Integration

John McCuan

April 9, 2020

Contents

1 Preliminaries

1.1 Dimension and Euclidean Models for Space

We say the real line $\mathbb{R} = \mathbb{R}^1$ has **dimension** one. This terminology extends to certain continuous images of $\mathbb{R}$ or intervals in $\mathbb{R}$. Thus, the curve

$$\Gamma = \{(\cos t, \sin t, t) : t \in \mathbb{R}\}$$

is a one dimensional object. This one dimensional object is realized as a subset of the three dimensional Euclidean space built from three copies of the line $\mathbb{R}$:

$$\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R} = \{(x, y, z) : x, y, z \in \mathbb{R}^1\}.$$

Three dimensional Euclidean space $\mathbb{R}^3$ provides the most common, and most intuitively natural, model for the physical space apparently observed and encountered in the world. Three dimensional subsets of $\mathbb{R}^3$, moreover, like the **ball**

$$B_r(\mathbf{p}) = \{\mathbf{x} \in \mathbb{R}^3 : |\mathbf{x} - \mathbf{p}| < r\}$$

are often used to model physical objects composed of "matter" apparently observed and encountered in the space of the world.

From the time of the famous Greek philosophers some people have felt the physical matter encountered in the world was composed of indivisible parts or "atoms." Whether the indivisibility was based upon composition or size was not known due primarily to observational limitations. Our current modeling of matter suggests **size** is the defining characteristic and the structure (or at least the most important characteristics determining macroscopic properties) of matter at various scales (mollecular, atomic, subatomic) is very different. This view of matter, in light of the mathematical requirements on an "object" M suitable to consider for integration, strongly suggests any very direct identification between material objects (consisting of matter composed of atoms) and "objets" M on which one integrates should be viewed with suspicion, and one should not be surprised if such an identification leads to assertions or predictions at odds with observation.

More broadly it is safer, I think, to put some distance between mathematics and the physical world. Indeed **mathematical modeling** involves an identification between mathematical concepts (or what I like to call "imaginary pictures") and observed physical phenomena. We have a function P satisfying a differential equation $P' = kP$, and we have a population, but those (as students of Malthus may observe) are not the same thing. The general philosophy then is that science is more about the **comparison** (and contrast) of imaginary pictures to observations than about the identification which is generally a rather loose one.

On the other hand, the identification between objects on which we integrate and material objects is often made. In such cases, one still has recourse to the mathematical concept of space (an imaginary picture) in which the model of matter will be a **subset**. Certainly we can say, in all cases where we consider

$$\int_M f,$$

the object M is a **set**. Some sets, generally those modeling space, are relatively simple, and other sets, for example those modeling matter, may be more complicated. We will begin with simpler sets.

1.2 Units, Particle Motion, Distrubution of Mass, and Motion of Mass

1.2.1 Point Mass

As suggesteed above, the modeling of matter in space is a delicate and difficult endeavor. Perhaps the simplest approach is to associate with a position $\mathbf{x}$ in space, say Euclidean three dimensional space $\mathbb{R}^3$, a certain positive number m. This may be called a **simple point mass**. Using this idealization, the situation when the position may change as a function of time and the mass remains constant is called **simple particle motion** and is of particular interest. The position may be described by a vector valued function of one variable $\mathbf{x} : [0, T] \to \mathbb{R}^3$ on a time interval starting at $t = 0$ or, more generally, by $\mathbf{r} : [a, b] \to \Sigma$ where Σ is a general space, which may for the moment be thought of as $\mathbb{R}^n$. If it makes sense to differentiate the position function with respect to time (twice), then one may posit Newton's second law of motion

$$\mathbf{f} = m\mathbf{r}''. \tag{1}$$

Generally, **mass** (in this case simply the number m) is considered to be measured by a **fundamental unit** denoted M. The symbol M here is not a number, but a

designation of the unit. We say that the mass m "has the units of mass," and write

$$[m] = M.$$

The square brackets here indicate the **specification of the units of a quantity**. This notation can be quite useful. Time is usually considered as another fundamental unit, and for a specific time t, we write $[t] = T$. That is, a specific time t has, naturally, the units of time. Each dimension, and one dimension in particular, in Euclidean space is considered to have a the fundamental unit of **length**. As we will see below, there is more to the notion of length than one might initially think, and some aspects of this concept may be illuminated and extended through the consideration of **one-dimensional Euclidean measures**. For now, however, let us simply note that a position $\mathbf{r}$ is generally considered to have units of length: $[\mathbf{r}] = L$. This may seem somewhat strange if one considers, for example, the motion of a particle moving along a path in $\mathbb{R}^3$ such as

