494/598 Fall 2008

WEEK 4: REVIEW OF ANALYSIS.

1. Convergence and Completeness in Metric Spaces

The reference for this section is [1].

Now that we have seen Fourier analysis in the finite-dimensional context of $\ell^2(\mathbb{Z}_n)$, we would like to move on to Fourier analysis on intervals and on the real line. Unfortunately, the pure linear algebra approach we used in finite-dimensional spaces will not work in function spaces on the real numbers, which are infinite-dimensional. The development of Fourier theory in such infinite-dimensional settings requires tools from analysis and integration theory, which the next several sections of notes will introduce.

For the most part, the results in this section bear a strong resemblance to the material covered in advanced calculus, but they will be more abstract. The real line was the original setting in which mathematicians studied distance, convergence, continuity, and other ideas. But in order to work spaces other than the real line, which we must in order to use Fourier analysis, we need more abstract ideas of distance— those presented in the section on metrics, norms, and inner products—and we need to understand the ideas of convergence and continuity in these contexts.

The proofs of most of the results listed below are similar to the proofs of the analogous facts about real numbers, so for conciseness we will largely omit proofs.

In this section we will consider a metric space A with a metric d, and denote this as a pair (A, d). Recall that the metric or distance d allows us to measure the 'distance' between two points in a space.

Definition 1.1. Let (A, d) be a metric space. Suppose $x \in A$ and $r \geq 0$. Denote $B_r(x) = \{y \in A : d(x,y) < r\}$ and $\bar{B}_r(x) = \{y \in A : d(x,y) \leq r\}$. We refer to $B_r(x)$ as the *open ball* of radius r centered at x, and we refer to $\bar{B}_r(x)$ as the *closed ball* of radius r centered at x.

Now that we defined an open ball, we want to extend the notion of 'openness' to any set. If we consider an open and closed balls, then the difference between them comes from the fact that the closed ball has a boundary. However, a boundary is another term which we would have to define for an abstract set, so we will consider a notion of 'openness' from a different point of view, namely, each point in the open ball comes with a neighborhood, in other words, each point comes with an open ball around it of some small radius. Note that this is not the case for the closed ball, a point which is on a boundary of the closed ball has only part of an open ball in the closed ball. So in terms of rigorous mathematics, this can be defined as follows:

Definition 1.2. Let (A, d) be a metric space. A set $G \subset A$ is an *open* set in A if, for every $x \in G$ there is an $r > 0$ such that $B_r(x) \subset G$.

Thus, every point $x \in G$ has an open ball $B_r(x)$ for some small r which is in G.

Our next question is if we have a sequence of point (we will assume that it is infinite), can this sequence 'approach' a point, or 'converge' to a point? In order to answer this question, we should define what 'convergence' would mean.

Definition 1.3. In a metric space A, we say that a sequence $\{x_n\}$ *converges* to $x \in A$ if for every $\epsilon > 0$ there is an $N \in \mathbb{N}$ such that for every $n \geq N$,

$$d(x, x_n) < \epsilon.$$

In this case, x is called the limit of $\{x_n\}$.

This is denoted by several ways, most commonly $x_n \to x$ or $\lim\limits_{n\to\infty} x_n = x$. Note that convergence means that if take any small neighborhood of x (for example, an open ball of a radius ϵ), then the tail of this sequence starting from x_N will be contained in this open ball.

How essential is it to have $d(x, x_n) < \epsilon$ in the above definition? The following remark will clarify it.

Remark 1.4. Since the choice of $\epsilon > 0$ was arbitrary, we could replace the strong inequality $d(x, x_n) < \epsilon$ with a weak inequality, meaning, with $\leq$ sign. That is, in order to prove convergence, it suffices to show that for each $\epsilon > 0$ there is $N \in \mathbb{N}$ such that whenever $n \geq N$, $d(x, x_n) \leq \epsilon$. Also note that the limit of a sequence, if it exists, is unique. This uniqueness can be proved by assuming that a sequence $\{x_n\}$ converges to both x and y, and then showing that for any $\epsilon > 0$ we must have $d(x, y) < \epsilon$, which implies that $d(x, y) = 0$, and hence, $x = y$.

