PRINCIPLE RESULTS OF INTRODUCTORY REAL ANALYSIS

Term Paper for Dr. Hooper's Real Analysis Course

by Victor Aguilar

December 19, 1994

ABSTRACT

Many books on real analysis contain extensive material on the real number system. This includes many results regarding sequences and series of real numbers as well as results from metric topology as they relate to the real number system. Royden is brief in this respect and I will omit it entirely since I feel it is introductory and should be prerequisite of a real analysis course. Also, I will omit any mention of most of the practical calculus techniques (particularly differential calculus) which are generally taught in introductory calculus courses. However, introductory real analysis does not extend beyond calculus of one variable, so multivariable calculus is not prerequisite of this essay.

I will begin by defining the Riemann integral and discuss some of the theorems relating to it. Throughout this section most of the sets mentioned will be intervals on the real number line. In the next section I will discuss "measurable" sets (still on the real number line) and functions. In the third section I will define the Lebesque integral and discuss some of the results of Lebesque integration. Finally, I will discuss how Lebesque integration extends Riemann integration, that is, why we are not content with an integral defined only in terms of intervals, which are more tractable than

measureable sets. Proofs will be omitted for the sake of brevity and ease of exposition.

Since this wordprocessor does not have an epsilon symbol, I will use alpha in a similar manner. ~E represents the complement of the set E. The symbol < means either "is less than" or "is a subset of." The context should make it clear which is meant.

Section 1

Riemann Integration

The term "almost everywhere" appears throughout real analysis. Almost everywhere in a set is everywhere except on a subset of measure zero. "Measure" will be defined in section 2 but it is possible (and useful) to define "measure zero" seperately. A set E is of measure zero (denoted $\mu E = 0$) if, for all $\alpha > 0$, there exists a countable sequence of open intervals I_n which cover E and $\Sigma|I_n| < \alpha$. If I is an interval then $|I|$ represents its length. Clearly, a countable union of sets, each of measure zero, is also of measure zero. Also, any countable set is of measure zero. This is remarkable considering that there is a countable set (the rational numbers) which is dense in the reals. The reals are of measure ∞. It may seem that all sets of measure zero would be countable. However, the Cantor set is uncountably infinite and of measure zero. This is also remarkable because, in probability, measure zero means (roughly) "impossible", though there may be uncountably many counterexamples.

Several concepts must be defined in preparation for defining the Riemann integral. Unless explicitly stated otherwise, all intervals and functions will be assumed to be bounded. $M[f;I] = \sup f(x)$ and $m[f;I] = \inf f(x)$ with x in the interval I. $\sigma = \langle x_k \rangle$ with $i = 0, 1, \cdots, n$ is a subdivision of [a,b] if $a = x_0 < x_1 < \cdots < x_n = b$. To simplify the notation, all integrals in this section, unless otherwise noted, will be over [a,b]. Also, σ and τ will always be subdivisions of this interval and I_k the interval $[x_{k-1}, x_k]$. With regards to a particular subdivision σ, the upper and lower sums of f in [a,b] are $U[f;\sigma] = \Sigma M[f;I_k] \cdot |I_k|$ and $L[f;\sigma] = \Sigma m[f;I_k] \cdot |I_k|$, respectively. Clearly, we have $L[f;\sigma] \leq U[f;\tau]$. Finally, the upper and lower Riemann integrals are defined as $\sup U[f;\sigma]$ and $\inf L[f;\sigma]$ over all subdivisions of [a,b]. From the previous result we have the lower Riemann integral less than or equal to the upper Riemann integral. If they are exactly equal then f is Riemann integrable over [a,b] and this common value is its Riemann integral, denoted

$$R\int f(x)dx$$

The R will be dropped as will will consider no other kind of integral in this section. In section 3 the Lebesque integral will be defined and, from there on, only it will be considered so there will be no need to denote it with an L.

It follows directly from the definition that $c \in (a,b)$ implies that the Riemann integrals over [a,c] and [c,b] exist and sum to the Riemann integral over [a,b]. Also, by defining the integral over [a,b] to be the negative of the integral over [b,a], the statement above is true regardless of the order of a, b, and c.

$$c \int f(x)dx + \int g(x)dx = \int cf(x) + g(x)dx$$

with the integrability of the function on the left implying that of the function on the right.

