}) - \frac{\left(z_1^{\text{meas}} - \dfrac{1}{r_1}x\right)^2}{2\sigma_1^2} - \text{Ln}(\sigma_2\sqrt{2\pi}) - \frac{\left(z_2^{\text{meas}} - \dfrac{1}{r_2}x\right)^2}{2\sigma_2^2}\right] \\
&= \min_x \left[\frac{\left(z_1^{\text{meas}} - \dfrac{1}{r_1}x\right)^2}{2\sigma_1^2} + \frac{\left(z_2^{\text{meas}} - \dfrac{1}{r_2}x\right)^2}{2\sigma_2^2}\right]
\end{aligned} \tag{12.14}$$

The minimum sought is found by

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\left(z_1^{\text{meas}} - \dfrac{1}{r_1}x\right)^2}{2\sigma_1^2} + \frac{\left(z_2^{\text{meas}} - \dfrac{1}{r_2}x\right)^2}{2\sigma_2^2}\right] \\
&\quad = \frac{-\left(z_1^{\text{meas}} - \dfrac{1}{r_1}x\right)}{r_1\sigma_1^2} - \frac{\left(z_2^{\text{meas}} - \dfrac{1}{r_2}x\right)}{r_2\sigma_2^2} = 0
\end{aligned}$$

giving

$$x^{\text{est}} = \frac{\left(\dfrac{z_1^{\text{meas}}}{r_1\sigma_1^2} + \dfrac{z_2^{\text{meas}}}{r_2\sigma_2^2}\right)}{\left(\dfrac{1}{r_1^2\sigma_1^2} + \dfrac{1}{r_2^2\sigma_2^2}\right)} \tag{12.15}$$

If one of the ammeters is of superior quality, its variance will be much smaller than that of the other meter. For example, if $\sigma_2^2 \ll \sigma_1^2$, then the equation for x^{est} becomes

$$x^{\text{est}} \simeq z_2^{\text{meas}} \times r_2$$

Thus, we see that the maximum likelihood method of estimating our unknown parameter gives us a way to weight the measurements properly according to their quality.

It should be obvious by now that we need not express our estimation problem as a maximum of the product of probability density functions. Instead, we can observe a direct way of writing what is needed by looking at Eqs. 12.8 and 12.14. In these equations, we see that the maximum likelihood estimate of our unknown parameter is always expressed as that value of the parameter that gives the minimum of the sum of the squares of the difference between each measured value and the true value being measured (expressed as a function of our unknown parameter) with each squared difference divided or "weighted" by the variance of the meter error. Thus, if we are estimating a single parameter, x, using N_m measurements, we would write the expression

$$\min_x J(x) = \sum_{i=1}^{N_m} \frac{[z_i^{\text{meas}} - \mathrm{f}_i(x)]^2}{\sigma_i^2} \tag{12.16}$$

where

f_i = function that is used to calculate the value being measured by the i^{th} measurement

σ_i^2 = variance for the i^{th} measurement

$J(x)$ = measurement residual

N_m = number of independent measurements

z_i^{meas} = i^{th} measured quantity

Note that Eq. 12.16 may be expressed in per unit or in physical units such as MW, MVAR, or kV.

If we were to try to estimate N_s unknown parameters using N_m measurements, we would write

$$\min_{\{x_1, x_2, \ldots, x_{N_s}\}} J(x_1, x_2, \ldots, x_{N_s}) = \sum_{i=1}^{N_m} \frac{[z_i - \mathrm{f}_i(x_1, x_2, \ldots, x_{N_s})]^2}{\sigma_i^2} \tag{12.17}$$

The estimation calculation shown in Eqs. 12.16 and 12.17 is known as a *weighted least-squares* estimator, which, as we have shown earlier, is equivalent to a maximum likelihood estimator if the measurement errors are modeled as random numbers having a normal distribution.

12.3.3 Matrix Formulation

If the functions $f_i(x_1, x_2, \ldots, x_{N_s})$ are linear functions, Eq. 12.17 has a closed-form solution. Let us write the function $f_i(x_1, x_2, \ldots, x_{N_s})$ as

$$f_i(x_1, x_2, \ldots, x_{N_s}) = f_i(\mathbf{x}) = h_{i1}x_1 + h_{i2}x_2 + \ldots + h_{iN_s}x_{N_s} \tag{12.18}$$

Then, if we place all the f_i functions in a vector, we may write

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_{N_m}(\mathbf{x}) \end{bmatrix} = [H]\mathbf{x} \tag{12.19}$$

where

$[H]$ = an N_m by N_s matrix containing the coefficients of the linear functions $f_i(\mathbf{x})$

N_m = number of measurements

N_s = number of unknown parameters being estimated

Placing the measurements in a vector:

$$\mathbf{z}^{\text{meas}} = \begin{bmatrix} z_1^{\text{meas}} \\ z_2^{\text{meas}} \\ \vdots \\ z_{N_m}^{\text{meas}} \end{bmatrix} \tag{12.20}$$

We may then write Eq. 12.17 in a very compact form.

$$\min_{\mathbf{x}} J(\mathbf{x}) = [\mathbf{z}^{\text{meas}} - \mathbf{f}(\mathbf{x})]^T[R^{-1}][\mathbf{z}^{\text{meas}} - \mathbf{f}(\mathbf{x})] \tag{12.21}$$

where

$$[R] = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_{N_m}^2 \end{bmatrix}$$

$[R]$ is called the *covariance matrix of measurement errors.* To obtain the general expression for the minimum in Eq. 12.21, expand the expression and substitute $[H]\mathbf{x}$ for $\mathbf{f}(\mathbf{x})$ from Eq. 12.19.

$$\min_{\mathbf{x}} J(\mathbf{x}) = \{\mathbf{z}^{\text{meas}^T}[R^{-1}]\mathbf{z}^{\text{meas}} - \mathbf{x}^T[H]^T[R^{-1}]\mathbf{z}^{\text{meas}}$$

$$-\mathbf{z}^{\text{meas}^T}[R^{-1}][H]\mathbf{x} + \mathbf{x}^T[H]^T[R^{-1}][H]\mathbf{x}\} \tag{12.22}$$

Similar to the procedures of Chapter 3, the minimum of $J(\mathbf{x})$ is found when $\partial J(\mathbf{x})/\partial x_i = 0$, for $i = 1, \ldots, N_s$; this is identical to stating that the gradient of $J(\mathbf{x})$, $\nabla J(\mathbf{x})$, is exactly zero.

The gradient of $J(\mathbf{x})$ is (see the appendix to this chapter)

$$\nabla J(\mathbf{x}) = -2[H]^T[R^{-1}]\mathbf{z}^{\text{meas}} + 2[H]^T[R^{-1}][H]\mathbf{x}$$

Then $\nabla J(\mathbf{x}) = \mathbf{0}$ gives

$$\mathbf{x}^{\text{est}} = [[H]^T[R^{-1}][H]]^{-1}[H]^T[R^{-1}]\mathbf{z}^{\text{meas}} \tag{12.23}$$

Note that Eq. 12.23 holds when $N_s < N_m$; that is, when the number of parameters being estimated is less than the number of measurements being made.

When $N_s = N_m$, our estimation problem reduces to

$$\mathbf{x}^{\text{est}} = [H]^{-1}\mathbf{z}^{\text{meas}} \tag{12.24}$$

There is also a closed-form solution to the problem when $N_s > N_m$, although in this case we are not estimating $\mathbf{x}$ to maximize a likelihood function since $N_s > N_m$ usually implies that many different values for $\mathbf{x}^{\text{est}}$ can be found that cause $f_i(\mathbf{x}^{\text{est}})$ to equal z_i^{meas} for all $i = 1, \ldots, N_m$ exactly. Rather, the objective is to find $\mathbf{x}^{\text{est}}$ such that the sum of the squares of x_i^{est} is minimized. That is,

$$\min_{\mathbf{x}} \sum_{i=1}^{N_s} x_i^2 = \mathbf{x}^T\mathbf{x} \tag{12.25}$$

subject to the condition that $\mathbf{z}^{\text{meas}} = [H]\mathbf{x}$. The closed-form solution for this case is

$$\mathbf{x}^{\text{est}} = [H]^T[[H][H]^T]^{-1}\mathbf{z}^{\text{meas}} \tag{12.26}$$

In power system state estimation, underdetermined problems (i.e., where $N_s > N_m$) are not solved, as shown in Eq. 12.26. Rather, "pseudo-measurements" are added to the measurement set to give a completely determined or overdetermined problem. We will discuss pseudo-measurements in Section 12.6.3. Table 12.1 summarizes the results for this section.

TABLE 12.1 Estimation Formulas

Case	Description	Solution	Comment
$N_s < N_m$	Overdetermined	$\mathbf{x}^{est} = [[H]^T[R^{-1}][H]]^{-1} \times \{H]^T[R^{-1}]\mathbf{z}^{meas}$	$\mathbf{x}^{est}$ is the maximum likelihood estimate of $\mathbf{x}$ given the measurements $\mathbf{z}^{meas}$
$N_s = N_m$	Completely determined	$\mathbf{x}^{est} = [H]^{-1}\mathbf{z}^{meas}$	$\mathbf{x}^{est}$ fits the measured quantities to the measurements $\mathbf{z}^{meas}$ exactly
$N_s > N_m$	Underdetermined	$\mathbf{x}^{est} = [H]^T[[H][H]^T]^{-1}\mathbf{z}^{meas}$	$\mathbf{x}^{est}$ is the vector of minimum norm that fits the measured quantities to the measurements exactly. (The norm of a vector is equal to the sum of the squares of its components)

12.3.4 An Example of Weighted Least-Squares State Estimation

We now return to our three-bus example. Recall from Figure 12.2 that we have three measurements to determine θ_1 and θ_2, the phase angles at buses 1 and 2. From the development in the preceding section, we know that the states θ_1 and θ_2 can be estimated by minimizing a residual $J(\theta_1, \theta_2)$ where $J(\theta_1, \theta_2)$ is the sum of the squares of individual measurement residuals divided by the variance for each measurement.

To start, we will assume that all three meters have the following characteristics.

Meter full-scale value: 100 MW

Meter accuracy: ± 3 MW

This is interpreted to mean that the meters will give a reading within ± 3 MW of the true value being measured for approximately 99% of the time. Mathematically, we say that the errors are distributed according to a normal probability density function with a standard deviation, σ, as shown in Figure 12.8.

Notice that the probability of an error decreases as the error magnitude increases. By integrating the PDF between -3σ and $+3\sigma$ we come up with a value of approximately 0.99. We will assume that the meter's accuracy (in our case ± 3 MW) is being stated as equal to the 3σ points on the probability density function. Then ± 3 MW corresponds to a metering standard deviation of $\sigma = 1 \text{ MW} = 0.01$ pu.

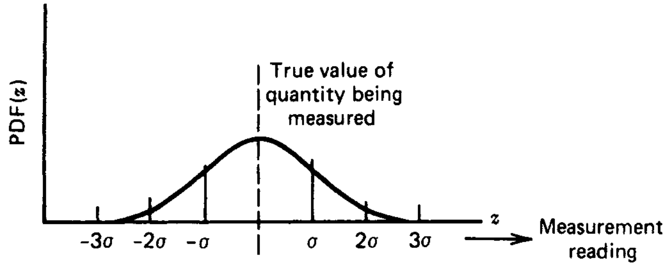

FIG. 12.8 Normal distribution of meter errors.

The formula developed in the last section for the weighted least-squares estimate is given in Eq. 12.23, which is repeated here.

$$\mathbf{x}^{\text{est}} = [[H]^T[R^{-1}][H]]^{-1}[H]^T[R^{-1}]\mathbf{z}^{\text{meas}}$$

where

$\mathbf{x}^{\text{est}}$ = vector of estimated state variables

$[H]$ = measurement function coefficient matrix

$[R]$ = measurement covariance matrix

$\mathbf{z}^{\text{meas}}$ = vector containing the measured values themselves

For the three-bus problem we have

$$\mathbf{x}^{\text{est}} = \begin{bmatrix} \theta_1^{\text{est}} \\ \theta_2^{\text{est}} \end{bmatrix} \tag{12.27}$$

To derive the $[H]$ matrix, we need to write the measurements as a function of the state variables θ_1 and θ_2. These functions are written in per unit as

$$\begin{aligned} M_{12} &= f_{12} = \frac{1}{0.2}(\theta_1 - \theta_2) = 5\theta_1 - 5\theta_2 \\ M_{13} &= f_{13} = \frac{1}{0.4}(\theta_1 - \theta_3) = 2.5\theta_1 \\ M_{32} &= f_{32} = \frac{1}{0.25}(\theta_3 - \theta_2) = -4\theta_2 \end{aligned} \tag{12.28}$$

The reference-bus phase angle, θ_3, is still assumed to be zero. Then

$$[H] = \begin{bmatrix} 5 & -5 \\ 2.5 & 0 \\ 0 & -4 \end{bmatrix}$$

The covariance matrix for the measurements, $[R]$, is

$$[R] = \begin{bmatrix} \sigma_{M12}^2 & & \\ & \sigma_{M13}^2 & \\ & & \sigma_{M32}^2 \end{bmatrix} = \begin{bmatrix} 0.0001 & & \\ & 0.0001 & \\ & & 0.0001 \end{bmatrix}$$

Note that since the coefficients of $[H]$ are in per unit we must also write $[R]$ and $\mathbf{z}^{\text{meas}}$ in per unit.

