Calculus Notes
Professor Sager

Section 1. Preliminaries (function, limit, continuity)

A. Function
Definition: A **function** is a rule that associates a unique output number with each input number.
Notation and terminology: x = input, $f(x)$ = output, $f: D \to \Re$, D is the domain, $\Re$ is the range (the real numbers), x is the independent variable, $y = f(x)$ is the dependent variable.

Functions are usually represented as formulas.

Ex1: Purchase $1000 promissory note, not at discount. After one year, how much will you have? Depends on interest rate r. f(r) = 1000 + 1000r

Ex2: If your promissory note pays an interest rate of .05 and compounds annually, how much will you have in n years? $f(n) = 1000(1.05)^n$

Ex3: How much would I pay now to receive $A in 5 years? Depends on my interest (discount) rate r. Let y be solution. I reason: I would pay whatever amount that would compound annually to become A: $y(1+r)^5 = A$. So $y = A(1+r)^{-5}$, or $f(r) = A(1+r)^{-5}$.

Ex4: The probability that a light bulb will burn out before x hours of use may be $f(x) = 1 - e^{-x/1000}$

Algebra of functions
Suppose that f and g are functions. New functions can be defined from f and g:

$(f+g)(x) = f(x) + g(x)$ (The sum of two functions is a new function.)

$(f-g)(x) = f(x) - g(x)$ (The difference of two functions is a new function.)

$(f*g)(x) = f(x)*g(x)$ (The product of two functions is a new function.)

$(f/g)(x) = f(x)/g(x)$ (The quotient of two functions is a new function.)

Composition (function of a function): $(f \circ g)(x) = f(g(x))$ (A function of a function is a new function.)

Ex1: Composition of functions. Suppose $f(x) = 1 + x^2$ and $g(x) = 2 - x$. Then $f \circ g(x) = 1 + (2-x)^2 = 5 - 4x + x^2$ (substitute $g(x)$ wherever there is an x in $f(x)$). Note that $f \circ g \neq g \circ f$ because $g \circ f(x) = 2 - (1 + x^2) = 1 - x^2$ (substitute $f(x)$ wherever there is an x in $g(x)$.

Ex2: CAPM (Capital Asset Pricing Model) A simple form of CAPM says that the value of a stock should be the present value of all future dividends. Suppose a stock is expected to pay $A per year. The PV of year i dividend is $\frac{A}{(1+r)^i}$. So the price of the stock should be $\sum_{i=1}^{\infty} \frac{A}{(1+r)^i}$. [Addition, division, composition of functions.]

Ex3: More sophisticated CAPM (Myron Gordon, 1959). Suppose dividend A compounds at a rate *s* per year. So the price of the stock should be $\sum_{i=1}^{\infty} \frac{A(1+s)^{i-1}}{(1+r)^i}$. [Addition, division, composition of functions.]

Ex4: Manufacturing business. Total revenue from producing q items may be f(q) = 15q – q^2. Then average revenue = f(q)/q = (15q – q^2)/q = 15 – q. [Subtraction, multiplication, division of functions.]

Ex5: Cost of running a manufacturing business. Total cost = variable cost + fixed cost. Suppose fixed cost = 125, variable cost = q^3 – 39.5q^2 + 120q, where q = number (quantity) of items produced. Then total cost of producing q items is f(q) = q^3 – 39.5q^2 + 120q + 125. The average cost of producing q = f(q)/q = (q^3 – 39.5q^2 + 120q + 125)/q. [Subtraction, addition, multiplication, division of functions.]

B. Limit

Definition: Let *f(x)* be a function defined for all *x* in an open interval containing x_0 (except possibly x_0). Let *L* be a number. We say the **limit** of *f(x)* as *x* approaches x_0 is *L* and write $\lim_{x\to x_0} f(x) = L$ if for each ε >0, there is a δ > 0 such that $0 < |x - x_0| < \delta$ implies $|f(x) - L| < \varepsilon$.

Sometimes we also write $f(x) \underset{x\to x_0}{\rightarrow} L$.

Discussion of the meaning of limit:

- The definition is the mathematical way of saying that $f(x)$ must be near *L* whenever *x* is near x_0.
- This says we can "box in" the function values *f(x)* near *L* whenever *x* is near x_0.
- The dimensions of the "box" are 2ε vertically and 2δ horizontally.
- The graph of the function must stay entirely within the box whenever *x* is within the box.
- You must be able to make such a box for every ε > 0, no matter how small. That is, if you are given <u>*any*</u> length ε > 0 of the box vertically, you must provide a horizontal width δ to "box in" the function <u>*completely*</u> within the box.
- You may choose a different δ for each different ε.
- When *x* is within δ of x_0, the function $f(x)$ must be within ε of *L*.

Ex1: $f(x)$ = 6. What is $\lim_{x\to 10} f(x)$?

Ex2: $f(x)$ = 1000(1+r). What is $\lim_{r\to .05} f(r)$?

Ex3: What is $\lim_{v\to 0}(2+v^2)$?

Ex4: $f(r) = \begin{Bmatrix} 1000(1+r), \text{if } r \ge 0 \\ 0 \quad \text{if } r < 0 \end{Bmatrix}$. What is $\lim_{r\to 0} f(r)$?

Ex5: What is $\lim_{v\to 1} \frac{1-v^2}{1-v}$?

Ex6: $f(x) = \sin\left(\frac{\pi}{x}\right)$, for x ≠ 0. What is $\lim_{x\to 0} f(x)$? For ε < 1, no box can trap the function.

Ex7: $f(x) = |x| \sin\left(\frac{\pi}{x}\right)$, for x ≠ 0. What is $\lim_{x \to 0} f(x)$?

The definition can be generalized to cover infinite limits:
Definition: (L is infinite) Let $f(x)$ be a function defined for all x in an open interval containing x_0 (except possibly x_0). We say the limit of $f(x)$ as x approaches x_0 is ∞ and write $\lim_{x \to x_0} f(x) = \infty$ if for each M > 0, there is a δ > 0 such that $0 < |x - x_0| < \delta$ implies $f(x)$ > M.
Sometimes we also write $f(x) \underset{x \to x_0}{\to} \infty$.

Ex: $f(x) = \frac{1}{x}$, for x > 0. What is $\lim_{x \to 0} f(x)$?

Definition: (x goes to infinity) Let $f(x)$ be a function defined for all x in an infinite open interval [like (a,∞)]. Let L be a number. We say the limit of $f(x)$ as x goes to infinity is L and write $\lim_{x \to \infty} f(x) = L$ if for each ε >0 there is a M > 0 such that x > M implies $|f(x) - L| < \varepsilon$.
Sometimes we also write $f(x) \underset{x \to \infty}{\to} L$.

Ex1: What is $\lim_{x \to \infty} \frac{1}{x}$?

Ex2: $f(x) = \begin{cases} 1 - e^{-x/1000}, \text{if } x \ge 0 \\ 0 \quad \text{if } x < 0 \end{cases}$. What is $\lim_{x \to \infty} f(x)$?

Ex3: \$1000 note, annual r interest. In one year you have 1000(1+r). If compound twice in year, you have $1000(1+r/2)^2$. If three times in year, you have $1000(1+r/3)^3$. What if compound continuously?
$\lim_{n \to \infty} 1000\left(1 + \frac{r}{n}\right)^n = 1000e^r$ (takes work to show).

Definition: (Right-side limit) Let $f(x)$ be a function defined for all x in an open interval (x_0 , a) (except possibly x_0). Let L be a number. We say the limit of $f(x)$ as x approaches x_0 from the right is L and write $\lim_{x \to x_0^+} f(x) = L$ if for each ε >0 there is a δ > 0 such that $0 < x - x_0 < \delta$ implies $|f(x) - L| < \varepsilon$.
Sometimes we also write $f(x) \underset{x \to x_0^+}{\to} L$.

Ex: You offer terms for payment: If bill paid within 30 days, discount 1%, 31-60 days net, more than 60 days 2% interest. Let $f(x)$ = discount if bill paid in x days. What is $\lim_{x \to 60^+} f(x)$?

Definition: (Left-side limit) Let $f(x)$ be a function defined for all x in an open interval (a, x_0) (except possibly x_0). Let L be a number. We say the limit of $f(x)$ as x approaches x_0 from the left is L and write $\lim_{x \to x_0^-} f(x) = L$ if for each ε >0 there is a δ > 0 such that $0 < x_0 - x < \delta$ implies $|f(x) - L| < \varepsilon$.
Sometimes we also write $f(x) \underset{x \to x_0^-}{\to} L$.

Ex: $f(r)=\begin{cases} 1000(1+r), \text{if } r \ge 0 \\ 0 \quad \text{if } r<0 \end{cases}$. What is $\lim_{r\to 0^-} f(r)$?

Algebra of limits

Suppose $\lim_{x\to x_0} f(x)=L_1$ and $\lim_{x\to x_0} g(x)=L_2$.

1) $\lim_{x\to x_0} c=c$ (The limit of a constant function is the constant.)
2) $\lim_{x\to x_0} cf(x)=cL_1$ (The limit of a constant times a function is the constant times the limit of the function.)
3) $\lim_{x\to x_0}[f(x)+g(x)]=L_1+L_2$ (The limit of the sum of two functions is the sum of the limits of the functions.)
4) $\lim_{x\to x_0}[f(x)-g(x)]=L_1-L_2$ (The limit of the difference of two functions is the difference of the limits of the functions.)
5) $\lim_{x\to x_0}[f(x)\cdot g(x)]=L_1\cdot L_2$ (The limit of the product of two functions is the product of the limits of the functions.)
6) $\lim_{x\to x_0}[f(x)/g(x)]=L_1/L_2$, provided $L_2 \neq 0$ (The limit of the quotient of two functions is the quotient of the limits of the functions.)

C. Continuity

Definition: A function $f(x)$ is **continuous** at $x = x_0$ if all three of the following are true:

i. $f(x_0)$ is defined
ii. $\lim_{x\to x_0} f(x)$ exists
iii. $\lim_{x\to x_0} f(x) = f(x_0)$.

f is a continuous function if f is continuous at every point x in its domain.

Discussion of the meaning of continuity:

- You can find the limit of a continuous function at x_0 by substituting x_0 for x in the formula for the function.
- You can draw the graph of a continuous function without needing to lift your pen from your paper.
- A discontinuous function may have a gap or break in its graph, or it may oscillate wildly near some points.

<Go through previous examples of limits and point out which ones are continuous.>

<Notice how concept of limit uses functions; the concept of continuity uses functions and limits; calculus will use all three.>

Algebra of continuous functions

Suppose f and g are continuous. Then

1) $f + g$ is continuous
2) $f - g$ is continuous
3) $f*g$ is continuous
4) f/g is continuous (except if $g(x) = 0$)
5) $f \circ g$ is continuous
6) $f(x) = |x|$ is continuous
7) polynomials are continuous
8) $f(x) = \sqrt{x}$ is continuous if $x \geq 0$
9) If g is continuous at L and $\lim_{x \to x_0} f(x) = L$, then $\lim_{x \to x_0} g(f(x)) = g(L)$. (This is like point 5 except that here in point 9 we do not assume that f is continuous at x_0.)
10) If g is a function of two variables and g is continuous at (L_1, L_2) and $\lim_{x \to x_1} f_1(x) = L_1$ and $\lim_{x \to x_2} f_2(x) = L_2$, then $\lim_{x \to x_1, x \to x_2} g(f_1(x), f_2(x)) = g(L_1, L_2)$. (This point is a very powerful and general result. All preceding points are actually special cases of this point.)

Ex1: What is $\lim_{v \to v_0} \frac{4v^2}{v^2 + 1}$?

Ex2: What is $\lim_{x \to 2} (|x - 2| + 1)$?

Ex3: What is $\lim_{v \to \infty} \frac{2v + 5}{v + 1}$?

Ex4: What is $\lim_{v \to 2} \frac{v^3 + v^2 - 4v - 4}{v^2 - 4}$? (Note that $v^3 + v^2 - 4v - 4 = (v - 2)(v^2 + 3v + 2)$ and $v^2 - 4 = (v - 2)(v + 2)$.)

Section 2. Derivative

<u>Definition</u>: Let y = $f(x)$ be a function. The **difference quotient** is $\frac{\Delta y}{\Delta x} = \frac{\Delta f(x)}{\Delta x} = \frac{f(x+\Delta x) - f(x)}{\Delta x}$.
This is the average rate of change of $f(x)$ on the interval from x to $x + \Delta x$.

Ex1: You can buy a promissory note for \$1000 that will pay \$1,276.28 in five years. What is the rate of return?

1) Average rate of return = $\frac{1276.28 - 1000}{5 - 0}$ = 55.26 per year, so AvPR = 5.526%.
2) Annual compounded rate: $1000(1+r)^5 = 1276.28$, so r = 5%.
3) Instantaneous rate: $1000e^{5r} = 1276.28$, so r = 4.88%

<Limit of AvPR as averaging period goes to zero is $\lim_{\Delta x \to 0} \frac{1000(1+.05\Delta x)^{1/\Delta x} - 1000}{\Delta x}$, where Δx is the fraction of a year between compounding events.>

Ex2: Fixed cost f(q) = 125 at q=10.
Ex3: Total revenue R = (15 – q)q at q=10.
Ex4: $f(x) = 3x^2 - 4$ at x_0.

<u>Definition</u>: Let $f(x)$ be a function. The **derivative** of $f(x)$ at x_0 is
$\lim_{\Delta x \to 0} \frac{\Delta f(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$, provided this limit exists. If the limit exists, then $f(x)$ is said to be **differentiable** at x_0.

Ex1: Fixed cost f(q) = 125 at q=10.
Ex2: Total revenue R = (15 – q)q at q=10.
Ex3: $f(x) = 3x^2 - 4$ at x_0.