For $f(x) \leq g(x)$ almost everywhere,

$$\int f(x)dx \leq \int g(x)dx \quad \text{and} \quad \left|\int f(x)dx\right| \leq \int |f(x)|dx$$

Roughly, a subdivision cannot be made fine enough that any of its component intervals are contained in the set of points where $g(x) < f(x)$.

This is also why the characteristic function of any dense subset of measure zero (or its complement) is not Riemann integrable. A subdivision cannot be made fine enough that any of its component intervals are contained in the defining set. This is clear from the definition of zero measure. For instance, the characteristic function of rational numbers in [a,b] is not Riemann integrable. It may seem that the difficulty is only that its points of discontinuity are dense in [a,b]. However, consider the function defined on [0,1] such that, for all rational numbers $x = n/d$ (in lowest terms), $f(x) = 1/d$ and, for all irrational numbers, $f(x) = 0$. Because the irrationals are dense in [0,1], there is one in every neighborhood of a rational number so, taking $\alpha = 1/(2d)$ for each rational number, they are points of discontinuity and are dense in [0,1]. Clearly, the lower Riemann integral of $f(x)$ is zero. Let σ_n be the subdivision with $x_k = k/n$ and $U[f;\sigma_n] = 1/n$, that is, $U[f;\sigma_n]$ is the area of n squares, each 1/n on a side. Sup $U[f;\sigma_n]$ for all n is 0, so the Riemann integral of f is defined and is zero. Thus, density of discontinuities alone is not sufficient. As the example in the next paragraph shows, discontinuities of measure zero alone is not sufficient.

f is Riemann integrable iff, for all $\alpha > 0$, there exists a subdivision σ such that $U[f;\sigma] < L[f;\sigma] + \alpha$. This is clear from the definition but it is not useful in practice because it would require finding such a subdivision. It is a fundamental result of integration theory that f is Riemann integrable iff f is almost everywhere continuous. Consider the characteristic function of the Cantor set. The Cantor set is the set of numbers in [0,1] with a ternary expansion that does not contain the digit 1. At every point in the Cantor set, this function is discontinuous. This is because, for any larger element, there is a number not in the Cantor set between them obtained by replacing the first 2 of the larger element with a 1. At every point of [0,1] not in the Cantor set, this function is continuous because, if its first 1 is in the n'th place, all the numbers equal to it in their first n places are not in the Cantor set and form an open interval $1/3^n$ wide. For this same reason, the Cantor set is not dense. For reasons given in section 2, the Cantor set is of measure zero and, hence, Riemann integrable. For the same reason that binary expansions are uncountably infinite, the Cantor set is also uncountably infinite. Thus, it has as many discontinuities as a function can have without losing Riemann integrability.

The derivative of f(x) is $f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$

Considering that f is Riemann integrable iff f is almost everywhere continuous, one might ask if this is true of derivatives as well. Differentiability implies continuity but the converse is not true. The absolute value function is continuous but is not differentiable at zero. Periodic triangular waveforms are everywhere continuous but are not differentiable on a countably infinite set. By taking a countable summation of such functions it is possible to construct an everywhere continuous but nowhere differentiable function. Let $f_1(x) = |x|$ in $[-\frac{1}{2},\frac{1}{2}]$ and $f_1(x+n) = f_1(x)$ for all real x and integer n. For n = 2, 3,···, let $f_n(x) = f_1(4^{n-1}x)/4^{n-1}$. $\Sigma f_n(x)$ converges by the Weierstrass M-test because $|f_n(x)| \leq 1/4^{n-1}$. Let $h_n = \pm 1/4^{n+1}$ so we have $|f_n(x+h_n) - f_n(x)| = |h_n|$. Also, $|f_m(x+h_n) - f_m(x)| = |h_n|$ for $m \leq n$ but zero for $m > n$. Thus, $(f(x+h_n) - f(x))/h_n$ is an even or odd integer exactly where n is even or odd and, consequently, the derivative of f(x) does not exist.

