Math and Numerical Methods Review

Michael Caracotsios, Ph.D.
Clinical Associate Professor
Chemical Engineering Department
University of Illinois at Chicago

1 Introduction

In the study of chemical engineering systems, we find it necessary to consider the variation of a particular quantity as dependent on the variation of another. For example, in the study of chemical kinetics the concentration of a species undergoing a chemical reaction varies with time. Here, the concentration profile of the reactant species is considered to be a function of time. In general, if to each value of a variable *x* (within a certain domain) corresponds a definitive value of another variable *y*, then we say that *y* is a function of *x* and, in mathematical terms we may write $y = y(x)$ or .

2 Derivative of a Function

Let the function $y = f(x)$ be defined in a certain interval $x \in [a,b]$. The derivative $f'(x)$ of the given function is defined by the limit:

$$y' = \frac{dy}{dx} = \frac{df(x)}{dx} = f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \qquad (2\text{-}1)$$

The specific value of the derivative for $x = x_0 \in [a,b]$ is denoted by $f'(x_0)$

Example 2-1 Given the function $f(x) = x^3$ ***find its derivative.***

From the definition of the derivative (Equation 2-1) we obtain the following:

$$f'(x) = \lim_{\Delta x \to 0} \frac{(x+\Delta x)^3 - x^3}{\Delta x} = \lim_{\Delta x \to 0} \frac{3x^2\Delta x + 3x\Delta x^2 + \Delta x^3}{\Delta x} \Rightarrow$$

$$f'(x) = \lim_{\Delta x \to 0} (3x^2 + 3x\Delta x + \Delta x^2) = 3x^2$$

Using the definition of the derivative we can easily prove the following important rules of differentiation:

- The derivative of a constant function is equal to zero
- The derivative of the sum of two functions $f_1(x)$ and $f_2(x)$ is equal to the corresponding sum of the derivatives of these functions (the rule also applies to the sum of a finite number of functions). In derivative notation we have:

$$y = f_1(x) + f_2(x) \Rightarrow y' = f_1'(x) + f_2'(x) \qquad (2\text{-}2)$$

- The derivative of a constant *C* times a function $f(x)$ is equal to the product of the constant times the derivative of the function, specifically

$$y = Cf(x) \Rightarrow y' = Cf'(x) \qquad (2\text{-}3)$$

- *Product Rule:* The derivative of the product of two differentiable functions $f_1(x)$ and $f_2(x)$ is equal to the derivative of the first function times the second plus the product of the first function times the derivative of the second function; in derivative notation we have:

$$y = f_1(x) f_2(x) \Rightarrow y' = f_1'(x) f_2(x) + f_1(x) f_2'(x) \qquad (2\text{-}4)$$

- The derivative of the inverse of a function $f(x)$ is equal to the negative fraction whose denominator is the square of the given function and numerator its derivative; specifically

$$y = \frac{1}{f(x)} \Rightarrow y' = -\frac{f'(x)}{f^2(x)} \quad f(x) \neq 0 \qquad (2\text{-}5)$$

- *Quotient Rule:* Using Equations (2-5) and (2-4) we can derive the derivative of the ratio of two functions from the formula:

$$y = \frac{g(x)}{f(x)} \Rightarrow y' = g'(x)\frac{1}{f(x)} - g(x)\frac{f'(x)}{f^2(x)} = \frac{g'(x)f(x) - g(x)f'(x)}{f^2(x)} \qquad (2\text{-}6)$$

3 The Derivative of a Composite Function

Let the argument x of the function $y = f(x)$, be another function of t, more specifically let $x = g(t)$; then y is considered to be a composite function of t. Mathematically we denote the composite function by the notation,

$$y = (g \circ f)(t) = f(g(t)) \tag{3-1}$$

The derivative of the function y with respect to the argument t is given by the "chain rule":

$$\frac{dy}{dt} = \frac{df}{dg}\frac{dg}{dt} \tag{3-2}$$

Example 3-1 Given the function $y = \exp\left(t^2\right)$ ***calculate the derivative of y with respect to t.***

Define the auxiliary functions $y = f(x) = \exp(x)$ and $x = g(t) = t^2$, and then by applying the Equation (3-2) we obtain

$$\frac{dy}{dt} = \frac{df}{dx}\frac{dx}{dt} = 2t\exp(x) = 2t\exp(t^2)$$

Example 3-1 Given the Arrhenius function $k(T) = k_0 \exp\left(-\frac{E}{RT}\right)$ ***calculate the derivative of*** $k(T)$ ***with respect to*** T ***and of*** $\ln\left[k(T)\right]$ ***with respect to*** T.

Define the auxiliary functions $k = f(x) = k_0\exp(x)$ and $x = g(T) = -\frac{E}{RT}$, and then by applying the Equation (3-2) we obtain:

$$\frac{dk}{dT} = \frac{dk}{dx}\frac{dx}{dT} = k_0\exp\left(x\right)\frac{E}{RT^2} = k_0\exp\left(-\frac{E}{RT}\right)\frac{E}{RT^2}$$

or by making the appropriate substitutions,

$$\frac{dk}{dT} = k\frac{E}{RT^2} \Rightarrow \frac{1}{k}\frac{dk}{dT} = \frac{E}{RT^2} \Rightarrow \frac{d\ln k}{dT} = \frac{E}{RT^2} \Rightarrow$$

$$\frac{d\ln k}{d(1/T)} = -\frac{E}{R}$$

4 The Geometric meaning of the Derivative

The derivative of a function has a geometric interpretation, which is very important in engineering calculations. To illustrate this interpretation consider the function $y = f(x)$ which is plotted in a Cartesian coordinate system in Figure 4-1

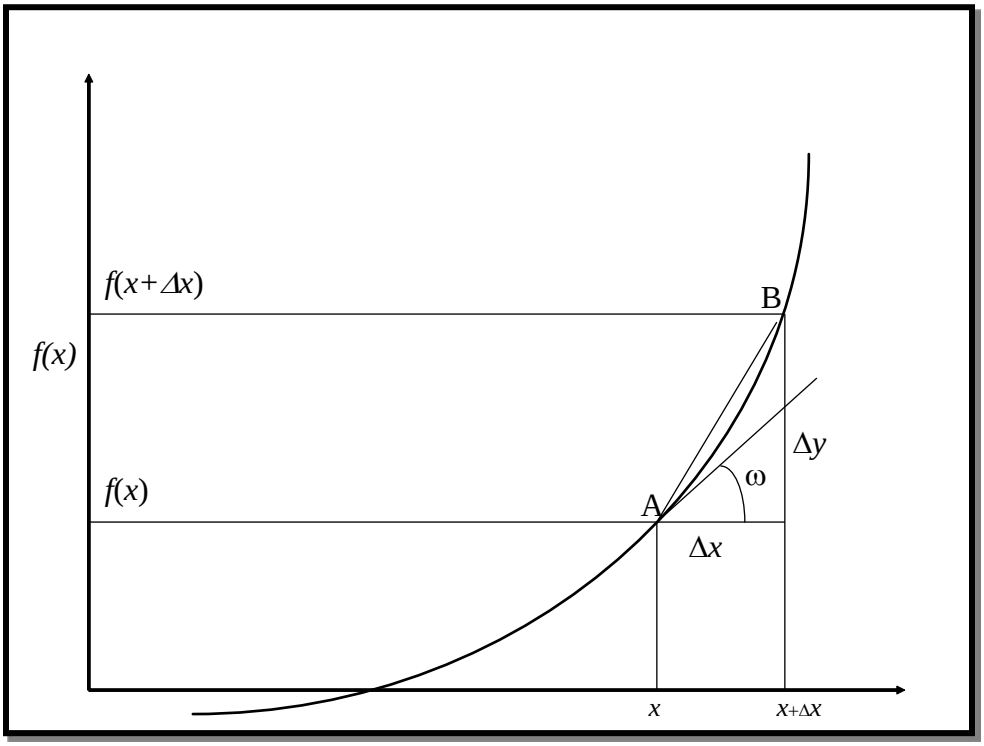

Figure 4-1 Geometric interpretation of a function derivative

At an arbitrary point *x* the value of the function is $f(x)$ and corresponds to the point *A* in the curve. If we increment the argument *x* by Δx, the new value of the function would be $y + \Delta y = f(x + \Delta x)$ and the corresponding point on the curve *B*. We draw the secant $\overline{AB}$ and denote by α the angle formed by this secant and the positive *x*-axis. Then from Figure 4-1 it follows that:

$$\tan(\alpha) = \frac{\Delta y}{\Delta x} \qquad (4\text{-}1)$$

We now let Δx approach zero; then the point *B* will be moving along the curve approaching the point *A*. Therefore the angle α will be changing as Δx approaches zero. At the limit the angle α approaches a certain value ω and the secant $\overline{AB}$ becomes the *tangent* of the curve at the point *A*. It is then easy to find the slope of this tangent line:

$$\tan(\omega) = \lim_{\Delta x \to 0} \tan(\alpha) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(x) \qquad (4\text{-}2)$$

which is to say that the value of the derivative of a function at an arbitrary point $x = x_0$ is equal to the tangent of the angle which is formed between the positive *x*-axis and the line tangent to the curve at the corresponding point $A(x, f(x_0))$.

To find the equation of the tangent line we make use of the equation of any straight line that passes through the point $(y_0, x_0)=(f(x_0), x_0)$:

$$y - y_0 = y - f(x_0) = m(x - x_0) \Rightarrow y = m(x - x_0) + f(x_0) \qquad (4\text{-}3)$$

where *m* is the slope of the line. Applying Equation (4-3) to the tangent line we obtain after some simple algebra:

$$y = f'(x_0)(x - x_0) + f(x_0) \qquad (4\text{-}4)$$

In a similar manner we can find the equation of the *normal* line at the point A on the curve. The normal line is perpendicular to the tangent and its slope is by definition equal to the negative inverse of the slope of the tangent. Consequently the equation for the normal line is given by:

$$y = -\frac{1}{f'(x_0)}(x - x_0) + f(x_0) \quad f'(x_0) \neq 0 \qquad (4\text{-}5)$$

5 Functions of Several Variables

In chemical engineering systems functions of several variables abound. For instance the concentration of a chemical species undergoing a chemical reaction may be a function of time and also of temperature. Then the time and temperature are called independent variables and the concentration is the dependent variable. In general if to every collection of variables $[x_1, x_2, \ldots, x_n]$ we assign a definite value of the variable *y*, we call *y* the function of the independent variables $[x_1, x_2, \ldots, x_n]$. Since the collection of the independent variables can also be represented by a column vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^{\mathrm{T}}$ we may call *y* the function of the vector ***x*** and write $y = y(\mathbf{x})$ or equivalently we may write $y = y(x_1, x_2, \ldots, x_n)$ or $y = f(\mathbf{x})$ or finally $y = f(x_1, x_2, \ldots, x_n)$.

6 Partial Derivatives, Gradient and Directional Derivative

The partial derivative of a function $y = f(\mathbf{x})$ of several variables $[x_1, x_2, \ldots, x_n]$ with respect to one of them, say x_1, is defined by the following limit:

$$\left.\frac{\partial f}{\partial x_1}\right)_{x_2,\ldots x_n} = \lim_{\Delta x_1 \to 0} \frac{f(x_1 + \Delta x_1, x_2, \ldots x_n) - f(x_1, x_2, \ldots x_n)}{\Delta x_1} \qquad (6\text{-}1)$$

Similar expressions can be written for the partial derivatives of the function $y = f(\mathbf{x})$ with respect to the other variables. In order to illustrate the concept of the gradient of a function $y = y(x_1, x_2)$ of two variables consider a contour of this function as depicted in Figure 6-1.

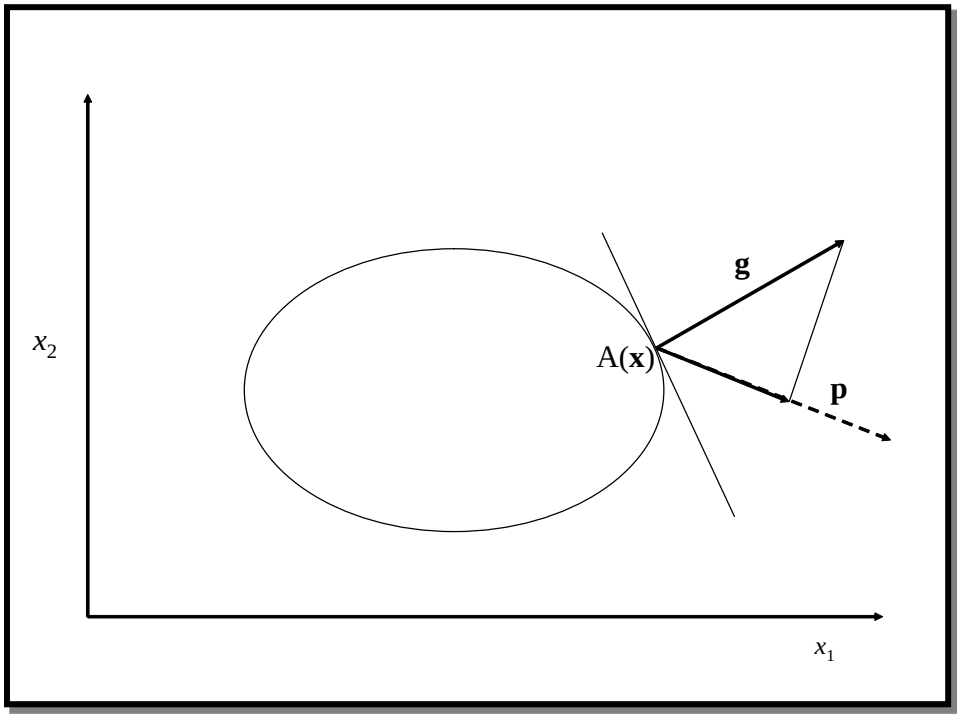

Figure 6-1 Geometric Illustration of a Gradient of a Function

and an arbitrary point A(**x**). The *gradient* of the multivariate function $y = f(x_1, x_2, \ldots x_n)$ is denoted by ∇f or ∇y and is defined as the vector:

$$\mathbf{g}(\mathbf{x}) = \nabla f(\mathbf{x}) := \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right)^T$$

The gradient is perpendicular to the tangent hyper-plane at the point $A(\mathbf{x})$. If we differentiate the gradient of a multivariate function one more time, then we obtain the **Hessian** matrix (or the matrix of the second derivatives). The Hessian matrix is symmetric and defined by the equation:

$$\nabla \mathbf{g}(\mathbf{x}) = \mathbf{H}(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

