# Putnam preparation

# Continuity, Derivatives and Integrals

We follow very closely (in many places we copy literally) other sources. I make no claim of originality.

**Limits of Functions.** For $x_0$ an accumulation point of the domain of a function $f$, we say that $\lim_{x\to x_0} f(x) = L$ if for every neighborhood $V$ of $L$, there is a neighborhood $U$ of $x_0$ such that $f(U) \subset V$. (You don't need to know what an accumulation point is to follow these notes.)

The usual presentation (when $x_0$ is a real number) of the above definition is as follows: we say that $\lim_{x\to x_0} f(x) = L$ if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that if $0 < |x - x_0| < \delta$, then $|f(x) - L| < \varepsilon$.

If $x_0 = +\infty$, then the definition of limit becomes: $\lim_{x\to+\infty} f(x) = L$ if for every $\varepsilon > 0$, there exists an $M > 0$, such that if $x > M$, then $|f(x) - L| < \varepsilon$.

And there are the obvious variations if $L$ is infinity, etc.

This definition is seldom used in applications. It is more customary to use operations with limits, the squeeze principle (if $f(x) \leq g(x) \leq h(x)$ for all $x$ and $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} h(x) = L$, then $\lim_{x\to x_0} g(x) = L$), continuity or L'Hospital's rule.

**Continuous functions.** A function is continuous at $x_0$ if it has a limit at $x_0$ and this limit is equal to $f(x_0)$. A function that is continuous at every point of its domain is simply called continuous.

It is not difficult to prove that $f$ is continuous at $x_0$ if and only if for every sequence $\{x_n\}$ converging to $x_0$, the sequence $\{f(x_n)\}$ converges to $f(x_0)$. In other words, $f$ is continuous at $x_0$ if and only if $\lim_{x\to x_0} f(x) = f(\lim_{x\to x_0} x)$, or in sequence language, $f$ is continuous at $x_0$ if and only if for every sequence $\{x_n\}$ converging to $x_0$, $\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n)$.

The sequential form for the definition of continuity of $f$ is used most often when one wishes to show a function is discontinuous at a point. E.g.

$$f(x) = \begin{cases} \sin\left(\frac{1}{x}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

is discontinuous at 0 because, for instance, the sequence $x_n = 2/(4n+1)\pi$ converges to 0, whereas $\{f(x_n)\} = \{\sin(2\pi n + \frac{1}{2}\pi)\}$ converges to 1 rather than to $f(0) = 0$.

Continuous functions can be pretty wild. E.g. a theorem of Peano asserts that there exists a continuous surjection $f : [0,1] \to [0,1] \times [0,1]$.

**The extreme and intermediate value properties.**

**Extreme Value Theorem.** If $f$ is a continuous function on $[a,b]$, then there are numbers $c$ and $d$ in $[a,b]$ such that $f(c) \leq f(x) \leq f(d)$ for all $x$ in $[a,b]$. (In other words, $f(d)$ is the maximum value for $f$ over $[a,b]$ and $f(c)$ is the minimum value for $f$ over $[a,b]$.)

**Intermediate Value Theorem.** If $f$ is a continuous function on $[a,b]$, and if $f(a) < y < f(b)$ (or $f(b) < y < f(a)$), then there is a number $c$ in $(a,b)$ such that $f(c) = y$.

This last theorem is easily remembered as "if you cross a river and there is no bridge, you get soaked in water."

If a real-valued function $f$ defined on an interval satisfies the conclusion of the intermediate value theorem, i.e. that whenever $a < b$ in the interval, and for every $y$ between $f(a)$ and $f(b)$, there exists $c$ between $a$ and $b$ such that $f(c) = y$, then $f$ is said to satisfy the intermediate value property. Equivalently, a real-valued function has the intermediate value property if it maps intervals to intervals. The higher-dimensional analogue requires the function to map connected sets to connected sets. Continuous functions satisfy both the one-dimensional case (the intermediate value theorem) and the higher-dimensional analogue. Derivatives of differentiable functions also satisfy the intermediate value property (even though they need not be continuous!).

We sketch a proof of the intermediate value theorem which uses the method of repeated bisection, which is useful in other cases.