For a function to have a derivative almost everywhere on [a,b], it must be, not just continuous, but absolutely continuous. That is, given $\alpha > 0$, there exists $\delta > 0$ such that $\Sigma|x_i'-x_i| < \delta$ implies $\Sigma|f(x_i')-f(x_i)| < \alpha$ with $\{(x_i'-x_i)\}$ a finite collection of non-overlapping intervals in [a,b]. Absolute continuity is equivalent to bounded variation, which is easier to verify. Bounded variation on [a,b] means that $\sup \Sigma|f(x_k)-f(x_{k-1})| < \infty$. Other than direct recourse to the definition of differentiability, there is no way to upgrade almost everywhere to everywhere. This can be a problem in practical applications such as Newton's Method, which assumes the existence of derivatives.

If a function is differentiable, will its derivative necessarily be continuous? No. Consider $f(0) = 0$ and $f(x) = x^2\sin(1/x)$ elsewhere. It is differentiable but its derivative is discontinuous at 0. If a function is continuous it is necessarily the derivative of some function, however. That is, if f is continuous at x then there exists a function F such that $F'(x) = f(x)$. Furthermore, not only can we prove the existence of such a function but we can give it a name.

$$F(x) = \int_a^x f(t)\,dt \qquad \text{with } x \in [a,b]$$

Notice that, while F(x) exists, it is not unique. It depends on one's choice of a. Furthermore, F may not be expressible in terms of the standard transcendental functions. For instance, $f(x) = \exp(-\frac{1}{2}x^2)/\sqrt{(2\pi)}$ is everywhere continuous and hence has an anti-derivative $F(x) = \frac{1}{2}\text{erf}(x/\sqrt{2})$, but there is no expression for the error function, F(x), other than as the indefinite integral of f(x).

Thus, the derivative is related to the integral. This is remarkable considering that their definitions make no mention of one another and do not look much alike. In fact, it is so remarkable that the theorem above is called the Fundamental Theorem of Calculus.

There is another (more useful) version of the Fundamental Theorem of Calculus which requires two preliminary theorems. If $f'(x) = 0$ on [a,b] then $f(x) = c$ on [a,b] and, immediately, if $f'(x) = g'(x)$ on [a,b], then $f(x) - g(x) = c$ on [a,b]. And so we have our most important result: If f is continous on [a,b] and F an anti-derivative of it, then

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

Differentiation is easy. Integration is harder but, having associated derivatives with integrals, we can use the chain rule to prove the very useful method of substitution. If g'(x) is continuous on [a,b] and f(x) is continuous on g([a,b]), then

$$\int_{g(a)}^{g(b)} f(x)dx = \int_a^b f(g(x))g'(x)dx$$

Integration by parts is derived from the product rule.

$$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx$$

Recall that all intervals and functions have been assumed to be bounded. If we relax this requirement for intervals we get improper integrals of the first kind. The theory of such integrals closely parallels that of sequences and series. Partial integrals play the part of partial sums and absolute convergence implies convergence but not the converse. For non-negative, non-increasing functions the improper integral converges iff the infinite series does. The concept of rearranging the terms in a series does not carry over to integration, however, because subdivisions only have a finite number of terms. Thus,

$$\int_\pi^\infty \frac{\sin(x)dx}{x} = \int_\pi^\infty \frac{\cos(x)dx}{x^2} + \frac{1}{\pi}$$

even though the integral on the left is conditionally convergent and the one on the right is absolutely convergent. This could not happen for an infinite series because, with appropriate rearrangement of its terms, a conditionally convergent series can be made to converge to any real number.

If we retain the requirement that intervals be bounded but relax the requirement that functions be bounded, we have improper integrals of the second kind. These are called Cauchy-Riemann integrals. If the function is unbounded on a (small) finite set of points in [a,b], one can partition [a,b] at those points (called singularities) and evaluate only over intervals with singularities at their boundary. If the singularity is on the left then one takes a right-hand limit

$$\lim_{\alpha \to 0^+} \int_{a+\alpha}^{b} f(x)\,dx$$

Similarly with left-hand limits for singularities on the right. These limits are handled in much the same way that improper integrals of the first kind are. In fact they can often be converted into one another as

$$\lim_{\alpha \to 0^+} \int_{\alpha}^{1} \frac{1}{x}dx = \lim_{\alpha \to 0^+} \int_{1}^{1/\alpha} \frac{1}{u}du$$

with u = 1/x. If there are a lot (or an infinity) of singularities this process is more difficult.