$$\mathbf{r}(t) = (\cos t, \sin t, t),$$

because the motion is described by three real valued functions of t. The designation $[\mathbf{r}] = L$ arises from consideration of the motion. In simple particle motion, though the motion itself can take place in several ambient dimensions, the actual **path of motion** is, in some sense, inherently one dimensional. This is clear, of course, for motions $\mathbb{R} : (a, b) \to \mathbb{R}^1$ along a line. More generally, the **average velocity** from time t_1 to time t_2 with $t_1 < t_2$ is defined by

$$\mathbf{v}_{avg} = \frac{\mathbf{r}(t_2) - \mathbf{r}(t_1)}{t_2 - t_1}.$$

And the average speed

$$|\mathbf{v}_{avg}| = \frac{1}{t_2 - t_1}|\mathbf{r}(t_2) - \mathbf{r}(t_1)|$$

is naturally considered to have **derived units** $[\,|\mathbf{v}_{avg}|\,] = L/T$. Working backwards the Euclidean norm should not change units, so $[\mathbf{v}_{avg}] = L/T$. That is,

$$[\mathbf{r}(t_2) - \mathbf{r}(t_1)] = L.$$

Finally, the difference of quantities having the same units should have the same unit. And conversely, if a difference of a certain quantity, e.g., $\mathbf{r}(t_2) - \mathbf{r}(t_1)$ has units L, then we should have $[\mathbf{r}] = L$. The limit as $t_2 \searrow t_1 = t$ is, of course, the **instantaneous**

velocity at t given by $\mathbf{v} = \mathbf{r}'$ or sometimes (as Newton himself[1] would write it) $\mathbf{v} = \dot{\mathbf{r}}$ with derived units $[\mathbf{v}] = L/T$ as mentioned above.

Similarly **acceleration** $\mathbf{r}''$ and **force** should have **derived units** given by

$$[\mathbf{r}''] = \frac{L}{T^2} \qquad \text{and} \qquad [\mathbf{f}] = \frac{ML}{T^2}$$

according to, and in accord with, Newton's second law.

It should be emphasized that this "law" of Newton is but one possibility for creating a framework in which to consider the motion of a point mass. The concepts of velocity and acceleration are, in a certain sense, more fundamental.[2] Newton's law may be considered here as giving us some familiar structure to consider the consequences of modeling with a point mass. There could be other possibilities, and in some modeling with point masses *there are*. Under some circumstances, however, Newton's law gives us a tolerably good approximation of the motion of, say, a baseball or a golf ball. Imagine a baseball sitting next to a point mass which pretends to model the baseball. Let's say they are sitting on the surface of a table as in Figure 1. Compared to the model baseball (the point mass), the real baseball is infinitely tall.

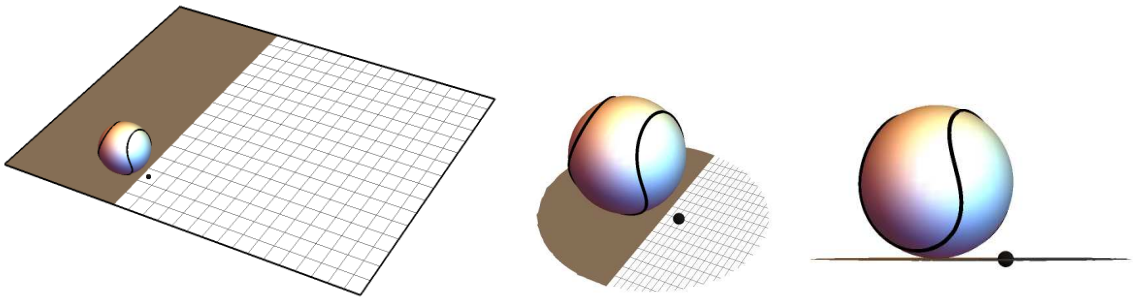

Figure 1: Baseball and a point mass as neighbors on a table

In fact, I've enlarged the representation of the point mass using a black dot, so we can even see it in the illustration. Notice that the **enlarged** graphic representation of the point mass extends through the surface of the model table top because the point mass itself is located on the surface of the table.