Our least-squares "best" estimate of θ_1 and θ_2 is then calculated as

$$\begin{bmatrix} \theta_1^{\text{est}} \\ \theta_2^{\text{est}} \end{bmatrix} = \left[\begin{bmatrix} 5 & 2.5 & 0 \\ -5 & 0 & -4 \end{bmatrix} \begin{bmatrix} 0.0001 & & \\ & 0.0001 & \\ & & 0.0001 \end{bmatrix}^{-1} \begin{bmatrix} 5 & -5 \\ 2.5 & 0 \\ 0 & -4 \end{bmatrix} \right]^{-1}$$

$$\times \begin{bmatrix} 5 & 2.5 & 0 \\ -5 & 0 & -4 \end{bmatrix} \begin{bmatrix} 0.0001 & & \\ & 0.0001 & \\ & & 0.0001 \end{bmatrix}^{-1} \begin{bmatrix} 0.62 \\ 0.06 \\ 0.37 \end{bmatrix}$$

$$= \begin{bmatrix} 312500 & -250000 \\ -250000 & 410000 \end{bmatrix}^{-1} \begin{bmatrix} 32500 \\ -45800 \end{bmatrix}$$

$$= \begin{bmatrix} 0.028571 \\ -0.094286 \end{bmatrix}$$

where

$$\mathbf{z}^{\text{meas}} = \begin{bmatrix} 0.62 \\ 0.06 \\ 0.37 \end{bmatrix}$$

From the estimated phase angles, we can calculate the power flowing in each transmission line and the net generation or load at each bus. The results are shown in Figure 12.9. If we calculate the value of $J(\theta_1, \theta_2)$, the residual, we get

$$J(\theta_1, \theta_2) = \frac{[z_{12} - f_{12}(\theta_1, \theta_2)]^2}{\sigma_{12}^2} + \frac{[z_{13} - f_{13}(\theta_1, \theta_2)]^2}{\sigma_{13}^2} + \frac{[z_{32} - f_{32}(\theta_1, \theta_2)]^2}{\sigma_{32}^2}$$

$$= \frac{[0.62 - (5\theta_1 - 5\theta_2)]^2}{0.0001} + \frac{[0.06 - (2.5\theta_1)]^2}{0.0001} + \frac{[0.37 + (4\theta_2)]^2}{0.0001}$$

$$= 2.14 \tag{12.29}$$

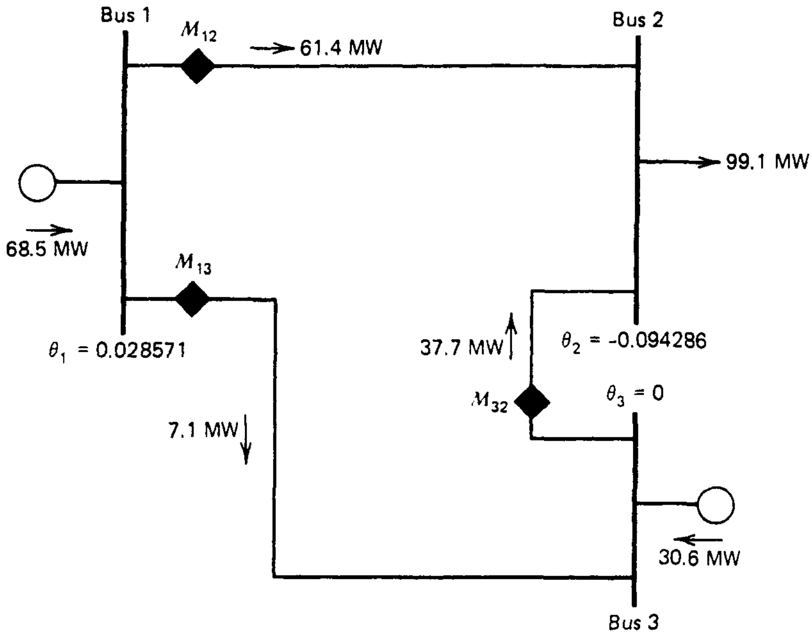

FIG. 12.9 Three-bus example with best estimates of θ_1 and θ_2.

Suppose the meter on the M_{13} transmission line was superior in quality to those on M_{12} and M_{32}. How will this affect the estimate of the states? Intuitively, we can reason that any measurement reading we get from M_{13} will be much closer to the true power flowing on line 1-3 than can be expected when comparing M_{12} and M_{32} to the flows on lines 1-2 and 3-2, respectively. Therefore, we would expect the results from the state estimator to reflect this if we set up the measurement data to reflect the fact that M_{13} is a superior measurement. To show this, we use the following metering data.

Meters M_{12} and M_{32}:	100 MW full scale ± 3 MW accuracy ($\sigma = 1$ MW $= 0.01$ pu)
Meter M_{13}:	100 MW full scale ± 0.3 MW accuracy ($\sigma = 0.1$ MW $= 0.001$ pu)

The covariance matrix to be used in the least-squares formula now becomes

$$[R] = \begin{bmatrix} \sigma_{M12}^2 & & \\ & \sigma_{M13}^2 & \\ & & \sigma_{M32}^2 \end{bmatrix} = \begin{bmatrix} 1 \times 10^{-4} & & \\ & 1 \times 10^{-6} & \\ & & 1 \times 10^{-4} \end{bmatrix}$$

We now solve Eq. 12.23 again with the new [R] matrix.

$$\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix} = \left[\begin{bmatrix}5 & 2.5 & 0\\ -5 & 0 & -4\end{bmatrix}\begin{bmatrix}1\times10^{-4} & & \\ & 1\times10^{-6} & \\ & & 1\times10^{-4}\end{bmatrix}^{-1}\begin{bmatrix}5 & -5\\ 2.5 & 0\\ 0 & -4\end{bmatrix}\right]^{-1}$$

$$\begin{bmatrix}5 & 2.5 & 0\\ -5 & 0 & -4\end{bmatrix}\begin{bmatrix}1\times10^{-4} & & \\ & 1\times10^{-6} & \\ & & 1\times10^{-4}\end{bmatrix}^{-1}\begin{bmatrix}0.62\\ 0.06\\ 0.37\end{bmatrix}$$

$$= \begin{bmatrix}6.5\times10^{6} & -2.5\times10^{5}\\ -2.5\times10^{5} & 4.1\times10^{5}\end{bmatrix}^{-1}\begin{bmatrix}1.81\times10^{5}\\ -0.458\times10^{5}\end{bmatrix}$$

$$= \begin{bmatrix}0.024115\\ -0.097003\end{bmatrix}$$

From these estimated phase angles, we obtain the network conditions shown in Figure 12.10. Compare the estimated flow on line 1-3, as just calculated, to the estimated flow calculated on line 1-3 in the previous least-squares estimate. Setting σ_{M13} to 0.1 MW has brought the estimated flow on line 1-3 much closer to the meter reading of 6.0 MW. Also, note that the estimates of flow on lines 1-2 and 3-2 are now further from the M_{12} and M_{32} meter readings, respectively, which is what we should have expected.

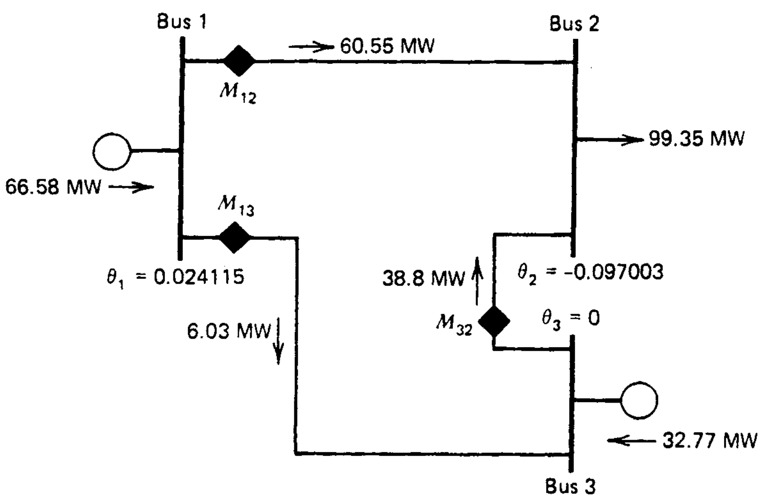

FIG. 12.10 Three-bus example with better meter at M_{13}.

12.4 STATE ESTIMATION OF AN AC NETWORK

12.4.1 Development of Method

We have demonstrated how the maximum likelihood estimation scheme developed in Section 12.3.2 led to a least-squares calculation for measurements from a linear system. In the least-squares calculation, we are trying to minimize the sum of measurement residuals:

$$\min_{\mathbf{x}} J(\mathbf{x}) = \sum_{i=1}^{N_m} \frac{[z_i - \mathrm{f}_i(\mathbf{x})]^2}{\sigma_i^2} \tag{12.30}$$

In the case of a linear system, the $\mathrm{f}_i(\mathbf{x})$ functions are themselves linear and we solve for the minimum of $J(\mathbf{x})$ directly. In an AC network, the measured quantities are MW, MVAR, MVA, amperes, transformer tap position, and voltage magnitude. The state variables are the voltage magnitude at each bus, the phase angles at all but the reference bus, and the transformer taps. The equation for power entering a bus is given in Eq. 4.21 and is clearly not a linear function of the voltage magnitude and phase angle at each bus. Therefore, the $\mathrm{f}_i(\mathbf{x})$ functions will be nonlinear functions, except for a voltage magnitude measurement where $\mathrm{f}_i(\mathbf{x})$ is simply unity times the particular x_i that corresponds to the voltage magnitude being measured. For MW and MVAR measurements on a transmission line from bus i to bus j we would have the following terms in $J(\mathbf{x})$:

$$\frac{\{\mathrm{MW}_{ij}^{\mathrm{meas}} - [|E_i|^2(G_{ij}) - |E_i|\,|E_j|(\cos(\theta_i - \theta_j)G_{ij} + \sin(\theta_i - \theta_j)B_{ij})]\}^2}{\sigma_{\mathrm{MW}_{ij}}^2} \tag{12.31}$$

and

$$\frac{\{\mathrm{MVAR}_{ij}^{\mathrm{meas}} - [-|E_i|^2(B_{\mathrm{cap}_{ij}} + B_{ij}) - |E_i|\,|E_j|(\sin(\theta_i - \theta_j)G_{ij} - \cos(\theta_i - \theta_j)B_{ij})]\}^2}{\sigma_{\mathrm{MVAR}_{ij}}^2} \tag{12.32}$$

A voltage magnitude measurement would result in the following term in $J(\mathbf{x})$:

$$\frac{(|E_i|^{\mathrm{meas}} - |E_i|)^2}{\sigma_{|E_i|}^2} \tag{12.33}$$

Similar functions can be derived for MVA or ampere measurements.

If we do not have a linear relationship between the states ($|E|$ values and θ values) and the power flows on a network, we will have to resort to an iterative technique to minimize $J(\mathbf{x})$. A commonly used technique for power system state estimation is to calculate the gradient of $J(\mathbf{x})$ and then force it to zero using Newton's method, as was done with the Newton load flow in Chapter 4. We will review how to use Newton's method on multidimensional problems before proceeding to the minimization of $J(\mathbf{x})$.

Given the functions $g_i(\mathbf{x})$, $i = 1, \ldots, n$, we wish to find $\mathbf{x}$ that gives $g_i(\mathbf{x}) = g_i^{\text{des}}$, for $i = 1, \ldots, n$. If we arrange the g_i functions in a vector we can write

$$\mathbf{g}^{\text{des}} - \mathbf{g}(\mathbf{x}) = 0 \tag{12.34}$$

by perturbing $\mathbf{x}$ we can write

$$\mathbf{g}^{\text{des}} - \mathbf{g}(\mathbf{x} + \Delta\mathbf{x}) = \mathbf{g}^{\text{des}} - \mathbf{g}(\mathbf{x}) - [\mathbf{g}'(\mathbf{x})]\Delta\mathbf{x} = 0 \tag{12.35}$$

where we have expanded $\mathbf{g}(\mathbf{x} + \Delta\mathbf{x})$ in a Taylor's series about $\mathbf{x}$ and ignored all higher-order terms. The $[\mathbf{g}'(\mathbf{x})]$ term is the Jacobian matrix of first derivatives of $\mathbf{g}(\mathbf{x})$. Then

$$\Delta\mathbf{x} = [\mathbf{g}'(\mathbf{x})]^{-1}[\mathbf{g}^{\text{des}} - \mathbf{g}(\mathbf{x})] \tag{12.36}$$

Note that if $\mathbf{g}^{\text{des}}$ is identically zero we have

$$\Delta\mathbf{x} = [\mathbf{g}'(\mathbf{x})]^{-1}[-\mathbf{g}(\mathbf{x})] \tag{12.37}$$

To solve for $\mathbf{g}^{\text{des}}$, we must solve for $\Delta\mathbf{x}$ using Eq. 12.36, then calculate $\mathbf{x}^{\text{new}} = \mathbf{x} + \Delta\mathbf{x}$ and reapply Eq. 12.36 until either $\Delta\mathbf{x}$ gets very small or $\mathbf{g}(\mathbf{x})$ comes close to $\mathbf{g}^{\text{des}}$.

Now let us return to the state estimation problem as given in Eq. 12.30:

$$\min_{\mathbf{x}} J(\mathbf{x}) = \sum_{i=1}^{N_m} \frac{[z_i - \mathrm{f}_i(\mathbf{x})]^2}{\sigma_i^2}$$

We first form the gradient of $J(\mathbf{x})$ as

$$\nabla_x \mathbf{J}(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial J(\mathbf{x})}{\partial x_1} \\ \dfrac{\partial J(\mathbf{x})}{\partial x_2} \\ \vdots \end{bmatrix}$$

$$= -2 \begin{bmatrix} \dfrac{\partial \mathrm{f}_1}{\partial x_1} & \dfrac{\partial \mathrm{f}_2}{\partial x_1} & \dfrac{\partial \mathrm{f}_3}{\partial x_1} & \cdots \\ \dfrac{\partial \mathrm{f}_1}{\partial x_2} & \dfrac{\partial \mathrm{f}_2}{\partial x_2} & \dfrac{\partial \mathrm{f}_3}{\partial x_2} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} \begin{bmatrix} \dfrac{1}{\sigma_1^2} & & \\ & \dfrac{1}{\sigma_2^2} & \\ & & \ddots \end{bmatrix} \begin{bmatrix} [z_1 - \mathrm{f}_1(\mathbf{x})] \\ [z_2 - \mathrm{f}_2(\mathbf{x})] \\ \vdots \end{bmatrix} \tag{12.38}$$

If we put the $f_i(\mathbf{x})$ functions in a vector form $\mathbf{f}(\mathbf{x})$ and calculate the Jacobian of $\mathbf{f}(\mathbf{x})$, we would obtain

$$\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \dfrac{\partial f_1}{\partial x_3} & \cdots \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \dfrac{\partial f_2}{\partial x_3} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} \tag{12.39}$$

We will call this matrix $[H]$. Then,

$$[H] = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \dfrac{\partial f_1}{\partial x_3} & \cdots \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \dfrac{\partial f_2}{\partial x_3} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} \tag{12.40}$$

And its transpose is

$$[H]^T = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_3}{\partial x_1} & \cdots \\ \dfrac{\partial f_1}{\partial x_2} & \dfrac{\partial f_2}{\partial x_2} & \dfrac{\partial f_3}{\partial x_2} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} \tag{12.41}$$

Further, we write

$$\begin{bmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \ddots \end{bmatrix} = [R] \tag{12.42}$$

Equation 12.38 can be written

$$\nabla_x \mathbf{J}(\mathbf{x}) = \left\{ -2[H]^T[R]^{-1} \begin{bmatrix} z_1 - f_1(\mathbf{x}) \\ z_2 - f_2(\mathbf{x}) \\ \vdots \end{bmatrix} \right\} \tag{12.43}$$

To make $\nabla_x \mathbf{J}(\mathbf{x})$ equal zero, we will apply Newton's method as in Eq. 12.37, then

$$\Delta \mathbf{x} = \left[\frac{\partial \nabla_x J(\mathbf{x})}{\partial \mathbf{x}}\right]^{-1} [-\nabla_x J(\mathbf{x})] \tag{12.44}$$

The Jacobian of $\nabla_x \mathbf{J}(\mathbf{x})$ is calculated by treating $[H]$ as a constant matrix:

$$\frac{\partial \nabla_x J(\mathbf{x})}{\partial \mathbf{x}} = \frac{\partial}{\partial \mathbf{x}} \left\{ -2[H]^T[R]^{-1} \begin{bmatrix} z_1 - \mathrm{f}_1(\mathbf{x}) \\ z_2 - \mathrm{f}_2(\mathbf{x}) \\ \vdots \end{bmatrix} \right\}$$

$$= -2[H]^T[R]^{-1}[-H]$$

$$= 2[H]^T[R]^{-1}[H] \tag{12.45}$$

Then

$$\Delta \mathbf{x} = \frac{1}{2}[[H]^T[R]^{-1}[H]]^{-1} \left\{ 2[H]^T[R]^{-1} \begin{bmatrix} z_1 - \mathrm{f}_1(\mathbf{x}) \\ \vdots \end{bmatrix} \right\}$$

$$= [[H]^T[R]^{-1}[H]]^{-1}[H]^T[R]^{-1} \begin{bmatrix} z_1 - \mathrm{f}_1(\mathbf{x}) \\ z_2 - \mathrm{f}_2(\mathbf{x}) \\ \vdots \end{bmatrix} \tag{12.46}$$

Equation 12.46 is obviously a close parallel to Eq. 12.23. To solve the AC state estimation problem, apply Eq. 12.46 iteratively as shown in Figure 12.11. Note that this is similar to the iterative process used in the Newton power flow solution.