The directional derivative of the function $f(\mathbf{x})$ along some direction of a unit vector **p** (refer to the Figure 6-1), is denoted by the symbol $\frac{\partial f}{\partial \mathbf{p}}$ and is equal to the projection of the gradient vector **g** onto the vector **p**; more specifically we may write that,

$$\frac{\partial f}{\partial \mathbf{p}} = \nabla f \bullet \mathbf{p} \qquad (6\text{-}2)$$

where the symbol $\bullet$ indicates the dot product of two vectors **u, w** defined by:

$$\mathbf{u} \bullet \mathbf{w} = \sum_{i=1}^{n} u_i w_i$$

In the three-dimensional space the unit vector **p** can be represented by the directional angles α, β, γ. In this particular case the directional derivative becomes:

$$\frac{\partial f}{\partial \mathbf{p}} = \cos(\alpha)\frac{\partial f}{\partial x} + \cos(\beta)\frac{\partial f}{\partial y} + \cos(\gamma)\frac{\partial f}{\partial z} \qquad (6\text{-}3)$$

For instance in the direction of the positive *x*-axis we have $\alpha = 0, \beta = \frac{\pi}{2}$ and $\gamma = \frac{\pi}{2}$ and therefore the directional derivative of the function $f(\mathbf{x})$ along the direction of the positive **x**-axis is equal to:

$$\frac{\partial f}{\partial \mathbf{p}} = \cos(0)\frac{\partial f}{\partial x} + \cos\left(\frac{\pi}{2}\right)\frac{\partial f}{\partial y} + \cos\left(\frac{\pi}{2}\right)\frac{\partial f}{\partial z} = \frac{\partial f}{\partial x}$$

It is evident from the definition of the directional derivative that the derivative along the direction of a vector that is perpendicular to the gradient vector is equal to zero. In addition, the derivative along the unit vector **p** has its maximum value if the direction of **p** coincides with the direction of the gradient vector.

7 Total Function Differentials and Implicit Derivatives

Frequently in engineering calculations we are interested in the calculation of total differentials. To illustrate these calculations consider the following function of two variables $f(x_1, x_2)$. The total differential df of this function is defined by the equation,

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 \qquad (7\text{-}1)$$

Equation (7-1) gives the absolute change in the function $f(x_1, x_2)$ due to small(infinitesimal) changes dx_1, dx_2 in the independent variables (x_1, x_2) respectively. Relative changes can also be calculated by simple manipulation of equation (7-1) as shown below:

$$\frac{df}{f} = \frac{x_1}{f}\frac{\partial f}{\partial x_1}\frac{dx_1}{x_1} + \frac{x_2}{f}\frac{\partial f}{\partial x_2}\frac{dx_2}{x_2} \Rightarrow \frac{df}{f} = \frac{\partial \ln f}{\partial \ln x_1}\frac{dx_1}{x_1} + \frac{\partial \ln f}{\partial \ln x_2}\frac{dx_2}{x_2} \qquad (7\text{-}2)$$

In the above equation the expressions $\left\{\frac{df}{f}, \frac{dx_1}{x_1}, \frac{dx_2}{x_2}\right\}$ represent the relative changes in the function and the independent variables respectively. If all terms are multiplied by a factor of one hundred, the changes become percentages.

Very often in engineering systems we cannot obtain an analytical expression of a function $y(x)$. Instead an implicit representation of the function is given via a formula $f(y, x)$. In this case in order to obtain the derivative of the function $y(x)$ with respect to *x* we must take the total derivative of the function $f(y, x)$ with respect to *x* and then solve for the derivative $y'(x)$ as shown below:

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} = 0 \Rightarrow \frac{dy}{dx} = -\frac{\partial f}{\partial x} \Big/ \frac{\partial f}{\partial y} \qquad (7\text{-}3)$$

Example 7-1 Calculate the derivative $\left.\frac{\partial V}{\partial T}\right)_P$ ***for a gas that follows the Van Der Waals equation of state.***

The Van Der Waals equation of state is given by the following function, under constant pressure:

$$f(V,T)=\left(P+\frac{a}{V^2}\right)(V-b)-RT=0$$

Applying equation (7-3) we obtain:

$$\left.\frac{\partial f}{\partial T}\right)_P+\left.\frac{\partial f}{\partial V}\frac{\partial V}{\partial T}\right)_P=0\Rightarrow -R+\left(P+\frac{a}{V^2}-\frac{2a(V-b)}{V^3}\right)\left.\frac{\partial V}{\partial T}\right)_P=0$$

Solving for the required derivative we obtain:

$$\left.\frac{\partial V}{\partial T}\right)_P=\frac{R}{\left(P+\frac{a}{V^2}-\frac{2a(V-b)}{V^3}\right)}\Rightarrow\boxed{\left.\frac{\partial V}{\partial T}\right)_P=\frac{R}{\left(P-\frac{a}{V^2}+\frac{2ab}{V^3}\right)}}$$

Alternatively, we can derive an expression for the derivative by differentiating the function $\left(P+\frac{a}{V^2}\right)(V-b)=RT$ as follows. First recall that $\left(\frac{1}{f}\right)'=-\frac{f'}{f^2}$ and denote $V'=\left.\frac{\partial V}{\partial T}\right)_P$. Then we have,

$$\left(P+\frac{a}{V^2}\right)'(V-b)+\left(P+\frac{a}{V^2}\right)(V-b)'=R\Rightarrow$$

$$-\frac{a}{V^4}\left(V^2\right)'(V-b)+\left(P+\frac{a}{V^2}\right)V'=R\Rightarrow$$

$$-\frac{a}{V^4}2VV'(V-b)+\left(P+\frac{a}{V^2}\right)V'=R\Rightarrow$$

$$V'\left[-\frac{a}{V^4}2V(V-b)+\left(P+\frac{a}{V^2}\right)\right]=R\Rightarrow$$

$$V'=\frac{R}{\left[-\frac{a}{V^4}2V(V-b)+\left(P+\frac{a}{V^2}\right)\right]}=\frac{R}{\left[-\frac{2a}{V^3}(V-b)+\left(P+\frac{a}{V^2}\right)\right]}=\frac{R}{\left[P-\frac{a}{V^2}+\frac{2ab}{V^3}\right]}$$

8 Vector Functions of Several Variables

In many instances we have to consider collections of functions $\left(f_1, f_2, \ldots f_m \right)$ of several variables $\left(x_1, x_2, \ldots, x_n \right)$. As an example in a chemical reaction we may wish to follow the concentration profiles of two or more species as a function of time and temperature. The notation $\mathbf{f}(\mathbf{x})$ is conventionally used to represent the collection of these functions. We now define the matrix $\mathbf{J}(\mathbf{x})$ by the following equation:

$$\mathbf{J}(\mathbf{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix} \tag{8-1}$$

The matrix $\mathbf{J}(\mathbf{x})$ is known as the **Jacobian matrix** of the collection of the functions $\mathbf{f}(\mathbf{x})$.

9 The Indefinite Integral

A function $F(x)$ is called the anti-derivative of the function $f(x)$ if $F'(x) = f(x)$. We subsequently define the sum $F(x) + c$ as the indefinite integral of the function $f(x)$, and C is an arbitrary constant. Mathematically we write:

$$\int f(x)dx = F(x) + C \Rightarrow \; F'(x) = f(x) \tag{9-1}$$

As an example we can easily prove that

$$\int x^n dx = \frac{x^{n+1}}{n+1} + C \quad n \neq -1$$

Some properties of the indefinite integral are easily verifiable and are presented here:

- the indefinite integral of an algebraic sum of a finite number of functions is equal to the algebraic sum of their integrals:

$$\int \left(f_1(x) + f_2(x) \right) dx = \int f_1(x)dx + \int f_2(x)dx \tag{9-2}$$

- a constant factor C may be taken outside the integral as shown in the expression below:

$$\int Cf(x)dx = C\int f(x)dx \qquad (9\text{-}3)$$

- integration by parts can be carried out according to the following formula:

$$\int udv + \int vdu = \int d(uv) = uv + C \qquad (9\text{-}4)$$

10 The Definite Integral.

The definite integral is one of the most fundamental concepts in mathematical analysis and a powerful research tool in physics and engineering. The integration of differential equations for example, reduces to the evaluation of a definite integral. Consider the function $y = f(x)$ depicted in Figure 10-1.