Several points about convergence are worth making. One is that, as with real sequences, the N which we choose depends on ϵ, in general, the smaller ϵ, the larger N we have to choose (i.e., further in the tail we should go). Another critical observation is that the definition of convergence in a metric space A requires that the limit point x must be an element of A. Recall that if (A, d) is a metric space and $B \subset A$, then $(B, d|_B)$ is also a metric space. However, suppose $\{b_n\}$ in $B \subset A$ converges in (A, d) to $a \in A \backslash B$, where $A \backslash B$ indicates the set difference of A and B, $\{x \in A : x \notin B\}$. Then clearly, since $a \notin B$, $\{b_n\}$ does not converge to a in $(B, d|_B)$. For a more concrete case, consider a sequence of rational numbers converging to a real number, e.g., the successive decimal approximations to π, which converge to π in the real numbers. This sequence does *not* converge in the rationals, because, if it had a rational limit r, that limit would have to be the limit of the sequence in the real numbers, and since π is not rational, this is impossible.

A major disadvantage of the definition of convergence: it explicitly depends on the limit. Often we do not know the limit, and we want to be able to tell whether a sequence converges by looking just at the terms of the sequence itself. This is not possible in every space; some spaces, like those mentioned above, have 'holes' in them, which means that sequences that 'look' like they should converge wind up not converging to anything in the space.

Definition 1.5. Let A be a metric space. We say a sequence $\{x_n\}$ is *Cauchy* in A if for each $\epsilon > 0$ there is an $N \in \mathbb{N}$ such that for any $n, m \geq N$,

$$d(x_n, x_m) < \epsilon.$$

Although it is not true in every metric space that Cauchy sequences must be convergent, the converse is true:

Proposition 1.6. *Let $\{x_n\}$ converge to x in a metric space (A, d). Then $\{x_n\}$ is Cauchy in A.*

Proof. Suppose $x_n \to x$ in a metric space (A, d). Let $\epsilon > 0$. There is an $N \in \mathbb{N}$ such that for every $k \geq N$, $d(x_k, x) < \frac{\epsilon}{2}$. Thus, for each $m, n \geq N$, the triangle inequality gives us

$$d(x_m, x_n) \leq d(x_n, x) + d(x_m, x) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

and thus, $\{x_n\}$ is Cauchy in A. ☐

That not every Cauchy sequence need to converge is easy to prove using the comments made before we discussed Cauchy sequences. Once again, take the example of the decimal approximations to π. Since the sequence converges in $\mathbb{R}$, it must be Cauchy in $\mathbb{R}$, and hence, also in $\mathbb{Q}$. But since limits are unique, the sequence cannot have a rational limit, and does not converge in $\mathbb{Q}$. Again, the central problem here is that $\mathbb{Q}$, in some sense, has 'holes' in it. One of the chief reasons for the development of the theory of the real numbers was to be able to work in a space that did not have such holes in it, an idea that we make precise below:

Definition 1.7. We say that a metric space (A, d) is *complete* if every Cauchy sequence in (A, d) converges in (A, d).

Most of the metric spaces we will work with in this class—the complex numbers, $\mathbb{C}^n$, and the function spaces we will work with—are complete. This proves to be an enormously valuable theoretical tool, because we can prove that sequences in these spaces converge without actually knowing the limit, which in most cases we do not. We simply need to be able to prove that the sequence is Cauchy, which is usually much easier to do.

It is worth noting that one way of defining the real numbers is to simply define them as the limits of all Cauchy sequences of rational numbers, the *completion of* $\mathbb{Q}$. This is a technique which can be adopted in general metric spaces, so that any metric space can be made complete using such a process. We will have no need, however, for this theoretical technique.

1.1. **Series in Normed spaces.**

We next want to sum elements in a vector space, but for that we need to have a norm. Thus, next we are going to deal with normed spaces: let X be a normed space with a norm $\|\cdot\|$. We will denote this pair by $(X, \|\cdot\|)$.