12.4.2 Typical Results of State Estimation on an AC Network

Figure 12.12 shows our familiar six-bus system with $P + jQ$ measurements on each end of each transmission line and at each load and generator. Bus voltage is also measured at each system bus.

To demonstrate the use of state estimation on these measurements, the base-case conditions shown in Figure 11.1 were used together with a random number generating algorithm to produce measurements with random errors. The measurements were obtained by adding the random errors to the base-case flows, loads, generations, and bus-voltage magnitudes. The errors were generated so as to be representative of values drawn from a set of numbers having a normal probability density function with zero mean, and variance as specified for each measurement type. The measurement variances used were

$P + jQ$ measurements: $\sigma = 5$ MW for the P measurement

$\sigma = 5$ MVAR for the Q measurement

Voltage measurement: $\sigma = 3.83$ kV

The base conditions and the measurements are shown in Table 12.2. The state estimation algorithm shown in Figure 12.11 was run to obtain estimates

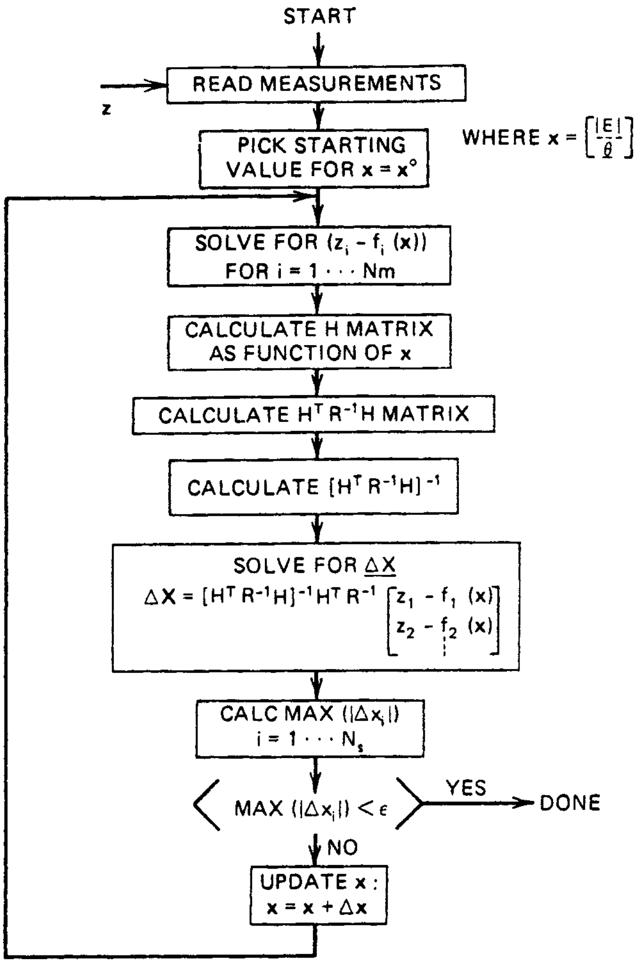

FIG. 12.11 State estimation solution algorithm.

for the bus-voltage magnitudes and phase angles given the measurements shown in Table 12.2. The procedure took three iterations with $\mathbf{x}^0$ initially being set to 1.0 pu and 0 rad for the voltage magnitude and phase angle at each bus, respectively. At the beginning of each iteration, the sum of the measurement residuals, $J(\mathbf{x})$ (see Eq. 12.30), is calculated and displayed. At the end of each iteration, the maximum $\Delta|E|$ and the maximum $\Delta\theta$ are calculated and displayed. The iterative steps for the six-bus system used here produced the results given in Table 12.3.

The value of $J(\mathbf{x})$ at the end of the iterative procedure would be zero if all measurements were without error or if there were no redundancy in the measurements. When there are redundant measurements with errors, the value of $J(\mathbf{x})$ will not normally go to zero. Its value represents a measure of the overall

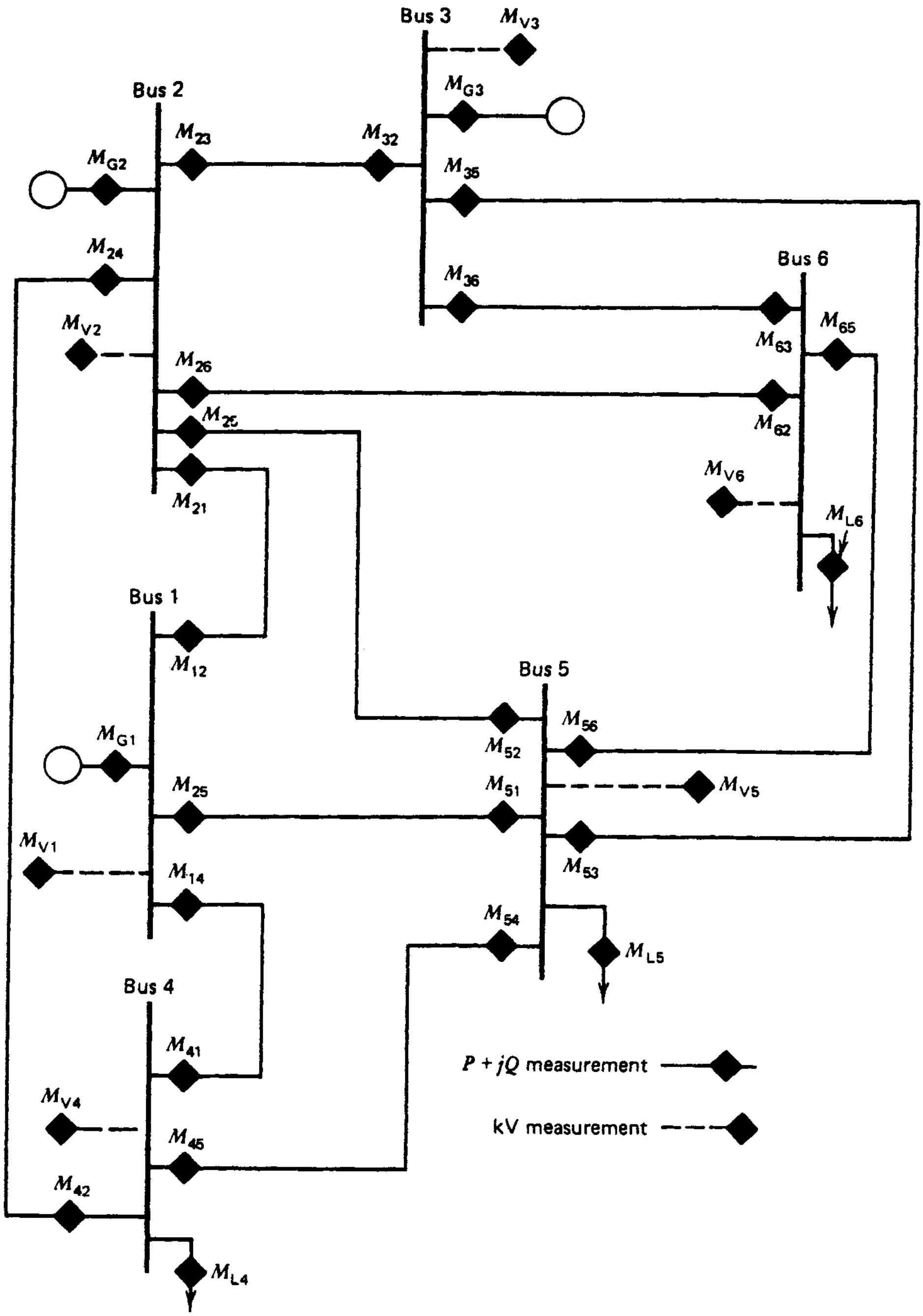

FIG. 12.12 Six-bus system with measurements.

TABLE 12.2 Base-Case Conditions

	Base-Case Value			Measured Value		
Measurement	kV	MW	MVAR	kV	MW	MVAR
M_{V1}	241.5			238.4		
M_{G1}		107.9	16.0		113.1	20.2
M_{12}		28.7	−15.4		31.5	−13.2
M_{14}		43.6	20.1		38.9	21.2
M_{15}		35.6	11.3		35.7	9.4
M_{V2}	241.5			237.8		
M_{G2}		50.0	74.4		48.4	71.9
M_{21}		−27.8	12.8		−34.9	9.7
M_{24}		33.1	46.1		32.8	38.3
M_{25}		15.5	15.4		17.4	22.0
M_{26}		26.2	12.4		22.3	15.0
M_{23}		2.9	−12.3		8.6	−11.9
M_{V3}	246.1			250.7		
M_{G3}		60.0	89.6		55.1	90.6
M_{32}		−2.9	5.7		−2.1	10.2
M_{35}		19.1	23.2		17.7	23.9
M_{36}		43.8	60.7		43.3	58.3
M_{V4}	227.6			225.7		
M_{L4}		70.0	70.0		71.8	71.9
M_{41}		−42.5	−19.9		−40.1	−14.3
M_{42}		−31.6	−45.1		−29.8	−44.3
M_{45}		4.1	−4.9		0.7	−17.4
M_{V5}	226.7			225.2		
M_{L5}		70.0	70.0		72.0	67.7
M_{54}		−4.0	−2.8		−2.1	−1.5
M_{51}		−34.5	−13.5		−36.6	−17.5
M_{52}		−15.0	−18.0		−11.7	−22.2
M_{53}		−18.0	−26.1		−25.1	−29.9
M_{56}		1.6	−9.7		−2.1	−0.8
M_{V6}	231.0			228.9		
M_{L6}		70.0	70.0		72.3	60.9
M_{65}		−1.6	3.9		1.0	2.9
M_{62}		−25.7	−16.0		−19.6	−22.3
M_{63}		−42.8	−57.9		−46.8	−51.1

TABLE 12.3 Iterative Results of State Estimator Solution

Iteration	$J(\mathbf{x})$ at Beginning of Iteration (pu)	Largest $\Delta\lvert E\rvert$ at End of Iteration (pu V)	Largest $\Delta\theta$ at End of Iteration (rad)
1	3696.86	0.1123	0.06422
2	43.67	0.004866	0.0017
3	40.33	0.0000146	0.0000227

fit of the estimated values to the measurement values. The value of $J(\mathbf{x})$ can, in fact, be used to detect the presence of bad measurements.

The estimated values from the state estimator are shown in Table 12.4, together with the base-case values and the measured values. Notice that, in general, the estimated values do a good job of calculating the true (base-case) conditions from which the measurements were made. For example, measurement M_{23} shows a P flow of 8.6 MW whereas the true flow is 2.9 MW and the estimator predicts a flow of 3.0 MW.

The example shown here started from a base case or "true" state that was shown in Table 12.2. In actual practice, we only have the measurements and the resulting estimate of the state, we never know the "true" state exactly and can only compare measurements with estimates. In the presentations to follow, however, we will leave the base-case or "true" conditions in our illustrations to aid the reader.

The results in Table 12.4 show one of the advantages of using a state estimation algorithm in that, even with measurement errors, the estimation algorithm calculates quantities that are the "best" possible estimates of the true bus voltages and generator, load, and transmission line MW and MVAR values.

There are, however, other advantages to using a state estimation algorithm. First, is the ability of the state estimator to detect and identify bad measurements, and, second, is the ability to estimate quantities that are not measured and telemetered. These are introduced later in the chapter.

12.5 STATE ESTIMATION BY ORTHOGONAL DECOMPOSITION

One problem with the standard least-squares method presented earlier in the chapter is the numerical difficulties encountered with some special state estimation problems. One of these comes about when we wish to drive a state estimator solution to match its measurement almost exactly. This is the case when we have a circuit such as shown in Figure 12.13. All of the actual flows and injections are shown in Figure 12.13 along with the values assumed for the measurements.

In this sample system, the measurement of power at bus 1 will be assumed to be zero MW. If the value of zero is dictated by the fact that the bus has no

TABLE 12.4 State Estimation Solution

	Base-Case Value			Measured Value			Estimated Value		
Measurement	kV	MW	MVAR	kV	MW	MVAR	kV	MW	MVAR
M_{V1}	241.5			238.4			240.6		
M_{G1}		107.9	16.0		113.1	20.2		111.9	18.7
M_{12}		28.7	−15.4		31.5	−13.2		30.4	−14.4
M_{14}		43.6	20.1		38.9	21.2		44.8	21.2
M_{15}		35.6	11.3		35.7	9.4		36.8	11.8
M_{V2}	241.5			237.8			239.9		
M_{G2}		50.0	74.4		48.4	71.9		47.5	70.3
M_{21}		−27.8	12.8		−34.9	9.7		−29.4	11.9
M_{24}		33.1	46.1		32.8	38.3		32.4	45.3
M_{25}		15.5	15.4		17.4	22.0		15.6	14.8
M_{26}		26.2	12.4		22.3	15.0		25.9	10.8
M_{23}		2.9	−12.3		8.6	−11.9		3.0	−12.6
M_{V3}	246.1			250.7			244.7		
M_{G3}		60.0	89.6		55.1	90.6		59.5	87.4
M_{32}		−2.9	5.7		−2.1	10.2		−3.0	6.2
M_{35}		19.1	23.2		17.7	23.9		19.2	22.9
M_{36}		43.8	60.7		43.3	58.3		43.3	58.3
M_{V4}	227.6			225.7			226.1		
M_{L4}		70.0	70.0		71.8	71.9		70.2	70.2
M_{41}		−42.5	−19.9		−40.1	−14.3		−43.6	−20.7
M_{42}		−31.6	−45.1		−29.8	−44.3		−30.9	−44.4
M_{45}		4.1	−4.9		0.7	−17.4		4.3	−5.1
M_{V5}	226.7			225.2			225.3		
M_{L5}		70.0	70.0		72.0	67.7		71.8	69.4
M_{54}		−4.0	−2.8		−2.1	−1.5		−4.2	−2.5
M_{51}		−34.5	−13.5		−36.6	−17.5		−35.6	−13.6
M_{52}		−15.0	−18.0		−11.7	−22.2		−15.1	−17.4
M_{53}		−18.0	−26.1		−25.1	−29.9		−18.1	−25.8
M_{56}		1.6	−9.7		−2.1	−0.8		1.3	−10.1
M_{V6}	231.0			228.9			230.1		
M_{L6}		70.0	70.0		72.3	60.9		68.9	65.8
M_{65}		−1.6	3.9		1.0	2.9		−1.2	4.4
M_{62}		−25.7	−16.0		−19.6	−22.3		−25.4	−14.5
M_{63}		−42.8	−57.9		−46.8	−51.1		−42.3	−55.7