Improper integrals of both the first and second kind can be partitioned so one first evaluates on the side with the singularity and then does the tail. Thus, it is never neccessary to deal with both improprieties at once, though both integrals must be finite for the combined integral to be finite.

Section 2

Lebesque Measure

In section 1 virtually all the sets considered were intervals and, in particular, all the integrals were evaluated over bounded intervals. Even for improper integrals of the first kind all the integrals considered were evaluated over bounded intervals. One just took a sequence of such integrals. In fact, for most applications involving real numbers one only considers intervals. However, particularly in probability, it is sometimes necessary to consider (and measure) more complicated sets. The measure of an interval is its length, so all that has been said in section 1 regarding intervals still holds in the more general setting of Lesbesque measure. We will not go beyond the real numbers, however, so "set" will always refer to a subset of the real number line.

The outer measure of a set is $m^* = \inf \Sigma|I_n|$ with $<I_n>$ any countable collection of intervals which cover the set. Clearly, the outer measure of the union of a countable collection of sets is less than or equal to the sum of the outer measures of all the sets, because they may overlap. Thus, $m^*A \leq m^*(A \cap E) + m^*(A \cap \tilde{}E)$. If, for all A, we also have $m^*A \geq m^*(A \cap E) + m^*(A \cap \tilde{}E)$, then E is measurable and

$\mu E = m^*E$. This definition is due to Carathéodory. Measurability is not much of a restriction. Royden, and every other author that I consulted, give the same highly artificial example of a set that is not measureable. It may be that there is no other example, or at least none that can be constructed without using the Axiom of Choice to pick from an uncountably infinite collection of sets, which is a somewhat dubious procedure. Thus, for practical applications, one can assume that all sets are measureable. In particular, those of measure zero considered in section 1 as well as unbounded intervals and all open and closed sets are measureable.

This is not the only way to define Lebesque measure. Because of how the general version of the Lebesque integral is defined, we really only need to consider sets of finite measure. Thus, one may begin by taking [a,b] rather than the entire real line as the universal set. This does not change the definition of outer measure, which can be abreviated as $m^*E = \inf |G|$ with G an open covering of E. Now we can define inner measure as $m_*E = \sup |F| = \sup (b - a - |\tilde{}F|)$, however, with F a closed set contained in E. Measureable sets are exactly those whose inner and outer measure are equal and this common value is their Lebesque measure. With this in mind, all sets mentioned from hereon will be of finite measure unless explicitely stated otherwise.

This definition leads directly to the Littlewood's First Principle, that the measurability of E implies the existence of a countable union of closed sets F and a countable intersection of open sets G such that F < E < G and E\F and G\E are both of measure zero. If "countable" is replaced with "finite" then we only have $\mu(E \setminus F) < \alpha$ and $\mu(G \setminus E) < \alpha$ for arbitrarily small α. Because defining both inner and outer measure leads so easily to Littlewood's First Principle, I prefer this definition rather than Carathéodory's definition (cited by Royden) which uses only outer measure. Carathéodory's definition requires proving that a relation holds for every possible set A, which does not seem very useful. There is even another definition which also involves only outer measure, that is, E is measureable iff, given $\alpha > 0$, there exists open sets G_1 and G_2 such that $E < G_1$, $\tilde{}E < G_2$, and $|G_1 \cap G_2| < \alpha$.