Definition 1.8. Let $(X, \|\cdot\|)$ be a normed vector space, and let $\{x_n\}$ be a sequence of elements of X. For each $N \in \mathbb{N}$, define

$$s_N = \sum_{n=1}^{N} x_n,$$

which is a well-defined finite sum. If there is an $s \in X$ such that $s_N \to s$, or $\lim_{N\to\infty} s_N = s$, then we say that the *series*

$$\sum_{n=1}^{\infty} x_n$$

converges to s.

It is clear from the definition that, as in the real line, a necessary (but not sufficient, as we know from the example of the harmonic sequence in $\mathbb{R}$) condition for the series $\sum x_n$ to converge, is for x_n to converge to the zero vector. One condition which can ensure convergence in a complete space, and even offers an alternative definition of completeness, is that of absolute convergence:

Definition 1.9. Let $(X, \|\cdot\|)$ be a normed vector space, and let $\{x_n\}$ be a sequence of elements of X. If the sequence of positive real numbers

$$\sum_{n=1}^{\infty} \|x_n\|$$

converges to some real number $M < \infty$, then we say that the sequence $\{x_n\}$ is *absolutely convergent.*

Proposition 1.10. *The space X is complete if and only if every absolutely convergent sequence $\{x_n\}$ converges in X.*

Proof. Exercise (see Assignment) ☐

One final concept related to series, which will prove important for our development of Fourier series, is that of summing a sequence of integers (sometimes referred to as a *bi-infinite* sequence):

Definition 1.11. Let $(X, \|\cdot\|)$ be a normed vector space, and let $\{x_n\} : n \in \mathbb{Z}$ be a sequence of elements of X, that is, a function from the integers (rather than just the natural numbers) to X. For each $N \in \mathbb{N}$, define

$$s_N = \sum_{n=-N}^{N} x_n,$$

again a well-defined finite sum. If there is an $s \in X$ such that $s_N \to s$, then we say that the series

$$\sum_{n=-\infty}^{\infty} x_n = \sum_{n\in\mathbb{Z}} x_n$$

converges to s.

Does it matter in which order we sum the sequence $\{x_n\}$? It really does. We can change the order of summation only if the corresponding (bi-infinite) sequence of real numbers converges when we replace x_n with $\|x_n\|$. In the context of Fourier series we will often not expect to have such absolute convergence.

1.2. **More on Metric, Normed and Inner product spaces.**

Before we wrap up this section, it is worth defining a few more ideas which should be familiar from advanced calculus in the context of generic metric spaces.

Definition 1.12. Let (A, d) be a metric space, and $B \subset A$. We say a subset B is *bounded* if there is an $M > 0$ and an $a \in A$ such that $B \subset B_M(a)$.

Definition 1.13. Let $\{x_n\}$ be a sequence in a metric space (A, d). We say $\{x_n\}$ is *bounded* if its image—that is, the set of all values taken by the sequence—is bounded. Equivalently, $\{x_n\}$ is bounded if there is an $M > 0$ and an $a \in A$ such that for all n, $x_n \in B_M(a)$.

Example 1.14. Consider a set $\{\sin n\}_{n=0}^{\infty}$. This is a bounded set, since $|\sin n| \leq 1$ for all n.

Definition 1.15. Let $\{x_n\}$ be a sequence in a metric space (A, d) and $\{n_j\}$ be a strictly increasing sequence from $\mathbb{N}$ to itself—in other words, for all $i, j \in \mathbb{N}$, $i > j$ implies that $n_i > n_j$. Then the sequence $\{x_{n_j}\}$ obtained from using the sequence $\{n_j\}$ to select elements of the sequence $\{x_n\}$ is called a *subsequence* of $\{x_n\}$.

Definition 1.16. A point x in a metric space X is a *cluster point* of x if, for each $\epsilon > 0$, the set $X \cap B_\epsilon(x) \backslash x$ is nonempty.

Definition 1.17. A subspace S of a metric space X is *closed* if it contains all its cluster points.

Definition 1.18. A metric space X, (or a subspace X of a metric space Y, or a subset X of Y) is *compact* if every bounded sequence in X has a convergent subsequence.

Example 1.19. A set $\{\sin n\}_{n=0}^{\infty}$ is not compact, but $\{\frac{1}{n}\}_{n=1}^{\infty}$ is.