So assume $f$ is a continuous function on the closed interval $[a, b]$, and suppose $f(a) < f(b)$ (a similar proof can be given if $f(a) > f(b)$.) Let $y \in [f(a), f(b)]$. We want to find an element $c \in [a, b]$ such that $f(c) = y$. The procedure goes as follows (a diagram will help.) Let $a_0 = a$, $b_0 = b$, and let $x_1$ denote the midpoint of the interval $[a, b]$ (the first bisection.) If $f(x_1) < y$, define $a_1 = x_1$, and $b_1 = b$, whereas if $y \leq f(x_1)$, define $a_1 = a$ and $b_1 = x_1$. In either case we have $f(a_1) \leq y \leq f(b_1)$, and the length of $[a_, b_1]$ is one-half the length of $[a, b]$.

Now let $x_2$ be the midpoint of $[a_1, b_1]$ (the second bisection.) If $f(x_2) < y$, define $a_2 = x_2$ and $b_2 = b_1$, and if $y \leq f(x_2)$, define $a_2 = a_1$ and $b_2 = x_2$. Again, it follows that $f(a_2) \leq y \leq f(b_2)$, and $b_2 - a_2 = (b - a)/4$.

Continue in this way. The result will be an infinite nested sequence of closed intervals

$$[a_0, b_0] \supset [a_1, b_1] \supset [a_2, b_2] \supset \cdots \supset \cdots$$

whose lengths converge to zero (in fact $b_i - a_i = (b - a)/2^i$.) These conditions imply that $\{a_i\}$ and $\{b_i\}$ each converge to the same real number in $[a, b]$: call this number $c$.

By continuity of $f$, $\lim_{i\to\infty} f(a_i) = f(c)$ and $\lim_{i\to\infty} f(b_i) = f(c)$. Furthermore, for each $i$, $f(a_i) \leq y \leq f(b_i)$, and therefore by the squeeze principle,

$$f(c) = \lim_{i\to\infty} f(a_i) \leq y \leq \lim_{i\to\infty} f(b_i) = f(c) \ .$$

It follows that $f(c) = y$ and the theorem is proved. $\square$

The following surprising theorem is due to Lebesgue.

**Theorem.** There exists a function $f : [0, 1] \to [0, 1]$ that has the intermediate value property and is discontinuous at every point.

**The derivative.** A function $f$ defined in an open interval containing the point $x_0$ is called differentiable at $x_0$ if

$$\lim_{h\to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists. In this case, the limit is called the derivative of $f$ at $x_0$ and is denoted by $f'(x_0)$ or $\frac{df}{dx}(x_0)$. If the derivative is defined at every point of the domain of $f$, then $f$ is called differentiable.

Note that if $f$ has a derivative at $x_0$, then it is continuous at $x_0$.

The derivative is the instantaneous rate of change. Geometrically, it is the slope of the tangent to the graph of the function. Because of this, where the derivative is positive, the function is increasing, where the derivative is negative, the function is decreasing, and on intervals where the derivative is zero, the function is constant. Moreover, the maxima and minima of a differentiable function on an open interval $(a, b)$, show up at points where the derivative is zero, the so-called critical points. Recall that, when checking maxima and minima of a function continuous defined in a closed interval $[a, b]$, you have to check the endpoints $a$ and $b$, the points where it is not differentiable, and among the points where it is differentiable (which are not the endpoints), the critical points.

Derivatives have an important application to the computation of limits:

**L'Hospital's rule.** For an open interval $I$, if the functions $f$ and $g$ are differentiable on $I \setminus \{x_0\}$, $g'(x) \neq 0$ for $x \in I$, $x \neq x_0$, and either $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = 0$ or $\lim_{x\to x_0} |f(x)| = \lim_{x\to x_0} |g(x)| = \infty$, and if additionally $\lim_{x\to x_0} \frac{f'(x)}{g'(x)}$ exists, then $\lim_{x\to x_0} \frac{f(x)}{g(x)}$ exists and

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f'(x)}{g'(x)} \ .$$

**Rolle's theorem.** If $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$, differentiable on $(a, b)$, and satisfies $f(a) = f(b)$, then there exists $c \in (a, b)$ such that $f'(c) = 0$.

**Cauchy's theorem.** If $f, g : [a, b] \to \mathbb{R}$ are two functions, continuous on $[a, b]$, differentiable on $(a, b)$, then there exists $c \in (a, b)$ such that

$$(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c) \ .$$

In the particular case that $g(x) = x$, we have the following

**The mean value theorem (Lagrange).** If $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on $(a, b)$, then there exists $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a} \ .$$

The mean value theorem has several important corollaries that are useful in practice, among which are:

**Corollaries:** Suppose $f$ and $g$ are continuous on $[a, b]$ and differentiable on $(a, b)$.

(a) If $f'(x) = 0$ for all $x$ in $(a, b)$, then $f$ is constant.