Discussion:
1) In the definition, x_0 is thought of as being a definite fixed point.
2) Δx varies and approaches 0, but Δx must not be set exactly equal to 0, if that would result in dividing by zero.
3) Δx may be negative as well as positive in the definition. The limit must exist for both positive and negative Δx.
4) If the **difference quotient** is continuous at x_0, then Δx may be set exactly equal to 0 to evaluate the limit. But it is not wise to assume automatically that the difference quotient is continuous.
5) The derivative is defined on a point-by-point basis. Differentiability is not necessarily a property of the entire function. That is, a function can be differentiable at some x and not at others.
   Ex: $f(x)$ **= |x|.**
6) The limit must exist from both the right and the left.
   Ex: $f(x)$ **= |x|.**

7) An equivalent limit formulation of the derivative is $\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$. Here, x_0 is fixed and x varies and approaches 0. To see that this definition is the same as above, let $\Delta x = x - x_0$, and write $\lim_{\Delta x \to 0} \frac{\Delta f(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \lim_{x - x_0 \to 0} \frac{f(x_0 + x - x_0) - f(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$

8) The derivative of $f(x)$ is also a function (on the domain for which it exists). So the derivative itself may have a derivative (called **second derivative**).

9) Notation: If it exists, the derivative of $f(x)$ is written $\frac{d}{dx} f(x)$ or $f'(x)$ or $\frac{dy}{dx}$.

10) For now, the symbol $\frac{dy}{dx}$ should be thought of as a single symbol, not as the ratio of *dy* divided by *dx*. Later, the concept of differential will allow us to separate *dy* and *dx*.

11) $f'(x)$ is the instantaneous rate of change of f at x.

12) $f'(x)$ is the slope of the tangent to f at x.

13) If $f'(x) > 0$, then f is increasing at x.

14) If $f'(x) < 0$, then f is decreasing at x.

15) In economics, $f(x)$ represents the **total function**, and $f'(x)$ represents the **marginal function**. Examples include total and marginal cost, total and marginal revenue.

16) If f is differentiable at x, then f is continuous at x.

Ex1: Find the equation of line tangent to $f(x) = \sqrt{x}$ at x = 2. $\frac{\sqrt{x} - \sqrt{2}}{x - 2} = slope = f'(2) = \frac{1}{2\sqrt{2}}$.

Ex2: The total cost of producing q units is C(q) = 1040 + 25q + $3q^2$. What is the marginal cost of producing one more unit? Exactly: C(q+1) – C(q) = 28 + 6q. But this can be approximated: Since $C'(q) \cong \frac{C(q+1) - C(q)}{(q+1) - q}$, then $C(q+1) - C(q) \cong C'(q)(q+1-q) = C'(q) = 25 + 6q$. Note that marginal cost increases with increasing production.

Ex: If total revenue is R = (15 – q)q, then marginal revenue is approximated by $\frac{dR}{dq} = 15 - 2q$. Note that marginal revenue declines with increasing production.

Section 3. Algebra of derivatives

These are the fundamental rules for calculating derivatives. Let *f* and *g* be functions that are differentiable at a given number *x*.

1) $\frac{d}{dx}[c]=0$ where c is a constant.
2) $\frac{d}{dx}[x^n]=nx^{n-1}$ where n is an integer.
3) $\frac{d}{dx}[x^c]=cx^{c-1}$ where c is a constant.
4) $\frac{d}{dx}[cf(x)]=cf'(x)$
5) $\frac{d}{dx}\left[f(x)+g(x)\right]=f'(x)+g'(x)$
6) $\frac{d}{dx}\left[f(x)-g(x)\right]=f'(x)-g'(x)$
7) $\frac{d}{dx}\left[f(x)*g(x)\right]=f(x)g'(x)+f'(x)g(x)$
8) $\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right]=\frac{f'(x)g(x)-f(x)g'(x)}{[g(x)]^2}$ if $g(x)\neq 0$.
9) $\frac{d}{dx}\left[\frac{1}{g(x)}\right]=\frac{-g'(x)}{[g(x)]^2}$ if $g(x)\neq 0$.

Examples:

Powers: $x^3, x^9, x^0, x^{-3}, x^{1/2}, 2x, 4x^3, 3x^{-2}$

Sum: $5x^3+9x^3, 7x^4+2x^3-3x+37$

Product: $(2x+3)(3x^2)$

Quotient: $\frac{2x-3}{x+1}, \frac{5x}{x^2+1}, \frac{ax^2+b}{cx}$

Ex: Average revenue and marginal revenue
Total Revenue (R) = 15q – q^2.
Average Revenue (AR) = R/q = 15 – q = Price.
Marginal Revenue (MR) = 15 – 2q.

In general, R = ARxQ = *f(Q)Q*, if *f(Q)* represents average revenue. So MR = $f(Q)+f'(Q)Q$. So MR – AR = MR – P = $Qf'(Q)\begin{Bmatrix} =0 & \text{under pure competition} \\ <0 & \text{under imperfect competition}\end{Bmatrix}$

Ex: Average cost and marginal cost
Total Cost = $C(q) = q^3 - 12q^2 + 60q$.
Average Cost = $C(q)/q = q^2 - 12q + 60$.
Marginal Cost = $3q^2 - 24q + 60$.

In general, $\frac{d}{dQ}AC = \frac{C'(Q)Q - C(Q)}{Q^2} = \frac{1}{Q}\left[C'(Q) - \frac{C(Q)}{Q}\right]$ = [MC – AC]/Q. So if average costs are declining, then MC < AC; and if average costs are increasing, then MC > AC.

<u>The chain rule:</u> If *g* is differentiable at *x* and *f* is differentiable at *g(x)*, then $f \circ g$ is differentiable at *x* and $\frac{d}{dx}[(f \circ g)(x)] = f'(g(x)) \cdot g'(x)$.

Discussion:

- The chain rule shows how to differentiate a function of a function.
- The chain rule can be written $\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}$.
- The chain rule extends to multiple compositions:
  $\frac{d}{dx}[(f \circ g \circ h)(x)] = f'(g(h(x))) \cdot g'(h(x)) \cdot h'(x)$, written $\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dw}\frac{dw}{dx}$.

Ex1: $u = g(x) = x^2$, $y = f(u) = u^3$. We might think $\frac{dy}{dx} = 3u^2 = 3(x^2)^2 = 3x^4$. However,

$y = u^3 = (x^2)^3 = x^6$, so $\frac{dy}{dx} = 6x^5$. By chain rule, $\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = 3u^2 \cdot 2x = 3(x^2)^2 \cdot 2x = 6x^5$.
Ex2: $z = 3y^2$, $y = 2x + 5$.
Ex3: $z = y - 3$, $y = x^3$
Ex4: z = $(x^2 + 3x - 2)^{17}$
Ex5: z = $((3x+2)^2 + (x-4)^3)^{17}$
Ex6: Marginal products of labor : R = f(Q), Q = g(L). Marginal revenue product of labor = marginal revenue x marginal physical product of labor $MRP_L = \frac{dR}{dL} = \frac{dR}{dQ}\frac{dQ}{dL} = MR \cdot MPP_L$

<u>Definition</u>: Suppose $f(x)$ is a function. If $x_1 = x_2$ whenever $f(x_1) = f(x_2)$, then *f* is said to be **one-to-one**, or **invertible**. If *f* is invertible, then we can define a function f^{-1} by $f^{-1}(y) = x$ whenever $f(x) = y$. f^{-1} is called the **inverse function** of *f*.

Ex1: $f(x) = x^2$ is not one-to-one. For example, f(3) = f(-3) = 9.
Ex2: $y = x^2$, for x > 0 only (negative x not permitted), is invertible and the inverse function is x = $\sqrt{y}$.
Ex3: y = 5x + 25
Ex4: $y = x^5 + 5$
Ex5: y = $\sqrt{x}$

The inverse function rule: If $f(x)$ is invertible and differentiable, then f^{-1} is differentiable, and $\frac{d}{dy}[f^{-1}(y)] = \frac{1}{f'(f^{-1}(y))}$.

<immediately preceding 5 examples>

Discussion:

- The argument of f^{-1} is $y = f(x)$, not *x*.
- Informally, the inverse function rule says $\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}$. But this equation should not be taken literally, because $\frac{dx}{dy}$ is a function of *y*, whereas $\frac{dy}{dx}$ is a function of *x*.

L'Hôpital's Rule: Suppose $\lim_{x \to x_0} f(x) = 0$, $\lim_{x \to x_0} g(x) = 0$, and $f'(x_0)$, and $g'(x_0)$ exist. Then $\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{f'(x)}{g'(x)}$, if the latter is finite.

Discussion:

- L'Hôpital's Rule is useful for finding limits of ratios when the limits of both numerator and denominator are zero.
- Here is a heuristic sketch of why L'Hôpital's Rule is true: When *x* is close to x_0, then $f(x) \cong f(x_0) + f'(x)(x - x_0) = 0 + f'(x)(x - x_0)$ and $g(x) \cong g(x_0) + g'(x)(x - x_0) = 0 + g'(x)(x - x_0)$. So $\frac{f(x)}{g(x)} \cong \frac{f'(x)(x - x_0)}{g'(x)(x - x_0)} = \frac{f'(x)}{g'(x)}$. Thus the limits on each side are equal. (Proof of this requires more careful work.)
- If you apply L'Hôpital's Rule and the result $\lim_{x \to x_0} \frac{f'(x)}{g'(x)}$ is still of the form $\frac{0}{0}$, then you can apply L'Hôpital's Rule again to the result and consider $\lim_{x \to x_0} \frac{f''(x)}{g''(x)}$.
- If $\lim_{x \to x_0} f(x) = \infty$ and $\lim_{x \to x_0} g(x) = \infty$, you can write $\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{1/g(x)}{1/f(x)}$, and apply L'Hôpital's Rule to the latter.
- In fact, L'Hôpital's Rule also applies without modification to $\lim_{x \to x_0} \frac{f(x)}{g(x)}$ when the ratio is of the form $\frac{\infty}{\infty}$, as well as $\frac{0}{0}$. And it applies when $x \to \infty$, as well as $x \to x_0$.

Ex1: $\lim_{v \to 1} \frac{1 - v^2}{1 - v} = \lim_{v \to 1} \frac{-2v}{-1} = 2$

Ex2: $\lim_{v \to 2} \frac{v^3 - 5v^2 + 8v - 4}{v^2 - 4v + 4} = \lim_{v \to 2} \frac{3v^2 - 10v + 8}{2v - 4} = \lim_{v \to 2} \frac{6v - 10}{2} = \frac{2}{2} = 1$

Ex3: $\lim_{x\to 0}\frac{x+\frac{2}{\sqrt{x}}}{3+\frac{5}{x}}=\lim_{x\to 0}\frac{1/\left(3+\frac{5}{x}\right)}{1/\left(x+\frac{2}{\sqrt{x}}\right)}=\lim_{x\to 0}\frac{x/(3x+5)}{x^{1/2}/(x^{3/2}+2)}$ = (apply L'Hôpital's Rule)

$$\lim_{x\to 0}\frac{(3x+5-3x)/(3x+5)^2}{(x^{-1/2}/2(x^{3/2}+2)-x^{1/2}3x^{1/2}/2)/(x^{3/2}+2)^2} =$$

$$\lim_{x\to 0}\frac{5/(3x+5)^2}{(\frac{1}{\sqrt{x}}-x)/(x^{3/2}+2)^2}=\frac{5/(3\cdot 0+5)^2}{(\lim_{x\to 0}\frac{1}{\sqrt{x}}-0)/(0^{3/2}+2)^2}=\frac{.2}{\infty}=0.$$ Alternatively,

$$\lim_{x\to 0}\frac{x+\frac{2}{\sqrt{x}}}{3+\frac{5}{x}}=\lim_{x\to 0}\frac{\frac{d}{dx}\left(x+\frac{2}{\sqrt{x}}\right)}{\frac{d}{dx}\left(3+\frac{5}{x}\right)}=\lim_{x\to 0}\frac{1-x^{-3/2}}{-5x^{-2}}=\lim_{x\to 0}\frac{x^2-x^{1/2}}{-5}=0.$$

Section 4: Partial Derivatives

The concept of derivative can be extended to functions of more than one variable.

Definition: Let $y = f(x_1, x_2, \ldots, x_n)$ be a function of n variables. The **difference quotient for x_1** is

$$\frac{\Delta y}{\Delta x_1} = \frac{f(x_1 + \Delta x_1, x_2, \ldots, x_n) - f(x_1, x_2, \ldots, x_n)}{\Delta x_1}.$$

Ex: $f(x_1, x_2) = 3x_1^2 + x_1 x_2 + 4x_2^2$. $\frac{\Delta y}{\Delta x_1} = 6x_1 + 3\Delta x_1 + x_2$. $\frac{\Delta y}{\Delta x_2} = x_1 + 8x_2 + 4\Delta x_2$.

Definition: Let $y = f(x_1, x_2, \ldots, x_n)$ be a function of n variables. The **partial derivative** of *f* with respect to x_1 (at the point $x_1, x_2, \ldots, x_n$) is $\lim_{\Delta x_1 \to 0} \frac{\Delta y}{\Delta x_1} = \lim_{\Delta x_1 \to 0} \frac{f(x_1 + \Delta x_1, x_2, \ldots, x_n) - f(x_1, x_2, \ldots, x_n)}{\Delta x_1}$, provided this limit exists.