In a similar manner as on the real line, it is possible to prove that a compact space is a space for which every open cover has a finite subcover. At this point, we should note that continuity can be defined on general metric spaces, in roughly the same manner that it is defined on the real numbers, in terms of sequences, $\epsilon - \delta$ interactions, or inverse images of open sets. The interaction between compactness and continuity is largely the same in general metric spaces as in the real numbers. Since we will have little use for these ideas in abstract spaces in this course (although they play an important role in graduate-level developments of the same material), we will elect not to cover these topics.

We conclude this section with two terms which will appear often in the remainder of the course, given that all of the metric spaces we will encounter are actually normed spaces, and many are also inner product spaces:

Definition 1.20. A normed space which is complete under the metric induced by its norm is called a *Banach space.*

Definition 1.21. An inner product space which is a Banach space under the norm induced by its inner product is called a *Hilbert space.* Equivalently, a complete inner product space is a Hilbert space.

For example, $\mathbb{R}^n$ is a Hilbert space under the usual dot product, and a Banach space under the associated Euclidean norm, but the vector space $\mathbb{Q}^n$, viewed as a vector space over the field $\mathbb{Q}$ (why not $\mathbb{R}$?) and equipped with the norm and inner product inherited from the Euclidean space $\mathbb{R}^n$, is not complete, hence, neither a Hilbert space nor a Banach space.

2. SURVEY OF LEBESQUE INTEGRATION

The references for this section are [2] and [3].

We assume that the reader is familiar with Riemann integration. Riemann integration is a useful tool, and the Fundamental Theorem of Calculus proved for Riemann integrals gives us a way to compute Riemann integrals. But the Riemann integral also several weaknesses.

The most obvious is that there are many relatively simple functions that we cannot integrate. For example, the function which is 1 on the rationals and 0 on the irrationals (such a function is called

Dirichlet function) is not Riemann integrable, but since the rationals are only a countable set, it seems logical that the integral ought to be 0. In fact, it can be shown that a function is Riemann integrable if and only if its discontinuities have measure zero, that is, they carry no n-dimensional mass (the idea of sets of measure zero will be elaborated upon in the exercises).

In many applications, of course, we are likely to work with functions that are continuous most of the time, so in some sense the weakness above is often not a problem. A more serious issue is that the Riemann integral need not behave well with respect to limits. Although the uniform limit of a sequence of Riemann integrable functions is Riemann integrable with integral equal to the limit of the integrals of that sequence, this result *only* holds for uniform limits. In many instances, we would like to be able to work with pointwise limits of functions, and the Riemann integral does not allow us to do this.

The Lebesgue integral eliminates both these weaknesses. The basic idea of the Lebesgue integral is that, instead of partitioning the function's domain into small pieces and estimating the volume with rectangles, we partition the function's range and measure the sets on which the function takes values within the different pieces of the range. Of course, this requires developing a theory of how to measure sets, and a theory telling us how to measure sets and how to determine which functions are 'measurable,' that is, which functions are regular enough for us to apply the above method to them (this is an issue because not all sets are measurable.) Measure theory is rather time-consuming to cover, so we will leave it to courses like MAT 473 and MAT 570. For our purposes, virtually any function will be measurable, as the construction of a non-measurable set, and hence a non-measurable function, requires use of the axiom of choice. Certainly anything we can easily write down will be measurable, as will be nearly function representing some physical phenomenon in the real world.

Before we begin listing properties of the Lebesgue integral, we list a few facts about measurability:

- The composition of a measurable function with a continuous or piecewise continuous function is measurable.
- As special cases of the above, if f and g are measurable, then fg, (f, g) (the function with f and g as component functions), $|f|$, $|f|^p$, and, if $g \neq 0$, f/g are.
- Linear combinations of measurable functions are measurable.
- If we make a differentiable or piecewise differentiable change of variables to a measurable function, the resulting function is measurable (e.g., a change to polar coordinates, a translation, rotation, or dilation, etc.).
- If $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{C}$, and for some $y \in \mathbb{R}^m$ we set $g(x) = f(x, y)$, then $g : \mathbb{R}^n \to \mathbb{R}$ will be measurable for almost every y (this is usually listed as part of the Fubini-Tonelli theorem, which we will present later). The function $g(\cdot) = f(\cdot, y)$ is called a *section* or a *slice* of f.
- The supremum, infimum, limit supremum, and limit infimum of a sequence of measurable functions are measurable. It follows that so is the pointwise limit if the sequence converges pointwise.