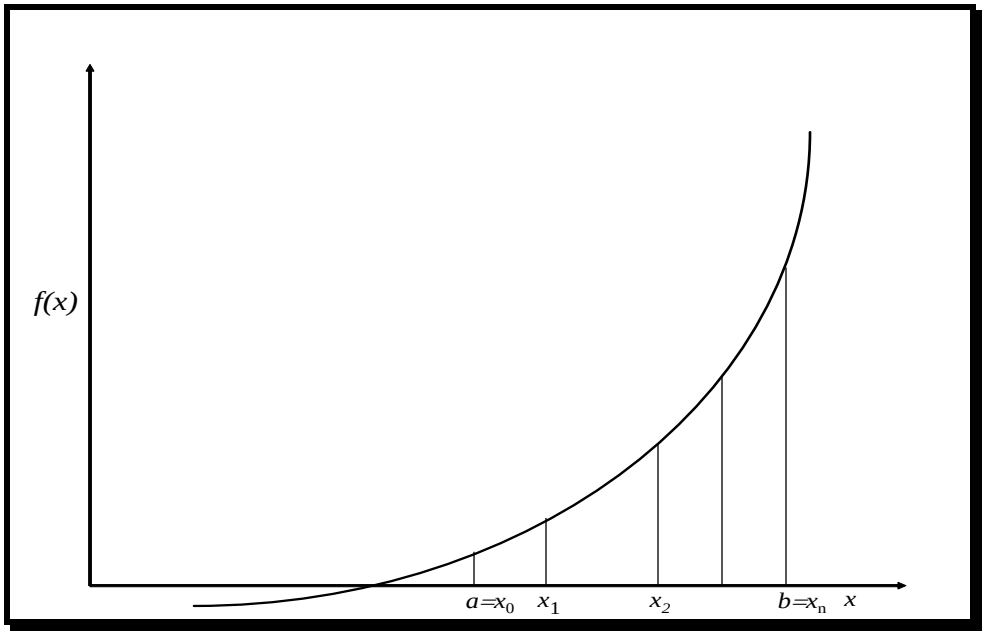

Figure 10-1 Definition of the Definite Integral

Consider the following division of the interval [*a*, *b*] into *n* subintervals (refer to Fig. 10-1):

$$a = x_0 < x_1 < x_2 .... < x_{n-1} < x_n = b \qquad (10\text{-}1)$$

and we denote by ξ_j a point in the interval $\left[x_j, x_{j+1}\right]$. Then the definite integral of the function $f(x)$ from *a* to *b* is defined by the limit:

$$\int_a^b f(x)dx = \lim_{\max(\Delta x_j)\to 0} \sum_{j=0}^{j=n} f(\xi_j)(x_{j+1} - x_j) \qquad (10\text{-}2)$$

where $\Delta x_j = x_{j+1} - x_j$. It is evident from Figure 10-1 that the definite integral of the function $f(x)$ is equal in value to the area of the curvilinear trapezoid bounded by the curve $y = f(x)$, the straight lines $x = a$ and $x = b$ and the *x*-axis. Some of the basic properties of the indefinite integrals also apply in the definite integrals:

- a constant factor may be taken out of a definite integral
- the definite integral of an algebraic sum of a finite number of functions is equal to the algebraic sum of their definite integrals.

For any numbers a, b, c we can write the following equation (provided that all definite integrals exist):

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx \qquad (10\text{-}3)$$

The *mean value theorem* states that if a function $f(x)$ is continuous on an interval [*a*, *b*] then, there exists a number ξ in this interval for which the following identity holds:

$$\int_a^b f(x)dx = (b-a)\, f(\xi) \qquad (10\text{-}4)$$

11 The Newton-Leibniz Formula

Consider the function

$$\Phi(x) = \int_a^x f(x)dx \qquad (11\text{-}1)$$

where *a* is a constant number. Then we have the following identity:

$$\Phi'(x) = \frac{d\Phi}{dx} = f(x) \qquad (11\text{-}2)$$

i.e., the derivative of a definite integral with respect to the upper limit is equal to the integrand in which the value of the upper limit replaces the variable of integration; since the variable of integration is a dummy variable we may equivalently write that

$$\Phi(x)=\int_a^x f(\xi)d\xi \;\Rightarrow\; \Phi'(x)=f(x)\Rightarrow f\left(x\right)=\frac{d}{dx}\int_a^x f\left(\xi\right)d\xi \qquad (11\text{-}3)$$

Equation (11-3) is the well-known Newton-Leibniz formula. It is also known as the "*second fundamental theorem of calculus*". Now suppose that we want to differentiate the integral:

$$I(t)=\int_{\varphi_1(t)}^{\varphi_2(t)} f(x,t)dx \qquad (11\text{-}4)$$

which is only function of the independent variable *t*. Leibniz has developed a formula for the derivative of *I*(*t*) with respect to *t*. This formula is given by the equation:

$$I'(t)=\int_{\varphi_1(t)}^{\varphi_2(t)} \frac{\partial f(x,t)}{\partial t}dx+f(\varphi_2(t),t)\frac{d\varphi_2(t)}{dt}-f(\varphi_1(t),t)\frac{d\varphi_1(t)}{dt} \qquad (11\text{-}5)$$

It is evident from Equation (11-5) that for constant functions $\varphi_1(t)=a$ and $\varphi_2(t)=b$ we obtain:

$$I'(t)=\int_b^a \frac{\partial f(x,t)}{\partial t}dx \qquad (11\text{-}6)$$

12 Solution of Nonlinear Algebraic Equations.

In this section we discuss numerical solution algorithms for nonlinear algebraic equation systems. Such systems occur widely in models of chemical and phase equilibrium, and in algebraic approximations of differential and integral equations. Some of these problems have multiple solutions, and most require iterative computation to get a solution. In addition solving nonlinear algebraic equations is the primary task of steady-state simulators. In both modular and equation-oriented modes, the unit operation models are combined with physical property and flowsheet topology equations to form systems of nonlinear algebraic equations that need to be solved in a robust and reliable manner. Graphical methods remain valuable for analysis of simple cases and for visualization of iteration schemes.

In the discussion of numerical algorithms we focus primarily, on the Newton method because of its very wide use in nonlinear computations. The one-dimensional Newton method is first introduced followed by the *n*-dimensional version. Derivative-free versions of Newton's method are also addressed with emphasis on the Broyden method and, at the end of the section we briefly investigate first order methods that include direct substitution and Wegstein acceleration.

The main limitation of Newton's method is its local character; one usually needs a good initial guess to obtain convergence. The quest for more robust methods is an active area of research. Many of the ideas discussed in this section are implemented in the mathematical solvers provided within the **Athena Visual Studio** software package.

13 Newton's Method in One Dimension

Consider a single algebraic equation in a single real variable,

$$f(x) = 0 \tag{13-1}$$

where $f(x)$ is a real-valued continuous function with a continuous first derivative. A graph of $f(x)$ as in Figure 13-1 shows the solution (or root) directly, and is also useful for illustrating iteration methods and particularly the Newton method.