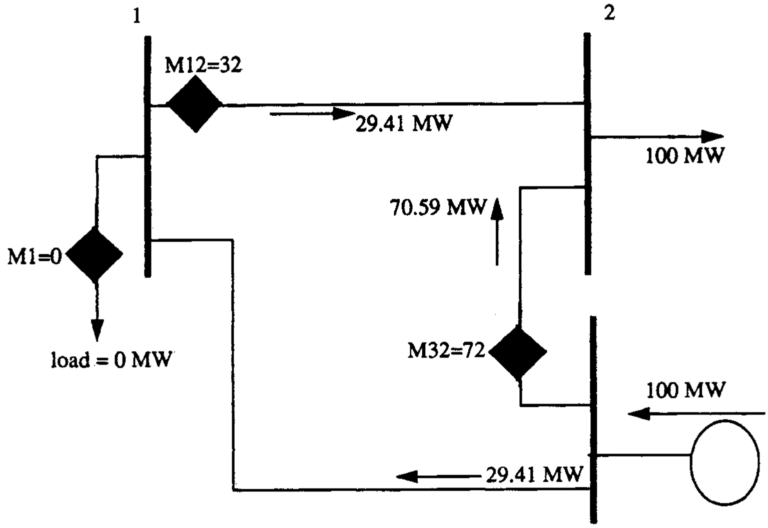

FIG. 12.13 Zero injection system example.

load or generation attached to it, then we know this value of zero MW with certainty and the concept of an error in its "measured" value is meaningless. Nonetheless, we proceed by setting up the standard state estimator equations and specifying the value of the measurement σ for M_1 as: $\sigma_{M1} = 10^{-2}$. This results in the following solution when using the state estimator equations as shown in Eq. 12.23:

$$P_{\text{flow}} \text{ estimate on line 1-2} = 30.76 \text{ MW}$$

$$P_{\text{flow}} \text{ estimate on line 3-2} = 72.52$$

$$\text{Injection estimate on bus 1} = 0.82$$

The estimator has not forced the bus injection to be exactly zero; instead, it reads 0.82 MW. This may not seem like such a big error. However, if there are many such buses (say 100) and they all have errors of this magnitude, then the estimator will have a large amount of load allocated to the buses that are known to be zero.

At first, the solution to this dilemma may seem to be simply forcing the σ value to a very small number for the zero injection buses and rerun the estimator. The problem with this is as follows. Suppose we had changed the zero injection σ to $\sigma_{M1} = 10^{-10}$. Hopefully, this would force the estimator to make the zero injection so dominant that it would result in the correct zero value coming out of the estimator calculation. In this case, the $[H^T R^{-1} H]$ matrix used in the standard least-squares method would look like this for the sample system:

$$[H] = \begin{bmatrix} 5.0 & -5.0 \\ 0 & -4.0 \\ 7.5 & -5.0 \end{bmatrix}$$

$$[R] = \begin{bmatrix} 10^{-4} & & \\ & 10^{-4} & \\ & & 10^{-20} \end{bmatrix}$$

then

$$[H^T R^{-1} H] = \begin{bmatrix} 56.25 \times 10^{20} & -37.5 \times 10^{20} \\ -37.5 \times 10^{20} & 25.0 \times 10^{20} \end{bmatrix}$$

Unfortunately, this matrix is very nearly singular. The reason is that the terms in the matrix are dominated by those terms which are multiplied by the 10^{20} terms from the inverse of the R matrix, and the other terms are so small by comparison that they are lost from the computer (unless one is using an extraordinarily long word length or extra double precision). When the above is presented to a standard matrix inversion routine or run into a Gaussian elimination solution routine, an error message results and garbage comes out of the estimator.

The solution to this dilemma is to use another algorithm for the least-squares solution. This algorithm is called the orthogonal decomposition algorithm and works as follows.

12.5.1 The Orthogonal Decomposition Algorithm

This algorithm goes under several different names in texts on linear algebra. It is often called the QR algorithm or the Gram–Schmidt decomposition. The idea is to take the state estimation least-squares equation, Eq. 12.23, and eliminate the R^{-1} matrix as follows: let

$$[R^{-1}] = R^{-1/2} R^{-1/2} \tag{12.47}$$

where

$$[R^{-1/2}] = \begin{bmatrix} \dfrac{1}{\sigma_{m1}} & & \\ & \dfrac{1}{\sigma_{m2}} & \\ & & \dfrac{1}{\sigma_{m3}} \end{bmatrix} \tag{12.48}$$

then

$$[H^T R^{-1} H]^{-1} = [H^T R^{-1/2} R^{-1/2} H]^{-1} = [H'^T H'] \tag{12.49}$$

with

$$[H'] = [R^{-1/2}][H] \tag{12.50}$$

Finally, Eq. 12.23 becomes

$$\mathbf{x}^{\text{est}} = [H'^T H']^{-1}[H'^T]\mathbf{z}'^{\text{meas}} \tag{12.51}$$

where

$$\mathbf{z}'^{\text{meas}} = [R^{-1/2}]\mathbf{z}^{\text{meas}} \tag{12.52}$$

The idea of the orthogonal decomposition algorithm is to find a matrix $[Q]$ such that:

$$[H'] = [Q][U] \tag{12.53}$$

(Note that in most linear algebra text books, this factorization would be written as $[H'] = [Q][R]$; however, we shall use $[Q][U]$ so as not to confuse the identity of the $[R]$ matrix.)

The matrix $[Q]$ has special properties. It is called an orthogonal matrix so that

$$[Q^T][Q] = [I] \tag{12.54}$$

where $[I]$ is the identity matrix, which is to say that the transpose of $[Q]$ is its inverse. The matrix $[U]$ is now upper triangular in structure, although, since the $[H]$ matrix may not be square, $[U]$ will not be square either. Thus,

$$[H'] = \begin{bmatrix} h'_{11} & h'_{12} \\ h'_{21} & h'_{22} \\ h'_{31} & h'_{32} \end{bmatrix} = [Q][U] = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \\ 0 & 0 \end{bmatrix} \tag{12.55}$$

Now, if we substitute $[Q][U]$ for $[H']$ in the state estimation equation:

$$\mathbf{x}^{\text{est}} = [U^T Q^T Q U]^{-1}[U^T][Q^T]\mathbf{z}' \tag{12.56}$$

or

$$\mathbf{x}^{\text{est}} = [U^T U]^{-1} U^T \hat{\mathbf{z}} \tag{12.57}$$

since

$$[Q^T Q] = I$$

and

$$\hat{\mathbf{z}} = [Q^T]\mathbf{z}' \tag{12.58}$$

Then, by rearranging we get

$$[U^T U]\mathbf{x}^{\text{est}} = [U^T]\hat{\mathbf{z}} \tag{12.59}$$

and we can eliminate U^T from both sides so that we are left with

$$[U]\mathbf{x}^{\text{est}} = \hat{\mathbf{z}} \tag{12.60}$$

or

$$\begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1^{\text{est}} \\ x_2^{\text{est}} \end{bmatrix} = \begin{bmatrix} \hat{z}_1 \\ \hat{z}_2 \\ \hat{z}_3 \end{bmatrix} \tag{12.61}$$

This can be solved directly since U is upper triangular:

$$x_2^{\text{est}} = \frac{\hat{z}_2}{u_{22}} \tag{12.62}$$

and

$$x_1^{\text{est}} = \frac{1}{u_{11}}(\hat{z}_1 - u_{12}x_2^{\text{est}}) \tag{12.63}$$

The Q matrix and the U matrix are obtained, for our simple two-state–three-measurement problem here, using the Givens rotation method as explained in reference 15.

For the Givens rotation method, we start out to define the steps necessary to solve:

$$[Q^T][H] = [U] \tag{12.64}$$

where $[H]$ is a 2×2 matrix:

$$\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$$

and $[U]$ is

$$\begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix}$$

The $[Q]$ matrix must be orthogonal, and when it is multiplied times $[H]$, it eliminates the h_{21} term. The terms in the $[Q]$ matrix are simply:

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix}$$

where

$$c = \frac{h_{11}}{\sqrt{h_{11}^2 + h_{21}^2}} \tag{12.65}$$

and

$$s = \frac{h_{21}}{\sqrt{h_{11}^2 + h_{21}^2}} \tag{12.66}$$

The reader can easily verify that the $[Q]$ matrix is indeed orthogonal and that:

$$\begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix} = \begin{bmatrix} 1 & (ch_{12} + sh_{22}) \\ 0 & (-sh_{12} + ch_{22}) \end{bmatrix} \tag{12.67}$$

When we solve the 3×2 $[H]$ matrix in our three-measurement–two-state sample problem, we apply the Givens rotation three times to eliminate h_{21}, h_{31}, and h_{32}. That is, we need to solve

$$[Q^T]\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \\ h_{31} & h_{32} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \\ 0 & 0 \end{bmatrix} \tag{12.68}$$

We will carry this out in three distinct steps, where each step can be represented as a Givens rotation. The result is that we represent $[Q^T]$ as the product of three matrices:

$$[Q^T] = [N_3][N_2][N_1] \tag{12.69}$$

These matrices are numbered as shown to indicate the order of application. In the case of the 3×2 $[H]$ matrix,

$$[N_1] = \begin{bmatrix} c & s & 0 \\ -s & c & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{12.70}$$

where c and s are defined exactly as before. Next, $[N_2]$ must be calculated so as to eliminate the 31 term which results from $[N_1][H]$. The actual procedure loads $[H]$ into $[U]$ and then determines each $[N]$ based on the current contents of $[U]$. The $[N_2]$ matrix will have terms like

$$[N_2] = \begin{bmatrix} c' & 0 & s' \\ 0 & 1 & 0 \\ -s' & 0 & c' \end{bmatrix} \tag{12.71}$$

where c' and s' are determined from $[N_1][H]$. Similarly for $[N_3]$:

$$[N_3] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c'' & s'' \\ 0 & -s'' & c'' \end{bmatrix} \tag{12.72}$$

For our zero injection example, we start with the $[H]$ and $[R]$ matrices as shown before:

$$[H] = \begin{bmatrix} 5.0 & -5.0 \\ 0 & -4.0 \\ 7.5 & -5.0 \end{bmatrix}$$

and

$$[R] = \begin{bmatrix} 10^{-4} & & \\ & 10^{-4} & \\ & & 10^{-20} \end{bmatrix}$$

Then, the $[H']$ matrix is

$$[H'] = \begin{bmatrix} 5.0 \times 10^{2} & -5.0 \times 10^{2} \\ 0 & -4.0 \times 10^{2} \\ 7.5 \times 10^{10} & -5.0 \times 10^{10} \end{bmatrix}$$

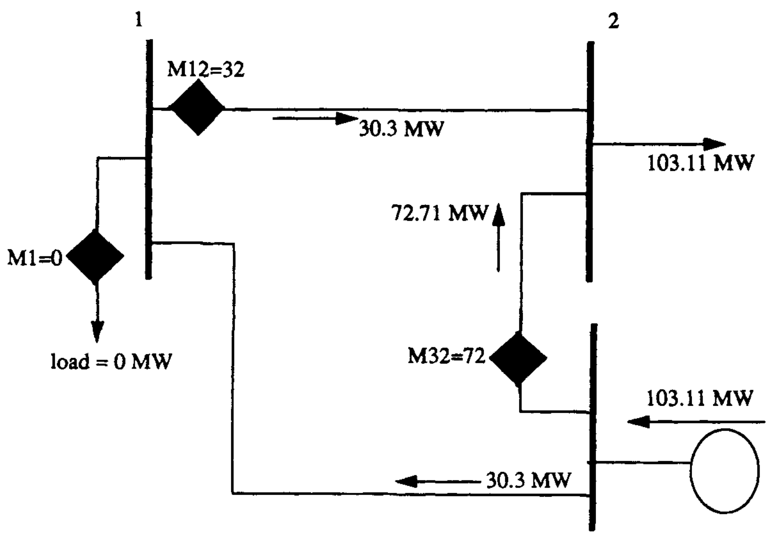

FIG. 12.14 State estimate resulting from orthogonal decomposition algorithm.

and the measurement vector is

$$\hat{\mathbf{z}} = \begin{bmatrix} 32 \\ 72 \\ 0 \end{bmatrix}$$

The resulting state estimate is shown in Figure 12.14. Note particularly that the injection at bus 1 is estimated to be zero, as we desired.

The orthogonal decomposition algorithm has the advantage that measurement weights can be adjusted to extreme values as demonstrated by the numerical example shown. As such, its robust numerical advantages have made it a useful algorithm for power system state estimators.

12.6 AN INTRODUCTION TO ADVANCED TOPICS IN STATE ESTIMATION

12.6.1 Detection and Identification of Bad Measurements

The ability to detect and identify bad measurements is extremely valuable to a power system's operations department. Transducers may have been wired incorrectly or the transducer itself may be malfunctioning so that it simply no longer gives accurate readings. The statistical theory required to understand and anlayze bad measurement detection and identification is straightforward but lengthy. We are going to open the door to the subject in this chapter. The serious student who wishes to pursue this subject should start with the chapter references. For the rest, we present results of these theories and indicate application areas.

To detect the presence of bad measurements, we will rely on the intuitive notion that for a given configuration, the residual, $J(\mathbf{x})$, calculated after the state estimator algorithm converges, will be smallest if there are no bad measurements. When $J(\mathbf{x})$ is small, a vector $\mathbf{x}$ (i.e., voltages and phase angles) has been found that causes all calculated flows, loads, generations, and so forth to closely match all the measurements. Generally, the presence of a bad measurement value will cause the converged value of $J(\mathbf{x})$ to be larger than expected with $\mathbf{x} = \mathbf{x}^{\text{est}}$.

What magnitude of $J(\mathbf{x})$ indicates the presence of bad measurements?

The measurement errors are random numbers so that the value of $J(\mathbf{x})$ is also a random number. If we assume that all the errors are described by their respective normal probability density functions, then we can show that $J(\mathbf{x})$

has a probability density function known as a *chi-squared distribution*, which is written as $\chi^2(K)$. The parameter K is called the degrees of freedom of the chi-squared distribution. This parameter is defined as follows:

$$K = N_m - N_s$$

where

N_m = number of measurements (note that a $P + jQ$ measurement counts as two measurements)

N_s = number of states = $(2n - 1)$

n = number of buses in the network

It can be shown that when $\mathbf{x} = \mathbf{x}^{\text{est}}$, the mean value of $J(\mathbf{x})$ equals K and the standard deviation, $\sigma_{J(\mathbf{x})}$, equals $\sqrt{2K}$.