(b) If $f'(x) = g'(x)$ for all $x$ in $(a, b)$, then there exists a constant $C$ such that $f(x) = g(x) + C$.

(c) If $f'(x) > 0$ for all $x$ in $(a, b)$, then $f$ is an increasing function. Similarly, if $f'(x) < 0$ ($f'(x) \geq 0$, $f'(x) \leq 0$) for all $x$ in $(a, b)$, then $f$ is a decreasing (nondecreasing, nonincreasing, respectively) function on $(a, b)$.

A curious property of the derivative is that it satisfies the so-called Darboux property, namely: a differentiable function $f : [a, b] \to \mathbb{R}$ satisfies the conclusion of the intermediate value theorem. I.e. if $d$ is any number between $f'(a)$ and $f'(b)$, then there is a number $c$ in the interval $(a, b)$ such that $f'(c) = d$. Notice that we are not assuming that $f'$ is continuous!!

**Indefinite integrals.** Integration is the inverse operation to differentiation. The fundamental methods for computing integrals are the backward application of the chain rule, which takes the form

$$\int f(u(x))u'(x)\,dx = \int f(u)\,du$$

and shows up in the guise of substitutions, and integration by parts

$$\int u\,dv = uv - \int v\,du \ ,$$

which comes from the product rule for derivatives. Otherwise there is the partial fraction decomposition method for computing integrals of rational functions, and the standard substitutions (trigonometric, etc.)

**Partial fractions.** A brief review of partial fractions is as follows. Assume you have a quotient of polynomials

$$\frac{P(x)}{Q(x)} \ .$$

First of all, by the division algorithm, we can reduce to the case that degree($P(x)$) is strictly smaller than degree($Q(x)$). Factor $Q(x)$ into irreducible factors over $\mathbb{R}$, so that we have, e.g. $Q(x) = A(x-r)(x-s)^3(x^2+bx+c)(x^2+dx+e)^2$. Then

$$\frac{P(x)}{Q(x)} = \frac{C_1}{x-r} + \frac{C_2}{x-s} + \frac{C_3}{(x-s)^2} + \frac{C_4}{(x-s)^3} + \frac{C_5x+C_6}{x^2+bx+c} + \frac{C_7x+C_8}{x^2+dx+e} + \frac{C_9x+C_{10}}{(x^2+dx+e)^2} \ , \tag{1}$$

where all the $C_i$ are real numbers to be found. The general rule is that one writes $\frac{C_i}{x-r_i}$ for each simple root $r_i$, then for multiple roots $s_i$ (say of multiplicity $m$) one writes $\frac{C_j}{x-s_i} + \frac{C_{j+1}}{(x-s_i)^2} + \cdots + \frac{C_{j+m-1}}{(x-s_i)^m}$ (i.e. as many fractions as the multiplicity of the root, increasing the exponent in the denominator one by one, until you reach the multiplicity of the root.) For quadratic irreducible factors $(x^2+bx+c)$ one writes the corresponding fraction with the quadratic irreducible factor in the denominator, and a first degree polynomial in the numerator, i.e. $\frac{C_kx+C_{k+1}}{x^2+bx+c}$, and if the quadratic factor is repeated with multiplicity $n$, one writes $n$ fractions, each of which with first degree polynomials in the numerator, and the denominators are powers of the quadratic factor, with exponent increasing one at a time, from exponent 1 until the exponent agrees with the multiplicity of the factor. E.g. for $(x^2+dx+e)^2$ one would write as above $\frac{C_7x+C_8}{x^2+dx+e} + \frac{C_9x+C_{10}}{(x^2+dx+e)^2}$.

Now one has to find all the constants involved in the numerator. There are several tricks to do this. Typically in all of them you first clear denominators in the right-hand side of (1), and then you compare the numerators of the left-hand side and the right-hand side. Since they are two polynomials, they have to be equal for every value of $x$, and their coefficients of the same degree of $x$ also have to be equal. So one trick is to use that their coefficients of the same degree of $x$ have to be equal. Another trick is to substitute for convenient values of $x$ (typically 0 or 1 or the like), and a variation of this last trick is to substitute for the values of $x$ that annihilate $Q(x)$, i.e. the roots of $Q(x)$. A third trick is to take derivatives of the numerators (once the denominators have been cleared) and realize that they have to be equal, and apply the above tricks.