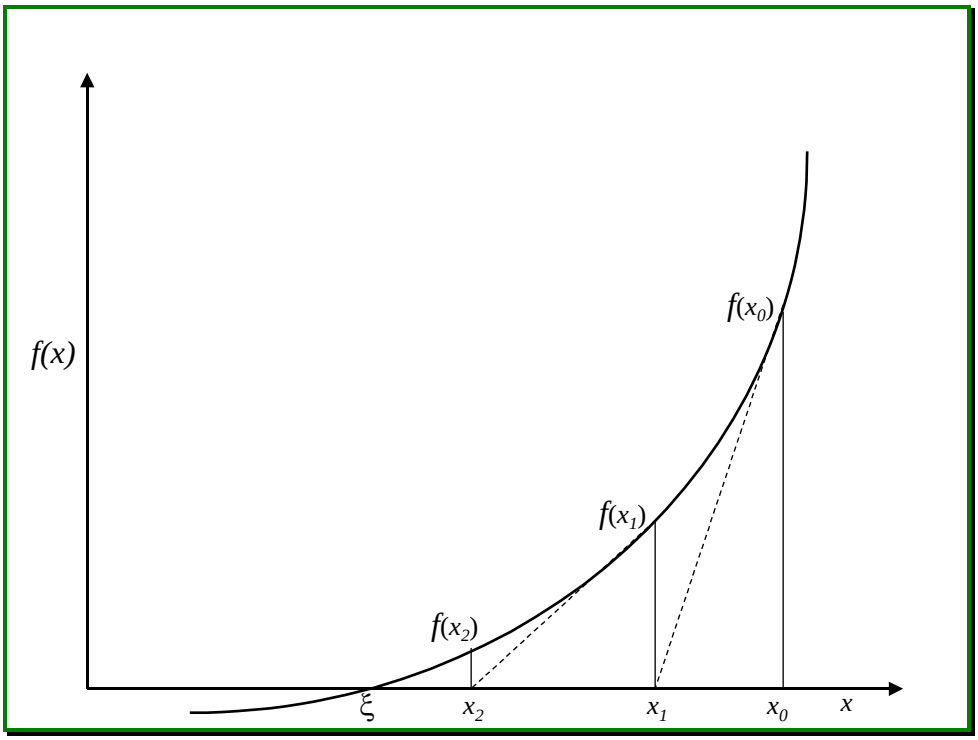

Figure 13-1 Illustration of Newton Iterations in One Dimension

Newton's method for finding a solution is iterative. One starts with a guess x_0, and linearizes Equation (13-1) around x_0 by expanding $f(x)$ in a truncated Taylor series. The guess should be good enough so that the resulting equation,

$$f(x_0) + f'(x_0)(x_1 - x_0) = 0 \tag{13-2}$$

for the new approximation x_1 is realistic; in particular, the derivative $f'(x_0)$ must be nonzero. Solving Equation (13-2) for x_1 we obtain the next approximation to the root of Equation (13-1):

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \qquad f'(x_0) \neq 0$$

Generalization of this procedure gives the Newton's formula

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \qquad f'(x_k) \neq 0 \tag{13-3}$$

for proceeding from the k_{th} approximation to the next. This formula gives the intercept x_{k+1} of a line drawn tangent to $f(x)$ from the current point $(x_k, f(x_k))$, as illustrated in the Figure 13-1. A good initial guess is usually needed to obtain convergence, and to obtain a particular root when other roots are present.

Example 13-1 Demonstration of Newton's Method in One Dimension

Application of Equation (13-3) to the solution of equation $x^2 - 3 = 0$ (whose solution is $\xi = \sqrt{3} = 1.73205808$) with the initial guess $x_0 = 3$, gives the following values:

k	x_k	$f(x_k) = x_k^2 - 3$	$f'(x_k) = 2x_k$	$x_{k+1} - x_k$	$e_k = x_k - \xi$
0	3	6	6	-1	1.267949192
1	2	1	4	-0.25	0.267949192
2	1.75	0.0625	3.5	-0.01785714	0.017949192
3	1.732142857	0.000318877	3.46428	-0.00009204	0.000092049
4	1.732050810	0.000000008	3.46410162	-0.00000000	0.000000002

Note that the correction $(x_{k+1} - x_k)$ approximates $-e_k$ very accurately in the later stages. A negative guess x_0 would give a corresponding iteration sequence converging toward $x_\infty = -\sqrt{3}$.

Convergence of the Newton iteration sequence is obtained when the relative error

$$\varepsilon = \frac{|x_{k+1} - x_k|}{\max\{1, |x_k|\}} \tag{13-4}$$

becomes less or equal to a user specified tolerance. Selection of the tolerance depends on the number of significant digits required for the solution. In order to illustrate the relation of relative error and significant digits we refer back to our numerical example to obtain the following table of approximations to the solution $x_\infty = \xi = \sqrt{3} = 1.732050807$

k	x_k	No. of Digits	$e_k = x_k - \xi$	$\varepsilon = \dfrac{\lvert e_k \rvert}{\max\{1, \lvert x_k \rvert\}}$
0	3	0	1.267949192	$0.4 \times 10^{+0}$
1	2	0	0.267949192	$0.1 \times 10^{+0}$
2	1.75	2	0.017949192	0.1×10^{-1}
3	1.732142857	4	0.000092049	0.5×10^{-4}
4	1.732050810	8	0.000000002	0.1×10^{-8}

An examination of the above table reveals that if two numbers agree to *k* decimal digits then the relative error will be approximately 0.5×10^{-k}.

14 Newton's Method in Multiple Dimensions

Consider now a system of *n* algebraic equations in *n* real variables:

$$\mathbf{F}(\mathbf{x}) = 0 \tag{14-1}$$

Here $\mathbf{x}$ and $\mathbf{F}(\mathbf{x})$ denote *n*-dimensional column vectors, with real-valued elements $\{x_1, x_2, ..., x_n\}$ and $\{F_1(\mathbf{x}), F_2(\mathbf{x}), ..., F_n(\mathbf{x})\}$ respectively. The functions $F_i(\mathbf{x})$ are assumed to have continuous derivatives through the first order.

Following the Newton method, we approximate each nonlinear function $F_i(\mathbf{x})$ by a linear Taylor expansion around a starting point $\mathbf{x}^0$:

$$F_i(\mathbf{x}) = F_i(\mathbf{x}^0) + \sum_{j=1}^{n} \left.\frac{\partial F_i}{\partial x_j}\right|_{\mathbf{x}^0} (x_j - x_j^0) + \cdots \qquad i = 1, 2, ...n \tag{14-2}$$

These expansions can be summarized more compactly as

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{x}^0) + \mathbf{J}(\mathbf{x}^0)(\mathbf{x} - \mathbf{x}^0) + \cdots \tag{14-3}$$

in which $\mathbf{J}(\mathbf{x}^0)$ denotes the value at $\mathbf{x}^0$ of the Jacobian matrix

$$\mathbf{J}(\mathbf{x}) = \frac{\partial \mathbf{F}}{\partial \mathbf{x}} = \begin{vmatrix} \dfrac{\partial F_1}{\partial x_1} & \cdots & \dfrac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F_n}{\partial x_1} & \cdots & \dfrac{\partial F_n}{\partial x_n} \end{vmatrix} \tag{14-4}$$

We assume this matrix to be nonsingular at each iteration point. Setting $\mathbf{F}(\mathbf{x}) = 0$ in Equation (14-3), and solving the resulting linear system, gives the first-stage Newton correction $\Delta\mathbf{x}^0 = \mathbf{x}^1 - \mathbf{x}^0$:

$$\mathbf{J}\left(\mathbf{x}^0\right)\Delta\mathbf{x}^0 = -\mathbf{F}\left(\mathbf{x}^0\right)$$
$$\mathbf{x}^1 = \mathbf{x}^0 + \Delta\mathbf{x}^0$$

