Analysis Concepts

Math 425A

October 26, 2006

1 Supremum and infimum

Completeness property of the real numbers:

Let S be a non-empty set of real numbers that is bounded above. Then S has a supremum (least upper bound).

Let S be a non-empty set of real numbers that is bounded below. Then S has an infimum (greatest lower bound).

2 Sequences

A real sequence $a_n, n \in \mathbf{N}$ *converges* to a limit a, that is, $\lim_{n\to\infty} a_n = a$, provided $\forall\epsilon > 0 \exists N \forall n \geq N |a_n - a| < \epsilon$.

A real sequence $a_n, n \in \mathbf{N}$ is *monotone increasing* proved that for each m, n with $m \leq n$ we have $a_m \leq a_n$.

A real sequence $a_n, n \in \mathbf{N}$ is *monotone decreasing* proved that for each m, n with $m \leq n$ we have $a_m \geq a_n$.

Theorem. Suppose $a_n, n \in \mathbf{N}$ is a monotone increasing real sequence whose values are bounded above. Then it converges to the supremum of these values.

Theorem. Suppose $a_n, n \in \mathbf{N}$ is a monotone decreasing real sequence whose values are bounded below. Then it converges to the infimum of these values.

Theorem. Suppose $a_n, n \in \mathbf{N}$ is a real sequence whose values are bounded (above and below). Then it has a subsequence $a_{n_j}, j \in \mathbf{N}$ that converges.

Proof: Say that n is a peak index if for each k with $k > n$ we have $a_k \leq a_n$.

If there are infinitely many peak indices, then we may take n_j to be the jth peak index. Then a_{n_j} is a monotone decreasing sequence that is bounded below, so it converges.

If there are finitely many peak indices, then there is a largest one p. Then for each $n > p$ there is a k with $k > n$ and $a_k > a_n$. Take $n_1 > p$ and construct the remaining indices so that for each n_j we have $n_{j+1} > n_j$ and $a_{n_{j+1}} > a_{n_j}$. Then the a_{n_j} is a monotone increasing sequence that is bounded above, so it converges.

3 Closed sets

A set S of real numbers is *closed* provided that whenever $a_n, n \in \mathbf{N}$ is a sequence with each a_n in S that converges to a limit a, then $a \in S$.

Remark: If a set S is non-empty and bounded, then it has a supremum and and infimum. If the set S is non-empty and closed and bounded, then it has a maximum and an minimum.

4 Sequentially compact sets

A set S of real numbers is *sequentially compact* provided that whenever $a_n, n \in \mathbf{N}$ is a real sequence with each a_n in S, then it has a subsequence $a_{n_j}, j \in \mathbf{N}$ that converges to a limit a in S.

Theorem. If a set S of real numbers is sequentially compact, then it is closed and bounded.

The proof is easy. One shows that if the S is not closed, then it is not sequentially compact. Then one shows that if the set is not bounded, then it is not sequentially compact.

Theorem. If a set S of real numbers is closed and bounded, then it is sequentially compact.

Given the convergent subsequence theorem, the proof is easy. Since the set S is bounded, if there is a sequence with values in S, then it has a convergent subsequence. Since the set S is closed, then limit of this convergent subsequence must be in S.

5 Continuous functions and sequential compactness

A real function f on a set D of real numbers is *continuous* on D if whenever $a_n, n \in \mathbf{N}$ is a sequence of numbers in D that converges to some a in D, then the sequence $f(a_n), n \in \mathbf{N}$ converges to $f(a)$.

Theorem: If f is continuous and D is sequentially compact, then the image of D under f is sequentially compact.

Proof: Suppose that b_n is a sequence of numbers in the image of D under f. Then there is a sequence a_n of numbers in D with $f(a_n) = b_n$. Since D is sequentially compact, a_n has a subsequence a_{n_j} that converges to some a in D. Since f is continuous, it follows that $b_{n_j} = f(a_{n_j})$ converges to $f(a)$ as $j \to \infty$, and $f(a)$ is in the image of D under f. This proves that the image of D under f is sequentially compact.

Corollary: If f is continuous and D is closed and bounded and non-empty, then the image of D under f is closed and bounded and non-empty, and so the image of D under f has a maximum and an minimum.