[1]The "prime" notation for derivatives $\mathbf{r}'$ is that used by Lagrange, and $d\mathbf{r}/dt$ is from Leibziz. The last one is especially important because it suggests the possibility of considering, for example, $d/dt : C^1 \to C^0$ as an **operator** independent of a specific function.

[2]And perhaps the contemplation of a timeless point mass is more fundamental yet.

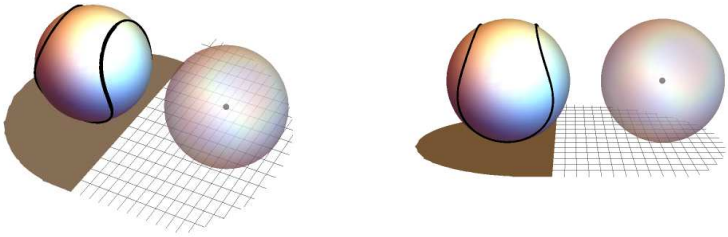

Figure 2: Baseball and a point mass with a set of occumpation as neighbors on a table

We may improve our model by assigning a subset of $\mathbb{R}^3$ to the point mass. For example, we can assign a ball $B_r(\mathbf{p}) = \{\mathbf{x} \in \mathbb{R}^3 : |\mathbf{x} - \mathbf{p}| < r\}$ of **exclusion** or a **set of occupation** to the point mass at $\mathbf{p}$. The model properties of this set may vary and the exclusion need not be absolute, but the simplest condition might be that nothing else (other than the point mass at the center) is allowed inside this sphere. Figure 2 illustrates the manner in which our model is improved with the inclusion of a set of occupation. This is an instance in which it is natural to consider the ratio

$$\rho = \frac{m}{4\pi r^3/3} \tag{2}$$

of mass to the volume of occupation. This ratio is, of course, a density ρ with units

$$[\rho] = \frac{M}{L^3}.$$

We will come back to the notion of density in more detail below.

More generally, a particle, or point mass, may be considered to have **time dependent mass** $m = m(t)$. Even in this case, the quantity $\mathbf{p} = m\mathbf{r}' = m\mathbf{v}$ is called **momentum** and $[\mathbf{p}] = ML/T$, though Newton's law (1) will no longer be valid as stated.

Exercise 1. *Give some alternative laws of motion (alternative to Newton's second law) giving interesting dynamics with application to video games.*

The **kinetic energy** of a point mass m is defined to be

$$\text{K.E.} = \frac{1}{2}m|\mathbf{v}|^2 = \frac{1}{2}m|\dot{\mathbf{r}}|^2.$$

1.2.2 Compound Point Mass

As we may see below, in some applications involving time dependent mass it can be useful to distinguish two components within a single time dependent point mass. Thus, we partition the positive number $m = m(t)$ as

$$m(t) = \nu(t) + \alpha(t)$$

where $\nu(t)$ is a positive (or at least non-negative) function considered to be the **native mass** at the position $\mathbf{r}$ of the mass and α is the **generated** or **derived** mass. These two functions have essentially identical properties and both are naturally considered as time dependent **singular measures** on the ambient space Σ. We will discuss general properties of measure later, but for now one can think of a measure as a function

$$\nu : \mathcal{P}(\Sigma) \to [0, \infty)$$

defined on the collection of **all subsets** $\mathcal{P}(\Sigma)$ of the space Σ having the property that for $E \in \mathcal{P}(\Sigma)$, i.e., for all $E \subset \Sigma$,

$$\nu(E) = \begin{cases} \nu\{\mathbf{r}\}, & \text{if } \mathbf{r} \in E \\ 0, & \text{if } \mathbf{r} \notin E. \end{cases}$$

The distinction between natural mass ν and generated mass α is basically useful for modeling mass added (or created) by a specific process. In general, the distinction between the native mass ν and the created mass α is not so important, but rather the rate of change of a specific portion α of m with respect to time is the quantity of interest:

$$\gamma(t) = \frac{d\alpha}{dt} \qquad \text{with} \qquad [\gamma] = \frac{M}{T}.$$

This rate of change γ , according to which

$$m(t) = \nu(t) + \alpha(0) + \int_0^t \gamma(\tau)\, d\tau$$

is called the **imposition** or **external forcing** on the mass.