We copy one example from the notes on generating functions (which in turn is taken from Wilf's Generatingfunctionology):

$$A(x) = \frac{1-2x+2x^2}{(1-x)^2(1-2x)} \ .$$

Now we find the partial fraction expansion of $A(x)$. We know that

$$\frac{1-2x+2x^2}{(1-x)^2(1-2x)} = \frac{A}{(1-x)^2} + \frac{B}{1-x} + \frac{C}{(1-2x)} \ , \tag{2}$$

for some constants $A, B, C$ which we have to find.

In order to find $A, B, C$ we proceed as follows. Multiply both sides of (2) by $(1-x)^2$ and let $x = 1$. This gives $A = -1$. Now multiply both sides of (2) by $(1-2x)$ and let $x = 1/2$. This gives $C = 2$. Now choose an easy value of $x$, say $x = 0$, and substitute into (2). Since we know $A$ and $C$, we find that $B = 0$. Thus

$$A(x) = \frac{1-2x+2x^2}{(1-x)^2(1-2x)} = \frac{(-1)}{(1-x)^2} + \frac{2}{(1-2x)} \ . \tag{3}$$

In the case of definite integrals, the limits of integration also play a role. Also, we have

**The fundamental theorem of calculus.** If $F(t)$ has a continuous derivative on an interval $[a, b]$, then

$$\int_a^b F'(t)\ dt = F(b) - F(a) \ .$$

In other words, differentiation followed by integration recovers the function up to a constant, in the sense that

$$F(x) = \int_a^x F'(t)\ dt + C \ ,$$

where $C = F(a)$.

**The fundamental theorem for derivatives of integrals** states that if $f$ is a continuous function in an interval $[a, b]$, then for any $x$ in $(a, b)$,

$$\frac{d}{dx}\int_a^x f(t)\ dt = f(x) \ .$$

In other words, integration followed by differentiation recovers the function exactly.

The definite integral of a function is the area under the graph of the function. In approximating the area under the graph by a family of rectangles, the sum of the areas of the rectangles, called a Riemann sum, approximates the integral. When these rectangles have equal width, the approximation of the integral by Riemann sums reads

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

where each $\xi_i$ is a number in the interval $[a + \frac{i-1}{n}(b-a), a + \frac{i}{n}(b-a)]$.

Since the Riemann sum depends on the positive integer $n$, it can be thought of as the term of a sequence. Sometimes the terms of a sequence can be recognized as the Riemann sums of a function, and this can prove helpful for finding the limit of the sequence.

**Example.** Find

$$\lim_{n\to\infty}\left(\frac{1}{n}+\frac{1}{n+1}+\cdots+\frac{1}{2n-1}\right) .$$

One way to think about this is to interpret the sum geometrically: construct rectangles on $[n,2n]$ in the following way (draw a picture.) The first rectangle has base $[n,n+1]$ and height $\frac{1}{n}$ (i.e. the value of the positive decreasing function $f(t)=\frac{1}{t}$ at the left-endpoint of the interval $[n,n+1]$.) The second rectangle has base $[n+1,n+2]$ and height $\frac{1}{n+1}$ (i.e. the value of the positive decreasing function $f(t)=\frac{1}{t}$ at the left-endpoint of the interval $[n+1,n+2]$.) And so on, until the last rectangle (the $n$th one), which has base $[2n-1,2n]$ and height $\frac{1}{2n-1}$ (i.e. the value of the positive decreasing function $f(t)=\frac{1}{t}$ at the left-endpoint of the interval $[2n-1,2n]$.) The sum of areas of these rectangles is obviously larger than the area under $f(t)=\frac{1}{t}$ in $[n,2n]$ so that

$$\frac{1}{n}+\frac{1}{n+1}+\cdots+\frac{1}{2n-1} > \int_n^{2n}\frac{1}{t}\,dt = \ln t\big|_n^{2n} = \ln(2n)-\ln(n)=\ln(2) .$$

Similarly, if one builds the rectangles with the same base, but with height the right-endpoint of the interval which is the base of the rectangle, it follows that

$$\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{2n} < \int_n^{2n}\frac{1}{t}\,dt = \ln(2) .$$