Ex: $f(x_1, x_2) = 3x_1^2 + x_1 x_2 + 4x_2^2$. $\frac{\partial y}{\partial x_1} = \lim_{\Delta x_1 \to 0} \frac{\Delta y}{\Delta x_1} = \lim_{\Delta x_1 \to 0} (6x_1 + 3\Delta x_1 + x_2) = 6x_1 + x_2$.

$$\frac{\partial y}{\partial x_2} = \lim_{\Delta x_2 \to 0} \frac{\Delta y}{\Delta x_2} = \lim_{\Delta x_2 \to 0} (x_1 + 8x_2 + 4\Delta x_2) = x_1 + 8x_2.$$

Discussion:

1) In the definition, $x_1, x_2, \ldots, x_n$ is thought of as being a definite fixed point.
2) The only coordinate that varies in the limit definition is the first coordinate (it varies by Δx_1).
3) Δx_1 varies and approaches 0, but Δx_1 must never be set exactly equal to 0 if that would result in dividing by zero.
4) In the definition, Δx_1 may be negative as well as positive. The limit must exist for both positive and negative Δx_1.
5) The partial derivative is defined on a point-by-point basis. Partial differentiability is not necessarily a property of the entire function. That is, a function can have a partial derivative at some values of x_1 and not at others.
6) The limit must exist from both the right and the left.
7) Notation: If it exists, the partial derivative of *f* with respect to x_1 is written as $\frac{\partial}{\partial x_1} f(x_1, x_2, \ldots, x_n)$, or $\frac{\partial y}{\partial x_1}$, or f_1.
8) Similarly, partial derivatives with respect to x_2, x_3, etc. can be defined, with analogous notation f_2, $\frac{\partial y}{\partial x_2}$; f_3, $\frac{\partial y}{\partial x_3}$; etc.

9) The partial derivative $\frac{\partial}{\partial x_1} f(x_1, x_2, \ldots, x_n)$ is also a function (on the domain for which it exists). So the partial derivative itself may have a partial derivative – and not only with respect to x_1 again $\left(\frac{\partial^2}{\partial x_1^{\,2}} f(x_1, x_2, \ldots, x_n) \right)$, but also cross partial derivatives with respect to other coordinates $\left(\text{for example } \frac{\partial^2}{\partial x_1 \partial x_2} f(x_1, x_2, \ldots, x_n) \right)$.

10) Graphically, $\frac{\partial}{\partial x_1} f(x_1, x_2, \ldots, x_n)$ is the instantaneous rate of change of $f(x_1, x_2, \ldots, x_n)$ at the point $x_1, x_2, \ldots, x_n$ in a direction parallel to the x_1 axis (holding $x_2, \ldots, x_n$ constant).

11) $\frac{\partial}{\partial x_1} f(x_1, x_2, \ldots, x_n)$ corresponds to the economist's qualifier "*ceteris paribus*" because $\frac{\partial y}{\partial x_1}$ is the rate of change in y as x_1 changes but all other variables remain the same.

Key to the technique of partial differentiation: To find the partial derivative of $f(x_1, x_2, \ldots, x_n)$ with respect to x_1, treat $x_2, \ldots, x_n$ as though they were constants and differentiate $f(x_1, x_2, \ldots, x_n)$ as though it were a function of x_1 alone.

Ex1: $f(x_1, x_2) = 3x_1^2 + x_1 x_2 + 4x_2^2$

Ex2: y = (u + 4)(3u + 2v)

Ex3: $y = \frac{3u - 2v}{u^2 + 3v}$

Ex4: Regression. Suppose we do a regression analysis and estimate demand Q as a function of price P and advertising A. Suppose that the regression equation is Q = 100 – 5P + 2A. Then $\frac{\partial Q}{\partial P} = -5$. This says that a unit increase in price results in a decrease in demand of 5 units, when advertising is held constant. When advertising is changed, the effect is to change the intercept of the price-demand relationship, but not the slope. That is, the intercept is 100 + 2A. Similarly, $\frac{\partial Q}{\partial A} = 2$. This says that a unit increase in advertising results in an increase in demand of 2 units, when price is held constant. When price is changed, the effect is to change the intercept of the advertising-demand relationship, but not the slope. That is, the intercept is 100 – 5P

Ex5: Cobb-Douglas production function. Let Q be the quantity produced, and let Q be a function of amount of labor L and capital K: Q = Q(K,L) = 96 $K^{0.3}L^{0.7}$. The marginal physical productivity of capital is $\frac{\partial Q}{\partial K} = 28.8K^{-.7}L^{.7} = 28.8\left(\frac{L}{K}\right)^{0.7}$. The marginal physical productivity of labor is $\frac{\partial Q}{\partial L} = 67.2K^{.3}L^{-.3} = 67.2\left(\frac{K}{L}\right)^{0.3}$. These two marginals depend on the ratio of capital to labor. Thus, if capital and labor are both doubled, the marginal physical productivities of capital and labor remain the same. So for the Cobb-Douglas model, productivity depends only on the proportions of capital and labor used, not on their absolute amounts.

Ex6: Market model. Consider a supply-demand model for a good, in which Q is quantity, P is price, a,b,c,d are positive parameters: The structural equations are $\begin{Bmatrix} Q = a - bP \\ Q = -c + dP \end{Bmatrix} = \begin{Bmatrix} \text{demand} \\ \text{supply} \end{Bmatrix}$. In matrix terms, $\begin{bmatrix} b & 1 \\ d & -1 \end{bmatrix}\begin{bmatrix} P \\ Q \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix}$. We solve for the reduced form:

$\begin{bmatrix} P \\ Q \end{bmatrix} = \begin{bmatrix} b & 1 \\ d & -1 \end{bmatrix}^{-1}\begin{bmatrix} a \\ c \end{bmatrix} = \begin{bmatrix} \dfrac{a+c}{b+d} \\ \dfrac{ad-bc}{b+d} \end{bmatrix} = \begin{bmatrix} P(a,b,c,d) \\ Q(a,b,c,d) \end{bmatrix}$. How do the solutions for P and Q respond to small changes in a, b, c, or d, holding the other parameters constant? For price, we calculate $\dfrac{\partial P}{\partial a} = \dfrac{\partial P}{\partial c} = \dfrac{1}{b+d} > 0$, and $\dfrac{\partial P}{\partial b} = \dfrac{\partial P}{\partial d} = \dfrac{-(a+c)}{(b+d)^2} < 0$.

Ex7: Leontief Input-Output model. Suppose we have a Leontief Input-Output model with an open sector. The solution for the amount each sector should produce to satisfy endogenous and exogenous demand is $x = (I - A)^{-1} d \equiv Bd = \begin{bmatrix} b_{11}d_1 + b_{12}d_2 + \cdots + b_{1n}d_n \\ b_{21}d_1 + b_{22}d_2 + \cdots + b_{2n}d_n \\ \vdots \\ b_{n1}d_1 + b_{n2}d_2 + \cdots + b_{nn}d_n \end{bmatrix} = x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$. The effect of an increase in demand for good j on the output required from industry i is $\dfrac{\partial x_i}{\partial d_j} = b_{ij}$.

Section 5. Differentials

<u>Definition</u>: Let $f(x)$ be differentiable at x_0. Then the **differentials** *dy* and *dx* with respect to *f* at x_0 are defined by $dy = f'(x_0)dx$.

Discussion:

1) The definition of differential permits us to separate the symbols *dy* and *dx* in $\frac{dy}{dx}$. The latter symbol has been considered indivisible until now. The reason we can separate *dy* and *dx* is that dividing both sides of $dy = f'(x_0)dx$ by *dx* yields $\frac{dy}{dx} = f'(x)$. That is, the single symbol $\left(\frac{dy}{dx}\right)$ can now be considered equal to the ratio $\frac{(dy)}{(dx)}$.
2) The justification for separating *dy* and *dx* is that if $y = f(x)$, then the change in *y* near x_0 is approximately $f'(x_0)$ times the change in *x* near x_0. That is, since $\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(x_0)$, then for small Δx, we have approximately $\frac{\Delta y}{\Delta x} \cong f'(x_0)$, so $\Delta y \cong f'(x_0)\Delta x$.
3) The previous point shows how to approximate a nonlinear function by a linear function when *x* is near x_0: In $\Delta y \cong f'(x_0)\Delta x$, substitute $\Delta x = x - x_0$ and $\Delta y = f(x) - f(x_0)$, to obtain $f(x) - f(x_0) \cong f'(x_0)(x - x_0)$. Rewrite this as $f(x) \cong f(x_0) - x_0 f'(x_0) + f'(x_0)x$. The right-hand-side is a linear function with slope $f'(x_0)$ and intercept $f(x_0) - x_0 f'(x_0)$ that approximates $f(x)$ near x_0. In fact, the equation $f(x) = f(x_0) - x_0 f'(x_0) + f'(x_0)x$ is the equation of the tangent to the graph of $f(x)$ at $x = x_0$.
4) The property $dy = f'(x_0)dx$ provides us with another interpretation of the derivative: $f'(x)$ is the function that converts a small change in *x* (*dx*) into the corresponding change in *y* (*dy*).
5) The same algebra of derivatives given in section 3 may be used to calculate differentials. (For example, the differential of the sum of two functions is the sum of their differentials. Etc.)

Ex1: Let $y = 3x^2 + 7x - 5$. Then $dy = f'(x)\,dx = (6x+7)\,dx$. Differentials may also be taken term by term: $dy = d(3x^2) + d(7x) + d(-5) = 6x\,dx + 7dx + 0 = (6x+7)dx$. The "algebra of derivatives" may be applied to calculate differentials.

Ex2: Let $y = (3x^2 - 2)(\frac{1}{x} + 4)$. Then by the product rule, $dy = (3x^2 - 2)d\left(\frac{1}{x} + 4\right) + \left(\frac{1}{x} + 4\right)d(3x^2 - 2)$

$= dy = (3x^2 - 2)\frac{-1}{x^2}dx + \left(\frac{1}{x} + 4\right)6x\,dx = \left[(3x^2 - 2)\frac{-1}{x^2} + \left(\frac{1}{x} + 4\right)6x\right]dx$.

Elasticity.

<u>Definition</u>. If $y = f(x)$, then the **elasticity of *y* with respect to *x*** is $\frac{dy}{y} / \frac{dx}{x} = \frac{dy}{dx} / \frac{y}{x}$.

We say that

$f(x)$ is **elastic** if the elasticity (in absolute value) exceeds 1;
$f(x)$ is **inelastic** if the elasticity (in absolute value) is less than 1;
$f(x)$ has **unit elasticity** if the absolute value of elasticity is 1.

Ex: Suppose the demand curve is $Q = 100 - 2P$, where Q is the quantity demanded at price P. The **price elasticity of demand** is $\varepsilon_d = \dfrac{\frac{dQ}{Q}}{\frac{dP}{P}}$. For small ΔP and ΔQ, this is approximately $\frac{\Delta Q}{Q} / \frac{\Delta P}{P}$. The denominator $\frac{\Delta P}{P}$ represents the percentage change in price; the numerator $\frac{\Delta Q}{Q}$ represents the corresponding percentage change in quantity. So the price elasticity of demand may be interpreted as the percentage change in Q resulting from a given percentage change in P. For example, suppose P=40 and is increased by 1, an increase of 2.5%. Then Q goes from 20 to 18, a decline of 10 percent. So at (P=40, Q=20), $\varepsilon_d = \frac{-2}{20} / \frac{1}{40} = -10\% / 2.5\% = -4$. In percentage terms, Q changes 4 times as fast as P. Demand is elastic at this point. Note that $\varepsilon_d = \dfrac{\frac{dQ}{Q}}{\frac{dP}{P}} = \dfrac{\frac{dQ}{dP}}{\frac{Q}{P}}$. So in this example, $\varepsilon_d =$ $-2 / \frac{Q}{P} = -2P/(100-2P)$. This exceeds 1 in magnitude if 25 < P < 50, which is therefore the set of P for which demand is elastic.

<u>Definition</u>: Let $y = f(x_1, x_2, \ldots, x_n)$ be partially differentiable with respect to x_1 at $(x_1, x_2, \ldots, x_n)$. Then the (f_1) **partial differentials** ∂y and ∂x_1 with respect to *f* at $(x_1, x_2, \ldots, x_n)$ are defined by $\partial y = f_1(x_1, x_2, \ldots, x_n) \partial x_1$.

Discussion: The partial differential converts a small change in x_1 into the corresponding change in the function when all other variables remain the same.

Ex: Suppose that the demand Q as a function of price P and advertising A is Q = 100 – 5P + 2A. Then $\partial Q = -5\,\partial P$. This says that when advertising is held constant, an increase in price results in 5 times as much decline in demand as the increase in price. Similarly, $\partial Q = 2\,\partial A$. This says that when price is held constant, an increase in advertising results in 2 times as much increase in demand as the increase in advertising.

Partial elasticity. By using partial differentials in the definition of elasticity, we may speak of partial elasticity.

Ex: Suppose that the demand Q as a function of price P and advertising A is Q = 100 – 5P + 2A. Then $\partial Q = -5\,\partial P$. So the partial price elasticity of demand (holding advertising constant) is $\frac{\partial Q}{\partial P} / \frac{Q}{P} = -5 / \frac{100-5P+2A}{P} = \frac{-5P}{100-5P+2A}$.

Definition: Let $y = f(x_1, x_2, \ldots, x_n)$ be differentiable with respect to each variable at the point ($x_1, x_2, \ldots, x_n$). Then the **total differential** of *f* is $dy = f_1 dx_1 + f_2 dx_2 + \cdots + f_n dx_n$, where each partial derivative f_i is evaluated at ($x_1, x_2, \ldots, x_n$).

Discussion:

1) Whereas the partial differential converts a change in one *x* into the corresponding function change *when all other variables are held constant*, the total differential converts *simultaneous* changes in all of the *x*'s into the corresponding function change.
2) You may use the algebra of total differentials to help evaluate a total differential. For example, if *f* and *g* are functions, then $d(f+g) = df + dg$ (the differential of the sum is the sum of the differentials), $d(f \cdot g) = f \cdot dg + g \cdot df$ (corresponds to the derivative rule for a product), $d(f \circ g) = (df \circ g) \cdot dg$ (corresponds to the chain rule for derivatives), etc.

