TEACHING NOTE 99-03:
MATHEMATICS REVIEW FOR FINANCIAL DERIVATIVES

Version date: March 29, 1999 C:\Class\Tn99-03.wpd

This material provides a light review of basic mathematical concepts that one should have before undertaking study of financial derivatives. If you have had some exposure to these concepts previously and the coverage in this review sounds familiar, you are in good shape. While you may find integration more difficult and may not have had exposure to differential equations, the extent of your required knowledge for the masters level course is fairly limited.

1. Summation notation

The summation symbol is often used in finance, for example,

$$\sum_{i=1}^{n} X_i = X_1 + X_2 + \ldots + X_n$$

The subscript is usually called the index. Any variable that appears on the right-hand side without the index of the summation can be brought to the left-hand side, as in the following examples:

$$\sum_{i=1}^{n} X_i Y = Y \sum_{i=1}^{n} X_i,$$
$$\sum_{i=1}^{n} X_i A_j = A_j \sum_{i=1}^{n} X_i.$$

We sometimes encounter a double summation, such as

$$\sum_{i=1}^{n} \sum_{j=1}^{q} X_i A_j,$$

which is like a double loop in a computer program. Start with the outer summation, letting i = 1. Then sum through the inner summation. Then increase the index in the outer summation and sum through the inner summation again, and continue in that manner. Thus, the above adds up to

$$\begin{aligned} &X_1A_1 + X_1A_2 + \ldots + X_1A_q \\ + &X_2A_1 + X_2A_2 + \ldots + X_2A_q \\ &+ \ldots \\ + &X_nA_1 + X_nA_2 + \ldots + X_nA_q. \end{aligned}$$

Since we can move the X_i outside the second summation, we could have written this as

$$\sum_{i=1}^{n} X_i \sum_{j=1}^{q} A_j = (X_1 + X_2 + \ldots. + X_n)(A_1 + A_2 + \ldots. + A_q).$$

Also note that the following operations are appropriate and often used in finance:

$$\sum_{i=1}^{n} (X_i + Y_i) = \sum_{i=1}^{n} X_i + \sum_{i=1}^{n} Y_i$$

$$\sum_{i=1}^{n} A_j = nA_j.$$

2. Product notation

$$\prod_{i=1}^{n} X_i = X_1X_2 \; . \; . \; . \; X_n$$

The following operations are appropriate:

$$\prod_{i=1}^{n} X_iY = Y^n \prod_{i=1}^{n} X_i,$$

$$\prod_{i=1}^{n} X_iA_j = A_j^n \prod_{i=1}^{n} X_i$$

$$\prod_{i=1}^{n} X_iY_i = \prod_{i=1}^{n} X_i \prod_{i=1}^{n} Y_i$$

$$\prod_{i=1}^{n} A_j = A_j^n.$$

Double products are rarely encountered in finance, but if you see one, you should be able to figure out what it means.

3. Logarithms and Exponentials

There are two primary types of logarithms, the base e, called the *natural log*, and base 10, called the *common log*. We write the former as $\log_e$ or sometimes ln, and we write the latter as $\log_{10}$. In financial applications we nearly always use natural logs. By definition the natural log of a value x is the power to which e = 2.71828 . . . must be raised to equal x. That is,

$$\ln x = c \text{ means that } e^c = x.$$

The notation we use for the natural log of x is sometimes written as lnx or ln(x) or ln[x]. Working with logarithms oftentimes greatly facilitate mathematical operations. The following are the major operations performed with logarithms:

$$\begin{aligned} \ln(xy) &= \ln x + \ln y \\ \ln(x^a) &= a\ln x \\ \ln\left(\frac{x}{y}\right) &= \ln x - \ln y \\ \ln e &= 1 \\ \ln e^x &= x \\ \ln 1 &= 0. \end{aligned}$$

Be aware that $\ln(x + y) \neq \ln x + \ln y$.

The mathematical definition of e is

$$\begin{aligned} e &= \lim_{n\to\infty} = \left(1 + \frac{1}{n}\right)^n \\ e^x &= \lim_{n\to\infty} = \left(1 + \frac{x}{n}\right)^n, \end{aligned}$$

and e^x can be approximated as

$$\begin{aligned} e^x &= \lim_{n\to\infty}\left(1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \ldots + \frac{1}{n!}x^n\right) \\ &= \sum_{i=0}^{\infty} x^i/i! \end{aligned}$$

The expression e^x is sometimes written exp(x).