As a consequence,

$$\ln(2) < \frac{1}{n}+\frac{1}{n+1}+\cdots+\frac{1}{2n-1} < \left(\frac{1}{n}-\frac{1}{2n}\right)+\ln(2) .$$

Now, as $n\to\infty$, it is apparent that

$$\lim_{n\to\infty}\left(\frac{1}{n}+\frac{1}{n+1}+\cdots+\frac{1}{2n-1}\right)=\ln(2) .$$

What we have done can be repeated for any positive decreasing function $f(t)$ defined on $[0,+\infty)$, and is the formal statement that gives the precise inequalities for the well-known expression

$$\sum_{n=1}^{N} f(n) \approx \int_1^N f(t)\,dt .$$

There is another way to find this limit which is also instructive. Rewrite the sum as

$$\sum_{k=0}^{n-1}\frac{1}{n+k}=\sum_{k=0}^{n-1}\left(\frac{1}{1+\frac{k}{n}}\right)\frac{1}{n} ,$$

which we recognize as a Riemann sum of the function $f : [0,1] \to \mathbb{R}$, $f(x) = \frac{1}{1+x}$ associated to the subdivision $x_0 = 0 < x_1 = \frac{1}{n} < x_2 = \frac{2}{n} < \cdots < x_n = \frac{n}{n} = 1$, with the intermediate points $\xi_i = \frac{i}{n} \in [x_i, x_{i+1}]$. It follows that

$$\lim_{n\to\infty} \left( \frac{1}{n} + \frac{1}{n+1} + \cdots + \frac{1}{2n-1} \right) = \int_0^1 \frac{1}{1+x}\, dx = \ln(1+x)\Big|_0^1 = \ln(2) \ .$$

A useful fact about integrals is recorded in Exercise 11(a).

**Inequalities for integrals.**

A very simple one (to state, not always simple to apply) is the following: if $f : [a,b] \to \mathbb{R}$ is a nonnegative continuous function, then

$$\int_a^b f(x)\, dx \geq 0 \ ,$$

with equality if and only if $f$ is identically equal to zero.

Some other fundamental inequalities are as follows. We will be imprecise as to the classes of functions to which they apply, in order to avoid the subtleties of Lebesgue's theory of integration. The novice mathematician should think of piecewise continuous, real-valued functions on some domain $D$ that is an interval of the real axis or some region in $\mathbb{R}^n$.

**The Cauchy-Schwarz inequality.** Let $f$ and $g$ be square integrable functions. Then

$$\left( \int_D f(x) g(x)\, dx \right)^2 \leq \left( \int_D f^2(x)\, dx \right) \left( \int_D g^2(x)\, dx \right) \ .$$

**Minkowski's inequality.** If $p > 1$, then

$$\left( \int_D |f(x) + g(x)|^p\, dx \right)^{\frac{1}{p}} \leq \left( \int_D |f(x)|^p\, dx \right)^{\frac{1}{p}} + \left( \int_D |g(x)|^p\, dx \right)^{\frac{1}{p}} \ .$$

**Hölder's inequality.** If $p, q > 1$, such that $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\int_D |f(x) g(x)|\, dx \leq \left( \int_D |f(x)|^p\, dx \right)^{\frac{1}{p}} \left( \int_D |g(x)|^q\, dx \right)^{\frac{1}{q}} \ .$$

**Chebyshev's inequality.** Let $f$ and $g$ be two increasing functions on $\mathbb{R}$. Then for any real numbers $a < b$,

$$(b-a) \int_a^b f(x) g(x)\, dx \geq \left( \int_a^b f(x)\, dx \right) \left( \int_a^b g(x)\, dx \right) \ .$$

**Taylor series.** Some functions, called analytic, can be expanded around each point of their domain in a Taylor series

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \ldots$$

If $a = 0$, the expansion is also known as the Maclaurin series. Rational functions, trigonometric functions, the exponential and the natural logarithm are examples of analytic functions. A sum, difference, multiplication, division and composition of analytic functions is analytic (in the case of division, away from the zeros of the denominator), and usually you can obtain the Taylor series of the composition by composing the corresponding Taylor series.