Ex1: Suppose that the demand Q as a function of price P and advertising A is Q = 100 – 5P + 2A. Then the total differential is $dQ = -5dP + 2dA$. This says that when price is increased by 1 unit and advertising is increased by 1 unit, the demand has a net decline of 3 units.

Ex2: $y = 5x_1^2 + 3x_2$. Then $dy = 10x_1 dx_1 + 3dx_2$.

Ex3: $y = 3x_1^2 + x_1 x_2^2$. Then $dy = \frac{\partial}{\partial x_1}(3x_1^2 + x_1x_2^2)\,dx_1 + \frac{\partial}{\partial x_2}(3x_1^2 + x_1x_2^2)\,dx_2 = (6x_1 + x_2^2)\,dx_1 + 2x_1x_2\,dx_2$.

Alternative method (take total differential of both terms): $dy = d(3x_1^2) + d(x_1x_2^2) =$

$$\left(\frac{\partial}{\partial x_1}(3x_1^2)\,dx_1 + \frac{\partial}{\partial x_2}(3x_1^2)\,dx_2\right) + \left(\frac{\partial}{\partial x_1}(x_1x_2^2)\,dx_1 + \frac{\partial}{\partial x_2}(x_1x_2^2)\,dx_2\right) =$$

$$(6x_1\,dx_1 + 0\,dx_2) + (x_2^2\,dx_1 + 2x_1x_2\,dx_2) = (6x_1 + x_2^2)\,dx_1 + 2x_1x_2\,dx_2.$$

Ex4: $y = \frac{x_1 + x_2}{2x_1^2}$. Take total differential of both terms: $dy = d\left(\frac{x_1+x_2}{2x_1^2}\right) = d\left(\frac{1}{2x_1}\right) + d\left(\frac{x_2}{2x_1^2}\right) =$

$$\left(\frac{-1}{2x_1^2}dx_1 + 0\,dx_2\right) + \left(\frac{-4x_1x_2}{(2x_1^2)^2}dx_1 + \frac{1}{2x_1^2}dx_2\right) = \frac{-x_1 - 2x_2}{2x_1^3}dx_1 + \frac{1}{2x_1^2}dx_2.$$

There are rules for calculating total, partial, and ordinary differentials of sums, products, quotients, compositions,etc. These rules are so similar to the rules for derivatives (see "algebra of derivatives" in earlier section) that they are not stated here.

Total Derivative.

Ex1: Suppose that $y = f(x,w) = 3x - w^2$, where $x = g(w) = 2w^2 + w + 4$. How to find the rate of change of *y* with respect to *w*?

One solution is to substitute: $y = 3(2w^2 + w + 4) - w^2 = 5w^2 + 3w + 12$, so that $\frac{dy}{dw}$ = 10*w* + 3.

Another solution is to convert the total differential into a **total derivative**: $dy = \frac{\partial y}{\partial x}dx + \frac{\partial y}{\partial w}dw$ = $3\,dx - 2w\,dw$ Then the total rate of change in *y* with respect to *w* is
$\frac{dy}{dw} = 3\frac{dx}{dw} - 2w = 3(4w+1) - 2w = 10w + 3$.

Ex2: Suppose that $z = x^2 - 8xy - y^3$, where *x* = 3*t* and *y* = 1 – *t*. Then the total differential is
$dz = \frac{\partial}{\partial x}(x^2 - 8xy - y^3)dx + \frac{\partial}{\partial y}(x^2 - 8xy - y^3)dy$ $= (2x - 8y)\,dx + (-8x - 3y^2)\,dy$. So
$\frac{dz}{dt} = (2x - 8y)\,\frac{dx}{dt} + (-8x - 3y^2)\,\frac{dy}{dt}$ = $(2x - 8y)\,3 + (-8x - 3y^2)\,(-1)$ = $14x - 24y + 3y^2$ =
$14(3t) - 24(1-t) + 3(1-t)^2$ = $-21 + 60t + 3t^2$.

Implicit Differentiation

Another use for differentials is implicit differentiation. Usually, a function is defined by writing the dependent variable (*y* or $f(x)$) on the left-hand-side of the equal sign and an expression involving the independent variable (*x*) on the right. Occasionally, you will see an equation with the dependent and independent variables mixed together in an equation that can be expressed as $F(y,x) = 0$. In such cases, implicit differentiation may be used to find derivatives, assuming they exist.

Ex1: Find $\frac{dy}{dx}$ if $yx + y + 1 = x$. This could be solved for *y* and rewritten as $y = \frac{x-1}{x+1}$, if *x* ≠ -1. Then $\frac{dy}{dx}$ = $\frac{2}{(x+1)^2}$. But we can also write it as $F(x,y) = yx + y + 1 - x = 0$ and take the (total) differential of both sides: $\frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy = (y-1)dx + (x+1)dy$. So $(x+1)dy = (1-y)dx$. So $\frac{dy}{dx} = \frac{1-y}{x+1}$, which =
$\frac{1 - \frac{x-1}{x+1}}{x+1} = \frac{x+1-(x-1)}{(x+1)^2} = \frac{2}{(x+1)^2}$.

Ex2: $y - 3x^4 = 0$. So $-12x^3dx + dy = 0$. So $\frac{dy}{dx} = 12x^3$.

Ex3: The equation of a circle $x^2 + y^2 - 9 = 0$. So $2x\,dx + 2y\,dy = 0$. So $\frac{dy}{dx} = -\frac{x}{y}$ if y ≠ 0 (must also restrict to one semicircle at a time to be a function.)

Ex4: Let $F(y,x,w)=y^3x^2+w^3+yxw-3=0$. Find $\dfrac{\partial y}{\partial x}$. Take partial differentials (holding *w* constant): $\dfrac{\partial F}{\partial x}\partial x+\dfrac{\partial F}{\partial y}\partial y=0$. So $(y^3 2x+yw)\,\partial x+(3y^2x^2+xw)\,\partial y=0$. So $\dfrac{\partial y}{\partial x}=-\dfrac{2xy^3+yw}{3y^2x^2+xw}$.

Section 6. Optimization

A major use of differential calculus is to find maxima and minima of functions.

Definition: Let $f(x)$ be a function defined on an interval [a, b].

- *f* is **increasing** on [a, b] if $f(x_1) < f(x_2)$ whenever $a \le x_1 < x_2 \le b$.
- *f* is **decreasing** on [a, b] if $f(x_1) > f(x_2)$ whenever $a \le x_1 < x_2 \le b$.
- *f* is **constant** on [a, b] if $f(x_1) = f(x_2)$ whenever $a \le x_1 < x_2 \le b$.
- *f* is **concave** on [a, b] if the straight line segment connecting $f(x_1)$ and $f(x_2)$ lies entirely below $f(x)$ whenever $a \le x_1 < x < x_2 \le b$.
- *f* is **convex** on [a, b] if the straight line segment connecting $f(x_1)$ and $f(x_2)$ lies entirely above $f(x)$ whenever $a \le x_1 < x < x_2 \le b$.
- If *f* is concave on [a, x_0] and convex on [x_0, b] (or the reverse) then x_0 is an **inflection point** of *f*.
- *f* has a **relative maximum** at x_0 if there is an interval (a, b) containing x_0 with $f(x_0) \ge f(x)$ whenever a < *x* < b. [Analogously for **relative minimum**.] Relative maxima and relative minima are called **relative extrema**.
- *f* has an **absolute maximum** at x_0 if $f(x_0) \ge f(x)$ for all *x*. [Analogously for **absolute minimum**.] Absolute maxima and absolute minima are called **absolute extrema**.
- The **critical points** of *f* are those *x* for which $f(x)$ exists and either $f'(x)$ = 0 or $f'(x)$ does not exist.

Theorem. Let *f* be continuous on [a, b] and differentiable on (a, b).

- If $f'(x)$ > 0 whenever a < *x* < b, then *f* is increasing on [a, b].
- If $f'(x)$ < 0 whenever a < *x* < b, then *f* is decreasing on [a, b].
- If $f'(x)$ = 0 whenever a < *x* < b, then *f* is constant on [a, b].
- If $f''(x)$ < 0 whenever a < *x* < b, then *f* is concave on [a, b].
- If $f''(x)$ > 0 whenever a < *x* < b, then *f* is convex on [a, b].

Ex1: Where does $f(x) = x^2 - 4x + 3$ increase and decrease? Where concave or convex? Any inflection points? $f'(x)$ = 2x – 4 is negative for x < 2 (*f* decreasing) and is positive for x > 2 (*f* increasing). $f''(x)$ = 2 > 0 so convex everywhere. No inflection points.

Ex2: Where does $f(x) = x^3 - 3x^2 + 1$ increase and decrease? Where concave or convex? Any inflection points? $f'(x)$ = $3x^2$ – 6x = 3x(x – 2) is positive for x < 0 or x > 2 (both increasing) and negative for 0 < x < 2 (decreasing). $f''(x)$ = 6x – 6 is negative if x < 1 and positive if x > 1 so concave if x < 1, convex if x > 1. x = 1 is an inflection point.

Theorem. If *f* has any relative extrema, they occur at critical points of *f*.

Ex1: What are the critical points of $f(x) = x^2 - 4x + 3$? $f'(x)$ = 2x – 4 = 0 at x = 2.

Ex2: What are the critical points of $f(x) = x^3 - 3x^2 + 1$? $f'(x)$ = 3x^2 – 6x = 3x(x – 2) = 0 at x = 0 and x = 2.

Ex3: What are the critical points of $f(x) = -5x + 10$? $f'(x)$ = -5 is never 0. No critical points.

Theorem. If *f* has any absolute extrema, they occur either at critical points of *f* or at endpoints of the domain of *f*.

Theorem (**First Derivative Test**): Suppose *f* is continuous at a critical point x_0. If there is an interval (a, b) containing x_0 and

a) $f'(x) > 0$ on (a, x_0) and $f'(x) < 0$ on (x_0, b), then *f* has a relative maximum at x_0.
b) $f'(x) < 0$ on (a, x_0) and $f'(x) > 0$ on (x_0, b), then *f* has a relative minimum at x_0.
c) $f'(x)$ has the same sign on (a, x_0) and (x_0, b), then *f* does not have a relative extremum at x_0.

Ex1: Classify the relative extrema of $f(x) = x^2 - 4x + 3$. $f'(x)$ = 2x – 4 is negative for x < 2 and is positive for x > 2. So x = 2 is a point of relative minimum.

Ex2: Classify the relative extrema of $f(x) = |x|$. $f'(x)$ = -1 for x < 0, and $f'(x)$ = 1 for x > 0. So x = 0 is a point of relative minimum. This applies for all x, so x = 0 is a point of absolute minimum. Note that this example shows that the first derivative test does not require that the derivative exist at the point of relative extremum, since $f'(0)$ does not exist.

Ex3: Classify the relative extrema of $f(x) = x^3 - 3x^2 + 1$. $f'(x)$ = 3x^2 – 6x = 3x(x – 2) is positive for x < 0 or x > 2 and negative for 0 < x < 2. So x = 0 is a point of relative maximum, x = 2 is a point of relative minimum.

Ex4: Classify the relative extrema of $f(x) = x^3 - 12x^2 + 36x + 8$. $f'(x) = 3x^2 - 24x + 36$ = $3(x-2)(x-6)$ is positive for x < 2 or x > 6 and negative for 2 < x < 6. So x = 2 is a point of relative maximum, x = 6 is a point of relative minimum.

Ex5: Let average cost as function of quantity be $f(Q) = Q^2 - 5Q + 8$. For what quantity is average cost a minimum? $f'(Q) = 2Q - 5$ is negative for Q < 2.5 and positive for Q > 2.5. So Q = 2.5 is a relative minimum. This applies for all Q > 0, so this is an absolute minimum.

Ex6: Analyze the shape of $f(x) = a - \dfrac{b}{c+x}$, for a, b, c > 0, x > 0. $\lim\limits_{x \to 0} f(x) = a - \dfrac{b}{c}$. $\lim\limits_{x \to \infty} f(x) = a - \lim\limits_{x \to \infty} \dfrac{b}{c+x} = a$. $f'(x) = \dfrac{b}{(c+x)^2}$, which is positive when x > 0. $f''(x) = \dfrac{-b}{(c+x)^3}$, which is negative when x > 0. So $f(x)$ is everywhere concave and increasing. $f(x)$ = 0 at x = b/a – c.

Theorem (**Second Derivative Test**): Suppose that $f''(x_0)$ exists.

a) If $f'(x_0)=0$ and $f''(x_0) > 0$ then *f* has a relative minimum at x_0.
b) If $f'(x_0)=0$ and $f''(x_0) < 0$ then *f* has a relative maximum at x_0.
c) If $f'(x_0)=0$ and $f''(x_0) = 0$ the test is inconclusive.

Ex1: Classify the relative extrema of $y=4x^2-x$. $\frac{dy}{dx} = 8x - 1 = 0$ at x = 0.125. $\frac{d^2y}{dx^2}=8>0$. So x = 0.125 is a point of relative minimum.

Ex2: Classify the relative extrema of $y=x^3-3x^2+2$. $\frac{dy}{dx} = 3x^2 - 6x = 0$ at x = 0 or x = 2. $\frac{d^2y}{dx^2}=6x-6$, which is negative at x = 0 and positive at x = 2. So x = 0 is a relative maximum, x = 2 is a relative minimum.