1.2.3 Two Particle Interaction

If two point masses are restricted to $\mathbb{R}^1$ (or other ambient spaces as well), then one can model their interaction in a variety of ways. One possibility is that they have

no interaction. In this case, one simply gets the dynamics of two (or more) isolated particles taking place at the same time. While the interaction is not interesting in this case, consideration of the superimposed dynamics can still be interesting.

At the other extreme is **completely elastic collision**: If m_1 moving with constant velocity v_1 meets a mass m_2 with velocity v_2 at $x = 0$ and a certain time, say time $t = 0$, then we postulate that each mass m_j exits the collision with a velocity w_j satisfying **conservation of momentum**

$$m_1 w_1 + m_2 w_2 = m_1 v_1 + m_2 v_2$$

and **conservation of kinetic energy**

$$\frac{1}{2} m_1 w_1^2 + \frac{1}{2} m_2 w_2^2 = \frac{1}{2} m_1 v_1^2 + \frac{1}{2} m_2 v_2^2.$$

Rearranging these equations and canceling common factors we have

$$m_1(w_1 - v_1) + m_2(w_2 - v_2) = 0$$

and

$$m_1(w_1 - v_1)(w_1 + v_1) + m_2(w_2 - v_2)(w_2 + v_2) = 0.$$

Thus, as long as $w_2 - v_2 \neq 0$,

$$w_1 + v_1 = w_2 + v_2.$$

Thus, we have a linear system of two equations for w_2 and w_1:

$$\begin{cases} m_1 w_1 + m_2 w_2 &= m_1 v_1 + m_2 v_2 \\ w_1 - w_2 &= -(v_1 - v_2). \end{cases}$$

Therefore,

$$w_1 = \frac{m_1 v_1 + m_2 v_2 - m_2(v_1 - v_2)}{m_1 + m_2} = \frac{(m_1 - m_2)v_1 + 2m_2 v_2}{m_1 + m_2}$$

and

$$w_2 = \frac{m_1(v_1 - v_2) + m_1 v_1 + m_2 v_2}{m_1 + m_2} = \frac{2m_1 v_1 + (m_2 - m_1)v_2}{m_1 + m_2}.$$

A third possibility is that the two point masses maintain some **field relation with each other** as one finds in the modeling of gravitational interaction or the modeling of charged particles. Finally, an interesting possibility is that the point masses maintain a fixed rigid distance from one another. Such particles are naturally considered as a single "rigid body."

Exercise 2. *Given two point masses m_1 and m_2 at positions $\mathbf{r}$ and $\mathbf{q}$ respectively and with $|\mathbf{q}-\mathbf{r}| = d > 0$ fixed, determine the center of mass and total linear momentum*

$$\mathbf{p} = m_1\mathbf{r}' + m_2\mathbf{q}'$$

of the associated rigid body. What can you say about restriction on the position $\mathbf{q} = \mathbf{q}(t)$ in terms of $\mathbf{r} = \mathbf{r}(t)$ and the total momentum $\mathbf{p}$? Hint: Square the relation $|\mathbf{q}-\mathbf{r}| = d$ and differentiate with respect to t

1.2.4 Many Particles

Extending our suggestion of a rigid or **lattice structure** between two particles to many more particles, we find some interesting possibilities. The example of a model baseball of mass m with an occupying sphere $B_r(\mathbf{p})$ mentioned above can be modified, and perhaps improved in some sense, if we imagine the single mass m at the center replaced by a large number n of point masses each of size m/n, more or less evenly dispersed inside $B_r(\mathbf{p})$ and having fixed and rigid geometric relation to one another and to $\mathbf{p}$. Consideration of this kind of mass dispersion in a ball and in occupying volumes $\mathcal{V}$ of other shapes leads naturally to the notion of a **density function**

$$\rho(\mathbf{x}) \equiv \rho\,\chi_{\mathcal{V}}(\mathbf{x}) = \begin{cases} \rho, & \mathbf{x} \in \mathcal{V} \\ 0, & \mathbf{x} \in \mathbb{R}^3 \backslash \mathcal{V} \end{cases}$$

considered as a candidate for integration on $\mathcal{V}$. Since we have not yet considered carefully such integration on volumes in $\mathbb{R}^3$ and the process we have suggested is complicated, let us briefly consider in detail the one dimensional analogue which is relatively simple.