A particular example of a Taylor series expansion is the Newton's binomial formula

$$(x+1)^a = \sum_{n=0}^{\infty} \binom{a}{n} x^n = \sum_{n=0}^{\infty} \frac{a(a-1)\cdots(a-n+1)}{n!} x^n ,$$

which holds for all real numbers $a$ and for $|x| < 1$. Note the usual convention that $\binom{a}{0} = 1$.

In the notes about generating functions, there are a few examples of typical Taylor series, which I advise you to memorize. Recall that if a power series

$$f(z) = \sum_{n=0}^{\infty} a_n z^n$$

has radius of convergence $R > 0$ (see the notes about generating functions or below for the definition), then, for $|z| < R$,

$$f'(z) = \sum_{n=0}^{\infty} n a_n z^{n-1}$$

(i.e. in order to take derivatives, we can take the derivative in each term of the sum), and

$$\int_0^x f(t)\, dt = \sum_{n=0}^{\infty} a_n \int_0^x t^n\, dt ,$$

(i.e. in order to take an integral, we can take the integral in each term of the sum.)

**Power series: a quick reminder of the analytic theory.** (Taken from the notes on generating functions.)

Suppose we are given a power series

$$f(z) = \sum_{n\geq 0} a_n z^n ,$$

where we now use the letter $z$ to encourage thinking about complex numbers. For what set of complex values of $z$ does the series $f$ converge?

A quick idea of convergence is that, say, $\sum_{n=1}^{\infty} a_n z^n$ converges for, say, $z = 2$, if, when plugging in $z = 2$, the corresponding numerical series $\sum_{n=1}^{\infty} a_n 2^n$ converges. This in turn means that the limit $\lim_{N\to\infty} \sum_{n=1}^{N} a_n 2^n$ exists, in which case that limit is denoted precisely as $\sum_{n=1}^{\infty} a_n 2^n$.

Let us note in passing that the case of infinite products is similar: $\prod_{j\geq 0}^{\infty} a_j$ means $\lim_{N\to\infty} \prod_{j\geq 0}^{N} a_j = a_1 \cdot a_2 \cdots a_N$, if such limit exists.

**Theorem.** There exists a number $R$, $0 \leq R \leq +\infty$, called the radius of convergence of the series $f$, such that the series converges for all values of $z$ with $|z| < R$ and diverges for all $z$ such that $|z| > R$. The number $R$ satisfies

$$R = \frac{1}{\limsup_{n\to\infty} |a_n|^{\frac{1}{n}}} \qquad (\frac{1}{0} = \infty; \frac{1}{\infty} = 0).$$

Let's recall the definition of *limit superior* of a sequence. Let $\{x_n\}$ be a sequence of real numbers, and let $L$ be a real number (possibly $\pm\infty$.)

**Definition.** We say that $L$ is the limit superior (or upper limit) of the sequence $\{x_n\}$ (and we write $L = \limsup_{n\to\infty}\{x_n\}$ or $L = \limsup\{x_n\}$) if

(a) $L$ is finite and

  (i) for every $\varepsilon > 0$ all but finitely many members of the sequence satisfy $x_n < L + \varepsilon$, and

  (ii) for every $\varepsilon > 0$, infinitely many members of the sequence satisfy $x_n > L - \varepsilon$, or

(b) $L = +\infty$ and for every $M > 0$, there is an $n$ such that $x_n > M$, or

(c) $L = -\infty$ and for every $M > 0$, there are only finitely many $n$ such that $x_n > -M$.

The limit superior has the following properties:

• Every sequence of real numbers has one and only one limit superior in the extended real number system (i.e. $\mathbb{R}$ and the two points $\pm\infty$.)

• If a sequence has a limit $L$, then $L$ is also the limit superior of the sequence.

• If $S$ is the set of cluster points of $\{x_n\}_{n=0}^{\infty}$, then $\limsup\{x_n\}$ is the least upper bound of the numbers in $S$.

**Theorem.** Suppose the power series $\sum a_n z^n$ converges for all $z$ in $|z| < R$, and let $f(z)$ denote its sum. Then $f(z)$ is an analytic function in $|z| < R$. If furthermore the series diverges for $|z| > R$, then the function $f(z)$ must have at least one singularity on the circle of convergence $|z| = R$.

In other words, a power series keeps on converging until something stops it, namely a singularity of the function that is being represented.

**Example.** The series $\sum z^n$ converges if $|z| < 1$ and diverges if $|z| > 1$. Hence the function that is represented must have a singularity somewhere on the circle $|z| = 1$. That function is $\frac{1}{1-z}$, and it has a singularity at $z = 1$.