Ex3: Profit Maximization. Suppose total revenue and total cost as functions of quantity produced are $R(q)=1200q-2q^2$ and $C(q)=q^3-61.25q^2+1528.5q+2000$. Then profit is $P(q)=R(q)-C(q)$ = $(1200q-2q^2)-(q^3-61.25q^2+1528.5q+2000)$ = $-q^3+59.25q^2-328.5q-2000$. $P'(q)=-3q^2+118.5q-328.5$ = $-3(q-3)(q-36.5)$, which yields critical points q = 3 and q = 36.5 when set equal to 0. $P''(q)=-6q+118.5$ is positive if q = 3 and negative if q = 36.5. So q = 3 is a point of relative minimum and q = 36.5 is a point of relative maximum. To determine that q = 36.5 is absolute maximum, compare it with boundary point: P(36.5) = 16318.44 vs. P(0) = -2000.

Ex4: Profit Maximization in General. For general R(q) and C(q), the critical points for profit occur at $P'(q)=R'(q)-C'(q)=0$. That is, $R'(q)=C'(q)$, which says marginal revenue = marginal cost. This (these) critical point(s) will be relative maxima if $P''(q)<0$, which occurs if $R''(q)<C''(q)$, i.e., marginal revenue is changing less rapidly than marginal cost.

Ex5: Classify the relative extrema of $f(x)=x^3$. $f'(x) = 3x^2$. So x = 0 is a critical point. $f''(x)=6x$, which is 0 at x = 0. So the second derivative test is inconclusive. By looking at a graph of $f(x)$, we see that there are no relative extrema.

Technique for finding the extrema of $f(x)$ defined on an interval [a, b]:

1. Find the critical points of $f(x)$ in (a, b).
2. Evaluate $f(x)$ at the critical points and *f(a)* and *f(b)*.
3. The largest *f*-value in step 2 is an absolute maximum; the smallest *f*-value in step 2 is an absolute minimum.

Ex1: Determine the maxima and minima of $f(x)=\frac{1}{x}+9x-4$ for x > 0. $f'(x)=-\frac{1}{x^2}+9$ = 0 if x = 1/3. $f''(x)=\frac{2}{x^3}$ > 0 for x > 0. So x = 1/3 is a relative minimum. $f(x)$ is convex, so x = 1/3 is absolute minimum. $\lim_{x\to 0} f(x)=\infty$, so there is no (relative or absolute) maximum.

Ex2: Suppose that average revenue is AR = $8000-23q+1.1q^2-0.018q^3$. Find marginal revenue and analyze the shape of marginal revenue function. Total revenue = AR*q = $8000q-23q^2+1.1q^3-0.018q^4$. So marginal revenue = $f(q)=8000-46q+3.3q^2-0.072q^3$. At q = 0, MR = 8000. $f'(q)=-46+6.6q-.216q^2$. Set equal to 0 and solve to find q = 10.76 and 19.8. $f''(q)=6.6-0.432q$ is positive if 0 < q < 15.28 and MR is convex there, and is negative if q > 15.28 and MR is concave there. So q = 15.28 is an inflection point. And q = 10.76 is a point of relative minimum, and q = 19.8 is a point of relative maximum.

<u>Definition</u>: A power series is a function of the form $f(x)=a_0+a_1x+a_2x^2+a_3x^3+\cdots=\sum_{i=0}^{\infty}a_ix^i$, provided the sum is finite.

Discussion:

1) The sum may not actually have infinitely many terms, since a_i could be zero for all but a few of them.
2) All polynomials clearly have this form. So all polynomials are power series.
3) If we differentiate both sides and assume that we can differentiate the right-hand side term-by-term, then $f'(x)=a_1+2a_2x+3a_3x^2+\cdots=\sum_{i=1}^{\infty}ia_ix^{i-1}$. Similarly,

   $f''(x)=2a_2+3\cdot 2a_3x^1+\cdots=\sum_{i=2}^{\infty}i(i-1)a_ix^{i-2}$. In general,

   $f^{(n)}(x)=n!a_n+(n+1)!a_{n+1}x^1+\cdots=\sum_{i=n}^{\infty}i!a_ix^{i-n}$. Thus $f'(0)=a_1$, $f''(0)=2a_2$, …,

   $f^{(n)}(0)=n!a_n$.

Ex: For every –1 < x < 1, $f(x)=\dfrac{1}{1-x}=1+x+x^2+x^3+x^4+\cdots$. Thus a_n = 1 for all n. So

$$f^{(n)}(0)=\frac{d^n}{dx^n}\left(\frac{1}{1-x}\right)\bigg|_{x=0}=n!.$$

<u>**Taylor's Theorem**</u>: An arbitrary function $f(x)$ that has n derivatives at x_0 can be approximated by a power series as $f(x)=\sum_{i=0}^{n}\dfrac{f^{(i)}(x_0)}{i!}(x-x_0)^i+R_n$, where R_n is a remainder term and for notational convenience we define $f^{(0)}(x)=f(x)$.

<u>**Lagrange Remainder**</u>: If we assume another derivative (n+1) at x_0, the remainder term has the form $R_n=\dfrac{f^{(n+1)}(x^*)}{(n+1)!}(x-x_0)^{n+1}$, where x^* is some number between x and x_0.

Theorem (**n^{th} Derivative Test**): Suppose that $f(x)$ has n continuous derivatives at x_0. Suppose that all of these derivatives satisfy $f^{(i)}(x_0) = 0$, i = 1, 2, …, n-1, except that $f^{(n)}(x_0) \neq 0$. Then

a) x_0 is a point of relative maximum if n is even and $f^{(n)}(x_0) < 0$.
b) x_0 is a point of relative minimum if n is even and $f^{(n)}(x_0) > 0$.
c) x_0 is an inflection point if n is odd.

Discussion:

1) To apply the n^{th} derivative test, you start taking derivatives and evaluating them at x_0: $f'(x_0)$, $f''(x_0)$, $f'''(x_0)$, … until you find the first one that is not zero.
2) The n^{th} derivative test works because by Taylor's Theorem, the first n-1 derivative terms in the Taylor's expansion are all zero since the first n-1 derivatives are all zero (note that the $zero^{th}$ term $f^{(0)}(x_0) = f(x_0)$ is not a derivative and is not necessarily assumed to be zero):
$f(x) = \sum_{i=0}^{n-1} \frac{f^{(i)}(x_0)}{i!}(x-x_0)^i + R_{n-1} = f(x_0) + R_{n-1} = f(x_0) + \frac{f^{(n)}(x^*)}{n!}(x-x_0)^n$. So
$f(x) - f(x_0) = \frac{f^{(n)}(x^*)}{n!}(x-x_0)^n$. Because of the assumed continuity of the n^{th} derivative and the fact that x* is very close to x_0, the sign of $f^{(n)}(x^*)$ will be the same as the sign of $f^{(n)}(x_0)$. Now suppose n is even. Then $(x-x_0)^n > 0$, whether x is to the right or left of x_0. So $f(x) - f(x_0)$ has the same sign on both sides of x_0, which makes x_0 a point of relative extremum. If n is odd, then a similar analysis shows that since $(x-x_0)^n$ switches signs as x goes from one side of x_0 to the other, then $f(x) - f(x_0)$ switches signs, so x_0 cannot be a point of relative extremum and must be an inflection point.
3) The n^{th} derivative test includes the first and second derivative tests as special cases.

Ex: Classify the extrema of $f(x) = x^3$. $f'(x) = 3x^2$, $f''(x) = 6x$, $f'''(x) = 6$. There is one critical point x = 0. But it fails the second derivative test. Apply the n^{th} derivative test: n=3. So x=0 is an inflection point and there are no extrema.

Ex: Classify the extrema of $f(x) = (x-2)^4$. $f'(x) = 4(x-2)^3$, $f''(x) = 12(x-2)^2$, $f'''(x) = 24(x-2)$, $f^{(4)}(x) = 24$. There is one critical point x=2. The second derivative test fails. By n^{th} derivative test, x=2 is a relative minimum (absolute minimum).

Section 7. Exponentials and Logarithms

When you exponentiate, you calculate the power of a number. For example, $10^2 = 100$. 100 is the result of the exponentiation; 10 is the base; and 2 is the exponent (or logarithm). When you take the logarithm of a number, you find the exponent that yields the number. For example, $\log_{10}(100) = 2$ because 2 is the exponent of 10 that yields 100.

Most hand calculators include exponential and logarithmic function keys both for the common base 10 and for the special base $e = 2.71828\ldots$ The exponential keys usually say "10^x" and "e^x". For base 10, the logarithm key usually says "log_{10}" or just "*log*". For base *e*, the logarithm key usually says "log_e" or "ℓn". To use the keys correctly, you need to understand that to exponentiate means to calculate a power of a number, and to take the logarithm means to find the exponent. For example, to calculate $10^2 = 100$, enter "2" and press "10^x"; to calculate $\log_{10}(100) = 2$, enter "100" and press "log_{10}".

<u>Definition</u>: The **base** for an exponential or logarithmic number system is a positive number *b*, usually 10 or the special number $e = 2.71828\ldots$ In such a system, any positive number *y* can be represented as $y = b^x$ for some choice of *x*. However, neither zero nor any negative number equals a positive number *b* raised to any power. Thus, neither zero nor negative numbers may be used for *y* in an exponential or logarithmic system. Any number, positive, negative, or zero may be used for *x*. When $y = b^x$, *x* is said to be the **logarithm** to the base *b* of *y*, written $x = \log_b y$. Without loss of generality, *b* may be taken to be greater than 1.

- Base 2 ($b = 2$) is the base for number representation and integer calculation by computers. $2^0=1$, $2^1=2$, $2^2=4$, $2^3=8$, $2^4=16$, with binary representations 1_2 10_2 100_2 1000_2 10000_2 respectively.
- Base 10 ($b = 10$) is used for our common number system of everyday life (common base). It is widely used in engineering applications.
- Base *e* ($b = e$) is used everywhere else (the so-called "natural" base). $e = 2.71828\ldots$ "ln" is often written instead of $\log_e$ for logarithms with base *e*.

Examples of logarithmic/exponential scales include Richter earthquake scale, decibel sound scale, pH acidity scale, Value-Line stock average, the magnitude scale for stellar brightness in astronomy.

Most of our work in this course will focus on base *e*.

Properties of exponentials and logarithms

- The exponential function $y = b^x$ is one-to-one and so is invertible. The inverse function is the logarithm function $x = \log_b y$. That is, the exponential and logarithm functions are inverses of each other. The following properties are true for any base, but are illustrated for base 10:

Exponential	**Logarithm**	**Exp example**	**Log example**
$10^0 = 1$	$\mathrm{Log}_{10}1 = 0$	$10^0 = 1$	$\log_{10}10 = 1$
$10^{x+y} = 10^x 10^y$	$\mathrm{Log}_{10}(xy) = \log_{10}x + \log_{10}y$	$10^{2+3}=10^2\ 10^3$	$\log_{10}(100*1000) = \log_{10}(100) + \log_{10}(1000)$
$10^{x-y} = 10^x/10^y$	$\log_{10}(x/y) = \log_{10}x - \log_{10}y$	$10^{2-3} = 10^2/10^3$	$\log_{10}10^{2-3} = \log_{10}10^2 - \log_{10}10^3$
$10^{ax} = (10^x)^a$	$\log_{10}x^a = a \log_{10}x$	$10^{0.5*2} = (10^2)^{0.5}$	$\log_{10}100^{0.5} = 0.5 \log_{10}100$
(Change of base) $c^x = b^{x \cdot \log_b c}$	$\log_c x = (\log_c b)(\log_b x)$	$2^3 = 10^{3\log_{10}2}$	$\log_2 3 = (\log_2 10)(\log_{10}3)$

Beware of false properties! It may seem that the following should be true, but they are not true:

- $b^{x+y} \neq b^x + b^y$
- $b^{x-y} \neq b^x - b^y$
- $(b^x)^a \neq b^{x^a}$
- $\log(x+y) \neq \log(x) + \log(y)$
- $\log(x-y) \neq \log(x) - \log(y)$
- $\log(xy) \neq \log(x)\log(y)$
- $\log(x^a) \neq [\log(x)]^a$

The derivative of e^x and log(x)

Fact: For $-0.4 \leq u \leq 0.4$, it is true that $u - u^2 \leq \ln(1+u) \leq u$.

Theorem: $\frac{d}{dx}\ln x = \frac{1}{x}$ for $x > 0$.

Proof: The difference quotient is $\frac{\ln(x+\Delta x) - \ln x}{\Delta x} = \frac{1}{x}\frac{\ln(1+\frac{\Delta x}{x})}{\frac{\Delta x}{x}}$. Use the preceding fact with $u = \frac{\Delta x}{x}$ to bound this: $\frac{1}{x}\frac{\frac{\Delta x}{x} - \left(\frac{\Delta x}{x}\right)^2}{\frac{\Delta x}{x}} \leq \frac{1}{x}\frac{\ln(1+\frac{\Delta x}{x})}{\frac{\Delta x}{x}} \leq \frac{1}{x}\frac{\frac{\Delta x}{x}}{\frac{\Delta x}{x}}$. So $\frac{1}{x}\left(1 - \frac{\Delta x}{x}\right) \leq \frac{1}{x}\frac{\ln(1+\frac{\Delta x}{x})}{\frac{\Delta x}{x}} \leq \frac{1}{x}$. Take the limit as $\Delta x \to 0$: $\lim_{\Delta x \to 0}\frac{1}{x}\left(1 - \frac{\Delta x}{x}\right) \leq \lim_{\Delta x \to 0}\frac{1}{x}\frac{\ln(1+\frac{\Delta x}{x})}{\frac{\Delta x}{x}} \leq \lim_{\Delta x \to 0}\frac{1}{x}$. So $\frac{1}{x} \leq \lim_{\Delta x \to 0}\frac{1}{x}\frac{\ln(1+\frac{\Delta x}{x})}{\frac{\Delta x}{x}} \leq \frac{1}{x}$. Hence $\frac{d}{dx}\ln x = \frac{1}{x}$.