6 Strictly increasing continuous functions and fixed point iteration

Theorem. Suppose that g is a strictly increasing continuous function on some open interval. Suppose that g has a fixed point with $g(r) = r$. Suppose that $x < r$ is equivalent to $g(x) > x$ and $x > r$ is equivalent to $g(x) < x$. Let x_1 be in the interval. Define $x_{n+1} = g(x_n)$. Then $x_n \to r$ as $n \to \infty$.

Proof: Suppose first that $x_1 < r$. Then $g(x_1) > x_1$. It is not difficult to prove by induction that $g(x_n) > x_n$. This is because $g(x_n) > x_n$ implies $x_{n+1} = g(x_n) > x_n$. Since g is strictly increasing, we have $g(x_{n+1}) > g(x_n) = x_{n+1}$. We see in this case that the sequence is strictly increasing and bounded above by r. Hence it converges to some t.

Instead suppose that $x_1 > r$. Then $g(x_1) < x_1$. It is not difficult to prove by induction that $g(x_n) < x_n$. This is because $g(x_n) < x_n$ implies $x_{n+1} = g(x_n) < x_n$. Since g is strictly increasing, we have $g(x_{n+1}) < g(x_n) = x_{n+1}$. We see in this case that the sequence is strictly decreasing and bounded below by r. Hence it converges to some t.

We have shown that for some t in the interval we have that $x_n \to t$ as $n \to \infty$. It follows that $x_{n+1} \to t$. Also, by continuity, $g(x_n) \to g(t)$. Since $x_{n+1} = g(x_n)$, we have $t = g(t)$. Thus $t = r$.

Note: This theorem gives a practical way of finding solutions of equations. For instance, say that we want to solve the equation $x = \ln(x)+2$ for $x > 1$. Let $g(x) = \ln(x) + 2$. If one starts the iteration of g at 2, one gets 2.00000, 2.69314, 2.99071, 3.09551, 3.12995, 3.14101, 3.14454, 3.14566, 3.14602, 3.14614, 3.14617, 3.14618, 3.14619, and so on. This is strictly increasing, and it converges to the solution. In particular, each value is a lower bound for the solution. If one instead starts at 4, then the result is 4.00000, 3.38629, 3.21973, 3.16929, 3.15351, 3.14851, 3.14693, 3.14642, 3.14626, 3.14621, and so on. This is a decreasing sequence that converges to the solution. Each value is an upper bound for the solution.

Graphical representation of iteration. Graph $y = g(x)$. Graph the diagonal $y = x$. The point where they intersect is the fixed point of g. The iteration $x_{n+1} = g(x_n)$ has a nice graphical representation. Pick a point (x_1, x_1) on the diagonal. Then repeat the following process. Start with (x_n, x_n) on the diagonal, draw a vertical line to (x_n, x_{n+1}) on the graph, then draw a horizontal line to (x_{n+1}, x_{n+1}) on the diagonal. Mark the resulting points on the diagonal; they represent the sequence that results from the iteration.

7 Continuous functions and the intermediate value theorem

Theorem. Say that f is a real continuous function on the closed bounded interval $[a, b]$. Suppose $f(a) < c < f(b)$. Then there exists a number t in (a, b) with $f(t) = c$.

Proof: Let S be the set of all x in $[a, b]$ with $f(x) < c$. Then S is non-empty, since $f(a) < c$. Furthermore, S is bounded above by b. Let t be the supremum of S. Then t is in the interval $[a, b]$.

Since t is the least upper bound for S, it follows that for each n there is a number x_n with $t - 1/n < x_n$ and $f(x_n) < c$. Since t is an upper bound for S, it follows that $x_n \leq t$. Thus $x_n \to t$ as $n \to \infty$. By continuity $f(x_n) \to f(t)$ as $n \to \infty$. Since each $f(x_n) < c$, it follows that $f(t) \leq c$. In particular, $t < b$.

Since t is an upper bound for S, it follows that $x > t$ implies $f(x) \geq c$. Hence $f(t + 1/n) \geq c$. By continuity, $f(t + 1/n) \to f(t)$ as $n \to \infty$. Hence $f(t) \geq c$. In particular, $t > a$.

We conclude that $a < t < b$ and $f(t) = c$.

Corollary. Say that f is a real continuous function on the closed bounded interval $[a, b]$. Then the image of f is also a closed bounded interval.