When one or more measurements are bad, their errors are frequently much larger than the assumed $\pm 3\sigma$ error bound for the measurement. However, even under normal circumstances (i.e., all errors within $\pm 3\sigma$), $J(\mathbf{x})$ can get to be large—although the chance of this happening is small. If we simply set up a threshold for $J(\mathbf{x})$, which we will call t_J, we could declare that bad measurements are present when $J(\mathbf{x}) > t_J$. This threshold test might be wrong in one of two ways. If we set t_J to a small value, we would get many "false alarms." That is, the test would indicate the presence of bad measurements when, in fact, there were none. If we set t_J to be a large value, the test would often indicate that "all is well" when, in fact, bad measurements were present. This can be put on a formal basis by writing the following equation:

$$\text{prob}(J(\mathbf{x}) > t_J \mid J(\mathbf{x}) \text{ is a chi-squared with } K \text{ degrees of freedom}) = \alpha \tag{12.73}$$

This equation says that the probability that $J(\mathbf{x})$ is greater than t_J is equal to α, given that the probability density for $J(\mathbf{x})$ is chi-squared with K degrees of freedom.

This type of testing procedure is formally known as *hypothesis testing*, and the parameter α is called the *significance level* of the test. By choosing a value for the significance level α, we automatically know what threshold t_J to use in our test. When using a t_J derived in this manner, the probability of a "false alarm" is equal to α. By setting α to a small number, for example $\alpha = 0.01$, we would say that false alarms would occur in only 1% of the tests made. A plot of the probability function in Eq. 12.73 is shown in Figure 12.15.

In Table 12.3, we saw that the minimum value for $J(\mathbf{x})$ was 40.33. Looking at Figure 12.12 and counting all $P + jQ$ measurements as two measurements,

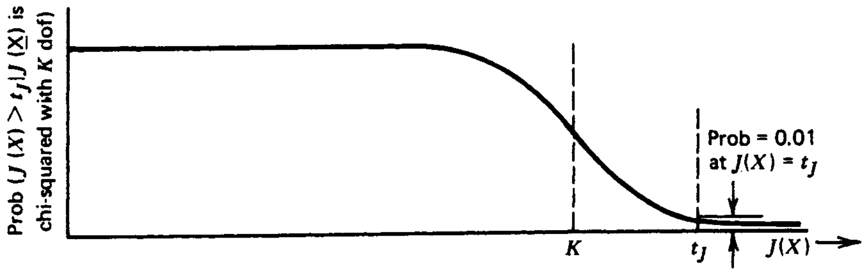

FIG. 12.15 Threshold test probability function.

we see that N_m is equal to 62. Therefore, the degrees of freedom for the chi-square distribution of $J(\mathbf{x})$ in our six-bus sample system is

$$K = N_m - N_s = N_m - (2n - 1) = 51$$

where

$$N_m = 62 \quad \text{and} \quad n = 6$$

If we set our significance level for this test to 0.01 (i.e., $\alpha = 0.01$ in Eq. 12.73), we get a t_J of 76.6.* Therefore, with a $J(\mathbf{x}) = 40.33$, it seems reasonable to assume that there are no "bad" measurements present.

Now let us assume that one of the measurements is truly bad. To simulate this situation, the state estimation algorithm was rerun with the M_{12} measurement reversed. Instead of $P = 31.5$ and $Q = -13.2$, it was set to $P = -31.5$ and $Q = 13.2$. The value of $J(\mathbf{x})$ and the maximum $\Delta|E|$ and $\Delta\theta$ for each iteration for this case are given in Table 12.5. The presence of bad data does not prevent the estimator from converging, but it will increase the value of the residual, $J(\mathbf{x})$.

The calculated flows and voltages for this situation are shown in Table 12.6. Note that the number of degrees of freedom is still 51 but $J(\mathbf{x})$ is now 207.94 at the end of our calculation. Since t_J is 76.6, we would immediately expect bad

TABLE 12.5 Iterative Results with Bad Measurement

Iteration	$J(\mathbf{x})$ at Beginning of Iteration (pu)	Largest $\Delta\|E\|$ at End of Iteration (pu V)	Largest $\Delta\theta$ at End of Iteration (rad)
1	3701.06	0.09851	0.06416
2	211.13	0.004674	0.001481
3	207.94	0.00002598	0.00004848

* Standard tables of $\chi^2(K)$ usually only go up to $K = 30$. For $K > 30$, a very close approximation to $\chi^2(K)$ using the normal distribution can be used. The student should consult any standard reference on probability and statistics to see how this is done.

TABLE 12.6 State Estimation Solution with Measurement M_{12} Reversed

	Base-Case Value			Measured Value			Estimated Value		
Measurement	kV	MW	MVAR	kV	MW	MVAR	kV	MW	MVAR
M_{V1}	241.5			238.4			240.6		
M_{G1}		107.9	16.0		113.1	20.2		99.3	21.9
M_{12}		28.7	−15.4		−31.5	+13.2		25.0	−12.2
M_{14}		43.6	20.1		38.9	21.2		40.6	21.9
M_{15}		35.6	11.3		35.7	9.4		33.7	12.3
M_{V2}	241.5			237.8			239.9		
M_{G2}		50.0	74.4		48.4	71.9		54.4	67.0
M_{21}		−27.8	12.8		−34.9	9.7		−24.4	9.2
M_{24}		33.1	46.1		32.8	38.3		35.0	44.1
M_{25}		15.5	15.4		17.4	22.0		16.3	14.7
M_{26}		26.2	12.4		22.3	15.0		25.1	11.3
M_{23}		2.9	−12.3		8.6	−11.9		2.3	−12.2
M_{V3}	246.1			250.7			244.6		
M_{G3}		60.0	89.6		55.1	90.6		61.4	86.3
M_{32}		−2.9	5.7		−2.1	10.2		−2.3	5.8
M_{35}		19.1	23.2		17.7	23.9		−20.5	22.2
M_{36}		43.8	60.7		43.3	58.3		43.2	58.2
M_{V4}	227.6			225.7			226.1		
M_{L4}		70.0	70.0		71.8	71.9		69.0	70.0
M_{41}		−42.5	−19.9		−40.1	−14.3		−39.6	−21.9
M_{42}		−31.6	−45.1		−29.8	−44.3		−33.5	−43.1
M_{45}		4.1	−4.9		0.7	−17.4		4.1	−5.0
M_{V5}	226.7			225.2			225.3		
M_{L5}		70.0	70.0		72.0	67.7		71.8	69.3
M_{54}		−4.0	−2.8		−2.1	−1.5		−4.1	−2.6
M_{51}		−34.5	−13.5		−36.6	−17.5		−32.7	−14.7
M_{52}		−15.0	−18.0		−11.7	−22.2		−15.8	−17.2
M_{53}		−18.0	−26.1		−25.1	−29.9		−19.3	−25.1
M_{56}		1.6	−9.7		−2.1	−0.8		0.1	−9.6
M_{V6}	231.0			228.9			230.0		
M_{L6}		70.0	70.0		72.3	60.9		66.9	66.7
M_{65}		−1.6	3.9		1.0	2.9		−0.1	3.9
M_{62}		−25.7	−16.0		−19.6	−22.3		−24.6	−15.0
M_{63}		−42.8	−57.9		−46.8	−51.1		−42.3	−55.6

measurements at our 0.01 significance level. If we had not known ahead of running the estimation algorithm that a bad measurement was present, we would certainly have had good reason to suspect its presence when so large a $J(\mathbf{x})$ resulted.

So far, we can say that by looking at $J(\mathbf{x})$, we can detect the presence of bad measurements. But if bad measurements are present, how can one tell which measurements are bad? Without going into the statistical theory, we give the following explanation of how this is accomplished.

Suppose we are interested in the measurement of megawatt flow on a particular line. Call this measured value z_i. In Figure 12.16(a) we have a plot of the normal probability density function of z_i. Since we assume that the error in z_i is normally distributed with zero mean value, the probability density function is centered on the true value of z_i. Since the errors on all the measurements are assumed normal, we will assume that the estimate, $\mathbf{x}^{\text{est}}$ is approximately normally distributed and that any quantity that is a function of $\mathbf{x}^{\text{est}}$ is also an approximately normally distributed quantity. In Figure 12.16(b), we show the probability density function for the calculated megawatt flow, f_i, which is a function of the estimated state, $\mathbf{x}^{\text{est}}$. We have drawn the density function of f_i as having a smaller deviation from its mean than the measurement z_i to indicate that, due to redundancy in measurements, the estimate is more accurate.

The difference between the estimate, f_i, and the measurement, z_i, is called the *measurement residual* and is designated y_i. The probability density function for y_i is also normal and is shown in Figure 12.16(c) as having a zero mean and a standard deviation of σ_{y_i}. If we divide the difference between the estimate f_i and the measurement z_i by σ_{y_i}, we obtain what is called a *normalized measurement residual.* The normalized measurement residual is designated y_i^{norm} and is shown in Figure 12.16(d) along with its probability density function, which is normal and has a standard deviation of unity. If the absolute value of y_i^{norm} is greater than 3, we have good reason to suspect that z_i is a bad measurement value. The usual procedure in identifying bad measurements is to calculate all f_i values for the N_m measurements once $\mathbf{x}^{\text{est}}$ is available from the state estimator. Using the z_i values that were used in the estimator and the f_i values, a measurement residual y_i can be calculated for each measurement. Also, using information from the state estimator, we can calculate σ_{y_i} (see references for details of this calculation). Using y_i and σ_{y_i}, we can calculate a normalized residual for each measurement. Measurements having the largest absolute normalized residual are labeled as prime suspects. These prime suspects are removed from the state estimator calculation one at a time, starting with the measurement having the largest normalized residual. After a measurement has been removed, the state estimation calculation (see Figure 12.11) is rerun. This results in a different $\mathbf{x}^{\text{est}}$ and therefore a different $J(\mathbf{x})$. The chi-squared probability density function for $J(\mathbf{x})$ will have to be recalculated, assuming that we use the same significance level for our test. If the new $J(\mathbf{x})$ is now less than the new value for t_J, we can say that the measurement that

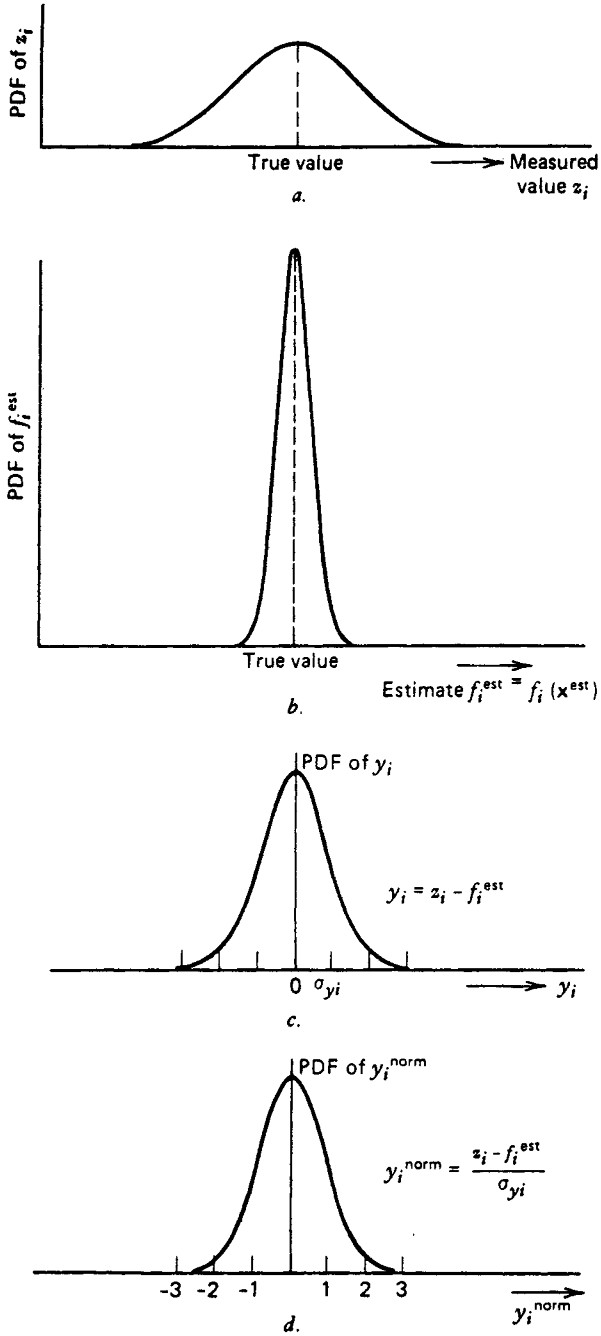

FIG. 12.16 Probability density function of the normalized measurement residual.

was removed has been identified as bad. If, however, the new $J(\mathbf{x})$ is greater than the new t_J, we must proceed to recalculate $f_i(\mathbf{x}^{est})$, σ_{y_i}, and then y_i^{norm} for each of the remaining measurements. The measurement with the largest absolute y_i^{norm} is then again removed and the entire procedure repeated successively until $J(\mathbf{x})$ is less than t_J. The references at the end of this chapter discuss a problem that the identification process may encounter, wherein several measurements may need to be removed to eliminate one "bad" measurement. That is, the identification procedure often cannot pinpoint a single bad measurement but instead identifies a group of measurements, one of which is bad. In such cases, the groups must be eliminated to eliminate the bad measurement.

The ability to detect (using the chi-squared statistic) and identify (using normalized residuals) are extremely useful features of a state estimator. Without the state estimator calculation using the system measurement data, those measurements whose values are not obviously wrong have little chance of being detected and identified. With the state estimator, the operations personnel have a greater assurance that quantities being displayed are not grossly in error.

12.6.2 Estimation of Quantities Not Being Measured

The other useful feature of a state estimator calculation is the ability to calculate (or estimate) quantities not being telemetered. This is most useful in cases of failure of communication channels connecting operations centers to remote data-gathering equipment or when the remote data-gathering equipment fails. Often data from some network substations are simply unavailable because no transducers or data-gathering equipment were ever installed.

An example of this might be the failure of all telemetry from buses 3, 4, 5, and 6 in our six-bus system. Even with the loss of these measurements, we can run the state estimation algorithm on the remaining measurements at buses 1 and 2, calculate the bus voltage magnitudes and phase angles at all six buses, and then calculate all network generations, loads, and flows. The results of such a calculation are given in Table 12.7. Notice that the estimate of quantities at the untelemetered buses are not as close to the base case as when using the full set of measurements (i.e., compare Table 12.7 to Table 12.4).

12.6.3 Network Observability and Pseudo-measurements

What happens if we continue to lose telemetry so that fewer and fewer measurements are available? Eventually, the state estimation procedure breaks down completely. Mathematically, the matrix

$$[[H]^T[R^{-1}][H]]$$

in Eq. 12.46 becomes singular and cannot be inverted. There is also a very interesting engineering interpretation of this phenomenon that allows us to alter the situation so that the state estimation procedure is not completely disabled.