We begin with a point mass m (perhaps $m = m(t)$) located at a point $\mathbf{p}$ in some set of occupation which we take to be an interval $I = (a, b)$ of length ℓ. Note that the interval I is a natural analogue of the ball mentioned above in connection with modeling a baseball. The analogue of the density (2) associated with this point mass m in the interval of occupation is

$$\rho = \frac{m}{\ell} \qquad \text{with} \qquad [\rho] = \frac{M}{L}.$$

Let us disperse this point mass into k point masses of equal size m/k located at points $x_1, x_2, \ldots, x_k \in I$ satisfying

$$a = x_0 < x_1 < x_2 < \cdots < x_k < x_{k+1} = b$$

with

$$x_{j+1} - x_j = \frac{\ell}{k+1}, \qquad j = 0, 1, \ldots, k.$$

Each such distrubution of point masses determines a **singular measure**

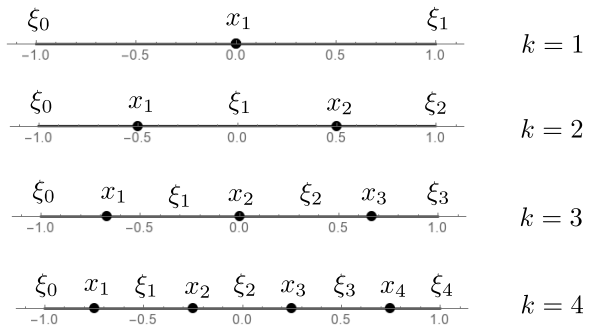

Figure 3: Distribution of point masses. At the k-th stage we may consider k point masses each with individual mass m/k. Note: This is a slightly different partitioning scheme and mass placement scheme for the interval than the one described. There are many different possibilities.

$$\nu_k : \mathcal{P} \to [0, m] \qquad \text{by} \qquad \nu_k(E) = \frac{m}{k} \, \#(\{x_1, x_2, \ldots, x_k\} \cap E) \tag{3}$$

where as usual $\mathcal{P}$ denotes the collection of all subsets of I, that is $\mathcal{P} = \{E : E \subset I\}$ and $\# : \mathcal{P} \to [0, \infty]$ is **counting measure** which assigns to any set the number of points in the set, i.e., the cardinality of the set. We will discuss measures and integration in more detail below, but what is being accomplished here should at least be intuitively clear. In particular, one recognizes here the well-known one dimensional Dirac measure

$$\delta_x : \mathcal{P} \to [0, 1] \qquad \text{given by} \qquad \delta_x(E) = \begin{cases} 1, & x \in E \\ 0, & x \notin E. \end{cases}$$

Thus we can also write

$$\nu_k = \sum_{j=1}^{k} \frac{m}{k} \, \delta_{x_j}.$$

Each of these (singular) measures ν_k has associated with it an integration according to which a smooth function $\phi : I \to \mathbb{R}$ with compact support in I, i.e., $\phi \in C_c^\infty(I)$ has an integral

$$\int \phi = \sum_{j=1}^{k} \frac{m}{k}\, \phi(x_j) \tag{4}$$

with respect to the measure ν_k. To emphasize the dependence on the measure, which should definitely be noted, it is customary to denote the integral appearing in (4) by

$$\int_I \phi \, d\nu_k.$$

Exercise 3. *Observe that the integral*

$$\int_I f \, d\nu_k,$$

with respect to the measure ν_k of **any** *pointwise well-defined function $f : I \to \mathbb{R}$ makes sense. Write down a formula for how to compute this integral, and show*

$$N_k : F \to \mathbb{R} \qquad \text{by} \qquad N_k[f] = \int_I f \, d\nu_k$$

defines a **linear functional**[3] *on the vector space F of real valued functions $f : I \to \mathbb{R}$.*

Exercise 4. *Restricting the functional N_k of the previous exercise to the the subspace $C_c^\infty(I)$, show that given any $\phi \in C_c^\infty(I)$, we have*

$$\lim_{k\to\infty} N_k[\phi] = \int_I \phi \, \frac{m}{\ell}$$

where the integral on the right may be taken as a Riemann integral:

$$\int_I \phi \, \frac{m}{k} = \int_a^b \phi(x) \, \frac{m}{\ell} \, dx \qquad \textit{where } I \textit{ has endpoints } a \textit{ and } b,$$

or as an integral of ϕ times the constant density function $\rho \equiv m/\ell$ with respect to Lebesgue measure, which is the "usual" measure on an interval of the real line discussed in more detail below.