Actually, it does not really matter how measureability is defined since, for all practical purposes, all sets are measureable. With this in mind, most of the theorems about measurable sets can be skimmed over as they establish only existence. Countable additivity of disjoint sets is important because it can be used to find the measure of sets. In particular, for nested sets $E_1 < E_2 < \cdots$, the measure of their union is the limit of their measures. Similarly, for nested sets $E_1 > E_2 > \cdots$, the measure of their intersection is the limit of their measures. This latter result can be used to show that the Cantor set is of measure zero. I suspect that most sets whose measure is actually known are the union or the intersection of nested sets. Also, having established the measure of a set, if its symetric difference with another set is of measure zero, then that set has the same measure. In other words, subsets of measure zero do not contribute to the measure of a set.

Measurable functions are ones such that, for all real c, x such that $f(x) \geq c$ is a measurable set. Here, $\geq$ can be replaced with $>$, $\leq$, or $<$. Combining $\geq$ and $\leq$ implies that, if a function is measureable, then x such that $f(x) = c$ is a measurable set. Clearly, the characteristic function of a non-measurable function is not measureable. Otherwise, the situation parallels that of measurable sets, that is, in practical situations one can assume that all functions are measurable. Skimming over the existence theorems, there are only two important result regarding measurable functions.

Littlewood's Second Principle states that, if f is measurable on [a,b] and finite almost everywhere, there is a step function and a continuous function that fall outside an α-strip only on a set of measure less than α. That is, the closer your tolerances, the less the continuous (or step) approximation fails you. (Why can't people be like that?)

Littlewood's Third Principle states that for bounded functions, if f_n-->f pointwise almost everywhere, then $<f_n>$ converges in measure to f. Convergence in measure means that, given tolerances on both the narrowness of the α-strip and on the measure of the set of failures, an N can be found such that f_n approximates f for $n \geq N$. This is equivalent to saying that every subsequence has a subsequence that is almost everywhere pointwise convergent. The converse of Littlewood's Third Principle is not true, that is, almost everywhere pointwise convergence cannot be upgraded from

subsequences of subsequences to the sequence itself.
Consider the following (pseudo) BASIC program:
For $i = 1$ to ∞ \ For $j = 0$ to $2^i - 1$ \ $n = 2^i + j$ \ Next j,i.
Let $f_n = 1$ in $[j/2^i, (j+1)/2^i]$ and 0 in the rest of $[0,1]$.
$<f_n>$ is convergent in measure but not almost everywhere pointwise convergent. Thus, convergence in measure is the weakest possible kind of convergence.

In summary, besides being the basis for Lebesque integration, Lebesque measure allows us to make the following statements about (practically) all sets and functions:

1) Sets can be approximated by an intersection of open sets or a union of closed sets.

2) Almost everywhere bounded functions can be approximated by continuous (or step) functions.

3) Sequences of bounded functions that only converge pointwise (and not even everywhere) would converge uniformly but for a small part of their domain.

As will be shown in section 4, the most important results of Lebesque integration, and the ones which cannot be proven using Riemann integration, are all fairly direct consequences of Littlewood's Three Principles.

Section 3

Lebesque Integration

Riemann integration can be thought of as approximating a function from above and below with step functions, integrating the step functions in the obvious way, and taking their common value (if they have a common value) as the Riemann integral. Lebesque integration is exactly the same thing except with simple functions. Simple functions are like step functions except, instead of being constant on intervals, they are constant on measurable sets. They are integrated in the same way that step functions are except that the horazontal component of area is not just the length of an interval but the measure of a (parhaps quite complicated) set. Since intervals are one kind of set, Riemann integration is a restriction of Lebesque integration. If a function is Riemann integrable, then it is also Lebesque integrable and the two integrals agree.

Another way to think of Lebesque integration is that one subdivides the range of the integrand rather than its domain as in Riemann integration. If the subdivision is fine enough then the height of the points in each component interval is constant. From the definition of measureability, x such that $f(x) \geq c$ is a measurable set. The measures of the sets associated with each interval times its width can be summed over all the component intervals in a subdivision to give an approximation of the area under the graph of the function.

Clearly, for bounded functions, the Lebesque integral over a set of measure zero is zero regardless of the integrand's value on that set. Thus, one of the advantages of Lebesque integration is that we can ignore values of a function as long as they occur only on a set of measure zero. For this reason, most of the results relating to Lebesque integration are conditional on statements which do not have to hold everywhere, only almost everywhere. This is useful because relations such as $f_n(x) \dashrightarrow f(x)$ or $f(x) \leq g(x)$ may have blips where they fail. If the measure of such failures is zero, the relation can still be used in a proof as though it were true everywhere.