The process is iterated according to the formula:

$$\mathbf{J}\left(\mathbf{x}^k\right)\Delta\mathbf{x}^k = -\mathbf{F}\left(\mathbf{x}^k\right)$$
$$\mathbf{x}^{k+1} = \mathbf{x}^k + \Delta\mathbf{x}^k \qquad (14\text{-}5)$$

to obtain the successive corrections $\Delta\mathbf{x}^k = \mathbf{x}^{k+1} - \mathbf{x}^k$. An iteration thus involves the computation of the vector $\mathbf{F}(\mathbf{x}^k)$ and the matrix $\mathbf{J}(\mathbf{x}^k)$, and the solution of Equation (14-5) for the correction vector (normally by the **LU** decomposition method). The Newton algorithm is terminated when the relative error

$$\varepsilon = \frac{\left\|\Delta\mathbf{x}^k\right\|}{\max\left\{1, \left\|\mathbf{x}^k\right\|\right\}} \quad \text{where} \quad \|\mathbf{x}\| = \max\left\{|x_1|, |x_2|, \ldots, |x_n|\right\}$$

becomes less or equal to a user specified tolerance. Evaluation and factorization of the Jacobian matrix at every Newton iteration can become very expensive, especially for large systems of algebraic equations. A shortcut, used in many implicit integrator codes that employ the Newton method to solve the discretized system of the resultant algebraic equations, is to leave $\mathbf{J}(\mathbf{x}^k)$ constant in the later iterations, where $\mathbf{J}(\mathbf{x}^k)$ changes very little. One thus avoids the **LU** factorization which entails $O(n^3)$ arithmetic operations each time, as well as the computation of $\mathbf{J}(\mathbf{x}^k)$ itself. A Newton iteration with constant Jacobian matrix, requires only the computation of the residual vector $\mathbf{F}(\mathbf{x}^k)$, plus $O(n^2)$ arithmetic operations to solve Equation (14-5), with $\mathbf{J}(\mathbf{x}^k)$ approximated by the old **LU** decomposition. This tactic reduces the order of convergence from quadratic to linear for an isolated root, so it should be used only when $\mathbf{x}^k$ is judged to be near to a solution, as in the predictor-corrector algorithms for differential equations.

Example 14-1 Demonstration of Newton's Method in Two Dimensions

Consider the solution of the following system of equations:

$$f_1(x_1, x_2) = x_1^2 + x_1 x_2 - 10 = 0$$
$$f_2(x_1, x_2) = x_2 + 3x_1 x_2^2 - 57 = 0$$

We consider the formulation of Newton's method from a starting point that is close to the solution. As a result we expect good performance and little difficulty with convergence. The Newton iteration is given by:

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \mathbf{J}^{-1}(\mathbf{x}^k)\mathbf{F}(\mathbf{x}^k)$$

where the Jacobian matrix and the residual vector are given as:

$$\mathbf{J}(\mathbf{x}) = \begin{bmatrix} 2x_1 + x_2 & x_1 \\ 3x_2^2 & 1 + 6x_1 x_2 \end{bmatrix} \qquad \mathbf{F}(\mathbf{x}) = \begin{bmatrix} x_1^2 + x_1 x_2 - 10 \\ x_2 + 3x_1 x_2^2 - 57 \end{bmatrix}$$

starting with the initial guess $\left[x_1^0 = 1.5, x_2^0 = 3.5\right]$, the reader can easily verify that we obtain the following values:

k	$\mathbf{x}^k$	$\|\mathbf{F}(\mathbf{x}_k)\|$	$\|\mathbf{x}^{k+1} - \mathbf{x}^k\|$
0	$\left[x_1^0 = 1.50000, x_2^0 = 3.50000\right]$	2.98172	
1	$\left[x_1^1 = 2.03603, x_2^1 = 2.84388\right]$	4.75664	0.84725
2	$\left[x_1^2 = 1.99870, x_2^2 = 3.00229\right]$	0.04978	0.16275
3	$\left[x_1^3 = 2.00000, x_2^3 = 3.00000\right]$	0.00002	0.00263

15 Derivative-Free Newton Algorithms

Numerical approximations to the Jacobian matrix are often used as alternatives to exact evaluation. The forward-difference formula

$$J_{ij}(\mathbf{x}) = \frac{\partial F_i}{\partial x_j} \approx \frac{F_i(x_1,...,x_j + \Delta x_j,...,x_n) - F_i(x_1,...,x_j,...,x_n)}{\Delta x_j} \tag{15-1}$$

allows approximation of $\mathbf{J}\left(\mathbf{x}^k\right)$ from (n+1) computations of vectors **F**(**x**), and can give good results when the step sizes are properly chosen to balance the discretization error and rounding errors. One evaluation of **J**(**x**) by this formula requires n+1 evaluations of **F**(**x**) if the step sizes are given, or about twice as many if step size optimization is included.

Another approach is to approximate $\mathbf{J}\left(\mathbf{x}^k\right)$ by a matrix $\mathbf{B}^k$, which is updated after each iteration to satisfy the divided-difference formula

$$\mathbf{B}^{k+1}(\mathbf{x}^{k+1} - \mathbf{x}^k) = \mathbf{F}(\mathbf{x}^{k+1}) - \mathbf{F}(\mathbf{x}^k) \tag{15-2}$$

for the changes across that iteration. The basis for this derivation can be seen by considering a single equation with a single variable. If we apply Newton's method to the system starting from x_0, we obtain the new point x_1 from the tangent to the curve at x_0:

$$\text{Newton step: } x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \tag{15-3}$$

where $f'(x)$ is the slope. If this derivative, $f'(x)$, is not readily available, we can approximate this term by a difference between two points. The secant formula to obtain x_3 can then be calculated by:

$$\boxed{\textbf{Secant step : } x_3 = x_2 - f(x_2)\left[\frac{x_2 - x_1}{f(x_2) - f(x_1)}\right]} \tag{15-4}$$

We can define a *secant relation* so that for some scalar, B_2, we have:

$$J_2(x_2 - x_1) = f(x_2) - f(x_1) \qquad x_3 = x_2 - J_2^{-1} f(x_2) \tag{15-5}$$

Extending Equation (15-5) into multiple dimensions we obtain the divided-difference formula

$$\mathbf{J}^{k+1}(\mathbf{x}^{k+1}-\mathbf{x}^{k})=\mathbf{F}(\mathbf{x}^{k+1})-\mathbf{F}(\mathbf{x}^{k}) \tag{15-6}$$

and the quasi-Newton step is obtained by the relation

$$\mathbf{x}^{k+1}=\mathbf{x}^{k}-\left(\mathbf{J}^{k}\right)^{-1}\mathbf{F}\left(\mathbf{x}^{k}\right) \tag{15-7}$$

Given a matrix $\mathbf{J}^{k}$, we calculate the *least change* for $\mathbf{J}^{k+1}$ from $\mathbf{J}^{k}$ that satisfies the secant formula. With this approach we obtain the *Broyden's formula:*

$$\mathbf{J}^{k+1}=\mathbf{J}^{k}-\frac{\left(\mathbf{F}\left(\mathbf{x}^{k+1}\right)-\mathbf{F}\left(\mathbf{x}^{k}\right)-\mathbf{J}^{k}\left(\mathbf{x}^{k+1}-\mathbf{x}^{k}\right)\right)\left(\mathbf{x}^{k+1}-\mathbf{x}^{k}\right)^{T}}{\left(\mathbf{x}^{k+1}-\mathbf{x}^{k}\right)^{T}\left(\mathbf{x}^{k+1}-\mathbf{x}^{k}\right)} \tag{15-8}$$