In general, nearly any operation we can write down explicitly will preserve measurability.

The Lebesgue integral first developed on nonnegative measurable functions, where it always exists, although it may take the value ∞. The integrals for general complex functions are then defined in terms of the integrals of the positive and negative pieces of their real and imaginary parts. This definition allows us to directly define the integral on all of $\mathbb{R}^n$ without having to go through the explicit limiting process that improper Riemann integrals require. On the other hand, this means that the direct Lebesgue integral is defined for an *absolutely integrable* function f, by which we mean a function for which $\int |f| < \infty$, whereas the improper Riemann integral can exist for functions which are not absolutely integrable but have cancelations between their positive and negative parts which make the limit exist. This occurs, for example, with the sinc function,

$$\operatorname{sinc}(x) = \frac{\sin(x)}{x},$$

which is not absolutely integrable but has finite Riemann integral. This is an important function in Fourier analysis. In the context of Lebesgue integration, an absolutely integrable function f is generally just called *integrable.*

The Lebesgue integral is defined on all of $\mathbb{R}^n$, but we can define the integral on measurable subsets $A \subset \mathbb{R}^n$ by simply letting $\int_A f = \int_{\mathbb{R}^n} f\chi_A$, where

$$\chi_A = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

is the *characteristic function* of A.

An important fact about the Lebesgue integral, which confirms that it agrees with our intuitive understanding integration, is given in the following proposition.

Proposition 2.1. *If f is Riemann integrable on a closed interval $[a, b]$, then the Lebesgue integral of f on $[a, b]$ is equal to the Riemann integral of f on $[a, b]$.*

This, among other things, means that we can use the familiar tricks involving antiderivatives to compute Lebesgue integrals of 'nice' functions.

Most of the other basic properties of the Lebesgue integral should be familiar from the setting of Riemann integration:

- If f is integrable, then $\left|\int f\right| \leq \int |f|$.
- Linear combinations of integrable functions are integrable, and the Lebesgue integral is linear. Using this fact, we can define a vector space L^1 of integrable functions, along with a norm $\|\cdot\|_1$ given by $\|f\|_1 = \int |f|$.
- The Lebesgue integral is continuous under translations. That is, if $R_k f(x) = f(x - k)$, then $\lim_{k\to 0} \int |f - R_k f| = 0$.
- $\int |f| = 0$ if and only if $f = 0$ except possibly on a set of measure zero.

The following theorems (convergence theorems), essentially the reason why Lebesgue integrals were developed in the first place, address the biggest weakness of the Riemann integrable, its lack of good behavior with respect to pointwise convergence.

Theorem 2.2. (FATOU LEMMA) *Let $\{f_n\}$ be a sequence of measurable functions with $f_n \geq 0$. Then*

$$\liminf_{n\to\infty} \int f_n \geq \int \liminf_{n\to\infty} f_n.$$

Note that if in the above statement the liminf is replaced with limsup, then the statement is false (can you find a counterexample?)

Next we state conditions under which the inequality sign can become equality, i.e., when it is possible to take limits under the integral sign.

Theorem 2.3. (LEBESGUE MONOTONE CONVERGENCE THEOREM) *Let $\{f_n\}$ be a sequence of measurable functions with $f_n \geq 0$. Suppose that $f_n \leq f_{n+1}$ for any index n. Then*

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} f_n = \int_{\mathbb{R}^d} f.$$

Before the next statement we need a definition:

Definition 2.4. If D is a set, we say that a nonnegative function $g : D \to \mathbb{C}$ *dominates* a collection of functions $\{f_\lambda : \lambda \in \Lambda\}$ if for each $\lambda \in \Lambda$ and each $x \in D$,

$$|f_\lambda(x)| \leq g(x).$$

The most common case is when the collection $\{f_\lambda\}$ is a sequence $\{f_n\}$.