For bounded functions on sets of finite measure, limits can be pulled out of integrals, that is, if $|f_n(x)| < M$ and $f_n \dashrightarrow f$ almost everywhere (Royden unneccessarily says everywhere) then

$$\int f(x)dx = \lim_{n \dashrightarrow \infty} \int f_n(x)dx$$

This is called the Bounded Convergence Theorem and is an application of Littlewood's Third Principle that sequences that only converge pointwise (and not even everywhere) would converge uniformly but for a small part of their domain. Uniform convergence would easily imply the result and, by Littlewood's Third Principle, pointwise convergence is close enough. This is the only significant result that requires functions be bounded and integrated on sets of finite measure. It is really a lemma for the more powerful convergence theorems which apply to general Lebesque integration.

By general Lebesque integration I mean integration of unbounded functions over infinite sets. The integral of a non-negative f over E is defined as the supremum of integrating h over E'. Here, h is any function that is bounded, dominated by f, and zero outside E' with $\mu E' < \infty$ and $E' < E$. In general, the positive part and the negative of the negative part must be done seperately. This definition allows us to consider only bounded functions on sets of finite measure, which is all that we actually know how to integrate, when proving theorem about general Lebesque integration. Thus, the remaining theorems in this section do not assume any bound on sets and will establish their own bounds on functions, when needed.

The Monotone Convergence Theorem also allows limits to be pulled out of integral. However, unlike the Bounded Convergence Theorem which requires $|f_n(x)| < M$, it requires $f_n(x)$ to be increasing and non-negative. This bounds $f_n(x)$ above by $f(x)$. It is this bound that is important so, if one can show that $f_n(x) \le f(x)$ for almost all x, $f_n(x)$ need not be increasing, though that is usually the easiest way to establish the bound. That this works is remarkable since f need not be Lebesque integrable, that is, its integral may not be finite. The Monotone Convergence Theorem has the very useful corollary that infinite summations can be pulled out of integrals.

$$\int f(x)dx = \sum_{n=1}^{\infty} \int u_n(x)dx$$

where f is the sum of $\langle u_n \rangle$. If the u_n can be integrated in closed form and U_n is dominated by a sequence whose sum converges, then f is Lebesque integrable. This result may be difficult or impossible to obtain without reversing the order of the summation and the integral. In the context of Riemann integration, this is only defined for proper integrals and requires that the partial sums of $u_n(x)$ converge uniformly to f(x).

The Lebesque [Dominated] Convergence Theorem also allows limits to be pulled out of integrals, but requires that $f_n(x)$ be bounded above and below by Lebesque integrable functions. This is stronger than the comparable statement in Riemann integration which is only defined for proper integrals and requires uniform convergence of the limit before it is pulled out of the integral. Consider $f_n(x) = \sqrt{x}$ in (1/n,2/n) and zero in the rest of [0,2]. This is bounded by the Lebesque integrable function $g(x) = \sqrt{2}/\sqrt{x}$ in (0,2] and zero at zero, so $\langle f_n \rangle$ satisfies the conditions of the Lebesque Convergence Theorem but does not uniformly converge to f(x) = 0.

The Monotone Convergence Theorem is implied by the Lebesque Convergence Theorem only if $f_n(x)$ converges to an Lebesque integrable function. The strong point of the Monotone Convergence Theorem is that its bound, f(x), need not be Lebesque integrable. The strong point of the Lebesque Convergence Theorem is that its bounds, $g_n(x)$, can be different and more tractable functions than $f_n(x)$.