**Example.** Let $f(z) = \frac{1}{2-e^z}$. What is its radius of convergence? According to the theorem, the series will converge in the largest disk $|z| < R$ in which $f$ is analytic. The function $f$ is not analytic only at points $z$ where $e^z = 2$, i.e. points of the form $z = \log 2 + 2k\pi i$, for integer $k$, and the nearest one to the origin is $\log 2$. So the radius of convergence is $R = \log 2$.

**Abel's limit theorem.** Let $r > 0$ and suppose that $\sum_{n=0}^{\infty} a_n r^n$ converges. Then $\sum_{n=0}^{\infty} a_n x^n$ converges absolutely for $|x| < r$ and

$$\lim_{x \to r^-} \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n r^n \ .$$

**Example.** Sum the infinite series

$$1 - \frac{1}{4} + \frac{1}{7} - \frac{1}{10} + \dots$$

We know that

$$\frac{1}{1+x^3} = 1 - x^3 + x^6 - x^9 + \dots, \qquad |x| < 1,$$

and thus

$$\int_0^x \frac{dt}{1+t^3} = x - \frac{x^4}{4} + \frac{x^7}{7} - \frac{x^{10}}{10} + \dots, \qquad |x| < 1.$$

Note that the series on the right side converges for $x = 1$ (by the alternating series test), and, therefore, by Abel's limit theorem,

$$1 - \frac{1}{4} + \frac{1}{7} - \frac{1}{10} + \dots = \lim_{x \to 1^-} \int_0^x \frac{dt}{1+t^3}.$$

Working this integral by partial fractions (and omitting the details), we get

$$\int_0^x \frac{dt}{1+t^3} = \frac{1}{3}\left[\log\left(\frac{1+x}{\sqrt{1-x+x^2}}\right) + \sqrt{3}\left[\arctan\left(\frac{2x-1}{\sqrt{3}}\right) - \arctan\left(\frac{-1}{\sqrt{3}}\right)\right]\right].$$

Thus the series sums to

$$\frac{1}{3}\left[\log(2) + \frac{\pi}{\sqrt{3}}\right].$$

Going back to our main topic, recall also

**Stirling's formula.**

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \cdot e^{\frac{\theta_n}{12n}} \ \text{ for some } 0 < \theta_n < 1 \ .$$

**Fourier series.** The Fourier series allows us to write an arbitrary oscillation as a superposition of sinusoidal oscillations. Mathematically, a function $f : \mathbb{R} \to \mathbb{R}$ that is continuous and periodic of period $T$ admits a Fourier series expansion

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos\left(\frac{2n\pi}{T} x\right) + \sum_{n=1}^{\infty} b_n \sin\left(\frac{2n\pi}{T} x\right) \ . \tag{4}$$

This expansion is unique, and

$$a_0 = \frac{1}{T}\int_0^T f(x)\ dx\ ,$$

$$a_n = \frac{2}{T}\int_0^T f(x)\cos\left(\frac{2n\pi}{T}x\right)\ dx\ ,$$

$$b_n = \frac{2}{T}\int_0^T f(x)\sin\left(\frac{2n\pi}{T}x\right)\ dx\ .$$

Of course, we can require $f$ to be defined only on an interval of length $T$, and then extend it periodically, but if the values of $f$ at the endpoints of the interval differ, then the convergence of the series is guaranteed only in the interior of the interval.

**Parseval's identity.** For a general Fourier expansion as in (4), one has

$$\frac{1}{T}\int_0^T |f(x)|^2\ dx = a_0^2 + \frac{1}{2}\sum_{n=1}^{\infty}(a_n^2 + b_n^2)\ .$$

**Finding maxima and minima of functions of several variables** (This section is taken literally from Thomas' Calculus.) The extreme values of $f(x, y)$ can occur only at

(i) boundary points of the domain of f,

(ii) critical points (interior points where $f_x = f_y = 0$ or points where $f_x$ or $f_y$ fails to exist).

If the first- and second-order partial derivatives of $f$ are continuous throughout a disk centered at a point $(a, b)$ and $f_x(a, b) = f_y(a, b) = 0$, the nature of $f(a, b)$ can be tested with the second derivative test:

(i) If $f_{xx} < 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at $(a, b)$, then $f$ has a local maximum at $(a, b)$.