A corresponding update is also available for the **LU** representation, so that only the initial matrix $\mathbf{J}^{0}$ needs to be factored. If we define $\mathbf{B}^{k}=\left(\mathbf{J}^{k}\right)^{-1}$ then we obtain:

$$\mathbf{B}^{k+1}=\mathbf{B}^{k}-\frac{\left(\mathbf{B}^{k}\left(\mathbf{F}\left(\mathbf{x}^{k+1}\right)-\mathbf{F}\left(\mathbf{x}^{k}\right)\right)-\left(\mathbf{x}^{k+1}-\mathbf{x}^{k}\right)\right)\left(\mathbf{x}^{k+1}-\mathbf{x}^{k}\right)^{T}\mathbf{B}^{k}}{\left(\mathbf{x}^{k+1}-\mathbf{x}^{k}\right)^{T}\mathbf{B}^{k}\left(\mathbf{F}\left(\mathbf{x}^{k+1}\right)-\mathbf{F}\left(\mathbf{x}^{k}\right)\right)} \tag{15-9}$$

The Broyden method has been used widely in process simulation, especially when the number of equations is fairly small. For instance, this approach is used for inside-out flash calculations and for recycle convergence in flowsheets with or without design specifications. The rank one update formulas for Broyden's method that approximate the Jacobian ensure fast convergence. In fact, this method converges superlinearly, as defined by:

$$\lim_{k\to\infty}\frac{\left\|\mathbf{x}^{k+1}-\boldsymbol{\xi}\right\|}{\left\|\mathbf{x}^{k}-\boldsymbol{\xi}\right\|}\to 0 \tag{15-10}$$

which is slower than Newton's method but significantly faster than steepest descent.

16 Direct Substitution and Acceleration Methods

In this section we offer a brief presentation of first-order methods. We consider these methods in a fixed point form: $\mathbf{x}=\mathbf{g}(\mathbf{x})$, where **x** and **g(x)** are vectors of *n* variables. These methods are most commonly used to converge recycle streams, and here **x** represents a guessed tear stream and **g(x)** is the calculated value after executing the units around the flowsheet.

The simplest fixed point method is the *direct substitution*. Here we define the following iteration scheme:

$$\mathbf{x}^{k+1} = \mathbf{g}\left(\mathbf{x}^{k}\right) \tag{16-1}$$

with an initial guess $\mathbf{x}^{0}$. The convergence properties for the *n* dimensional case can be derived from the contraction mapping theorem. For the fixed point function, consider the Taylor series expansion:

$$\mathbf{g}\left(\mathbf{x}^{k}\right) = \mathbf{g}\left(\mathbf{x}^{k-1}\right) + \left(\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right)^{T}_{\mathbf{x}=\boldsymbol{\xi}} \left(\mathbf{x}^{k} - \mathbf{x}^{k-1}\right) \tag{16-2}$$

if we combine Equations (16-1) and (16-2) and apply Taylor's theorem we obtain:

$$\mathbf{x}^{k+1} - \mathbf{x}^{k} = \mathbf{g}\left(\mathbf{x}^{k}\right) - \mathbf{g}\left(\mathbf{x}^{k-1}\right) = \left(\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right)^{T}_{\mathbf{x}=\boldsymbol{\xi}} \left(\mathbf{x}^{k} - \mathbf{x}^{k-1}\right) \tag{16-3}$$

if we define

$$\mathbf{x}^{k+1} - \mathbf{x}^{k} = \Delta\mathbf{x}^{k+1} = \boldsymbol{\Gamma}\,\Delta\mathbf{x}^{k} \text{ with } \boldsymbol{\Gamma} = \left(\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right)^{T}_{\mathbf{x}=\boldsymbol{\xi}} \tag{16-4}$$

we obtain the following inequality:

$$\left\|\Delta\mathbf{x}^{k+1}\right\| \le \left\|\boldsymbol{\Gamma}\right\| \left\|\Delta\mathbf{x}^{k}\right\|. \tag{16-5}$$

From this expression we can show a linear convergence rate, but the speed of the iterations is related to $\left\|\boldsymbol{\Gamma}\right\|$. If we use the Euclidean norm, then $\left\|\boldsymbol{\Gamma}\right\| \le \left|\lambda\right|^{\max}$, which is the largest eigenvalue of the matrix $\boldsymbol{\Gamma}$ in magnitude. Now by recurring the iterations for *k* we can develop the following equation:

$$\left\|\Delta\mathbf{x}^{k}\right\| \le \left(\left|\lambda\right|^{\max}\right)^{k} \left\|\Delta\mathbf{x}^{0}\right\|. \tag{16-6}$$

and a necessary and sufficient condition for convergence is that $\left|\lambda\right|^{\max} < 1$. This equation is known as a *contraction mapping* if $\left|\lambda\right|^{\max} < 1$. Furthermore, the speed of convergence depends on how close $\left|\lambda\right|^{\max}$ is to zero. We can now estimate the number of iterations (n_{iter}) to reach $\left\|\Delta x^{n}\right\| \le \varepsilon$, from the relation:

$$n_{\text{iter}} \ge \frac{1}{\ln\left|\lambda\right|^{\max}} \ln\left[\frac{\varepsilon}{\left\|\Delta\mathbf{x}^{0}\right\|}\right] \tag{16-7}$$

For example, if we choose, $\varepsilon = 10^{-4}$ and $\left\|\Delta \mathbf{x}^0\right\| = 1$, we have the following maximum iteration counts, for different values of the maximum eigenvalue:

$$\begin{aligned} \left|\lambda\right|^{\max} &= 0.1 \quad n = 4 \\ \left|\lambda\right|^{\max} &= 0.1 \quad n = 14 \\ \left|\lambda\right|^{\max} &= 0.99 \quad n = 916 \end{aligned} \tag{16-8}$$

For problems where $\left|\lambda\right|^{\max}$ is close to one, the direct substitution method is limited and converges slowly. Fortunately, we can alter the fixed point function $\mathbf{g}(\mathbf{x})$ so that it reduces $\left|\lambda\right|^{\max}$. The general idea is to modify this function according to the equation show below:

$$\mathbf{x}^{k+1} = \mathbf{h}\left(\mathbf{x}^k\right) \equiv \omega\, \mathbf{g}\left(\mathbf{x}^k\right) + \left(1-\omega\right)\mathbf{x}^k \tag{16-9}$$

where ω (an acceleration factor) is chosen adaptively depending on the changes in $\mathbf{x}$ and $\mathbf{g}(\mathbf{x})$. The most commonly used fixed point method for flowsheet recycle convergence is the *Wegstein* acceleration method. Here we obtain the acceleration factor by applying a secant method independently to each component of the unknown vector $\mathbf{x}$. From the secant formula we have for component x_i:

$$x_i^{k+1} = x_i^k - f_i\left(x^k\right)\frac{x_i^k - x_i^{k-1}}{f_i\left(x^k\right) - f_i\left(x^{k-1}\right)} \tag{16-10}$$

Now, by defining $f_i\left(x^k\right) = x_i^k - g_i\left(x^k\right)$ and $s_i = \left[g_i\left(x^k\right) - g_i\left(x^{k-1}\right)\right] \Big/ \left[x_i^k - x_i^{k-1}\right]$, we have:

$$\begin{aligned} x_i^{k+1} &= x_i^k - \frac{f_i\left(x^k\right)\left[x_i^k - x_i^{k-1}\right]}{\left[f_i\left(x^k\right) - f_i\left(x^{k-1}\right)\right]} \\ &= x_i^k - \frac{\left\{x_i^k - g_i\left(x^k\right)\right\}\left[x_i^k - x_i^{k-1}\right]}{\left[x_i^k - g_i\left(x^k\right) - x_i^{k-1} + g_i\left(x^{k-1}\right)\right]} \\ &= x_i^k - \frac{\left\{x_i^k - g_i\left(x^k\right)\right\}}{\left[1-s_i\right]} = \frac{g_i\left(x^k\right)}{\left[1-s_i\right]} + \frac{s_i x_i^k}{\left[1-s_i\right]} \\ &= \omega_i g_i\left(x^k\right) + \left(1-\omega_i\right)x_i^k \end{aligned} \tag{16-11}$$

where $\omega_i = 1 \big/ \left[1-s_i\right]$. This approach works well on flowsheets where the components do not interact strongly (e.g., single recycle without reactors). On the other hand, interacting recycle loops and components can cause difficulties for this method.