Proof: We know from the compactness argument that the image has a minimum and a maximum point. Let s and t be such that $f(s)$ is the minimum and $f(t)$ is the maximum. If $f(s) = f(t)$, then the function is constant, and the image consists of just one point. Otherwise $f(s) < f(t)$. Then by the intermediate value theorem f assumes all values between $f(s)$ and $f(t)$. So the image is again the interval $[f(s), f(t)]$.

The bisection method. This gives a way of computing the root. Let f be continuous on $[a, b]$. The problem is to solve the equation $f(t) = c$, where $f(a) < c < f(b)$.

Start with an interval $[a, b]$ with $f(a) < c < b$. Repeat the following process. Compute $m = (a + b)/2$. If $f(m) > c$, then replace b by m. If $f(m) \leq c$, then replace a by m. After each such repetition the interval is half as long. After n repetitions it is only $1/2^n$ as long. Furthermore, we still have $f(a) \leq c$ and $f(b) > c$. So the result is that we can find a very small interval $[a, b]$ such that $a \leq t < b$, and $f(t) = c$. Thus we know t to high accuracy.

8 Uniform continuity

A function $f : S \to \mathbf{R}$ is uniformly continuous on S if whenever u_n, v_n are sequences in S such that $u_n - v_n \to 0$ as $n \to \infty$, then $f(u_n) - f(v_n) \to 0$ as $n \to \infty$.

Remark: To make sense of this definition, look at the u, v plane. Let $\epsilon > 0$ be small. Look at the set $|u - v| < \epsilon$. This is the same as the set where $-\epsilon < u - v < \epsilon$, or $u - \epsilon < v < u + \epsilon$. This is a small strip about the diagonal line $v = u$. To say that $u_n - v_n \to \infty$ as $n \to 0$ is to say that for every $\epsilon > 0$ the points u_n, v_n eventually stay in the ϵ strip about the diagonal. However they can wander around in a crazy way inside this strip.

Theorem. If $f : S \to \mathbf{R}$ is uniformly continuous, then $f : S \to \mathbf{R}$ is continuous.

Theorem. If $f : S \to \mathbf{R}$ is continuous, and S is compact, then $f : S \to \mathbf{R}$ is uniformly continuous.

Proof: Suppose $f : S \to \mathbf{R}$ is continuous and S is compact. Suppose that $f : S \to \mathbf{R}$ is not uniformly continuous. Then there are sequences u_n and v_n with $u_n - v_n \to 0$ as $n \to \infty$, yet for some $\epsilon > 0$ we have $|f(u_n) - f(v_n)| \geq \epsilon$ for infinitely many n.

It follows easily (by passing to a subsequence) that there are sequences u_n and v_n with $u_n - v_n \to 0$ as $n \to \infty$, yet $|f(u_n) - f(v_n)| \geq \epsilon$ for all n.

By compactness, there is a subsequence u_{n_j} that converges to some u in S. It follows that v_{n_j} also converges to u. So by continuity $f(u_{n_j}) - f(v_{n_j})$ converges to $f(u) - f(u) = 0$. This is a contradiction. The only way out is for f to be uniformly continuous.

9 ϵ and δ

The $\epsilon - N$ criterion for the limit of a sequence at infinity involves three alternating quantifiers: $\forall \epsilon > 0 \exists N \forall n (n \geq N \Rightarrow |a_n - a| < \epsilon)$.

The $\epsilon - \delta$ criterion for continuity of f at a also involves three alternating quantifiers: $\forall \epsilon > 0 \exists \delta > 0 \forall x (|x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon)$.

The $\epsilon - \delta$ criterion for continuity of f on S is

$$\forall a \in S \forall \epsilon > 0 \exists \delta > 0 \forall x \in S (|x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon).$$

The $\epsilon - \delta$ criterion for uniform continuity of f on S is

$$\forall \epsilon > 0 \exists \delta > 0 \forall a \in S \forall x \in S (|x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon).$$

What is the relation between these concepts? Look at where the $\forall a \in S$ quantifier appears. In the criterion for continuity the $\delta > 0$ can depend on $\epsilon > 0$ and on $a \in S$. However in the criterion for uniform continuity the $\delta > 0$ must depend only on the $\epsilon > 0$.

10 Monotone functions

Say that f is a monotone increasing function if $x < y$ implies $f(x) \leq f(y)$. Consider a monotone increasing function defined on an open interval (a, b). If t is a number in the interval, then $f(t-)$ is defined as the supremum of $f(x)$ for $x < t$, while $f(t+)$ is defined as the infimum of the $f(x)$ for $x > t$. Then $f(t-) \leq f(t) \leq f(t+)$.