Note that if an L^1 function g dominates a measurable function f, then f is L^1.

Theorem 2.5. (Lebesgue Dominated Convergence Theorem) *Let* $\{f_n\}$ *be a sequence of measurable functions* $\mathbb{R}^d \to \mathbb{C}$ *converging pointwise almost everywhere to a function* $f : \mathbb{R}^d \to \mathbb{C}$, *and suppose there is a nonnegative* L^1 *function* g *which dominates* $\{f_n\}$. *Then* $f \in L^1(\mathbb{R}^d)$, *and*

$$\lim_{n\to\infty} \int\limits_{\mathbb{R}^d} f_n = \int\limits_{\mathbb{R}^d} f.$$

Another important theorem tells us that we can evaluate integrals iteratively. For a function $f(x, y)$ on $\mathbb{R}^n \times \mathbb{R}^m$, the section $= f(\cdot, y)$ is measurable for almost every y. Fubini's theorem states that we can use this fact to evaluate Lebesgue integrals of integrable functions iteratively.

First, we make a remark on notation: when possible, it is customary to omit the argument in the integrand, and write the integral in the form $\int f$, as we have done above. Often, when we wish to make the variable of integration explicit, we write the integral using the familiar Leibniz notation, $\int f(x)dx$. Sometimes we also wish the measure to be made explicit, in which case we write $\int f(x)d\mu(x)$, where μ is the measure with respect to which we are integrating. This is sometimes necessary because it is possible to define Lebesgue integrals for measures other than the standard Lebesgue measure (we will have little need for that in this course, other than when using surface measure in polar coordinates).

Theorem 2.6. (Fubini's Theorem) *Let* $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{C}$ *be* L^1. *Then for almost every* y, *the section* $f(\cdot, y)$ *is* L^1, *and*

$$\int\limits_{\mathbb{R}^n\times\mathbb{R}^m} f(z)dz = \int_{-\infty}^{\infty} \int\limits_{\mathbb{R}^m} f(x, y)dydx,$$

where $z \in \mathbb{R}^n \times \mathbb{R}^n$, $x \in \mathbb{R}^n$, *and* $y \in \mathbb{R}^m$. [1]

An important property of the Lebesgue integral is needed to interpret Fubini's theorem: note that the inner integral need only be defined for *almost every* y, and yet we can still take the outer integral. This is a general fact about Lebesgue integration. A function need only be defined almost everywhere for us to take the integral, because the values on a set of measure zero, with no n-dimensional mass, do not affect the integral at all.

The following theorem extends Leibniz's rule for taking limits and derivatives inside integrals to Lebesgue integrals. Note that it requires a dominance condition similar to that in the dominated convergence theorem (which makes sense, since it is proved using the dominated convergence theorem)

Theorem 2.7. *Let* I *be an interval. Suppose* $f : \mathbb{R} \times I \to \mathbb{C}$, *with* $f(., t) \in L^1(\mathbb{R}^n)$ *for each* $t \in I$. *Let* $g \in L^1(\mathbb{R}^n)$, *and for each* t *set* $F(t) = \int_{-\infty}^{\infty} f(x, t)dx$.

i. *If* g *dominates* $f(., t)$ *for each* t *and if* $lim_{t\to t_0} f(x, t) = f(x, t_0)$ *for all* x, *then* $lim_{t\to t_0} \int_{-\infty}^{\infty} f(x, t) = F(t_0)$. *Note that if* $f(x, .)$ *is continuous for each* x, *then this implies that* F *is also continuous.*

ii. *If* $\partial f/\partial t$ *exists and* g *dominates* $\partial f/\partial t$ *for all* t, *then* $\int_{-\infty}^{\infty} f(x, .)$ *is differentiable with* $F'(t) = \int_{-\infty}^{\infty}(\partial f/\partial t)(x, t)dx$

2.1. **Change-of-Variables.**

We now state two theorems which justify the substitutions often used in calculus. The first covers just the simple but important case of a linear change of variables.