Theorem: $\frac{d}{dx}e^x = e^x$.

Proof: Because x = ln(y) and $y = e^x$ are inverses, use the inverse function theorem to get
$\frac{d}{dx}e^x = \frac{1}{\frac{d}{dy}\ln y} = \frac{1}{1/y} = y = e^x$.

Discussion: From the second derivatives, we see that the exponential function is everywhere convex $\left(\frac{d^2}{dx^2}e^x = e^x > 0\right)$, and the logarithm function is everywhere concave $\left(\frac{d^2}{dx^2}\ln x = \frac{-1}{x^2} < 0\right)$.

Theorem: $\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = e^x$.

Proof: Because ln is continuous, $\ln\left(\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n\right) = \lim_{n\to\infty}\ln\left(1+\frac{x}{n}\right)^n = \lim_{n\to\infty}\frac{\ln(1+x/n)}{1/n}$. The limits of both numerator and denominator are 0, so L'Hôpital's Rule applies:

$$\lim_{n\to\infty}\frac{\ln(1+x/n)}{1/n} = \lim_{n\to\infty}\frac{\frac{d}{dn}\ln(1+x/n)}{\frac{d}{dn}(1/n)} = \lim_{n\to\infty}\frac{\frac{1}{1+x/n}\frac{-x}{n^2}}{\frac{-1}{n^2}} = \lim_{n\to\infty}\frac{x}{1+x/n} = x = \ln(e^x).$$ This says ln(left-hand-side of the theorem) = ln(right-hand-side of the theorem). Thus both sides of the theorem are equal.

Theorem: $e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = \sum_{n=0}^{\infty}\frac{x^n}{n!}$.

Proof: Since $f(x) = e^x$ is its own derivative, so are its higher derivatives: $\frac{d^n}{dx^n}e^x = e^x$. Thus $f^n(0) = 1$ for all n. Substitute this into the general form for power series (Taylor's series) expansion: $f(x) = \sum_{n=0}^{\infty} f^{(n)}(0)\frac{x^n}{n!}$ to finish the argument.

Theorem: $\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots = \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}x^n$.

Proof: $\frac{d}{dx}\ln(1+x) = \frac{1}{1+x} = (1+x)^{-1}$. So $\frac{d^2}{dx^2}\ln(1+x) = -(1+x)^{-2}$, $\frac{d^3}{dx^3}\ln(1+x) = 2(1+x)^{-3}$, …, $\frac{d^n}{dx^n}\ln(1+x) = (-1)^{n+1}(n-1)!(1+x)^{-n}$, … Thus in the Taylor's series expansion, we have $f^{(n)}(0) = (-1)^{n+1}(n-1)!$ for n > 0, and $f^{(0)}(0) = \ln(1+0) = 0$. Substitute these into the general form for power series (Taylor's series): $f(x) = \sum_{n=0}^{\infty} f^{(n)}(0)\frac{x^n}{n!}$ to finish the argument.

Discussion: On the basis of the power series expansions, if x is small, the approximations $e^x \cong 1 + x$ and $\ln(1+x) \cong x$ are often used. (For a little more accuracy, use $e^x \cong 1 + x + \frac{x^2}{2}$ and $\ln(1+x) \cong x - \frac{x^2}{2}$.)

Ex: $e^{0.10}$ = 1.10517, compared with the approximation 1 + .10 = 1.10
Ex: ln (.95) = -.05129, compared with the approximation ln (.95) = ln (1 - .05) ≅ -.05

Exponentials and Growth

1) A quantity Q that grows at a compounded rate of r per year will become $Q(1+r)^t$ after t years.
2) If Q compounds more frequently than once per year, say n times per year, then after t years Q becomes $Q\left(1+\frac{r}{n}\right)^{nt}$.
3) If n is large, Q compounds almost continuously. Continuous compounding is the limit as n goes to infinity: Qe^{rt}.
4) The instantaneous rate of change for continuous compounding is $\frac{d}{dt}Qe^{rt} = rQe^{rt}$.
5) So the relative (percentage) increase in Q at any time is $\frac{\frac{d}{dt}Qe^{rt}}{Qe^{rt}} = \frac{rQe^{rt}}{Qe^{rt}} = r$, i.e., the rate of increase is constant.
6) If the value of a quantity undergoing compounded growth over time t is described by the general formula Qe^{rt}, then we immediately know that the initial value was Q and the growth rate is r.

Ex: If the value at year t is $573e^{0.05t}$ then we know the starting amount was 573 and the growth rate is 5% per year.

Ex: If the initial deposit is 1274 and the continuous compounding rate is 7.8% per year, then how much will the account be worth in 5 years, if no additional deposits or withdrawals are made? Solution: $1274e^{0.078*5} = 1274e^{0.39} = 1881.67$

7) If r is small, discrete growth $Q(1+r)^t$ may be approximated by continuous growth $Q(e^r)^t = Qe^{rt}$ because $e^r \cong 1+r$.
8) All of this discussion applies to decreases ($r < 0$) as well as to increases ($r > 0$).
9) The only function that has a constant percentage growth rate is the exponential. That is, if y grows at a rate proportional to y, then $\frac{dy}{dt} = cy$, where c is a constant of proportionality. This is a differential equation. The objective of differential equation analysis is to find all functions y that satisfy the differential equation. It can be shown that the only equation that satisfies this differential equation is $y(t) = Ae^{ct} + B$, where A and B are constants that are usually determined by "initial conditions", such as the starting value at time t=0.

Examples of derivatives for exponential and logarithmic functions

Ex1: $y = e^{-4x}$
Ex2: $y = \ln(x^2)$
Ex3: $y = x^3 \ln(x^2)$
Ex4: $y = 12^{1-x}$
Ex5: $y = 2^x$ (second derivative)
Ex6: $y = \ln(3x^2)$ (second derivative)
Ex7: $y = \frac{x^2}{(x+3)(2x+1)}$ (use logs)
Ex8: $y = x^2 e^{4x-1}$ (use logs)

Ex9: y = x^x

Ex10: $y = 2e^{3x^2}\ln(4x^3)$

Ex11: Find $\lim_{x\to 0}(x\ln x)$ (use L'Hôpital's Rule)

Ex12: Find $\lim_{x\to 0}(1+x)^{1/x}$ (take log and use L'Hôpital's Rule)

Ex13: Find $\lim_{x\to 0} x^n e^{1/x}$ (use L'Hôpital's Rule repeatedly)

Ex14: Suppose that a vintner owns a supply of wine worth K now. If the vintner holds the wine for t years, it will be worth $Ke^{\sqrt{t}}$. The instantaneous interest rate is r. So the present value of the wine's t-years ahead value is A = $Ke^{\sqrt{t}}/e^{rt}$ = $Ke^{\sqrt{t}-rt}$. At what time should the vintner sell in order to maximize the present value? Take logs, then find critical points: $\ln A = \ln K + \sqrt{t} - rt$. $\frac{1}{A}\frac{dA}{dt} = \frac{1}{2\sqrt{t}} - r = 0$. So $t = \frac{1}{4r^2}$ is the critical point. Since $A > 0$, inspection of the derivative $\frac{dA}{dt} = A\left(\frac{1}{2\sqrt{t}} - r\right)$ shows that the derivative changes sign from positive to negative at $t = \frac{1}{4r^2}$, which is therefore a relative maximum by the first derivative test. (Also note that the rate of growth of the wine's value is $\frac{1}{2\sqrt{t}}$, which is a decreasing function of t. So the best selling time is when the wine value's growth rate declines to the level of the interest rate, at which point it is more profitable to invest in CD's.)

<u>Definition</u>: Suppose $y = f(t)$ is a function of time. The rate of growth of y is $\frac{f'(t)}{f(t)}$.

Note: $\frac{d}{dt}\ln f(t) = \frac{f'(t)}{f(t)}$ provides an alternative and often easier way to find the rate of growth.

Ex1: $V = Ae^{rt}$.

Ex2: $y = 4^t$.

Properties of rates of growth:

- $r_{(uv)} = r_u + r_v$ (The rate of growth of the product of two functions is the sum of their rates of growth.)
- $r_{(u/v)} = r_u - r_v$ (The rate of growth of the ratio of two functions is the difference of their rates of growth.)
- $r_{(u+v)} = \frac{u}{u+v}r_u + \frac{u}{u+v}r_v$ (The rate of growth of the sum of two functions is a weighted average of their rates of growth.)
- $r_{(u-v)} = \frac{u}{u-v}r_u - \frac{u}{u-v}r_v$ (The rate of growth of the difference of two functions is a weighted average of the difference of their rates of growth.)

The point elasticity of y = $f(x)$ can be expressed as $\varepsilon_{yx} = \frac{dy}{y}/\frac{dx}{x} = \frac{d(\ln y)}{d(\ln x)}$.

Ex: Q = k/P. ln Q = ln k – ln P. So d(ln Q) = - d(ln P). So $\frac{d(\ln Q)}{d(\ln P)} = -1$.

Section 8. Optimization of Functions of Two Variables

Sufficient conditions for relative extrema of *z = f(x,y)*:

1) If $f_x(x_0,y_0)=0$ and $f_y(x_0,y_0)=0$ then (x_0,y_0) is a critical point of *f(x,y)*.
2) If also,
   a) $f_{xx}(x_0,y_0)<0$, $f_{yy}(x_0,y_0)<0$, $f_{xx}(x_0,y_0)f_{yy}(x_0,y_0)-f_{xy}(x_0,y_0)^2>0$ then (x_0,y_0) is a relative maximum of *f(x,y)*.
   b) $f_{xx}(x_0,y_0)>0$, $f_{yy}(x_0,y_0)>0$, $f_{xx}(x_0,y_0)f_{yy}(x_0,y_0)-f_{xy}(x_0,y_0)^2>0$ then (x_0,y_0) is a relative minimum of *f(x,y)*.

Discussion:

1) For a function $z = f(x)$ of one variable, the second derivative test can be reformulated in terms of differentials as $dz = 0$, and $d^2z < 0$ (relative maximum) [or $d^2z > 0$ (relative minimum)]. This is because if x_0 satisfies $dz = 0$, and $d^2z < 0$, then since $dz = f'(x_0)dx$ and $d^2z = f''(x_0)dx^2$, we have $f'(x_0)=0$ and $f''(x_0)<0$. $d^2z < 0$ says that the function slopes away from $f(x_0)$ in both directions.
2) To have a maximum at (x_0,y_0) for a function $z = f(x,y)$ of two variables, we want the function to be flat in *__all__* directions at (x_0,y_0) and to slope away from (x_0,y_0) in *__all__* directions. The function will be flat in all directions if the total differential $dz = 0$. The function will slope away in all directions if the total differential decreases in all directions, i.e., $d(dz) = d^2z < 0$.
3) For a function of two variables, the first condition $dz = 0$ is $f_x(x_0,y_0)\,dx+f_y(x_0,y_0)\,dy=0$. This is true for all (small) *dx* and *dy* if and only if $f_x(x_0,y_0)=0$ and $f_y(x_0,y_0)=0$.
4) For a function of two variables, the second condition $d(dz) = d^2z < 0$ is $\frac{\partial}{\partial x}(dz)\,dx+\frac{\partial}{\partial y}(dz)\,dy<0$, or $\frac{\partial}{\partial x}\left(f_x\ dx+f_y\ dy\right)dx+\frac{\partial}{\partial y}\left(f_x\ dx+f_y\ dy\right)dy<0$. Since f_x and f_y are functions of *x* and/or *y*, and *dx* and *dy* are small constants, then this is $\left(f_{xx}\ dx+f_{xy}\ dy\right)dx+\left(f_{yx}\ dx+f_{yy}\ dy\right)dy = f_{xx}\ dx^2+f_{xy}\ dy\,dx+f_{yx}\ dx\,dy+f_{yy}\ dy^2 =$ $f_{xx}\ dx^2+2f_{xy}\ dy\,dx+f_{yy}\ dy^2<0$.
5) The second condition may be formulated as $\begin{bmatrix} dx & dy \end{bmatrix}\begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}\begin{bmatrix} dx \\ dy \end{bmatrix}<0$. This matrix expression is a **quadratic form** in the “variables” *dx* and *dy*.

6) A **quadratic form** in two variables u and v is an expression of the form $au^2 + buv + cv^2$, for some constants a, b, c. Note that $au^2 + buv + cv^2 = \begin{bmatrix} u & v \end{bmatrix} \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$. It is known that a quadratic form in two variables is always negative if and only if both of its eigenvalues are negative (and is always positive if and only if both eigenvalues are positive). It is also known that the eigenvalues are both negative if the first principal minor $|a| < 0$ and the second principal minor $\begin{vmatrix} a & b/2 \\ b/2 & c \end{vmatrix} = ac - b^2/4 > 0$. [Similarly, the quadratic form is always positive if the first principal minor $|a| > 0$ and the second principal minor $\begin{vmatrix} a & b/2 \\ b/2 & c \end{vmatrix} = ac - b^2/4 > 0$.]

7) So by (5) and (6), the second condition ($d^2z < 0$) for a relative maximum of a function of two variables is equivalent to $f_{xx} < 0$ and $f_{xx} f_{yy} - f_{xy}{}^2 > 0$. This also requires $f_{yy} < 0$ in order for $f_{xx} f_{yy} > 0$.

8) For a relative minimum, the second order condition $d^2z > 0$ is equivalent to $f_{xx} > 0$ and $f_{xx} f_{yy} - f_{xy}{}^2 > 0$, which therefore also requires $f_{yy} > 0$ in order for $f_{xx} f_{yy} > 0$.