An example of a sequence of Riemann integrable functions which is monotonic and non-negative but converges to a function that is not Riemann integrable will now be given. Let I_1, I_2, $\cdots$, I_n be all the open intervals removed by the n'th iteration of the construction of the generalized Cantor set with $\mu C = \alpha$. Define $g_{n,k}(x)$ on [0,1] to be 1 outside I_k, 0 in $[a + (b-a)/2^n, -(b-a)/2^n + b]$ where a and b are the left and right endpoints of I_k, and linear between these intervals. Let $f_n(x)$ be the product of $g_{n,k}(x)$ for k = 1, 2, $\cdots$, n. $f_n(x)$ is piecewise linear and hence Riemann integrable. However $f_n(x)$ converges to the characteristic function of the generalized Cantor set which is discontinuous on a set of positive measure and hence is not Riemann integrable. This shows that the Monotone Convergence Theorem cannot be proven in the context of Riemann integrals. Furthermore, because the characteristic function of the generalized Cantor set is Lebesque integrable, the Lebesque Convergence Theorem cannot be proven in the context of Riemann integrals either.

The monotone and Lebesque convergence theorems each bound $f_n(x)$ in different ways. If $f_n(x)$ is completely unbounded above then we cannot integrate f(x) but only bound its integral by the limit inferior of the integrals of $f_n(x)$. That is,

$$\int f(x)\,dx \le \underline{\lim}_{n \to \infty} \int f_n(x)\,dx$$

This may be sufficient to establish the integrability of f, however. This result is called Fatou's Lemma. Since establishing Lebesque integrability is required for many of the results of Lebesque integration, Fatou's Lemma is an important result.

Also for completely unbounded functions we have a theorem similar to the statement that, for bounded functions, the Lebesque integral over a set of measure zero is zero regardless of the integrand's value on that set. If a function is unbounded but Lebesque integrable then its integral can be made arbitrarily small by reducing the measure of the set that it is integrated over. That is, given $\alpha > 0$ there exists $\delta > 0$ such that, for $\mu E < \delta$, we have

$$\int_E f(x)dx < \alpha$$

provided that f is Lebesque integrable. This is an easy (but useful) result of the Bounded Convergence Theorem. It actually has to be proven before the monotone and Lebesque convergence theorem but is stated now for expository purposes. Furthermore, this theorem can be strengthened to state that, for Lebesque integrable functions with $f(x) = g(x)$ almost everywhere, their integrals are equal. Confusingly, Royden never actually states this rather fundamental theorem, stating only the weaker version above, which is really just a lemma for the Lebesque Convergence Theorem.

Section 4

Conclusion

Littlewood's Third Principle states that, for bounded functions, if f_n-->f pointwise almost everywhere, then $<f_n>$ converges in measure to f. Thus, convergence in measure is the weakest kind of convergence. Littlewood's Second Principle states that, for an almost everywhere bounded function, there is a step function and a continuous function that fall outside an α-strip only on a set of measure less than α. With these two principles we have Fatou's Lemma, The Monotone Convergence Theorem, and the Lebesque Convergence Theorem all true assuming only convergence in measure. This is in stark contrast to the comparable statements in Riemann integration which are defined only for proper integrals and require uniform convergence, the strongest kind. Furthermore, the functions $f_n(x)$ need not be bounded by some M, as they must be in Riemann integration. Fatou's Lemma assumes no bound at all while the monotone and Lebesque convergence theorems assume much weaker bounds on $f_n(x)$.

For Lebesque integrable functions, the Lebesque integral over a set of measure zero is zero regardless of the integrand's value on that set. This conveniently ignores blips in functions that would prohibit the application of

Riemann integration. In particular, it makes possible integration over quite complicated sets of the type found in probability. Littlewood's First Principle states that the measurability of E implies the existence of a countable union of closed sets F and a countable intersection of open sets G such that F < E < G and E\F and G\E are both of measure zero. Thus, for nested sets $E_1 < E_2 < \cdots$, the measure of their union is the limit of their measures. Similarly, for nested sets $E_1 > E_2 > \cdots$, the measure of their intersection is the limit of their measures. This establishes not only the existence but the actual value of many complicated integrals that cannot be done with Riemann integration. In computing these integrals we have

$$c \int f(x)dx + \int g(x)dx = \int cf(x) + g(x)dx$$

and, for $f(x) \le g(x)$ almost everywhere,

$$\int f(x)dx \le \int g(x)dx \quad \text{and} \quad \left|\int f(x)dx\right| \le \int |f(x)|dx$$

This requires only that the functions be Lebesque integrable, which is a much stronger result than the comparable statements in Riemann integration which require that the functions be proper Riemann integrals.