4. Series Formulas

The following formulas for the sums of various finite and infinite series can be useful.

$$\sum_{i=0}^{n-1} x^i = \frac{1-x^n}{1-x}$$

$$\sum_{i=1}^{n} i = \frac{n(n + 1)}{2}$$

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$

$$\sum_{i=1}^{\infty} i x^i = \frac{x}{(1-x)^2} \quad for\ i < x < 1$$

$$\sum_{k=0}^{\infty} ar^k = \frac{a}{1-r}$$

$$\sum_{i=1}^{n} = x\left(\frac{1-x^n}{1-x}\right)$$

$$\sum_{i=0}^{\infty} x^i = \frac{1}{1-x}$$

$$\sum_{i=1}^{\infty} \frac{x^i}{i!} = \ln\left(\frac{1}{1-x}\right)$$

5. Derivatives

The derivative of a function tells us something about the rate at which the function changes. Given a function, y = f(x), the first derivative is denoted as dy/dx or f′(x) and is defined as

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}.$$

An alternative notation for the first derivative is y_x, but we do not use this notation much in finance because we tend to use subscripts for other things, such as an indication of a point in time. Thus, an expression like y_t, which in math usually means dy/dt, in finance often means the value of y at time t.

In words the first derivative is the limit of the slope of the line tangent to a function at a specific point. It is the rate of change of the function at that point. The terms Δy and Δx are called *differentials* and the ratio of differentials when Δx → 0 is the derivative. It is possible that the derivative does not exist at a certain point, such as a derivative that contains division by x when x = 0.

The first derivative is a function itself. There are various rules for determining the derivative of a function. The most commonly used ones are summarized below.

5a. Polynomial Functions

$$\begin{aligned}
&If\ y = c\ (a\ constant),\\
&\frac{dy}{dx} = 0.\\
&If\ y = cu\ where\ u = f(x),\\
&\frac{dy}{dx} = c\frac{du}{dx},\\
&If\ y = u + v\ where\ u = f(x)\ and\ v = g(x),\\
&\frac{dy}{dx} = \frac{du}{dx} + \frac{dv}{dx}
\end{aligned}$$

5b. Algebraic Functions

$$\begin{aligned}
&If\ y = uv\ where\ u = f(x)\ and\ v = g(x),\\
&\frac{dy}{dx} = u\frac{dv}{dx} + v\frac{dv}{u}\quad (the\ product\ rule)\\
&If\ y = \frac{u}{v}\ where\ u = f(x),\ v = g(x),\ and\ v \neq 0,\\
&\frac{dy}{dx} = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}\quad (the\ quotient\ rule)\\
&If\ y = a^x,\\
&\frac{dy}{dx} = a^x \ln a\\
&If\ y = u^n\ where\ u = f(x),\\
&\frac{dy}{dx} = nu^{n-1}\frac{du}{dx}
\end{aligned}$$

Note in the last case, the simple case where u = x gives dy/dx = nu^{n-1}(dx/dx) = nx^{n-1}.

5c. Logarithmic Functions

$$\text{If } y = \log_a u \text{ where } u = f(x),$$
$$\frac{dy}{dx} = \frac{\log_a e}{u}\frac{du}{dx}$$
$$\text{If } y = \ln u,$$
$$\frac{dy}{dx} = \frac{1}{u}\frac{du}{dx}$$
$$\text{If } y = \ln x,$$
$$\frac{dy}{dx} = \frac{1}{x}$$

5d. Exponential Functions

$$\text{If } y = a^u \text{ where } u = f(x),$$
$$\frac{dy}{dx} = a^u \ln a \frac{du}{dx}$$
$$\text{If } y = e^u \text{ where } u = f(x),$$
$$\frac{dy}{dx} = e^u \frac{du}{dx}$$
$$\text{If } y = u^v \text{ where } u = f(x) \text{ and } v = g(x),$$
$$\frac{dy}{dx} = vu^{v-1}\frac{du}{dx} + u^v \ln u \frac{dv}{dx}$$

5e. Trigonometric Functions

$$\text{If } y = \sin u \text{ where } u = f(x),$$
$$\frac{dy}{dx} = \cos u \frac{du}{dx}$$
$$\text{If } y = \cos u \text{ where } u = f(x),$$
$$\frac{dy}{dx} = -\sin u \frac{du}{dx}$$
$$\text{If } y = \tan u \text{ where } u = f(x),$$
$$\frac{dy}{dx} = \sec^2 u \frac{du}{dx}$$

5f. Inverse Functions

$$If\ y = f(x)\ and\ x = g(y)\ are\ inverse\ functions,$$
$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}}$$

5g. The Chain Rule.