9) $\begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix}$ is called a **Hessian determinant** (or simply **Hessian**).

10) This method can be extended to optimization of functions of three or more variables.

Ex1: $z = f(x, y) = 8x^3 + 2xy - 3x^2 + y^2 + 1$

Ex2: $z = f(x, y) = x + 2ey - e^x - e^{2y}$

Ex3: Suppose that the amount Q that is produced as a function of capital K and labor L is $Q(K, L) = -4K^2 - L^2 + 2KL + 100K + 80L + 1000$. Find the combination of capital and labor inputs that maximizes production.

Section 9. Optimization Subject to a Linear Constraint

Often we must optimize functions subject to constraints. For example, we may wish to choose capital and labor inputs to maximize production, but we may have a budget constraint that limits the amount of money we can spend on capital and labor. In this section of the Notes, I will show how to optimize a function of two variables $f(x,y)$ subject to a constraint of the form $g(x,y)=c$. The key to the solution is to replace the constrained problem by an unconstrained problem, for which the techniques of the previous section of these Notes can be applied. The cost of doing this is to increase the number of variables by one.

<u>Theorem</u>: Optimization of $z = f(x,y)$ subject to the constraint $g(x,y)=c$ is equivalent to unconstrained optimization of $Z = f(x,y)+\lambda[c-g(x,y)]$, assuming that $f(x,y)$ and $g(x,y)$ have continuous first and second partial derivatives.

Proof: I will prove the theorem for maximization (minimization is analogous). Here are the two problems:

Problem A: Maximize $z = f(x,y)$ subject to the constraint $g(x,y)=c$.

Problem B: Maximize $Z = f(x,y)+\lambda[c-g(x,y)]$ (no constraint).

I will show that the solutions for the two problems are the same. To do this, suppose that (x^*, y^*) is a solution for Problem A, and (x_0, y_0, λ_0) is a solution for Problem B. (Note that Problem B is a function of three variables, so its solution requires three values.)

Since (x^*, y^*) is a solution that provides the maximum for Problem A, then (x^*, y^*) satisfies the constraint $g(x^*, y^*)=c$ and

[†] $\quad f(x^*, y^*) \geq f(x,y)$ for all (x,y) that satisfy $g(x,y)=c$.

Since (x_0, y_0, λ_0) provides the maximum for Problem B, then $f(x_0,y_0)+\lambda_0[c-g(x_0,y_0)] \geq f(x,y)+\lambda[c-g(x,y)]$ for all (x,y,λ) - in particular, for (x^*, y^*, λ_0). That is,

[††] $\quad f(x_0,y_0)+\lambda_0[c-g(x_0,y_0)] \geq f(x^*,y^*)+\lambda_0[c-g(x^*,y^*)]$.

Also, since (x_0, y_0, λ_0) maximizes Problem B, then (x_0, y_0, λ_0) satisfies the first partial derivative condition (dZ = 0): $\left\{\begin{aligned} \frac{\partial Z}{\partial \lambda} &= c-g(x,y)=0 \\ \frac{\partial Z}{\partial x} &= f_x - \lambda\, g_x = 0 \\ \frac{\partial Z}{\partial y} &= f_y - \lambda\, g_y = 0 \end{aligned}\right\}$. From the first of these conditions, we see that (x_0, y_0) satisfies the constraint $g(x_0,y_0)=c$. By [†] this means that $f(x^*,y^*) \geq f(x_0,y_0)$. But by definition, (x^*, y^*) also satisfies the constraint $g(x^*,y^*)=c$. Thus, $\lambda_0[c-g(x_0,y_0)]=0$ and $\lambda_0[c-g(x^*,y^*)]=0$. Substituting these into [††], we have $f(x_0,y_0) \geq f(x^*,y^*)$. Putting together the two inequalities $f(x^*,y^*) \geq f(x_0,y_0)$ and $f(x_0,y_0) \geq f(x^*,y^*)$, we have $f(x^*,y^*) = f(x_0,y_0)$. This shows that the solutions for Problems A and B produce the same maximum value of $f(x,y)$ and must therefore be the same.

Discussion:

1) To set up the method, just add a multiple $\lambda[c-g(x,y)]$ of the constraint to the function $f(x,y)$ that you want to optimize.
2) The method becomes a little more understandable after you realize that the added amount $\lambda[c-g(x,y)]$ is zero for any (x,y) that is a solution. Adding zero does not change the value of the function. Since the added amount is zero, it does not matter whether you add it or subtract it.
3) The auxiliary variable λ is called a **Lagrange multiplier**.
4) The essence of this method for solving optimization problems subject to a linear constraint is to use a Lagrange multiplier to convert the constrained problem in two variables into an unconstrained problem in three variables.

Now let us consider the common special case in which the constraint is a linear function.

Theorem: Sufficient conditions for relative extrema of $z = f(x,y)$ subject to the linear constraint $g(x,y)=c$. Let $Z = f(x,y)+\lambda[c-g(x,y)]$.

1) (First-order condition) Suppose that (x_0, y_0, λ_0) satisfies $\begin{Bmatrix} \dfrac{\partial Z}{\partial \lambda} = c - g(x,y) = 0 \\ \dfrac{\partial Z}{\partial x} = f_x - \lambda\, g_x = 0 \\ \dfrac{\partial Z}{\partial y} = f_y - \lambda\, g_y = 0 \end{Bmatrix}$. (Then (x_0, y_0, λ_0) is a critical point.)
2) (Second-order condition) If also, the **bordered Hessian** $\begin{vmatrix} 0 & g_x & g_y \\ g_x & f_{xx} - \lambda\, g_{xx} & f_{xy} - \lambda\, g_{xy} \\ g_y & f_{xy} - \lambda\, g_{xy} & f_{yy} - \lambda\, g_{yy} \end{vmatrix} < 0$ then (x_0, y_0) is a relative minimum of $f(x,y)$ subject to the constraint. If the bordered Hessian > 0, then (x_0, y_0) is a relative maximum of $f(x,y)$ subject to the constraint.

Discussion:

1) For the special case that $g(x,y)$ is a linear function, the bordered Hessian $\begin{vmatrix} 0 & g_x & g_y \\ g_x & f_{xx} - \lambda\, g_{xx} & f_{xy} - \lambda\, g_{xy} \\ g_y & f_{xy} - \lambda\, g_{xy} & f_{yy} - \lambda\, g_{yy} \end{vmatrix} = \begin{vmatrix} Z_{\lambda\lambda} & Z_{\lambda x} & Z_{\lambda y} \\ Z_{\lambda x} & Z_{xx} & Z_{xy} \\ Z_{\lambda y} & Z_{xy} & Z_{yy} \end{vmatrix}$, and g_x and g_y will be constants equal to the coefficients of x and y in the linear function $g(x,y)$.
2) For the special case that $g(x,y)$ is a linear function, the second-order condition for optimization can be relaxed somewhat, compared with the general second-order condition for optimization given in the previous section of the Notes. The relaxed version is given in the above theorem.

Ex1: Find the largest rectangular area that can be enclosed within a fence if the budget for fencing is \$40,000 and the price of fencing is \$100 per meter. Let x = width, y = height. Then maximize area = A(x,y) = xy, subject to cost constraint 2x + 2y = 400. A(x, y, λ) = xy + λ[400 – 2x – 2y].

Ex2: Suppose that the amount Q that is produced as a function of capital K and labor L is $Q(K,L) = -4K^2 - L^2 + 2KL + 100K + 80L + 1000$. Suppose that the price of capital is \$1000 per unit, and the price of labor is \$50 per unit. The budget for production is \$30,000. Find the combination of capital and labor inputs that maximizes production subject to the budgetary constraint.

$Q(K,L,\lambda) = -4K^2 - L^2 + 2KL + 100K + 80L + 1000 + \lambda[30{,}000 - 1000K - 50L]$.

Section 10. Integration

Mathematically, integration is the inverse of differentiation in the same way that square root is the inverse of square. That is, if you differentiate a function and then integrate the result, you will recover the original function. Or vice-versa, in either order.

Ex: (square and square root) $9^2 = 81$, and $\sqrt{81} = 9$; also $\sqrt{9} = 3$ and $3^2 = 9$.

Ex: (derivative and integral) $\frac{d}{dx}\left(\frac{1}{12}x^2\right) = \frac{1}{6}x$ and $\int \frac{1}{6}x\,dx = \frac{1}{12}x^2$; also $\int \frac{1}{12}x^2\,dx = \frac{1}{36}x^3$ and $\frac{d}{dx}\left(\frac{1}{36}x^3\right) = \frac{1}{12}x^2$

This simple inverse relationship between derivatives and integrals is called the **Fundamental Theorem of Calculus**. (Anent which, more shortly).

The most important skill to learn in integration is how to integrate functions, just as the most important skill to learn in differentiation is how to differentiate functions. If you know how to differentiate most functions, then you should know how to integrate most functions – just reverse the process!

Ex: Rule #3 from the "Algebra of derivatives" [Section 3 of these Notes] is $\frac{d}{dx}[x^c] = cx^{c-1}$ where c is a constant. In words: "Multiply by the exponent, then reduce the exponent by 1." Then if you are required to integrate a function like $f(x) = \frac{1}{6}x$ that you recognize could be the end product of a differentiation, reverse the process: "Increase the exponent by 1, then divide by the exponent." Thus $\int f(x)\,dx = \frac{\frac{1}{6}x^{1+1}}{1+1} = \frac{1}{12}x^2$. You can check the correctness of this approach by differentiating: $\frac{d}{dx}\left(\frac{1}{12}x^2\right) = \frac{1}{6}x = f(x)$. You should get back the function that you were required to integrate.

There are actually two kinds of integral: the **indefinite integral** and the **definite integral**. The indefinite integral is a function; the definite integral is a number that results from evaluating the indefinite integral function at a particular point.

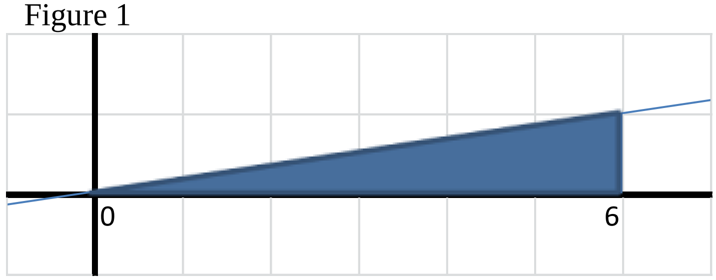
Ex: $\int \frac{1}{6}x\,dx = \frac{1}{12}x^2$ is a function. It is an indefinite integral. If we evaluate this function at x = 6, then we have a definite integral, which could be written $\int_0^6 \frac{1}{6}x\,dx = \frac{1}{12}6^2 = 3$.

There is an additional reason that the indefinite integral is called "indefinite." That is because the indefinite integral is not a unique function. For example, $\int \frac{1}{6}x\,dx = \frac{1}{12}x^2$ + 3 also works, since its derivative $\frac{d}{dx}\left(\frac{1}{12}x^2+3\right) = \frac{1}{6}x = f(x)$. In fact, $\int \frac{1}{6}x\,dx = \frac{1}{12}x^2$ + c works for any constant c since the derivative of a constant is zero. Fortunately, constants are the only exceptions to the uniqueness of indefinite integrals. So we can say that the indefinite integral is unique *up to a constant*.

The corresponding distinction is not made with derivatives. That is, there is no concept of indefinite derivative and definite derivative. If there were, then we would say that with $f(x) = \frac{1}{12}x^2$, then $f'(x) = \frac{1}{6}x$ is an indefinite derivative and $f'(6) = \frac{1}{6}6$ = 1 is a definite derivative. Furthermore, the derivative of a function is unique.

Let us examine the definite and indefinite integrals further. The integral of a function $f(x)$ calculates the area between the graph of the function and the *x*-axis from one point on the *x*-axis to another point. The area is a signed number – areas above the *x*-axis are positive and areas below the *x*-axis are negative. The following figure illustrates the situation:

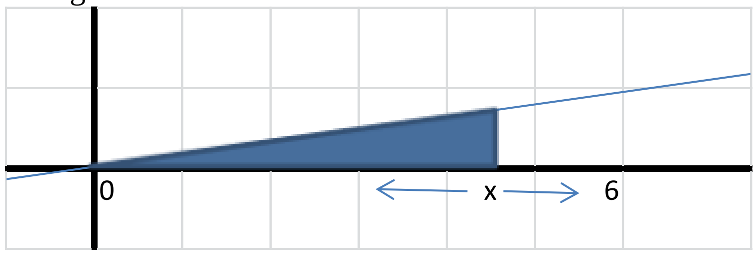
Figure 1



The graph of the function $f(x) = \frac{1}{6}x$ is shown, with the area between the function and the *x*-axis from 0 to 6 shaded in blue. By the properties of triangles, this area = ½ * base * height = ½ * 6 * 1 = 3. This result is also the value of a definite integral, obtained by evaluating the indefinite integral of $f(x) = \frac{1}{6}x$ at *x* = 6: $\int \frac{1}{6}x\,dx = \frac{1}{12}x^2$ and $\frac{1}{12}6^2$ = 3. The definite integral may be written in the same form as the indefinite integral, but with the low and high boundary points written in order at the bottom and top of the integral sign: $\int_0^6 \frac{1}{6}x\,dx = \frac{1}{12}6^2$ = 3.

If we change the upper boundary of the figure from 6 to some other number, then the area changes. The area becomes a function of the upper boundary. The situation is illustrated in Figure 2.