Theorem 2.8. (Linear Change-of-Variables) *If* $f\mathbb{R}^n \to \mathbb{C}$ *is measurable and* T *is an invertible linear transformation on* $\mathbb{R}^n$, *then* $f \circ T$ *is measurable, with*

$$\int_{-\infty}^{\infty} f \circ T = |\det T|^{-1} \int_{-\infty}^{\infty} f.$$

---

[1] There similar theorem, called Tonelli's (actually, used to prove Fubini's) stating that the result always holds for nonnegative functions, whether or not they are L^1. Some authors put the two together and call it the Fubini-Tonelli theorem.

Note that, in the special case where T is a rotation, this theorem states that Lebesgue integration is invariant under rotations. Similarly, from the case where T is a dilation by t, $T(x) = tx$, then we get that

$$\int_{-\infty}^{\infty} f(tx)dx = \frac{1}{t}\int_{-\infty}^{\infty} f(x)dx.$$

To give the general change-of-variables formula, we first need a definition:

Remark 2.9. When dealing with derivatives of functions $f : \mathbb{R}^m \to \mathbb{R}^n$, $D_x f(x_0)$ denotes the derivative of f at x_0, defined as the best affine approximation to f near x_0 (assuming it exists). By best affine approximation, we mean the liner map which, after adjusting by $f(x_0)$, approximates f most closely. We identify this map with the $n \times m$ matrix which represents it, sometimes called the *Jacobian matrix*; in practice we find this matrix by noting that its ijth entry will be $(\partial/\partial_{x_j} f_i$, where f_i is the ith component function of f. Unfortunately, existence of the partial derivatives of f at x_0 is not sufficient to guarantee differentiability of f at x_0, but it can be shown that if all of the partial derivatives of f exist and are continuous at x_o, then $D_x f(x_0)$ exists and is continuous.

Definition 2.10. If $U, V \subset \mathbb{R}^n$, then a function $G : U \to V$ is a *diffeomorphism* if it is a continuously differentiable bijection with continuously differentiable inverse.

Theorem 2.11. (Change-of-Variables) *Let $f : V \to \mathbb{C}$ be measurable and let G be a diffeomorphism $U \to V$. then $f \circ G$ is measurable, with*

$$\int_V f = \int_U \int_{-\infty}^{\infty} |\det D_x G|\, f \circ G(x)dx.$$

Note that, in the special case where G is a translation, so that its derivative is just the identity, we find that Lebesgue integration is invariant under translations. This theorem also justifies the substitutions often made in calculus.

2.2. **Polar coordinates and surface measure.**

Another special case is in shifting to polar coordinates. To be precise, let S^n be the unit sphere in R^n. We want to represent a point $x \in \mathbb{R}^n$ by a pair $(r, \omega) \in \mathbb{R} \times S^{n-1}$, where $x = r\omega$. Of course, this is not a bijection, because $r\omega = (-r)(-\omega.)$ But if we restrict r to $[0, \infty]$, then it is a bijection everywhere except at the origin, which is a set of measure zero and hence can be ignored in integration. We then define a surface measure σ on S^n, such that that if g is a function $S^n \to \{0, 1\}$, then $\int g d\sigma$ gives the surface area of $g^{-1}(1,)$ This measure is unique, and has a formula in terms of trigonometric functions—in the familiar cases of polar coordinates in $\mathbb{R}^2$ this is where the term $\sin\theta$ comes from—but we will not worry about details here.

It can then be shown, by the change-of-variables formula, and some measure theory, that for any $f \in L^1(\mathbb{R}^n)$,

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^{\infty} \int_{S^1} f(r\omega) r^{n-1} d\sigma(\omega)dr.$$

References

[1] Charles Epstein. *Introduction to the Mathematics of Medical Imaging.* Prentice Hall, Upper Saddle River, 2nd edition, 2007.

[2] Walter Rudin. *Real and Complex Analysis.* McGraw-Hill, New York, 3rd edition, 1987.

[3] Walter Rudin. *Real Analysis: Modern Techniques and Their Applications.* John Wiley and Sons, New York, 2nd edition, 1999.