17 Solution of Linear Differential Equations

In the study of chemical engineering systems, the application of macroscopic and microscopic balances for mass, momentum, energy and mechanical energy often leads to systems of linear differential equations. A differential equation simply relates variables in terms of derivatives. Differential equations describe the evolution of concentration of species as a function of time in reactive lumped systems, as well as the dependence of physical variables in space for distributed systems. In this section we review first and second order linear differential equations and describes methodologies for obtaining analytical solutions. Analytical solutions, among other things, are useful in understanding the behavior of systems in the presence of external disturbances.

18 First Order Linear Differential Equations

The simplest differential equation is of the form:

$$a(u)\frac{du}{dt} = b(t) \tag{18-1}$$

where $a(u)$ and $b(t)$ are known functions. Direct integration of this equation leads to the solution of the differential equation (18-1) in the following form:

$$\int a(u)du = \int b(t)dt + C \tag{18-2}$$

where *C* is an integration constant. This constant can be calculated by imposing a constraint on the solution at a specific point in time. This constraint is conventionally known as initial condition. The integrals in the above equation may or may not be expressed in terms of tabulated functions.

Example 18-1 The evolution of a chemical species undergoing first order chemical reaction is described the following equation and initial condition.

$$\frac{du}{dt} = -u$$
$$t = 0 \; u = 1$$

Applying the solution equation (18-2) with $a(u) = 1/u$ and $b(t) = -1$ we obtain $\ln u(t) = -t + C$. Using the initial condition the integration constant must be equal to zero and therefore the solution to this equation is $u(t) = e^{-t}$

A more general form of a first order differential equation that appears very frequently in chemical engineering systems is given by:

$$\frac{du}{dt}+a(t)u=b(t) \tag{18-3}$$

along with the appropriate initial condition. To solve this differential equation we first define an auxiliary function $p(t)$ from the following differential equation:

$$\frac{d}{dt}(pu)=pb$$

Carrying out the differentiation we obtain:

$$p\frac{du}{dt}+u\frac{dp}{dt}=pb\Rightarrow\frac{du}{dt}+\left(\frac{1}{p}\frac{dp}{dt}\right)u=b(t) \tag{18-4}$$

We now define:

$$\frac{1}{p}\frac{dp}{dt}=a(t)$$

and solving for the function $p(t)$ we obtain,

$$p=\exp\left(\int a(t)\,dt\right) \tag{18-5}$$

and therefore

$$u(t)=\frac{1}{p(t)}\int p(t)b(t)\,dt+\frac{C}{p(t)} \tag{18-6}$$

for the special case where the function $a(t)$ is constant, equation (18-6) yields

$$\begin{aligned} p(t)&=e^{at}\\ u(t)&=e^{-at}\int e^{at}b(t)\,dt+Ce^{-at}\end{aligned} \tag{18-7}$$

Example 18-2 As an example consider the differential equation that describes the concentration of a chemical A undergoing a first order reaction in a single CSTR. This equation is given by

Equation	$\frac{dC_A}{dt} + 2C_A = C_{A0}$
Initial Condition	$t = 0 \quad C_A = 0$

Using equation (18-7) with $a(t) = 2$ and $b(t) = C_{A0}$ we obtain the solution as:

$$C_A(t) = e^{-2t}\int e^{2t} C_{A0} dt + Ce^{-2t} \Rightarrow C_A(t) = \frac{C_{A0}}{2} + Ce^{-2t}$$

Applying the initial condition we obtain

$$C = -\frac{C_{A0}}{2} \Rightarrow C_A(t) = \frac{C_{A0}}{2}\left(1 - e^{-2t}\right)$$

19 Second Order Linear Differential Equations

Second order linear differential equations describe first order systems in series and are used very frequently in control studies. The general form of a second order linear differential equation is:

$$\frac{d^2u}{dt^2} + a_1\frac{du}{dt} + a_0 u = b(t) \qquad (19\text{-}1)$$

where $\{a_1, a_0\}$ are constant coefficients and $b(t)$ is called a forcing function. The general solution of the differential equation (19-1) comprises of two parts, a homogeneous u_H and a particular solution u_P. Two initial or boundary conditions are required for the solution of the second order differential equation. The homogeneous solution satisfies the differential equation when $b(t) = 0$ and it is given by an expression that depends on the roots of the characteristic equation:

$$r^2 + a_1 r + a_0 = 0 \qquad (19\text{-}2)$$

The following table shows the different type of homogeneous solutions based on the type of roots of the quadratic equation:

Types of Roots	**Roots**	**Solution**
Two real distinct	ρ_1, ρ_2	$u_H(t) = C_1 \exp(\rho_1 t) + C_2 \exp(\rho_2 t)$
One double	$\rho = \rho_1 = \rho_2$	$u_H(t) = (C_1 + C_2 t)\exp(\rho t)$
Two complex	$\rho_1 + i\rho_2, \rho_1 - i\rho_2$	$u_H(t) = \exp(\rho_1 t)(C_1 \sin(\rho_2 t) + C_2 \cos(\rho_2 t))$

A particular solution can be found analytically for special forms of the function $b(t)$ according to the table below:

Terms in function $b(t)$	Terms in particular solution $u_P(t)$
t^m	$\{t^m, t^{m-1}, \ldots t, 1\}$
$\sin(qt)$	$\sin(qt), \cos(qt)$
$\cos(qt)$	$\sin(qt), \cos(qt)$
e^{pt}	e^{pt}

Once a particular and homogeneous solution is found, the solution to the differential equation is given by:

$$u(t) = u_H(t) + u_P(t) \tag{19-3}$$

The following methodology can be applied in determining a particular solution for the family type functions given in the above table:

1. Construct the family of each term, of which $b(t)$ is a linear combination
2. If any family has a member that appears in the homogeneous solution, then replace that family with a new one in which each member of the original family is multiplied by *t*. This rule does not apply to the presence of the exponential and trigonometric functions in the homogeneous solution.
3. Assume as a particular solution a linear combination of all members of the resultant families

Example 19-1 As an example we are going to solve the differential equation that describes two CSTR in series. The form of this equation is:

Equation	$\frac{d^2C_A}{dt^2}+4\frac{dC_A}{dt}+4C_A=C_{A0}$
Boundary Conditions	$t=0 \quad C_A=\frac{dC_A}{dt}=0$

(19-4)

The roots of the characteristic polynomial $r^2+4r+4=0$ are equal $\rho_1=\rho_2=-2$, and therefore the homogeneous solution is given by

$$C_{AH}(t)=(C_1+C_2t)\exp(-2t)$$

The particular solution is easily calculated because the right hand side of the differential equation is a constant number and therefore $C_{AP}(t)=C_{A0}/4$. Consequently the solution to the differential equation is given by:

$$C_{AH}(t)=\frac{C_{A0}}{4}+(C_1+C_2t)\exp(-2t)$$

The integration constants are calculated by proper boundary conditions which when applied they give

$$C_1=-\frac{C_{A0}}{4} \quad C_2=2C_1$$