A result of Riemann integration is that, while derivatives are not always continuous, if a function is continuous it is necessarily the derivative of some function. That is, if f is continuous at x then there exists a function F such that $F'(x) = f(x)$ and that function is

$$F(x) = \int_a^x f(t)dt \qquad \text{with } x \in [a,b]$$

Furthermore, because f is Riemann integrable iff f is almost everywhere continuous, F'(x) = f(x) almost everywhere. Lebesque integration can upgrade this result by replacing the requirement that f be continuous with the much weaker requirement that f be Lebesque integrable. This is primarily a result of Littlewood's Second Principle that, for an almost everywhere bounded function, there is a continuous function that falls outside an α-strip only on a set of measure less than α. Since Lebesque integrable functions are neccessarily Riemann integrable, we automatically have F'(x) = f(x) almost everywhere, that is

$$\frac{d}{dx}\int_a^x f(t)dt = f(x) \quad \text{almost everywhere}$$

Since the expression above would not make sense if f(x) were not Lebesque integrable, this is not much of a restriction. Thus, the Fundamental Theorem of Calculus is significantly stronger in the context of Lebesque integration than in Riemann integration. However, it still holds only almost everywhere, which is considerably weaker than if it held everywhere. For instance, there is a function with a dense set of failures. Consider the function defined on [0,1] such that, for all rational numbers x = n/d (in lowest terms), f(x) = 1/d and, for all irrational numbers, f(x) = 0. In section one it was shown that the integral of this function is zero, so the equality above holds only for irrational numbers and fails on the set of rational numbers, which is dense in [0,1].

The Fundamental Theorem of Calculus puts a condition on

$f(x)$ to establish the existence of an indefinite integral $F(x)$. While F may not be expressible in terms of the standard transcendental functions, given a function that is, what condition must be put on it to establish the existence of a derivative? In other words, how do we know if F is the indefinite integral of some f? Recall that, for a function to have a derivative almost everywhere on [a,b], it must be, not just continuous, but absolutely continuous. That is, given $\alpha > 0$, there exists $\delta > 0$ such that $\Sigma|x_i'-x_i| < \delta$ implies $\Sigma|F(x_i')-F(x_i)| < \alpha$ with $\{(x_i'-x_i)\}$ a finite collection of non-overlapping intervals in [a,b]. Absolute continuity is equivalent to bounded variation, which is easier to verify. Bounded variation on [a,b] means that sup $\Sigma|F(x_k)-F(x_{k-1})| < \infty$. Since the stronger version of the Fundamental Theorem of Calculus (in Lebesque integration) provides that $F'(x) = f(x)$ almost everywhere, $F(x)$ is an indefinite integral iff it is absolutely continuous or of bounded variation. That is,

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

iff sup $\Sigma|F(x_k)-F(x_{k-1})| < \infty$. No comparable statement can be made in the theory of Riemann integration.

Finally, looking ahead, define square integrable to mean that $f^2(x)$ is Lebesque integrable over [a,b]. The equivalence classes of square integrable functions that are equal almost everywhere are a metric space and the norm of a square integrable function is

$$\| f \|_2 = \left(\int f^2(x)dx \right)^{\frac{1}{2}}$$

It is a fundamental result that this metric space is complete, a result that cannot be proven in the context of the Riemann integral. Furthermore, the set of continuous functions on [a,b] is dense in this metric space. This leads into the subject of orthonormal expansions, including Fourier series and Tschebysheff polynomials, which are a principle topic of numerical analysis, as well as Hermite polynomials used in quantum mechanics.

For functions of more than one variable we have Fubini's Theorem that double integrals of Lebesque integrable functions of two (or more) variables can be iterated. Furthermore, by the Tonelli-Hobson Theorem, the order of integration can be switched if either of the iterated integrals converges absolutely. Of interest to electrical engineers is the theorem that the Laplace transform of the convolotion of two functions is the product of their Laplace transforms.