TABLE 12.7 State Estimation Solution with Measurement at Buses 1 and 2 Only

	Base-Case Value			Measured Value			Estimated Value		
Measurement	kV	MW	MVAR	kV	MW	MVAR	kV	MW	MVAR
M_{V1}	241.5			238.4			238.8		
M_{G1}		107.9	16.0		113.1	20.2		112.4	20.5
M_{12}		28.7	−15.4		31.5	−13.2		30.6	−13.4
M_{14}		43.6	20.1		38.9	21.2		44.7	19.4
M_{15}		35.6	11.3		35.7	9.4		37.1	14.6
M_{V2}	241.5			237.8			237.6		
M_{G2}		50.0	74.4		48.4	71.9		48.2	71.7
M_{21}		−27.8	12.8		−34.9	9.7		−29.6	11.1
M_{24}		33.1	46.1		32.8	38.3		30.5	40.2
M_{25}		15.5	15.4		17.4	22.0		16.1	16.8
M_{26}		26.2	12.4		22.3	15.0		22.4	15.2
M_{23}		2.9	−12.3		8.6	−11.9		8.8	−11.7
M_{V3}	246.1						241.4		
M_{G3}		60.0	89.6					27.2	94.9
M_{32}		−2.9	5.7					−8.7	5.5
M_{35}		19.1	23.2					15.1	25.3
M_{36}		43.8	60.7					20.9	64.0
M_{V4}	227.6						225.0		
M_{L4}		70.0	70.0					67.6	61.2
M_{41}		−42.5	−19.9					−43.6	−18.9
M_{42}		−31.6	−45.1					−29.3	−39.7
M_{45}		4.1	−4.9					5.3	−2.6
M_{V5}	226.7						221.4		
M_{L5}		70.0	70.0					71.9	76.7
M_{54}		−4.0	−2.8					−5.2	−4.8
M_{51}		−34.5	−13.5					−35.9	−15.9
M_{52}		−15.0	−18.0					−15.5	−19.0
M_{53}		−18.0	−26.1					−14.0	−28.0
M_{56}		1.6	−9.7					−1.4	−9.0
M_{V6}	231.0						226.2		
M_{L6}		70.0	70.0					40.5	77.2
M_{65}		−1.6	3.9					1.4	3.4
M_{62}		−25.7	−16.0					−21.9	−18.8
M_{63}		−42.8	−57.9					−20.0	−61.8

If we take the three-bus example used in the beginning of Section 12.2, we note that when all three measurements are used, we have a redundant set and we can use a least-squares fit to the measurement values. If one of the measurements is lost, we have just enough measurements to calculate the states. If, however, two measurements are lost, we are in trouble. For example, suppose M_{13} and M_{32} were lost leaving only M_{12}. If we now apply Eq. 12.23 in a straightforward manner, we get

$$M_{12} = f_{12} = \frac{1}{0.2}(\theta_1 - \theta_2) = 5\theta_1 - 5\theta_2$$

Then

$$[H] = [5 \quad -5]$$

$$[R] = [\sigma_{M12}^2] = [0.0001]$$

and

$$\begin{bmatrix} \theta_1^{\text{est}} \\ \theta_2^{\text{est}} \end{bmatrix} = \left[\begin{bmatrix} 5 \\ -5 \end{bmatrix}[0.0001]^{-1}[5 \quad -5]\right]^{-1}[5 \quad -5][0.0001]^{-1}(0.55)$$

$$= \begin{bmatrix} 2500 & -2500 \\ -2500 & 2500 \end{bmatrix}^{-1}[5 \quad -5][0.0001]^{-1}(0.55) \qquad (12.74)$$

The matrix to be inverted in Eq. 12.74 is clearly singular and, therefore, we have no way of solving for θ_1^{est} and θ_2^{est}. Why is this? The reasons become quite obvious when we look at the one-line diagram of this network as shown in Figure 12.17. With only M_{12} available, all we can say about the network is that

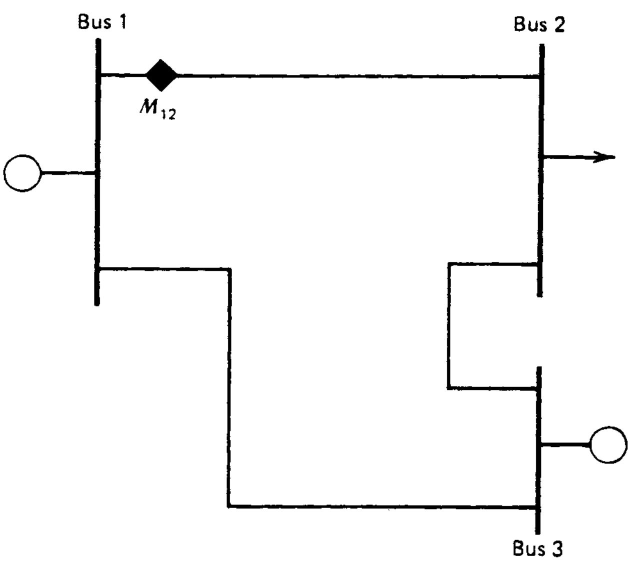

FIG. 12.17 "Unobservable" measurement set.

the phase angle across line 1-2 must be 0.11 rad, but with no other information available, we cannot tell what relationship θ_1 or θ_2 has to θ_3, which is assumed to be 0 rad. If we write down the equations for the net injected power at bus 1 and bus 2, we have

$$\begin{aligned} P_1 &= 7.5\theta_1 - 5\theta_2 \\ P_2 &= -5\theta_1 + 9\theta_2 \end{aligned} \tag{12.75}$$

If measurement M_{12} is reading 55 MW (0.55 pu), we have

$$\theta_1 - \theta_2 = 0.11 \tag{12.76}$$

and by substituting Eq. 12.75 into Eq. 12.76 and eliminating θ_1, we obtain

$$P_2 = 1.6P_1 - 1.87 \tag{12.77}$$

Furthermore,

$$P_3 = -P_1 - P_2 = -0.6P_1 + 1.87 \tag{12.78}$$

Equations 12.77 and 12.78 give a relationship between P_1, P_2, and P_3, but we still do not know their correct values. The technical term for this phenomenon is to say that the network is *unobservable*; that is, with only M_{12} available, we cannot observe (calculate) the state of the system.

It is very desirable to be able to circumvent this problem. Often a large power-system network will have missing data that render the network unobservable. Rather than just stop the calculations, a procedure is used that allows the estimator calculation to continue. The procedure involves the use of what are called *pseudo-measurements*. If we look at Eqs. 12.77 and 12.78, it is obvious that θ_1 and θ_2 could be estimated if the value of any one of the bus injections (i.e., P_1, P_2, or P_3) could be determined by some means other than direct measurement. This value, the pseudo-measurement, is used in the state estimator just as if it were an actual measured value.

To determine the value of an injection without measuring it, we must have some knowledge about the power system beyond the measurements currently being made. For example, it is customary to have access to the generated MW and MVAR values at generating stations through telemetry channels (i.e., the generated MW and MVAR would normally be measurements available to the state estimator). If these channels are out and we must have this measurement for observability, we can probably communicate with the operators in the plant control room by telephone and ask for the MW and MVAR values and enter them into the state estimator calculation manually. Similarly, if we needed a load-bus MW and MVAR for a pseudo-measurement, we could use historical records that show the relationship between an individual load and the total system load. We can estimate the total system load fairly accurately by knowing the total power being generated and estimating the network losses. Finally, if we have just experienced a telemetry failure, we could use the most recently

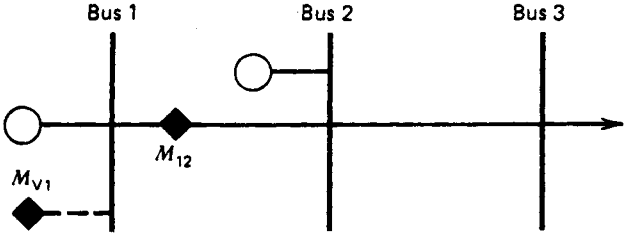

FIG. 12.18 Unobservable system showing importance of location of pseudo-measurements.

estimated values from the estimator (assuming that it is run periodically) as pseudo-measurements. Therefore, if needed, we can provide the state estimator with a reasonable value to use as a pseudo-measurement at any bus in the system.

The three-bus sample system in Figure 12.18 requires one pseudo-measurement. Measurement M_{12} allows us to estimate the voltage magnitude and phase angle at bus 2 (bus 1's voltage magnitude is measured and its phase angle is assumed to be zero). But without knowing the generation output at the generator unit on bus 2 or the load on bus 3, we cannot tell what voltage magnitude and phase angle to place on bus 3; hence, the network is unobservable. We can make this three-bus system observable by adding a pseudo-measurement of the net bus injected MW and MVAR at bus 2 or bus 3, but not at bus 1. That is, a pseudo-measurement at bus 1 will do no good at all because it tells nothing about the relationship of the phase angles between bus 2 and bus 3.

When adding a pseudo-measurement to a network, we simply write the equation for the pseudo-measurement injected power as a function of bus voltage magnitudes and phase angles as if it were actually measured. However, we do not wish to have the estimator treat the pseudo-measurement the same as a legitimate measurement, since it is often quite inaccurate and is little better than a guess. To circumvent this difficulty, we assign a large standard deviation to this measurement. The large standard deviation allows the estimator algorithm to treat the pseudo-measurement as if it were a measurement from a very poor-quality metering device.

To demonstrate the use of pseudo-measurements on our six-bus test system, all measurements were removed from buses 2, 3, 4, 5, and 6 so that bus 1 had all remaining measurements. This rendered the network unobservable and required adding pseudo-measurements at buses 2, 3, and 6. In the case, the pseudo-measurements were just taken from our base-case power flow. The results are shown in Table 12.8. Notice that the resulting estimates are quite close to the measured values for bus 1 but that the remaining buses have large measurement residuals. The net injections at buses 2, 3, and 6 do not closely match the pseudo-measurements since the pseudo-measurements were weighted much less than the legitimate measurements.

TABLE 12.8 State Estimation Solution with Measurements at Bus 1 and Pseudo-measurements at Buses 2, 3, and 6

	Base-Case Value			Measured Value			Estimated Value		
Measurement	kV	MW	MVAR	kV	MW	MVAR	kV	MW	MVAR
M_{V1}	241.5			238.4			238.4		
M_{G1}		107.9	16.0		113.1	20.2		111.4	19.5
M_{12}		28.7	−15.4		31.5	−13.2		33.3	−12.5
M_{14}		43.6	20.1		38.9	21.2		40.7	21.9
M_{15}		35.6	11.3		35.7	9.4		37.4	10.1
M_{V2}	241.5						236.2		
M_{G2}		50.0	74.4	Pseudo:	50.0	74.4		37.5	67.7
M_{21}		−27.8	12.8					−32.1	10.5
M_{24}		33.1	46.1					19.5	44.9
M_{25}		15.5	15.4					14.1	11.5
M_{26}		26.2	12.4					30.0	12.7
M_{23}		2.9	−12.3					6.0	−11.9
M_{V3}	246.1						240.5		
M_{G3}		60.0	89.6	Pseudo:	60.0	89.6		52.6	86.6
M_{32}		−2.9	5.7					−6.0	5.7
M_{35}		19.1	23.2					14.3	19.5
M_{36}		43.8	60.7					44.2	61.4
M_{V4}	227.6						223.8		
M_{L4}		70.0	70.0					51.9	73.3
M_{41}		−42.5	−19.9					−39.6	−21.8
M_{42}		−31.6	−45.1					−18.3	−44.6
M_{45}		4.1	−4.9					6.0	−6.9
M_{V5}	226.7						224.0		
M_{L5}		70.0	70.0					63.9	55.5
M_{54}		−4.0	−2.8					−5.9	−0.4
M_{51}		−34.5	−13.5					−36.3	−11.8
M_{52}		−15.0	−18.0					−13.7	−14.4
M_{53}		−18.0	−26.1					−13.6	−22.9
M_{56}		1.6	−9.7					5.5	−5.9
M_{V6}	231.0						224.9		
M_{L6}		70.0	70.0	Pseudo:	70.0	70.0		77.9	73.4
M_{65}		−1.6	3.9					−5.5	0.3
M_{62}		−25.7	−16.0					−29.3	−15.6
M_{63}		−42.8	−57.9					−43.2	−58.1

12.7 APPLICATION OF POWER SYSTEMS STATE ESTIMATION

In this last section, we will try to present the "big picture" showing how state estimation, contingency analysis, and generator corrective action fit together in a modern operations control center. Figure 12.19 is a schematic diagram showing the information flow between the various functions to be performed in an operations control center computer system. The system gets its information about the power system from remote terminal units that encode measurement transducer outputs and opened/closed status information into digital signals that are transmitted to the operations center over communications circuits. In addition, the control center can transmit control information such as raise/lower commands to generators and open/close commands to circuit breakers and switches. We have broken down the information coming into the control center as breaker/switch status indications and analog measurements. The analog measurements of generator output must be used directly by the AGC program (see Chapter 9), whereas all other data will be processed by the state estimator before being used by other programs.

In order to run the state estimator, we must know how the transmission lines are connected to the load and generation buses. We call this information the *network topology*. Since the breakers and switches in any substation can cause the network topology to change, a program must be provided that reads the telemetered breaker/switch status indications and restructures the electrical model of the system. An example of this is shown in Figure 12.20, where the opening of four breakers requires two electrical buses to represent the substation instead of one electrical bus. We have labeled the program that reconfigures the electrical model as the *network topology program.** The network topology program must have a complete description of each substation and how the transmission lines are attached to the substation equipment. Bus sections that are connected to other bus sections through closed breakers or switches are designated as belonging to the same electrical bus. Thus, the number of electrical buses and the manner in which they are interconnected can be changed in the model to reflect breaker and switch status changes on the power system itself.

As seen in Figure 12.20, the electrical model of the power system's transmission system is sent to the state estimator program together with the analog measurements. The output of the state estimator consists of all bus voltage magnitudes and phase angles, transmission line MW and MVAR flows calculated from the bus voltage magnitude and phase angles, and bus loads and generations calculated from the line flows. These quantities, together with the electrical model developed by the network topology program, provide the basis for the economic dispatch program, congtingency analysis program, and generation corrective action program. Note that since the complete electrical model of the transmission system is available, we can directly calculate bus penalty factors as shown in Chapter 4.

* Alternative names that are often used for this program are "system status processor" and "network configurator."