Theorem. Suppose that f is a monotone increasing function defined on an open interval (a, b). Let t be a number in the interval. Then either $f(t-) = f(t+)$ and the function is continuous at t, or else $f(t-) < f(t+)$ and the entire open interval from $f(t-)$ to $f(t+)$ is not in the image of the function, except possibly for the number $f(t)$.

Proof: Start from the fact that f is monotone increasing. We know that for $x < t$ we have $f(x) \leq f(t-)$ and for $x > t$ we have $f(x) \geq f(t+)$. So the only value that can be assumed between $f(t-)$ and $f(t+)$ is $f(t)$.

Suppose $f(t-) = f(t) = f(t+)$. Let $\epsilon > 0$. Since $f(t)$ is the least upper bound of the $f(x)$ for $x < t$, it follows that $f(t) - \epsilon$ is not an upper bound. So there is a $\delta_1 > 0$ such that $f(t - \delta_1) > f(t) - \epsilon$. Similarly, since $f(t)$ is the greatest lower bound of the $f(x)$ for $x > t$, it follows that $f(t) + \epsilon$ is not a lower bound. So there is a $\delta_2 > 0$ such that $f(t + \delta_2) < f(t) + \epsilon$. Let δ be the smaller of δ_1 and δ_2. Then from the fact that f is monotone we see that $t - \delta < x < t + \delta$ implies $f(t) - \epsilon < f(x) < f(t) + \epsilon$. This is continuity.

Say that a set of real numbers has a gap if there is an open interval not in the set that is bounded below and above by points in the set.

Corollary. Suppose that f is a monotone increasing function defined on an open interval (a, b). Suppose its image has no gaps. Then f is continuous.

Application to probability. Consider a random variable with a certain probability distribution. The probability that the random variable has a value $\leq t$ is $f(t)$. This is called the cumulative probability distribution function of the random variable. It is a monotone increasing function.

The probability that the random variable is $< t$ is $f(t-)$, while the probability that the random variable is $\leq t$ is $f(t+)$. So with the above definition we have for each t that $f(t) = f(t+)$.

Each point t is either a continuity point or a point where there is a gap. If there is a gap of a certain size at t, this represents the probability $f(t+) - f(t-)$ that the random variable has a value exactly equal to t. The corresponding probability is the length of the gap.

11 Differentiation

The derivative of f at x is

$$f'(x) = \lim_{t\to 0} \frac{f(x+t)-f(x)}{t} = \lim_{z\to x} \frac{f(z)-f(x)}{z-x}.$$

Examples: $\sin' = \cos$, $\cos' = -\sin$, $\exp' = \exp$.

The sum of two functions is given by $(f+g)(x) = f(x)+g(x)$. The sum rule says that $(f+g)' = f'+g'$. The additive identity is 0. The additive inverse is $-f$. Clearly $(-f)' = -f'$.

The product of two functions is given by $(fg)(x) = f(x)g(x)$. The product rule says that $(f+g)' = fg' + f'g$. The multiplicative identity is 1. The multiplicative inverse is $1/f$. Since $1 = f(1/f)$, we have $0 = 1' = (f(1/f))' = f(1/f)' + f'/f$, and derivative is $(1/f)' = -f'/f^2$.

The composition of two functions is given by $(f \circ g)(x) = f(g(x))$. This is not a commutative operation! The chain rule says that $(f\circ g)' = (f' \circ g)g'$. The composition identity is $\#$, the identity function. The composition inverse is f^{-1}. Since $\# = f \circ f^{-1}$, we have $1 = \#' = (f\circ f^{-1})' = (f' \circ f^{-1})(f^{-1})'$, and its derivative is $(f^{-1})' = 1/(f' \circ f^{-1})$.

Example: Since $\exp' = \exp$ and $\ln = \exp^{-1}$, we have $\ln' = 1/\exp' \circ \ln = 1/\exp \circ \ln = 1/\#$ on $(0, +\infty)$.

Example: Let r be a real number. Consider $\#^r$ defined on $(0, +\infty)$ by $\#^r = \exp \circ (r \ln)$. Then $(\#r)' = r\#^{r-1}$. This is proved by calculating $(\#r)' = (\exp' \circ (r\ln))r\ln' = r(\exp\circ(r\ln))1/\# = r\exp\circ(r\ln))\exp(-\ln) = r\exp\circ((r-1)\ln) = r\#^{r-1}$.