(ii) If $f_{xx} > 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at $(a, b)$, then $f$ has a local minimum at $(a, b)$.

(iii) If $f_{xx}f_{yy} - f_{xy}^2 < 0$ at $(a, b)$, then $f$ has a saddle point at $(a, b)$. (E.g. $f(x, y) = x^2 - y^2$ at the point $(a, b) = (0, 0)$.)

(iv) If $f_{xx}f_{yy} - f_{xy}^2 = 0$ at $(a, b)$, then the test is inconclusive.

**Finding maxima and minima of functions of several variables with constraints: Lagrange multipliers** (This section is taken literally from Thomas' Calculus).

Suppose that $f(x, y, z)$ and $g(x, y, z)$ are differentiable and $\nabla g \neq \mathbf{0}$ when $g(x, y, z) = 0$. To find the local maximum and minimum values of $f$ subject to the constraint $g(x, y, z) = 0$ (if these exist), find the values of $x$, $y$, $z$, and $\lambda$ that simultaneously satisfy the equations

$$\nabla f = \lambda \nabla g \qquad \text{and} \qquad g(x, y, z) = 0\ .$$

For functions of two independent variables, the condition is similar, but without the variable $z$. The number $\lambda$ is the so-called Lagrange multiplier.

If you have two constraints, say $g_1(x,y,z)=0$ and $g_2(x,y,z)=0$, and $g_1$ and $g_2$ are differentiable, with $\nabla g_1$ not parallel to $\nabla g_2$, we find the constrained local maxima and minima of $f$ by introducing two Lagrange multipliers $\lambda$ and $\mu$. I.e. we locate the points $P(x,y,z)$ where $f$ takes on its constrained extreme values (maxima and minima) by finding the values of $x$, $y$, $z$, $\lambda$, and $\mu$ that simultaneously satisfy the equations

$$\nabla f = \lambda \nabla g_1 + \mu \nabla g_2 \ , \qquad g_1(x,y,z)=0 \ , \qquad g_2(x,y,z)=0 \ .$$

**Example.** The plane $x+y+z=1$ cuts the cylinder $x^2+y^2=1$ in an ellipse. Figure the points on the ellipse closest and farthest from the origin.

We find the extreme values of

$$f(x,y,z)=x^2+y^2+z^2$$

(the square of the distance from $(x,y,z)$ to the origin) subject to the constraints

$$g_1(x,y,z)=x^2+y^2-1=0 \ , \tag{5}$$

$$g_2(x,y,z)=x+y+z-1=0 \ . \tag{6}$$

The gradient equation $\nabla f = \lambda \nabla g_1 + \mu \nabla g_2$ gives

$$(2x,2y,2z)=(2\lambda x+\mu, 2\lambda y+\mu, \mu) \ ,$$

or

$$2x=2\lambda x+\mu, \qquad 2y=2\lambda y+\mu, \qquad 2z=\mu \ . \tag{7}$$

As a consequence,

$$2x=2\lambda x+2z \Longrightarrow (1-\lambda)x=z \ ,$$

$$2y=2\lambda y+2z \Longrightarrow (1-\lambda)y=z \ .$$

These last equations are satisfied simultaneously if either $\lambda=1$ and $z=0$ or $\lambda\neq 1$ and $x=y=\frac{z}{(1-\lambda)}$.

If $z=0$, then solving equations (5) and (6) simultaneously to find the corresponding points on the ellipse gives the two points $(1,0,0)$ and $(0,1,0)$.

If $x=y$, then equations (5) and (6) give

$$x^2+x^2-1=0 \Longrightarrow x=\pm\frac{\sqrt{2}}{2}$$

and

$$x+x+z-1=0 \Longrightarrow z=1\mp\sqrt{2} \ .$$

The corresponding points on the ellipse are

$$P_1=\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2},1-\sqrt{2}\right) \qquad \text{and} \qquad P_2=\left(-\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2},1+\sqrt{2}\right) \ .$$

Although $P_1$ and $P_2$ both give local maxima of $f$ on the ellipse, $P_2$ is farther from the origin than $P_1$.

The points on the ellipse closest to the origin are $(1,0,0)$ and $(0,1,0)$. The point on the ellipse farthest from the origin is $P_2$.

Disclaimer: Most (if not all) the material here is from outside sources, I am not claiming any originality.