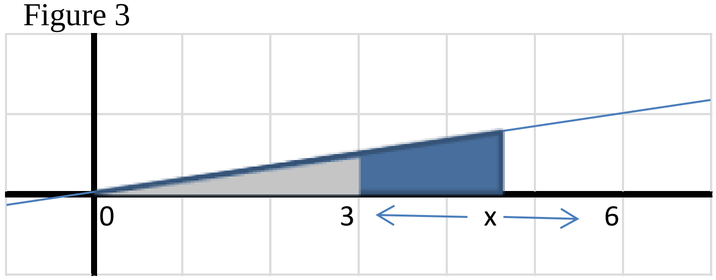
Figure 2



The area can be calculated for any $x > 0$ by calculating ½ * base * height = ½ * x * $\frac{1}{6}x$ = $\frac{1}{12}x^2$ or by plugging the value of x into the indefinite integral $\int \frac{1}{6}x\,dx = \frac{1}{12}x^2$. The definite integral version of the latter may be written $\int_0^x \frac{1}{6}x\,dx = \frac{1}{12}x^2$, or $\int_0^x \frac{1}{6}u\,du = \frac{1}{12}x^2$. Note that the meaning of $\int_0^x \frac{1}{6}u\,du = \frac{1}{12}x^2$ is "Find the indefinite integral of $\frac{1}{6}u$ (namely, $\frac{1}{12}u^2$), then substitute the upper boundary point x for the argument of the function (namely, u), to yield $\frac{1}{12}x^2$." Thus, it is immaterial which letter we use for the argument of the function being integrated, as long as we properly substitute the upper boundary point into the indefinite integral.

To find the area of the blue region in Figure 3 below, we can find the area of the triangle from 0 to x, then subtract the area of the gray triangle from 0 to 3 = $\frac{1}{12}x^2$ - $\frac{1}{12}3^2$ = $\frac{1}{12}x^2$ - 3/4. This can be written $\int_0^x \frac{1}{6}u\,du - \int_0^3 \frac{1}{6}u\,du = \int_3^x \frac{1}{6}u\,du$.

Figure 3



In general, if the area from 0 to x is $F(x)$, then the area from a to x, with $0 < a < x$, is $F(x) - F(a)$. To handle negative x, we define $\int_0^x \frac{1}{6}u\,du = -\int_x^0 \frac{1}{6}u\,du = -\frac{1}{12}x^2$ so that the two boundary points are listed in the integral sign in the proper order, from low to high. This means that we must also *interpret* $\int_6^3 \frac{1}{6}u\,du$ as $\int_0^3 \frac{1}{6}u\,du - \int_0^6 \frac{1}{6}u\,du = \frac{1}{12}3^2 - \frac{1}{12}6^2$ = -2.25, so that areas above the x-axis become negative when the integration runs from high end boundary point to low end boundary point, rather than in the normal order from low to high. Then the area from a to x is always $F(x) - F(a)$, regardless of the sign or order of a and x.

In general, this discussion about definite and indefinite integrals of $f(x)=\frac{1}{6}x$ extends to other functions: Suppose that $f(x)$ is a function that has an indefinite integral $F(x)$. Then the area between $f(x)$ and the *x*-axis from *a* to *b* on the *x*-axis is $F(b)-F(a)$. Moreover, for any *x*, $F(x)=\int_0^x f(u)\,du$ and $\int_a^b f(x)\,dx = F(b)-F(a)$. Implicit in these facts is the …

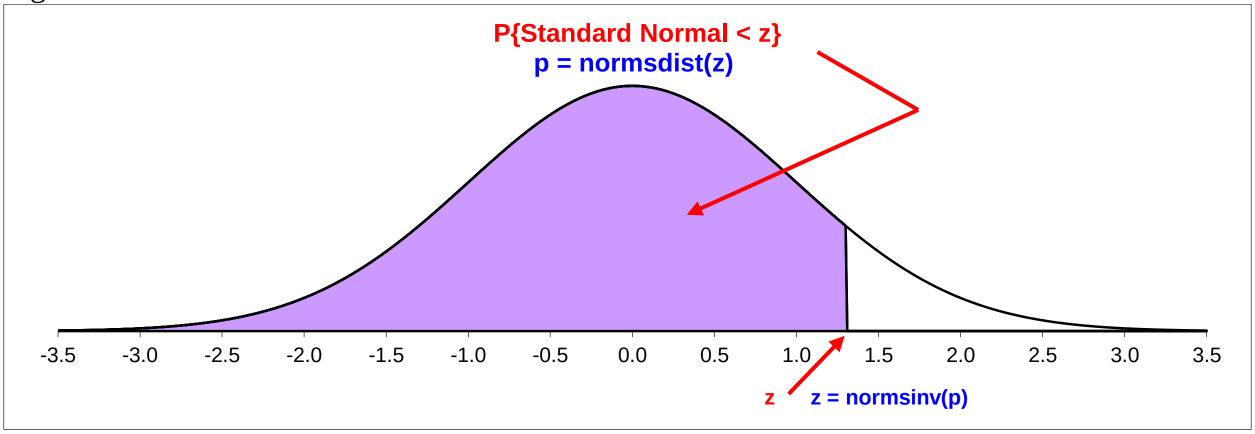
Fundamental Theorem of Calculus. Suppose that $f(x)$ is a function that has an indefinite integral $F(x)$. Then $F'(x)=f(x)$ and $F(x)=\int_0^x f(u)\,du$.

The Fundamental Theorem of Calculus establishes the connection between integration and differentiation of functions. It says that the derivative of the indefinite integral is the function and the integral of the function is the indefinite integral. The Fundamental Theorem may seem circularly tautologous in the context presented here. In my discussion I have tacitly assumed that the inverse of differentiation yields area. My discussion is intended to be correct, although not mathematically rigorous. In a rigorous context, the derivative would be defined as the rate of change of a function in a limit process (see Section 2) and the integral would be defined as the area under a curve as a limit process. In that context, proof of the Theorem requires some work.

Ex: The generation of normal probability tables illustrates a relevant application of the Fundamental Theorem. Figure 4 shows the curve of the standard normal probability density function: $f(z)=\frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}$. The purple region is the area between this curve and the *z*-axis from $-\infty$ to *z* = 1.282. Therefore, the purple region is the definite integral $\int_{-\infty}^{1.282}\frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}dz$, which equals 0.90. The purple region is the probability that a random variable that has a standard normal distribution is less than 1.282. This means that about 90% of the data values observed from a standard normal random variable should be less than 1.282. If the upper boundary point were changed from 1.282 to a different value, the probability would change. Because of the importance of the standard normal distribution in probability and statistics, there is considerable usefulness in having a function or algorithm to compute definite integrals of $f(z)$ and to compile tables of the values. This requires evaluating $\Phi(z)=\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{\frac{-u^2}{2}}du$ for any given value *z*. Although $\Phi'(z)=\frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}$, there is no closed-form elementary version of the indefinite integral $\Phi(z)$ – but approximating algorithms exist. One of them is implemented in the Excel function NORMSDIST(Z), illustrated in Figure 4.

The preceding example has an important generalization in probability theory. Every random variable (RV) has a probability density function (pdf). There are two major kinds of RV: discrete and continuous. For discrete RVs, probabilities are computed by summing values of the pdf. For continuous RVs, probabilities are computed by integrating the pdf. The indefinite integral of a pdf is called the **cumulative distribution function** (**CDF**). The CDF computes the probability less than a given value. For example, suppose that the uncertain lifetime of a light bulb (in hours) is represented by X. How can we find the probability that the light bulb will burn out before its warranteed lifespan of 1000 hours? Suppose that the pdf of X is $f(x) = \frac{1}{1000}e^{-x/1000}$, for $x > 0$. Then the probability that the light bulb will last less than 1,000 hours is $P(X < 1000)$ = $\int_0^{1000} \frac{1}{1000}e^{-x/1000}dx \; = \; -e^{-x/1000}\Big|_{x=0}^{x=1000} \; = 1 - e^{-1}$. In general, the probability that the light bulb will burn out before x hours is given by the CDF $F(x) = \int_0^x \frac{1}{1000}e^{-u/1000}du \; = \; -e^{-u/1000}\Big|_{u=0}^{u=x} \; = 1 - e^{-x/1000}$. Note that $F'(x) = f(x) = \frac{1}{1000}e^{-x/1000}$, as the Fundamental Theorem of Calculus asserts.

Figure 4



Section 11. The Algebra of Integrals

Each of the rules presented in Section 3 (Algebra of Derivatives) can be reversed to provide a useful rule for integration. Here are the nine rules from Section 3, as reversed for integration:

1) $\int 0\, dx = \int \frac{d}{dx}[c]\, dx = c$ where c is a constant.

2) $\int nx^{n-1}\, dx = \int \frac{d}{dx}[x^n]\, dx = x^n$ where n is an integer.

3) $\int cx^{c-1}\, dx = \int \frac{d}{dx}[x^c]\, dx = x^c$ where c is a constant. Special case: (c = 1) Then $\int 1\, dx = x$

4) $\int cf'(x)\, dx = \int \frac{d}{dx}[cf(x)]\, dx = cf(x)$

5) $\int \left[f'(x) + g'(x)\right] dx = \int \frac{d}{dx}\left[f(x) + g(x)\right] dx = f(x) + g(x)$

6) $\int \left[f'(x) - g'(x)\right] dx = \int \frac{d}{dx}\left[f(x) - g(x)\right] dx = f(x) - g(x)$

7) $\int \left[f(x)g'(x) + f'(x)g(x)\right] dx = \int \frac{d}{dx}\left[f(x) * g(x)\right] dx = f(x) * g(x)$

8) $\int \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}\, dx = \int \frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] dx = \frac{f(x)}{g(x)}$ if $g(x) \neq 0$.

9) $\int \frac{-g'(x)}{[g(x)]^2}\, dx = \int \frac{d}{dx}\left[\frac{1}{g(x)}\right] dx = \frac{1}{g(x)}$ if $g(x) \neq 0$.

Rule 7 is the inverse of the product rule for derivatives. It is usually written as $\int f(x)g'(x)\, dx = f(x) * g(x) - \int f'(x)g(x)\, dx$ and called **integration by parts**.

Ex: Find $\int x\log(x)\, dx$. Let $f(x) = \log(x)$ and $g'(x) = x$. Then $\int x\log(x)\, dx = \int f(x)g'(x)\, dx$ = (by Rule 7 – integration by parts) $\log(x)\frac{x^2}{2} - \int \frac{1}{x}\frac{x^2}{2} dx$ = $\log(x)\frac{x^2}{2} - \frac{x^2}{4}$.

Here is the Chain Rule from Section 3.
<u>The chain rule:</u> If g is differentiable at x and f is differentiable at $g(x)$, then $f \circ g$ is differentiable at x and $\frac{d}{dx}[(f \circ g)(x)] = f'(g(x)) \cdot g'(x)$.

The Chain Rule can be reversed to yield a useful integration rule:

$$\int f'(g(x)) \cdot g'(x)\, dx = \int \frac{d}{dx}[(f \circ g)(x)]\, dx = (f \circ g)(x)$$

Ex: Find $\int xe^{-x^2}dx$. Let $g(x)=-x^2$ and $f(x)=e^x$. Then $(f\circ g)(x)=e^{-x^2}$, so by the Chain Rule, $\frac{d}{dx}[(f\circ g)(x)]=f'(g(x))\cdot g'(x)$ = $e^{-x^2}(-2x)$. So by reversing the Chain Rule, $\int xe^{-x^2}dx$ = $-\frac{1}{2}\int -2xe^{-x^2}dx$ = $-\frac{1}{2}e^{-x^2}$. An alternative solution is to integrate by parts. (Can you see how?)

The reverse of the Chain Rule has another name: the **change of variable technique**. Here is how the preceding example would conventionally be presented as a change of variable:

Ex: Find $\int xe^{-x^2}dx$. Let the new variable $u=-x^2$. Then the differential is $du=-2x\,dx$. So $dx=-\frac{1}{2x}du$. [We recognize the change of variable to *u* as useful since the accompanying change of differential to *du* will "cancel out" the extra *x* in the original integrand – leaving a simple function e^u to integrate.] Substitute into $\int xe^{-x^2}dx$ = $\int xe^u(-\frac{1}{2x})du$ = $-\frac{1}{2}\int e^u du$ = $-\frac{1}{2}e^u$ = and substitute back for *u* to get $-\frac{1}{2}e^{-x^2}$.

The change of variable technique is easy to remember and to apply, as given in the preceding example. But what is really going on is that the integrand is being rewritten as a composite function of a function, in which the "new" variable "*u*" is *g*(*x*), the composite integrand is $f'(u)=f'(g(x))$, "*du*" is $g'(x)\,dx$, and the reverse of the Chain Rule is applied.

Here is the inverse function rule from Section 3.

__**The inverse function rule:**__ If $f(x)$ is invertible and differentiable, then f^{-1} is differentiable, and $\frac{d}{dy}[f^{-1}(y)]=\frac{1}{f'(f^{-1}(y))}$.

The inverse function can be reversed to yield a useful integration rule:

$$\int \frac{1}{f'(f^{-1}(y))}dy=\int \frac{d}{dy}[f^{-1}(y)]\,dy=f^{-1}(y).$$

Ex: Find $\int \frac{1}{5(x-5)^{4/5}}dx$. Observe that this integral can be written as $\int \frac{1}{5y^4}dx$ where $y=(x-5)^{1/5}$. So the inverse function is $f^{-1}(x)=(x-5)^{1/5}$, and the direct function is $f(y)=y^5+5$, and $f'(y)=5y^4$, which is the denominator of the integral. Thus, $\int \frac{1}{5(x-5)^{4/5}}dx$ = $\int \frac{1}{5y^4}dx$ = $\int \frac{1}{f'(f^{-1}(x))}dx$, which by the reverse of the inverse function rule is $f^{-1}(x)=(x-5)^{1/5}$.