Note: Say that $r = n/k$ is a rational number, where n, k are integers, and k is odd. Then $\#^r$ may be extended in a natural way to be an odd function if n is odd, and it may be extended in a natural way to be an even function if n is even.

First derivative test: If t is a local maximizer or a local minimizer for h, and if h is differentiable at t, then $h'(t) = 0$.

Rolle's theorem. If h is continuous on $[a, b]$ and differentiable on (a, b) and $h(a) = h(b)$, then there is a t in (a, b) with $h'(t) = 0$.

The Lagrange mean value theorem. If f is continuous on $[a, b]$ and differentiable on (a, b), then there is a t in (a, b) with $f'(t) = m$, where m is the slope from $a, f(a)$ to $b, f(b)$, that is, $f(b) - f(a) = m(b - a)$.

The Lagrange mean value theorem follows from Rolle's theorem by taking $h(x) = f(x) - mx$.

The Cauchy mean value theorem. If f and g are each continuous on $[a, b]$ and differentiable on (a, b), and if $f(b) - f(a) = m(g(b) - g(a)$, then there is a t in (a, b) with $f'(t) = mg'(t)$ A more revealing way to state this is to say that the vectors $[f'(t), g'(t)]$ and $[f(b) - f(a), g(b) - g(a)]$ in $\mathbf{R}^2$ are linearly dependent.

The Cauchy mean value theorem follows from Rolle's theorem by taking $h(x) = f(x) - mg(x)$.

Second derivative test. If $f'(t) = 0$ and $f''(t) > 0$, then t is a local minimizer. If $f'(t) = 0$ and $f''(t) < 0$, then t is a local maximizer.

Proof: Say for instance that $f''(t) > 0$. Then $(f'(u) - f'(t))/(u - t) > 0$ for all u near t with $u \neq t$. Since $f'(t) = 0$, this says that $f'(u)$ has the same sign as $u - t$. Consider z near t with $z \neq t$. By the Lagrange mean value theorem $f(z) - f(t) = f'(c)(z - t)$ for some c between t and z. Then $c - t$ has the same sign as $z - t$, so $f(z) - f(t) = f'(c)(z - t) > 0$. This says that $f(z) > f(t)$.

Effect of composition on first derivative test. If $g(x) = t$ and $f'(t) = 0$, then $(f \circ g)'(x) = 0$.

Proof: $(f \circ g)'(x) = f'(g(x))g'(x) = f'(t)g'(x) = 0$.

Effect of composition on second derivative test. If $g(x) = t$ and $g'(x) \neq 0$ and if $f'(t) = 0$ and $f''(t) > 0$ or $f''(t) < 0$, then $(f \circ g)''(x) > 0$ or $(f \circ g)''(x) < 0$.

Proof: $(f \circ g)''(x) = f''(g(x))g'(x)^2 + f'(g(x))g''(x) = f''(t)g'(x)^2 + f'(t)g''(x) = f''(t)g'(x)^2$.

12 Leibniz notation

In the following a function is said to be smooth on an open interval if it has derivatives of arbitrarily high order. A smooth function from an open interval to another open interval is said to be a diffeomorphism if it has an inverse function that is also smooth.

A one-manifold M is a set together with a collection of one-to one functions $u : M \to \mathbf{R}$, each with an image consisting of an open interval. These functions are called coordinate functions. It is required that if h is a diffeomorphism and u is a coordinate function, then $v = h \circ u$ is also a coordinate function. Furthermore, it is required that if u and v are coordinate functions, then the function $h = v \circ u^{-1}$ is a diffeomorphism.

A function $y : M \to \mathbf{R}$ is called a variable quantity if it is of the form $y = f \circ u$, where f is a smooth function and u is a coordinate function. We usually write this as $y = f(u)$.

If $y = f(u)$ is a variable quantity, then we write

$$\frac{dy}{du} = f'(u). \tag{1}$$

Each side of this equation is a variable quantity.