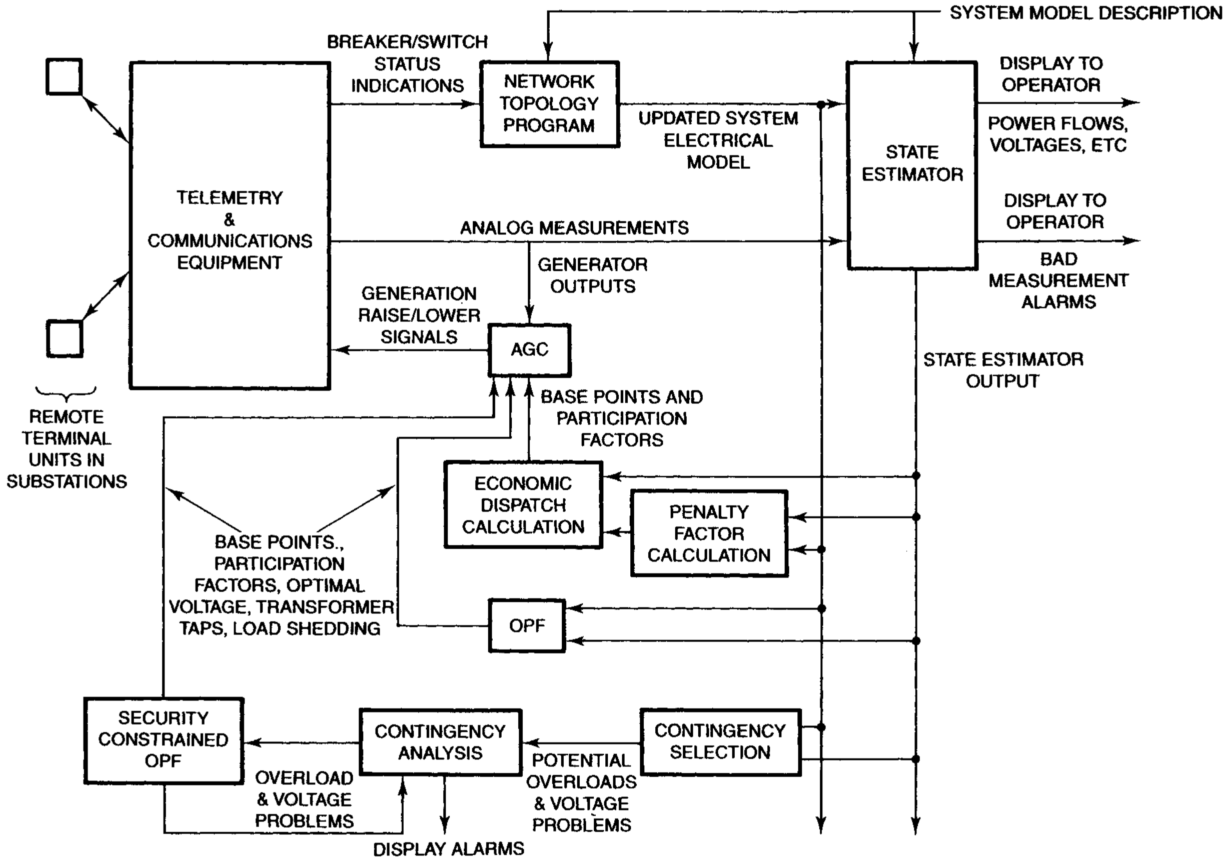

FIG. 12.19 Energy control center system security schematic.

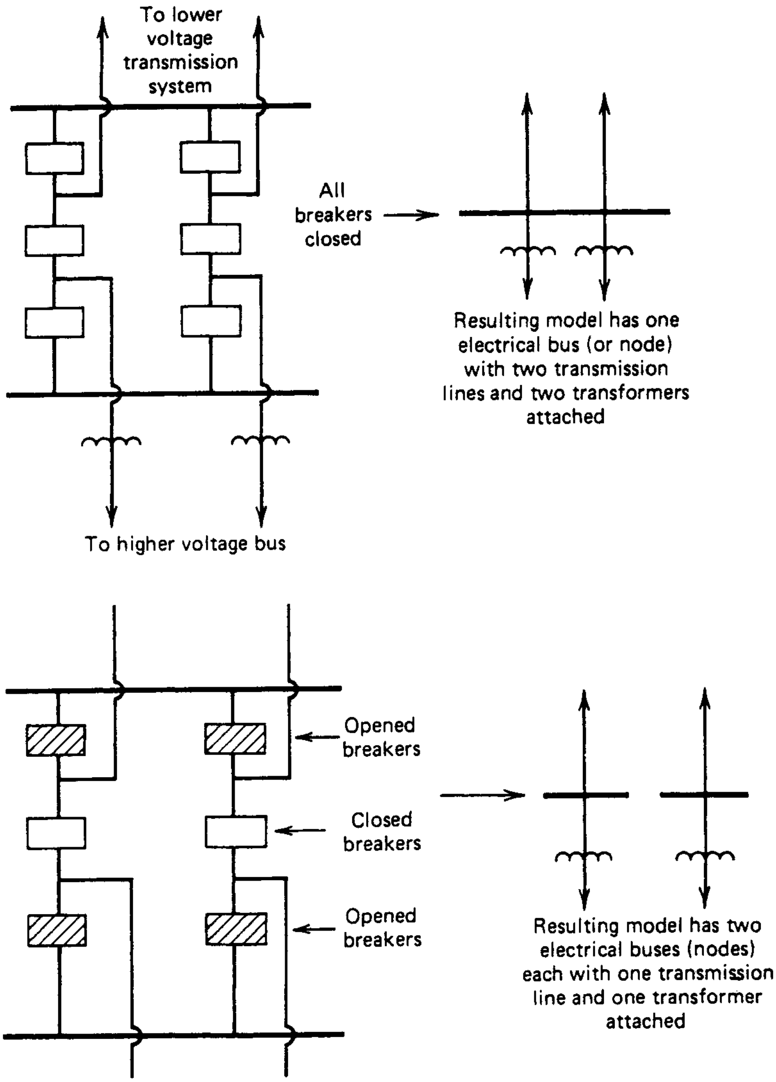

FIG. 12.20 Example of network topology updating.

APPENDIX
Derivation of Least-Squares Equations

One is often confronted with problems wherein data have been obtained by making measurements or taking samples on a process. Furthermore, the quantities being measured are themselves functions of other variables that we wish to estimate. These other variables will be called the state variables and designated $\mathbf{x}$, where the number of state variables is N_s. The measurement values will be called $\mathbf{z}$. We will assume here that the process we are interested in can be modeled using a linear model. Then we say that each measurement z_i is a linear function of the states x_i; that is,

$$z_i = h_i(\mathbf{x}) = h_{i1}x_1 + h_{i2}x_2 + \ldots + h_{iN_s}x_{N_s} \tag{12A.1}$$

We can also write this equation as a vector equation if we place the h_{ij} coefficients into a vector $\mathbf{h}$; that is,

$$\mathbf{h}_i = \begin{bmatrix} h_{i1} \\ h_{i2} \\ \vdots \\ h_{iN_s} \end{bmatrix} \tag{12A.2}$$

Then Eq. 12A.1 becomes

$$z_i = \mathbf{h}_i^T \mathbf{x} \tag{12A.3}$$

where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N_s} \end{bmatrix}$$

Finally, we can write all the measurement equations in a compact form

$$\mathbf{z} = [H]\mathbf{x} \tag{12A.4}$$

where

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{N_m} \end{bmatrix}$$

$$[H] = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1N_s} \\ h_{21} & h_{22} & \cdots & \\ \vdots & & & \\ h_{N_m1} & & \cdots & h_{N_mN_s} \end{bmatrix}$$

where row i of $[H]$ is equal to vector $\mathbf{h}_i^T$ (see Eq. 12A.2).

With N_m measurements we can have three possible cases to solve. That is,

N_s, the number of states, is either less than N_m, equal to N_m, or greater than N_m. We will deal with each case separately.

The Overdetermined Case ($N_m > N_s$)

In this case, we have more measurements or samples than state variables; therefore, we can write more equations, $h_i(\mathbf{x})$, than we have unknowns x_j. One way to estimate the x_i values is to minimize the sum of the squares of difference between the measurement values z_i and the estimate of z_i that is, in turn, a function of the estimates of x_i. That is, we wish to minimize

$$J(\mathbf{x}) = \sum_{i=1}^{N_m} [z_i - h_i(x_1, x_2, \ldots, x_{N_s})]^2 \tag{12A.5}$$

Equation 12A.5 can be written as

$$J(\mathbf{x}) = \sum_{i=1}^{N_m} (z_i - \mathbf{h}_i^T\mathbf{x})^2 \tag{12A.6}$$

and this can be written in a still more compact form as

$$J(\mathbf{x}) = (\mathbf{z} - [H]\mathbf{x})^T(\mathbf{z} - [H]\mathbf{x}) \tag{12A.7}$$

If we wish to find the value of $\mathbf{x}$ that minimizes $J(\mathbf{x})$, we can take the first derivative of $J(\mathbf{x})$ with respect to each x_j ($j = 1, \ldots, N_s$) and set these derivatives to zero. That is,

$$\frac{\partial J(\mathbf{x})}{\partial x_j} = 0 \quad \text{for } j = 1 \ldots N_s \tag{12A.8}$$

If we place these derivatives into a vector, we have what is called the gradient of $J(\mathbf{x})$, which is written $\nabla_x \mathbf{J}(\mathbf{x})$. Then,

$$\nabla_x \mathbf{J}(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial J(\mathbf{x})}{\partial x_1} \\ \dfrac{\partial J(\mathbf{x})}{\partial x_2} \\ \vdots \end{bmatrix} \tag{12A.9}$$

Then the goal of forcing each derivative to zero can be written as

$$\nabla_x \mathbf{J}(\mathbf{x}) = \mathbf{0} \tag{12A.10}$$

where $\mathbf{0}$ is a vector of N_s elements, each of which is zero. To solve this problem, we will first expand Eq. 12A.7:

$$\begin{aligned} J(\mathbf{x}) &= (\mathbf{z} - [H]\mathbf{x})^T(\mathbf{z} - [H]\mathbf{x}) \\ &= \mathbf{z}^T\mathbf{z} - \mathbf{x}^T[H]^T\mathbf{z} - \mathbf{z}^T[H]\mathbf{x} + \mathbf{x}^T[H]^T[H]\mathbf{x} \end{aligned} \tag{12A.11}$$

The second and third term in Eq. 12.A.11 are identical, so that we can write

$$J(\mathbf{x}) = \mathbf{z}^T\mathbf{z} - 2\mathbf{z}^T[H]\mathbf{x} + \mathbf{x}^T[H]^T[H]\mathbf{x} \tag{12A.12}$$

Before proceeding, we will derive a few simple relationships.

The gradient is always a vector of first derivatives of a scalar function that is itself a function of a vector. Thus, if we define $F(\mathbf{y})$ to be a scalar function, then its gradient $\nabla_y \mathbf{F}$ is:

$$\nabla_y \mathbf{F} = \begin{bmatrix} \dfrac{\partial F}{\partial y_1} \\ \dfrac{\partial F}{\partial y_2} \\ \vdots \\ \dfrac{\partial F}{\partial y_n} \end{bmatrix} \tag{12A.13}$$

Then, if we define F as follows:

$$F = \mathbf{y}^T\mathbf{b} = [y_1 \quad y_2 \quad \cdots] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \end{bmatrix} \tag{12A.14}$$

where $\mathbf{b}$ is a vector of constants b_i, $i = 1, \ldots, n$, then, F can be expanded as

$$F = y_1 b_1 + y_2 b_2 + y_3 b_3 + \ldots \tag{12A.15}$$

and the gradient of F is

$$\nabla_y \mathbf{F} = \begin{bmatrix} \dfrac{\partial F}{\partial y_1} \\ \dfrac{\partial F}{\partial y_2} \\ \vdots \\ \dfrac{\partial F}{\partial y_n} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \mathbf{b} \tag{12A.16}$$

It ought to be obvious that writing F with $\mathbf{y}$ and $\mathbf{b}$ reversed makes no difference. That is,

$$F = \mathbf{b}^T\mathbf{y} = \mathbf{y}^T\mathbf{b} \tag{12A.17}$$

and, therefore, $\nabla_y(\mathbf{b}^T\mathbf{y}) = \mathbf{b}$.

Suppose we now write the vector $\mathbf{b}$ as the product of a matrix $[A]$ and a vector $\mathbf{u}$.

$$\mathbf{b} = [A]\mathbf{u} \tag{12A.18}$$

Then, if we take F as shown in Eq. 12A.14,

$$F = \mathbf{y}^T\mathbf{b} = \mathbf{y}^T[A]\mathbf{u} \tag{12A.19}$$

we can say that

$$\nabla_y\mathbf{F} = [A]\mathbf{u} \tag{12A.20}$$

Similarly, we can define

$$\mathbf{b}^T = \mathbf{u}^T[A] \tag{12A.21}$$

If we can take F as shown in Eq. 12A.17,

$$F = \mathbf{b}^T\mathbf{y} = \mathbf{u}^T[A]\mathbf{y}$$

then

$$\nabla_y\mathbf{F} = [A]^T\mathbf{u} \tag{12A.22}$$

Finally, we will look at a scalar function F that is quadratic, namely,

$$\begin{aligned} F &= \mathbf{y}^T[A]\mathbf{y} \\ &= [y_1 \quad y_2 \quad \cdots \quad y_n]\begin{bmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & \vdots & \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} y_i a_{ij} y_j \end{aligned} \tag{12A.23}$$

Then

$$\begin{aligned} \nabla_y\mathbf{F} &= \begin{bmatrix} \dfrac{\partial F}{\partial y_1} \\ \dfrac{\partial F}{\partial y_2} \\ \vdots \\ \dfrac{\partial F}{\partial y_n} \end{bmatrix} = \begin{bmatrix} 2a_{11}y_1 + 2a_{12}y_2 + \cdots \\ 2a_{21}y_1 + 2a_{22}y_2 + \cdots \\ \vdots \end{bmatrix} \\ &= 2[A]\mathbf{y} \end{aligned} \tag{12A.24}$$

Then, in summary:

$$
\boxed{
\begin{array}{ll}
1.\ F = \mathbf{y}^T\mathbf{b} & \nabla_y \mathbf{F} = \mathbf{b} \\
2.\ F = \mathbf{b}^T\mathbf{y} & \nabla_y \mathbf{F} = \mathbf{b} \\
3.\ F = \mathbf{y}^T[A]\mathbf{u} & \nabla_y \mathbf{F} = [A]\mathbf{u} \\
4.\ F = \mathbf{u}^T[A]\mathbf{y} & \nabla_y \mathbf{F} = [A]^T\mathbf{u} \\
5.\ F = \mathbf{y}^T[A]\mathbf{y} & \nabla_y \mathbf{F} = 2[A]\mathbf{y}
\end{array}
}
\tag{12A.25}
$$

We will now use Eq. 12A.25 to derive the gradient of $J(\mathbf{x})$, that is $\nabla_x \mathbf{J}$, where $J(\mathbf{x})$ is shown in Eq. 12A.12. The first term, $\mathbf{z}^T\mathbf{z}$ is not a function of $\mathbf{x}$, so we can discard it. The second term is of the same form as (4) in Eq. 12A.25, so that,

$$\nabla_x(-2\mathbf{z}^T[H]\mathbf{x}) = -2[H]^T\mathbf{z} \tag{12A.26}$$

The third term is the same as (5) in Eq. 12A.25 with $[H]^T[H]$ replacing $[A]$; then,

$$\nabla_x(\mathbf{x}^T[H]^T[H]\mathbf{x}) = 2[H]^T[H]\mathbf{x} \tag{12A.27}$$

Then from Eqs. 12A.26 and 12A.27 we have

$$\nabla_x \mathbf{J} = -2[H]^T\mathbf{z} + 2[H]^T[H]\mathbf{x} \tag{12A.28}$$

But, as stated in Eq. A.10, we wish to force $\nabla_x \mathbf{J}$ to zero. Then

$$-2[H]^T\mathbf{z} + 2[H]^T[H]\mathbf{x} = 0$$

or

$$\mathbf{x} = [[H]^T[H]]^{-1}[H]^T\mathbf{z} \tag{12A.29}$$

If we had wanted to put a different weight, w_i, on each measurement, we could have written Eq. 12A.6 as

$$J(\mathbf{x}) = \sum_{i=1}^{N_m} w_i(z_i - \mathbf{h}_i^T\mathbf{x})^2 \tag{12A.30}$$

which can be written as

$$J(\mathbf{x}) = (\mathbf{z} - [H]\mathbf{x})^T[W](\mathbf{z} - [H]\mathbf{x})$$

where $[W]$ is a diagonal matrix. Then

$$J(\mathbf{x}) = \mathbf{z}^T[W]\mathbf{z} - \mathbf{x}^T[H]^T[W]\mathbf{z} - \mathbf{z}^T[W][H]\mathbf{x} + \mathbf{x}^T[H]^T[W][H]\mathbf{x}$$

If we once again use Eq. 12A.25, we would obtain

$$\nabla_x \mathbf{J} = -2[H]^T[W]\mathbf{z} + 2[H]^T[W][H]\mathbf{x}$$

and

$$\nabla_x \mathbf{J} = 0$$

gives

$$\mathbf{x} = ([H]^T[W][H])^{-1}[H]^T[W]\mathbf{z} \tag{12A.31}$$

The Fully-Determined Case ($N_m = N_s$)

In this case, the number of measurements is equal to the number of state variables and we can solve for $\mathbf{x}$ directly by inverting $[H]$.

$$\mathbf{x} = [H]^{-1}\mathbf{z} \tag{12A.32}$$

The Underdetermined Case ($N_m < N_s$)

In this case, we have fewer measurements than state variables. In such a case, it is possible to solve for many solutions $\mathbf{x}^{\text{est}}$ that cause $J(\mathbf{x})$ to equal zero. The usual solution technique is to find $\mathbf{x}^{\text{est}}$ that minimizes the sum of the squares of the solution values. That is, we find a solution such that

$$\sum_{j=1}^{N_s} x_j^2 \tag{12A.33}$$

is minimized while meeting the condition that the measurements will be solved for exactly. To do this, we treat the problem as a constrained minimization problem and use Lagrange multipliers as shown in Appendix 3A.