If we have a function y = f(x) and x is also a function such as x = f(g), then we can take the derivative dy/dg in the following manner:

$$\frac{dy}{dg} = \left(\frac{dy}{dx}\right)\left(\frac{dx}{dg}\right),$$

or in other words, as the product of two derivatives. This is called the *chain rule*.

5h. Higher Order Derivatives

The above are examples of the first derivative, the rate of change of the function at the indicated point. There is also a second derivative, which is the rate of change of the first derivative. For example, consider y = f(x) and its first derivative, dy/dx. Its second derivative is written as follows:

$$\frac{d\left(\frac{dy}{dx}\right)}{dx} = \frac{d^2y}{dx^2},$$

and sometimes as $f''(x)$. The second derivative provides further information, beyond that of the first derivative, about the characteristics of the function. For example, a positive first derivative and positive second derivative describes a function that is upward sloping and the slope is increasing at any increasing rate. A positive first derivative and negative second derivative describes a function that is increasing at a decreasing rate. A negative first derivative and positive second derivative is a function that is decreasing at a decreasing rate. A negative first derivative and negative second derivative is a function that is decreasing at an increasing rate. A zero first derivative and positive second derivative is a function that is at a minimum while a zero first derivative and negative second derivative is a function that is at a maximum. A zero second derivative is a function that is changing its slope from positive to negative or vice versa. This is called an inflection point.

There are also third and higher order derivatives, but we rarely need them in finance.

5i. Partial Derivatives

A function containing more than one variable can be differentiated with respect to one of the variables by treating the other variables as constants. This derivative is called a *partial derivative*. Consider y = f(x,z). Then the partial derivatives are written as $\partial y/\partial x$ and $\partial y/\partial z$.

The rules for taking partial derivatives are the same as the rules for taking ordinary derivatives. For example, consider the function $y = 2x^3 + 4z^2 + c$ where c is a constant. Then

$$\frac{\partial y}{\partial x} = 6x^2$$
$$\frac{\partial y}{\partial z} = 8z.$$

Likewise, there are second partial derivatives. Hence,

$$\frac{\partial\left(\frac{\partial y}{\partial x}\right)}{\partial x} = \frac{\partial^2 y}{\partial x^2} = 12x$$
$$\frac{\partial\left(\frac{\partial y}{\partial z}\right)}{\partial z} = \frac{\partial^2 y}{\partial z^2} = 8.$$

It is common in mathematical notation that given a function y(x,z), we use an expression such as y_1 to be the partial derivative of y with respect to the first variable indicated in the expression, i.e., $\partial y/\partial x$, and y_2 to be the partial derivative with respect to the second variable, i.e., $\partial y/\partial z$. Likewise y_{11} is $\partial^2 y/\partial x^2$ and y_{22} is $\partial^2 y/\partial z^2$. We do not use this notation much in finance, owing to the use of subscripts as other indications such as a point in time.

Consider the following function: $y = 4x^3z^2$. The partial derivatives are $\partial y/\partial x = 12x^2z^2$ and $\partial y/\partial z = 8x^3z$. You can also take the partial derivative of each of these derivatives with respect to the other variable. In other words,

$$\frac{\partial\left(\frac{\partial y}{\partial x}\right)}{\partial z} = \frac{\partial^2 y}{\partial x \partial z} = 24x^2 z$$

$$\frac{\partial\left(\frac{\partial y}{\partial z}\right)}{\partial x} = \frac{\partial^2 y}{\partial z \partial x} = 24x^2 z$$

Note that they are the same. It does not matter which order you do the differentiation.

In mathematical notation cross partial derivatives often are indicated with such symbols as y_{12} and y_{21}, but as noted above we do not use this notation much in finance.

5j. Taylor's Theorem

Consider a function, f(x), observed at two points, x and x + Δx. In other words, the function is f(x) at point x and f(x + Δx) at point x + Δx. *Taylor's Theorem* is that we can take the value of the function at f(x) and add a series of terms to approximate its value at f(x + Δx):

$$f(x+\Delta x) = f(x) + \frac{f'(x)}{1!}\Delta x + \frac{f''(x)}{2!}\Delta x^2 + \frac{f'''(x)}{3!}\Delta x^3 + \ . \ . \ .$$

$$f(x+\Delta x) - f(x) = \frac{f'(x)}{1}\Delta x + \frac{f''(x)}{2}\Delta x^2 + \frac{f'''(x)}{6}\Delta x^3 + \ . \ . \ .$$

$$\Delta f(x) = f'(x)\Delta x + \frac{f''(x)}{2}\Delta x^2 + \frac{f'''(x)}{6}\Delta x^3 + \ . \ . \ .$$

This is also called a *Taylor series expansion*. A Taylor series expansion is useful when we need to study the properties of the rate of change of a function, rather than its specific value at a point.