If $y = f(u)$ is a variable quantity and if $u = g(x)$, then we have the chain rule

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}. \tag{2}$$

This is because $y = f(g(x)) = (f \circ g)(x)$, and so

$$(f \circ g)'(x) = f'(g(x))g'(x) = f'(u)g'(x). \tag{3}$$

We can differentiate again by the chain rule and get

$$\frac{d^2y}{dx^2} = \frac{d^2y}{du^2}\left(\frac{du}{dx}\right)^2 + \frac{dy}{du}\frac{d^2u}{dx^2}. \tag{4}$$

Let P be a point in M where $\frac{dy}{du}(P) = 0$. Then $\frac{dy}{dx}(P) = 0$. This shows the coordinate invariance of the first derivative test.

Suppose in addition that $\frac{d^2y}{du^2}(P)$ is strictly positive or strictly negative at the point where the first derivative is zero. Then since we assume that $\frac{du}{dx} \neq 0$ we get that $\frac{d^2}{dx^2}(P)$ has the same sign. This shows the coordinate invariance of the second derivative test at a point where the first derivative test applies.

The trouble with this Leibniz notation is that there is not an accepted way of speaking of the value of a variable quantity on a point in M. Of course if $y = f(u)$ we can write the value $f(a)$, where a is a real number. But this is a value of f, not of y. Perhaps a solution is to introduce a notation like $u = a$ for the point P in M with $u(P) = a$. Then we have $y(u = a) = y(P) = f(u(P)) = f(a)$. Similarly, if $dy/du = f'(u)$, then $dy/du(u = a) = f'(a)$.

Note: It would be better to follow the convention used in many computer programming languages and use an assignment sign such as $u \leftarrow a$ rather than an equality sign $u = a$. The equality sign does not mean equality in this context; $u = a$ is not the same as $a = u$. For example, $u^2 + a^2$ with $u = a$ is $2a^2$, while $u^2 + a^2$ with $a = u$ is $2a^2$. Even some programming languages promote this confusion by using equality for assignment. Assignment is not equality. Nevertheless, mathematical convention is conservative, and so in the following we continue to use the equal sign for this kind of assignment.

Example: Consider the problem of finding a fence of length 40 to border a river. The set M of all fence configurations has coordinates u with $0 < u < 20$ and v with $0 < v < 40$. These coordinates satisfy $2u + v = 40$. If we want to enclose the maximum possible area $y = uv$, we need to maximize y on M.

If we express y in terms of v we get $y = u(40 - 2u)$. Then $dy/du = 40 - 4u$ and $d^2y/du^2 = -4$. The maximum is at the point $u = 10$.

If we express y in terms of u we get $y = (20 - v/2)v$ and $dy/du = 20 - v$ and $d^2y/dv^2 = -1$. The maximum is at the point $v = 20$. This is exactly the same fence configuration as with the first solution.

13 Differentials

If $y = f(u)$, then since $dy/dx = f'(u)du/dx$, and it does not matter which coordinate x is used, it is common to write

$$dy = f'(u)\,du \tag{5}$$

It is important to understand that these are not variable quantities of the kind we had before. It makes no sense to talk about the numerical values of such differentials. Nevertheless, it makes sense to talk about whether they are zero or non-zero. This does not depend on the choice of coordinate. In particular, the first derivative test can be carried out using differentials.

At a point P where $dy = 0$, we can write

$$d^2y = f''(u)(du)^2. \tag{6}$$

Again these are not variable quantities of the kind we had before. It makes no sense to talk about the numerical values of such second differentials. Nevertheless, at a point where the first derivative test applies it makes sense to talk about whether the second differential is zero, strictly positive, or strictly negative. This does not depend on the choice of coordinate. In particular, the second derivative test can be carried out using differentials.

Example: Consider the problem of finding a fence of length 40 to border a river. The set M of all fence configurations has coordinates u with $0 < u < 20$ and v with $0 < v < 40$. These coordinates satisfy $2u + v = 40$. If we want to enclose the maximum possible area $y = uv$, we need to maximize y on M.

We can express $dy = v\,du + u\,dv$. Also we have $2\,du + dv = 0$. At the configuration where $dy = 0$ we have $v = 2u$. At this configuration we have $d^2y = -4(du)^2 = -(dv)^2$. The fact that this is negative indicates a local maximum.

14 Integration

Let f be a bounded function on $[a, b]$. Let P be a partition of $[a, b]$. Thus $P = \{x_0, x_1, \ldots, x_{n-1}, x_n\}$ with $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$. We call the maximum value of $x_i - x_{i-1}$ the gap of the partition.