We formulate the problem as

$$\begin{aligned} &\textbf{Minimize:} && \sum_{j=1}^{N_s} x_j^2 \\ &\textbf{Subject to:} && z_i = \sum_{j=1}^{N_s} h_{ij} x_j \quad \text{for } i = 1, \ldots, N_m \end{aligned} \tag{12A.34}$$

This optimization problem can be written in vector–matrix form as

$$\begin{aligned} &\min && \mathbf{x}^T\mathbf{x} \\ &\text{subject to} && \mathbf{z} = [H]\mathbf{x} \end{aligned} \tag{12A.35}$$

The Lagrangian for this problem is

$$\mathscr{L} = \mathbf{x}^T\mathbf{x} + \boldsymbol{\lambda}^T(\mathbf{z} - [H]\mathbf{x}) \tag{12A.36}$$

Following the rules set down in Appendix 3A we must find the gradient of $\mathscr{L}$ with respect to $\mathbf{x}$ and with respect to $\boldsymbol{\lambda}$. Using the identities found in Eq. 12A.25 we get

$$\nabla_x \mathscr{L} = 2\mathbf{x} - [H]^T\boldsymbol{\lambda} = \mathbf{0}$$

which gives

$$\mathbf{x} = \frac{1}{2}[H]^T\boldsymbol{\lambda}$$

and

$$\nabla_\lambda \mathscr{L} = \mathbf{z} - [H]\mathbf{x} = 0$$

which gives

$$\mathbf{z} = [H]\mathbf{x}$$

Then

$$\mathbf{z} = \frac{1}{2}[H][H]^T\boldsymbol{\lambda}$$

or

$$\boldsymbol{\lambda} = 2[[H][H]^T]^{-1}\mathbf{z}$$

and finally,

$$\mathbf{x} = [H]^T[[H][H]^T]^{-1}\mathbf{z} \qquad (12A.37)$$

The reader should be aware that the matrix inversion shown in Eqs. 12A.29, 12A.32, and 12A.37 may not be possible. That is, the $[[H]^T[H]]$ matrix in Eq. 12A.29 may be singular, or $[H]$ may be singular in Eq. 12A.32, or $[[H][H]^T]$ may be singular in Eq. 12A.37. In the overdetermined case ($N_m > N_s$) whose solution is Eq. 12A.29, and the fully determined case ($N_m = N_s$) whose solution is Eq. 12A.32, the singularity implies what is known as an "unobservable" system. By unobservable we mean that the measurements do not provide sufficient information to allow a determination of the states of the system. In the case of the underdetermined case ($N_m < N_s$) whose solution is Eq. 12A.37, the singularity simply implies that there is no unique solution to the problem.

PROBLEMS

12.1 Using the three-bus sample system shown in Figure 12.1, assume that the three meters have the following characteristics.

Meter	Full Scale (MW)	Accuracy (MW)	σ (pu)
M_{12}	100	± 6	0.02
M_{13}	100	± 3	0.01
M_{32}	100	± 0.6	0.002

a. Calculate the best estimate for the phase angles θ_1 and θ_2 given the following measurements.

Meter	Measured Value (MW)
M_{12}	60.0
M_{13}	4.0
M_{32}	40.5

b. Calculate the residual $J(\mathbf{x})$. For a significance level, α, of 0.01, does $J(\mathbf{x})$ indicate the presence of bad data? Explain.

12.2 Given a single transmission line with a generator at one end and a load at the other, two measurements are available as shown in Figure 12.21. Assume that we can model this circuit with a DC load flow using the line reactance shown. Also, assume that the phase angle at bus 1 is 0 rad. Given the meter characteristics and meter readings telemetered from the meters, calculate the best estimate of the power flowing through the transmission line.

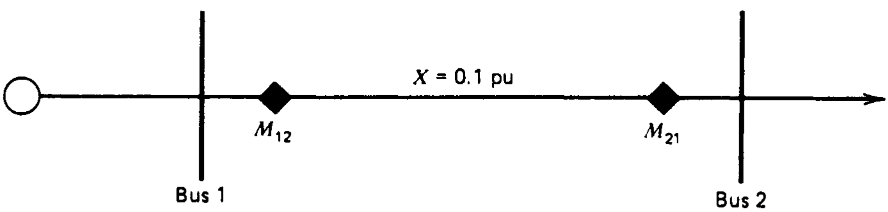

FIG. 12.21 Measurement configuration for Problem 12.2.

Meter	Full Scale (MW)	Meter Standard Deviation (σ) in Full Scale	Meter Reading (MW)
M_{12}	200	1	62
M_{21}	200	5	−52

Note: M_{12} measures power flowing from bus 1 to bus 2; M_{21} measures power flowing from bus 2 to bus 1.

Use 100 MVA as base.

12.3 You are given in the following network with meters at locations as shown in Figure 12.22.

Branch Impedances (pu)
$X_{12} = 0.25$
$X_{13} = 0.50$
$X_{24} = 0.40$
$X_{34} = 0.10$

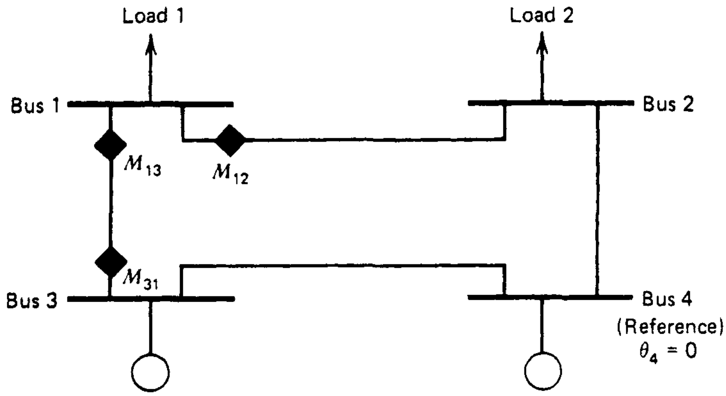

FIG. 12.22 Four-bus system with measurements for Problem 12.3.

Measurement Values	Measurement Errors
$M_{13} = -70.5$	$\sigma_{13} = 0.01$
$M_{31} = 72.1$	$\sigma_{31} = 0.01$
$M_{12} = 21.2$	$\sigma_{12} = 0.02$

a. Is this network observable? Set up the least-squares equations and try to invert $[H^T R^{-1} H]$.

b. Suppose we had a measurement of generation output at bus 3 and included it in our measurement set. Let this measurement be the following:

$$M_{3\,\text{gen}} = 92 \text{ MW} \quad \text{with } \sigma = 0.015$$

Repeat part a including this measurement.

12.4 Given the network shown in Figure 12.23, the network is to be modeled with a DC power flow with line reactances as follows (assume 100-MVA base):

$$x_{12} = 0.1 \text{ pu}$$
$$x_{23} = 0.25 \text{ pu}$$

The meters are all of the same type with a standard deviation of $\sigma = 0.01$ pu for each. The measured values are:

$$M_3 = 105 \text{ MW}$$
$$M_{32} = 98 \text{ MW}$$
$$M_{23} = -135 \text{ MW}$$
$$M_2 = 49 \text{ MW}$$
$$M_{21} = 148 \text{ MW}$$

a. Find the phase angles which result in a best fit to the measured values.

b. Find the value of the residual function J.

c. Calculate estimated generator output of each generator and the estimated power flow on each line.

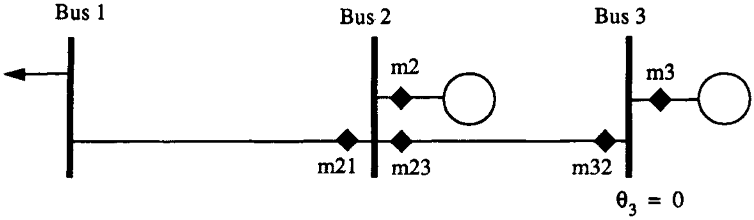

FIG. 12.23 Network for Problem 12.4.

d. Are there any errors in the measurements? If you think so, explain which meters are apt to be in error and why. Remove the suspected bad measurement and try to resolve the state estimator.

12.5 You are to purchase and install a set of programs that are to act as a monitor for the system security of a major utility company. You have solicited bids from major manufacturers of computer systems and are responsible for reviewing the technical contents of each bid. One of the manufacturers proposes to install a system with the flowchart and description given in Figure 12.24.

Bidders design:

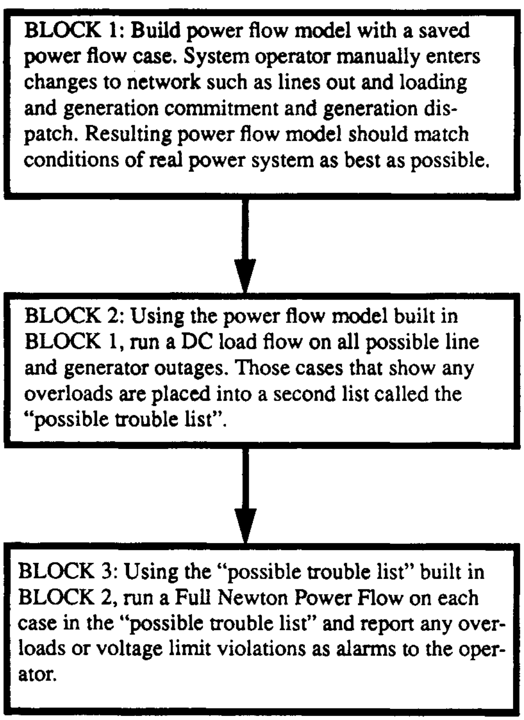

FIG. 12.24 Diagram for Problem 12.5.

a. Write down as many of the design flaws that you can find in this bidder's design.

b. Create a new design that you think will be a state-of-the-art system.

FURTHER READING

State estimation originated in the aerospace industry and only came to be of interest to power systems engineers in the late 1960s. Since then, state estimators have been installed on a regular basis in new energy control centers and have proved quite useful. References 1–4 provide a good introduction to this topic. Reference 4, in particular, is a carefully written overview with a good bibliography of literature up to 1974. References 5 and 6 show the variety of algorithms used to solve the state-estimation problem.

The remaining references cover some of the subtopics of state estimation. The use of the state estimator to detect bad measurements and model parameter errors is covered in references 7–10. Network observability determination is covered in references 11 and 12. Methods of automatically updating the network model topology to match switching status are covered in references 13 and 14. Finally, orthogonal decomposition methods are covered in references 15 and 16.

1. Schweppe, F. C., Wildes, J., "Power System Static State Estimation, Part I: Exact Model," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-89, January 1970, pp. 120–125.
2. Larson, R. E., Tinney, W. F., Peschon, J., "State Estimation in Power Systems, Part I: Theory and Feasibility," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-89, March 1970, 345–352.
3. Larson, R. E., Tinney, W. F., Hajdu, L. P., Piercy, P. S., "State Estimation in Power Systems, Part II: Implementation and Applications," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-89, March 1970, pp. 353–359.
4. Schweppe, F. C., Handschin, E., "Static State Estimation in Power Systems," *IEEE Proceedings*, Vol. 62, July 1974, pp. 972–982.
5. Dopazo, J. F., Klitin, O. A., Stagg, G. W., VanSlyck, L. S., "State Calculation of Power Systems from Line Flow Measurement," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-89, September/October 1970, pp. 1698–1708.
6. Dopazo, J. F., Klitin, O. A., VanSlyck, L. S., "State Calculation of Power Systems from Line Flow Measurements, Part II," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-91, January/February 1972, pp. 145–151.
7. Dopazo, J. F., Klitin, O. A., Sasson, A. M., "State Estimation for Power Systems: Detection and Identification of Gross Measurement Errors," Proceedings 8th PICA Conference, Minneapolis, June 1973.
8. Merrill, H. M., Schweppe, F. C., "On-Line System Model Error Correction," IEEE Winter Power Meeting, 1973, Paper C73-106-2.
9. Debs, A. S., "Estimation of Steady-State Power System Model Parameters," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-93, No. 5, 1974.
10. Handschin, E., Schweppe, F. C., Kohlar, J., Fiechter, A., "Bad Data Analysis for Power System State Estimation," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-94, March/April 1975, pp. 329–337.

11. Clements, K. A., Wollenberg, B. F., "An Algorithm for Observability Determination in Power System State Estimation," IEEE Summer Power Meeting, 1975, Paper A75-447-3.

12. Krumpholz, G. R., Clements, K. A., Davis, P. W., "Power System Observability: A Practical Algorithm using Network Topology," *IEEE Transactions on Power Apparatus and Systems*, Vol. 99, July/August 1980, pp. 1534–1542.

13. Sasson, A. M., Ehrman, S. T., Lynch, P., VanSlyck, L. S., "Automatic Power System Network Topology Determination," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-92, March/April 1973, pp. 610–618.

14. DyLiacco, T. E., Ramarao, K., Weiner, A., "Network Status Analysis for Real Time Systems," Proceedings 8th PICA Conference. Minneapolis, June 1973.

15. Simoes-Costa, A., Quintana, V., "An Orthogonal Row Processing Algorithm for Power System Sequential Estimation," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-100, No. 8, August 1981, pp. 3791–3800.

16. Vempati, N., Slutsker, I. W., Tinney, W. F., "Enhancements to Givens Rotation for Power System State Estimation," *IEEE Transactions on Power Systems*, Vol. 6, No. 2, May 1991, pp. 842–849.