If the function has more than one variable, we can apply Taylor's Theorem to a given variable at a time, holding the other variables constant. The derivatives above are then partial derivatives. Or we could do a Taylor series expansion around more than one variable. For example, let y = f(x,z). Then a Taylor series expansion around both x and z would be written as

$$\begin{aligned} \Delta y &= \frac{\partial y}{\partial x}\Delta x + \frac{\partial y}{\partial z}\Delta z \\ &+ \frac{1}{2}\frac{\partial^2 y}{\partial x^2}\Delta x^2 + \frac{1}{2}\frac{\partial^2 y}{\partial z^2}\Delta z^2 + \frac{\partial^2 y}{\partial x \partial z}\Delta x \Delta z \\ &+ \ . \ . \ . \end{aligned}$$

and the additional terms would represent higher order partial derivatives and all possible cross-partial combinations.

A Taylor series expansion is exact only in the limit, i.e., where the number of terms on the right-hand side is infinite. In applications we often do a Taylor series expansion to the second order, i.e., to the second derivative, and then change the equals sign to approximation ≈.

5k. Total Differential

Consider a function y = f(x,z). Suppose both x and z change. If we want to know how much the function changes by, we can obtain it using the *total differential,* which is

$$dy = \frac{\partial y}{\partial x}dx + \frac{\partial y}{\partial z}dz.$$

This expression basically says that the change in y is the change in x multiplied by the rate at which y changes if x changes plus the change in z multiplied by the rate at which y changes if z changes. It is accurate only if the changes in x and z are very small.

Note that the total differential looks somewhat like a first-order Taylor series expansion. A first-order Taylor series expansion, however, is just an approximation. The total differential is an exact expression. The distinction lies in the fact that the Taylor series expansion applies to discrete changes in x and z (Δx and Δz) and uses partial derivatives to approximate the discrete change in y (Δy). If we let Δx and Δz be infinitesimal changes, we denote them as dx and dz. Then we re-state the Taylor series expansion around two variables that we presented above as

$$\begin{aligned} dy &= \frac{\partial y}{\partial x}dx + \frac{\partial y}{\partial z}dz \\ &+ \frac{1}{2}\frac{\partial^2 y}{\partial x^2}dx^2 + \frac{1}{2}\frac{\partial^2 y}{\partial z^2}dz^2 + \frac{\partial^2 y}{\partial x \partial z}dxdz \\ &+ \ . \ . \ . \\ &= \frac{\partial y}{\partial x}dx + \frac{\partial y}{\partial z}dz, \end{aligned}$$

which is the total differential and which holds by virtue of the fact that in general any product such as $dxdx = dx^2$ or dydx is zero because it is a differential and is an infinitesimally small value, which when squared it moves closer to zero. In general $dx^k \rightarrow 0$ if $k > 1$.

5l. Total Derivative

As previously noted a derivative is a ratio of differentials. Working with the total differential, we can divide by one of the variables and obtain the *total derivative* with respect to that variable. In other words,

$$\frac{dy}{dx} = \frac{\partial y}{\partial x} + \frac{\partial y}{\partial z}\frac{dz}{dx}$$
$$\frac{dy}{dz} = \frac{\partial y}{\partial x}\frac{dx}{dz} + \frac{\partial y}{\partial z}.$$

6. Integration

Integration is closely related to differentiation, but people usually find it to be much harder. There are two general ways to classify an integral. One, called the *indefinite integral*, is the opposite of differentiation and is sometimes called the *antiderivative*. Given a derivative, indefinite integration attempts to find the function that when differentiated obtains the given derivative. For example, suppose we were given the expression 12x. The indefinite integral is $6x^2 + c$ where c is an unknown constant. This is written as

$$\int 12x = 6x^2 + c,$$

which is true because if $y = 6x^2$, then dy/dx = 12x. But what if $y = 6x^2$ plus some constant c. Then again, dy/dx = 12x. So we have to allow for a constant, which is why it is called the "indefinite" integral.