Let m_i be the infimum of f on $[x_{i-1}, x_i]$, and let M_i be the supremum of f on $[x_{i-1}, x_i]$. The lower and upper sums are defined by

$$L(f, P) = \sum_{i=1}^{n} m_i(x_i - x_{i-1}) \tag{7}$$

and

$$U(f, P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1}). \tag{8}$$

We always have $L(f, P) \leq U(f, P)$. Define the lower integral by

$$L(f) = \sup_P L(f, P) \tag{9}$$

and

$$U(f) = \inf_P U(f, P). \tag{10}$$

Then we always have $L(f) \leq U(f)$.

If $L(f) = U(f)$, then the integral of f on $[a, b]$ is defined by $I(f) = L(f) = U(f)$.

If f is a monotone increasing function, then $m_i = f(x_{i-1})$ and $M_i = f(x_i)$. If furthermore we take the regular partition in which each $x_i - x_{i-1} = (b-a)/n$, then

$$U(f, P) - L(f, P) = [f(b) - f(a)](b - a)/n. \tag{11}$$

So in this case the integral exists.

If f is a monotone decreasing function, then $m_i = f(x_i)$ and $M_i = f(x_{i-1})$. If furthermore we take the regular partition in which each $x_i - x_{i-1} = (b-a)/n$, then

$$U(f, P) - L(f, P) = [f(a) - f(b)](b - a)/n. \tag{12}$$

So in this case the integral also exists.

Some properties of the integral are easy to verify from the definition. For instance, if $a < b$ and $f \leq g$, then $I(f) \leq I(g)$.

Other properties are more subtle. Say that we want to prove that $I(f+g) = I(f)+I(g)$. The problem is that all we can say is that $m_i(f)+m_i(g) \leq m_i(f+g)$ and $M_i(f+g) \leq M_i(f) + M_i(g)$. However using this we get that

$$L(f,P) + L(g,P) \leq L(f+g,P) \leq U(f+g,P) \leq U(f,P) + U(g,P). \quad (13)$$

However if $I(f)$ and $I(g)$ exist, then $U(f,P) - L(f,P)$ and $U(g,P) - L(g,P)$ can be made small, and so $U(f+g,P) - L(f+g,P)$ can be made small. This shows that $I(f+g)$ exists and $I(f) + I(g) \leq I(f+g) \leq I(f) + I(g)$, which establishes what was wanted.

Theorem. If f is continuous on $[a,b]$, then $I(f)$ exists.

Proof: Let $f(v) - f(u)$ be the largest of the $M_i - m_i$. Then

$$U(f,P) - L(f,P) \leq [f(v) - f(u)](b-a), \quad (14)$$

where u and v are in the same partition interval. Now let P_n be a sequence of partitions with gap approaching zero. In particular $v_n - u_n$ approach zero. Write

$$U(f,P_n) - L(f,P_n) \leq [f(v_n) - f(u_n)](b-a). \quad (15)$$

By uniform continuity the right hand side approaches zero.

If f is defined on some larger interval, then we denote its integral on the interval $[a,b]$ with $a < b$ by $I_a^b(f)$. This is the setting for the first and second Fundamental Theorem of Calculus.

FTC1. If f is continuous and bounded on (a,b) and F is continuous on $[a,b]$ with $F' = f$ on (a,b), then $I_a^b(f) = F(b) - F(a)$.

FTC2. If f is continuous on $[a,b]$, and $F = I_a(f)$ regarded as a function of the upper limit of integration, then $F' = f$ on (a,b).

If $a < b$ we set $I_b^a(f) = -I_a^b(f)$, and we take $I_a^a(f) = 0$.

Substitution. Suppose $f : [c,d] \to \mathbf{R}$ is continuous, and $g : [a,b] \to \mathbf{R}$ is continuous, and that $g : (a,b) \to (c,d)$ has a derivative g' that is continuous and bounded. Then $I_a^b((f \circ g)g') = I_{g(a)}^{g(b)}(f)$.

15 Integration of differential forms

Let M be a one-dimensional manifold. Thus M is a set that resembles an open interval, except that it is not a set of numbers. Rather, it is a set of points. There are various coordinate functions on M, each coordinate function gives a correspondence of M with some open interval of numbers. This story was described in detail in an earlier section.

As we know, if u is a variable quantity defined on M, then expressions like like $y = F(u)$ or $dy/du = F'(u)$ also define variable quantities defined on M. However an expression like $dy = F'(u)\,du$ is not a variable quantity, it is a *differential form*. The way that a differential form becomes a variable quantity is that it is compared with the change of a coordinate, as in $dy/dt = F'(u)\,du/dt$.

If $u = c$ and $u = d$ are two points on the manifold, then the integral

$$\int_{u=c}^{u=d} f(u)\,du = I_a^b(f). \tag{16}$$

is defined.

This integral depends on the differential form and on the two points in M. However there is the possibility of using other coordinates to express the same differential form. So, for example, if $u = g(x)$ expresses the variable quantity u in terms of the variable quantity x, then we have

$$f(u)\,du = f(g(x))g'(x)\,dx. \tag{17}$$

Say that $c = g(a)$ and $d = g(b)$, so that $x = a$ and $x = b$ are the same points of M as $u = c$ and $u = d$. Then

$$\int_{x=a}^{x=b} f(g(x))g'(x)\,dx = \int_{u=g(a)}^{u=g(b)} f(u)\,du. \tag{18}$$

Example: The quantities need not be geometrical. Say that $pv = C$ is the equation of state for a gas at constant temperature. The manifold consists of the different configurations of the system, some with higher pressure and smaller volume, some with lower pressure and larger volume.

Then $p\,dv + v\,dp = 0$, so $p\,dv = -v\,dp$. The work done by the gas as it expands from one state to another is

$$w_{12} = \int_{v=v_1}^{v=v_2} p\,dv = -\int_{p=p_1}^{p=p_2} v\,dp = C\ln(v_2/v_1) = -C\ln(p_2/p_1). \tag{19}$$

Example: In this example $r > 0$ is fixed. Consider a quarter circle of radius r. This is the manifold M. Notice that it is a geometrical object. Let P and Q be the lower right hand and upper left hand end points of the quarter circle.

Various coordinates are possible. For example, one can take the angle θ with $0 < \theta < \pi/2$. Or one can take the coordinate $x = r\cos(\theta)$ with $0 < x < r$. Or one can take the coordinate $y = r\sin(\theta)$ with $0 < y < r$.

There are corresponding relations between differential forms. Since $x^2+y^2 = r^2$, we have $x\,dx + y\,dy = 0$. Furthermore we have $dx = -r\sin(\theta)\,d\theta$ and $dy = r\cos(\theta)\,d\theta$.

The area may be calculated by integrating various differential forms from P to Q. Some of these differential forms are $x\,dy = r^2\cos^2(\theta)\,d\theta$ and $-y\,dx = r^2\sin^2(\theta)\,d\theta$ and $(1/2)r^2\,d\theta$.

By using the identities $\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta)$ and $\cos^2(\theta) + \sin^2(\theta) = 1$, we get $2\cos^2(\theta) = 1 + \cos(2\theta)$ and $2\sin^2(\theta) = 1 - \cos(2\theta)$. Hence we get $x\,dy = (1/2)r^2\,d\theta + (1/2)r^2\cos(2\theta)\,d\theta$ and $-y\,dx = (1/2)r^2\,d\theta - (1/2)r^2\cos(2\theta)\,d\theta$.

But $\cos(2\theta)\,d\theta = (1/2)d(\sin(2\theta))$. Notice that the integral of $d(\sin(2\theta))$ from P to Q is zero. Thus the area is

$$\int_{y=0}^{y=r} x\,dy = -\int_{x=r}^{x=0} y\,dx = \frac{1}{2}r^2\int_{\theta=0}^{\theta=\pi/2} d\theta = \frac{1}{4}\pi r^2. \tag{20}$$

Summary: The integral

$$\int_P^Q f(g(x))g'(x)\,dx = \int_P^Q f(u)\,du \tag{21}$$

is the integral of a differential form $f(g(x)\,g'(x)\,dx = f(u)\,du$ over a manifold from the point P in M to the point Q in M. If $u = c$ at P and $v = d$ at Q, then this is

$$\int_P^Q f(u)\,du = I_c^d(f). \tag{22}$$

In applications variables such as x and u have independent meanings. One may be a function of another, for instance $u = g(x)$. However it is also possible that $u = h(t)$, where t is some other coordinate. The variable quantity u is not the same as the function g or the function h. However it can have numerical values, such as for example $u(x = a) = g(a)$ or $u(t = z) = h(z)$.