The other interpretation of integration is as the *definite integral*, which is the area under a curve between two points on the x-axis. For example, suppose we have a function f(x) and wish to know the area under the curve between the points where x = a and x = b. We write this as

$$\int_a^b f(x)dx = F(b) - F(a),$$

where the F function is obtained upon integration, which we shall demonstrate soon. This type of integral is defined specifically as the limit of the sum of an infinite series of rectangles drawn under the curve:

$$\lim_{n\to\infty} = \sum_{i=1}^{n} f(x_i)\Delta x_i = \int_a^b f(x)dx$$

where the curve between x = a and x = b has been partitioned into n rectangles. Note that the expression $f(x_i)\Delta x_i$ is the area of a rectangle with a base of length Δx_i and a height of $f(x_i)$. The expression $f(x_i)$ is simply the y value associated with a value of x. The value x_1 is x = a and the value x_n is x = b and the range of x from x = a to x = b has been partitioned into an infinite number of sub-ranges. The above definition is called a Riemann integral (pronounced Ree-mahn).

The actual process of finding the area under the curve involves determining the integral and evaluating it at the end points. We shall do that later. First let us review the main formulas and rules for integration. For indefinite integrals, the major ones are

$$\begin{aligned}
\int x^n dx &= \frac{1}{n+1}x^{n+1} + c \\
\int e^x dx &= e^x + c \quad (\textit{Note: because } de^x/dx = e^x) \\
\int \frac{1}{x}dx &= \ln x + c \\
\int [f(x) + g(x)]dx &= \int f(x)dx + \int g(x)dx \\
\int cf(x)dx &= c\int f(x)dx \\
\int f(u)\frac{du}{dx}dx &= \int f(u)du \\
\int vdu &= uv - \int udv \quad (\textit{integration by parts})
\end{aligned}$$

Tables have been constructed to provide the integrals for numerous functions, though not all functions can be integrated.

Definite integration uses these same rules to determine the integrals but without the addition of the constant. In addition the following rules are useful:

$$\int_b^a f(x)dx = -\int_a^b f(x)dx$$

$$\int_a^a f(x)dx = 0$$

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx, \quad a \le b \le c$$

$$\int_a^b -f(x)dx = -\int_a^b f(x)dx$$

$$\int_a^b cf(x)dx = c\int_a^b f(x)dx \quad where\ c\ is\ a\ constant$$

$$\int_a^b [f(x) + g(x)]dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

Now let us illustrate an example of finding the area under the curve. Suppose we have the following problem:

$$\int_a^b 12x\,dx.$$

The problem is worked by determining the function that represents the integral of 12x. This would be $F(x) = 6x^2$. Then we determine F(b) - F(a). The process is written in the following manner:

$$\int_a^b 12x\,dx = 6x^2\Big]_a^b = 6b^2 - 6a^2.$$

7. Differential Equations

A differential equation is an equation that contains a derivative. The objective of solving a differential equation is to determine the original function whose derivative is given by the differential equation. Differential equations that contain only ordinary derivatives are called *ordinary differential equations*. Differential equations that contain partial derivatives are called *partial differential equations*.

Taking a derivative creates a differential equation. For example, the expression

$$\frac{dy}{dx} = 3ax^4$$

is a differential equation. Oftentimes it is written as

$$dy = 3ax^4 dx.$$

The objective is to “solve” the differential equation, meaning to find the original function whose derivative is the differential equation. Solving differential equations can be very difficult. The one above is quite simple: we can use indefinite integration to obtain.

$$\int dy = \int 3ax^4 dx = 3a\int x^4 dx$$
$$y = \left(\frac{3a}{5}\right)x^5 + c.$$

Note, however, that without knowing the value of c we cannot be very specific about the solution, and the solution can vary widely depending on the value of c. To determine a more specific solution, we often impose a condition on the value of c, called an *initial condition* or a *boundary condition*. This represents a value we know. For example, if we know that at x = 0, the function value is 50, then we can substitute zero for x and obtain $50 = (3a/5)0^2 + c$ so c = 50. Or if we know that at x = 100, y = 5,200, then we know that $5{,}200 = (3a/5)100^2 + c$, giving us a value for c in terms of a.

Differential equations that contain partial derivatives are naturally called *partial differential equations* or PDEs. They are usually much harder to solve than are ordinary differential equations. Much of the process of solving differential equations involves classifying the equation into certain categories and then following known rules, hints and suggestions for solutions of equations in that category. Much advantage is taken of any knowledge of what the solution might possibly look like.

References

Chang, A. *Fundamental Methods of Mathematical Economics*, 3rd ed. New York: McGraw-Hill (1984).

Copeland, T. E. and J. F. Weston. *Financial Theory and Corporate Policy*, 3rd ed. Reading, Mass.: Addison-Wesley (1988), Appendix D.

Martin, J. D., S. H. Cox, and R. D. MacMinn. *The Theory of Finance: Evidence and Applications*. Chicago: Dryden Press (1988), Appendix.