[3]This just means a linear function with codomain the real numbers

The last exercise shows

$$\lim_{k\to\infty} \nu_k = \frac{m}{\ell}\,\mu \qquad \text{where } \mu \text{ is Lebesgue measure.}$$

Notice how the convergence works: Associated with each measure ν_k is a functional $N_k : C_c^\infty(I) \to \mathbb{R}$ and, associated to the scaling $\rho\mu = m\mu/\ell$ of Lebesgue meausure, there is a functional $\rho M : C_c^\infty(I) \to \mathbb{R}$. Thus, we are "naturally" led to the consideration of **limits of measures**, and the content of what it means for a sequence of measures to converge (at least weakly) to a given measure is that the values of the associated functionals converge for every $\phi \in C_c^\infty(I)$.

Exercise 5. *Show that given a non-negative continuous*[4] *function* $\rho \in C^0(\overline{I})$ *with*

$$\int \rho = m$$

it is possible to define a sequence of distributed Dirac measures:

$$\nu_k = \sum_{j=1}^{k} g_j\,\delta_{x_j}$$

with $\nu_k(I) = \sum g_j = m$ *such that*

$$\lim_{k\to\infty} \nu_k = \rho\,d\mu \qquad \textit{where } \rho\,d\mu \textit{ is the measure associated with } \phi \mapsto \int \phi\,\rho\,d\mu.$$

Hint: Use the partitioning scheme indicated in Figure 3.

Exercise 6. *Try to show that given a functional* $\mathcal{N} : C_c^\infty(I) \to \mathbb{R}$ *which is given by integration against a measure* ν*:*

$$\mathcal{N}[\phi] = \int \phi\,d\nu$$

the measure $\nu(J)$ *of every interval* $J \subset I$ *can be* extracted *from the values of the functional* N*. If this is not true, what goes wrong?*

[4]One might attempt to consider a less regular function here. Perhaps a natural choice would be a non-negative function $\rho \in L^1(I)$ the space of Lebesgue integrable functions considered below. In this case, it is much more difficult to extract the appropriate values g_j from the function ρ.

1.3 Extensive Quantities and Flux

Any quantity, like mass, which can be obtained by integrating a density function is called an **extensive quantity**.

A one dimensional **spatial flux function** along an interval of the real line is a function $\phi_1 = \phi_1(x,t)$ giving the **quantity per time** flowing in a certain direction at the point x and at time t. To better understand this idealization/model, it may be helpful to imagine the real line as the x-axis in Euclidean three dimensional space $\mathbb{R}^3$ enlarged to a solid cylinder. We do not mean a "pipe" per se, but rather a thin column composed of whatever quantity is flowing. If the quantity is a volume of fluid, for example, then we can imagine the cylinder as composed of a fluid whose cross-sections are homogeneous but may vary in density with position x along the cylinder and may also move along the length of the cylinder. A certain cross-section, imagined as either moving or stationary, may also be imagined to change in density with time, but the crucial assumption here is that every cross-section remains **homogeneous**. Such changes may create a change in the **spatial flux** ϕ_1 whose dimensions in the

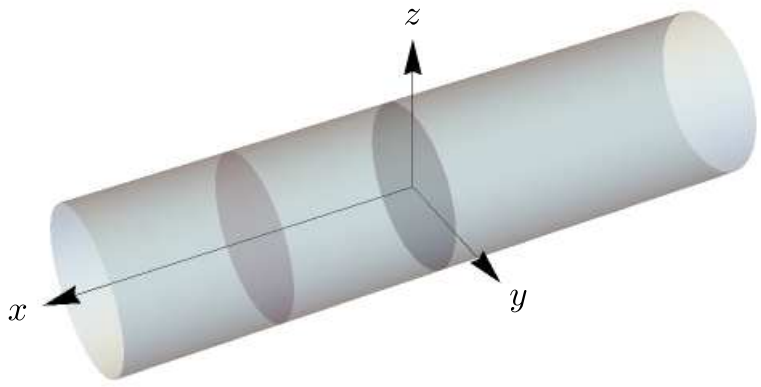

Figure 4: Fluid cylinder composed of homogeneous cross-sections of distinct densities and possibly moving

example of a fluid are typically mass per time:

$$[\phi_1] = \frac{M}{T}.$$

In our enlarged cylinder, the flux ϕ_1 may be further decomposed in terms of an areal

mass density flux $\phi = \phi_1/A$ where A is the cross-sectional area of the cylinder with

$$[\phi] = \frac{M}{L^2T}.$$

If the corresponding volumetric mass density $\rho = \rho(x,t)$ with units $[\rho] = M/L^3$ is given similarly in terms of a lineal mass density $\rho_1 = \rho A$, then the total mass between $x = a$ and $x = b$ at time t is

$$\int_a^b \rho_1(x,t)\,dx. \tag{5}$$

If $\phi > 0$ gives a rate of mass transfer, i.e., flux, in the direction of increasing x, and $\phi < 0$ indicates mass transfer in the opposite direction, then **conservation of mass** with respect to the region between the planes $x = a$ and $x = b$ is expressed by

$$\frac{d}{dt}\int_a^b \rho_1(x,t)\,dx = -\phi_1(b,t) + \phi_1(a,t). \tag{6}$$

More generally, we may assume a certain quantity of mass is internally generated or eliminated. This mass creation/depletion may be modeled by a **volumetric mass density per time** $Q = Q(x,t)$ with

$$[Q] = \frac{M}{L^3T}.$$

Again, the assumption of homogeneity of generation/elimination with respect to the cross-sections leads to a lineal density $Q_1 = QA$. If $Q > 0$ quantifies the generation/addition of mass at position x and time t while $Q < 0$ quantifies the elimination of mass, then the rate of change of the mass between the planes $x = a$ and $x = b$ with respect to time becomes

$$\frac{d}{dt}\int_a^b \rho_1(x,t)\,dx = -\phi_1(b,t) + \phi_1(a,t) + \int_a^b Q_1(x,t)\,dx. \tag{7}$$

The flow of many other extensive quantities may be described by a density/flux relation. Perhaps the simplest possibility to consider is the flux of **volume** itself. This corresponds to the motion of an **incompressible fluid** with constant density, and the relation (7) simplifies to

$$\phi_1(b,t) = \phi_1(a,t) + \int_a^b Q_1(x,t)\,dx$$

which says that the volume per time exiting across the plane $x = b$ is equal to the volume per time entering the region between the planes plus the volume injected/created.

Positive flux ϕ_1 on $\mathbb{R}^1$ generally models motion to the right, but that is merely a matter of convention.

On the line, we can think of the flux as a vector field tangent to the line so that $\mathbf{e}_1 = (1, 0) \in \mathbb{R}^2$ or $\mathbf{e}_1 = (1, 0, 0)$ and the **vector flux** is given in terms of the **scalar flux** ϕ_1 by

$$\phi_1 \mathbf{e}_1.$$

This interpretation lends itself to the consideration of a flux along a curve or path in a higher dimensional ambient space.

Before we leave this section, we mention in regard to the expression for the total mass in (5) that we will consider integrals over higher dimensional sets below. In that context, the mass associated with a volume $\mathcal{V}$ modeling a mass distributed with density ρ is given by

$$\int_{\mathcal{V}} \rho.$$

And if $\mathcal{V}$ is portion of the cylinder shown in Figure 4 between $x = a$ and $x = b$, then this volume integral may be expressed as an iterated integral

$$\int_{\mathcal{V}} \rho = \int_a^b \left(\int_{B_r(\mathbf{0})} \rho \right) \, dx = \int_a^b \rho \, A \, dx = \int_a^b \rho_1 \, dx$$

where $B_r(\mathbf{0}) = \{(y, z) : y^2 + z^2 < r^2\}$ is a disk of radius r representing the cross-section of the cylinder and having area A. Notice that ρ comes out of the inner integral in the second expression owing to the assumed homogeneity in each cross-section, and

$$\int_{B_r(\mathbf{0})} \rho = \rho \int_{B_r(\mathbf{0})} 1 = \rho \, A = \rho_1.$$

2 Riemann Integration in One Dimension

2.1 Riemann Sums

2.2 Riemann Integrals

2.3 Intepretations

2.3.1 Accumulated Displacement

2.3.2 Density

2.3.3 Time Dependent Integrals

3 Riemann Integration on Objects

In the second semester of calculus (Calc II) one learns about the Riemann integral on intervals:

$$\int_a^b f(x)\,dx.$$

The construction used to define the Riemann integral on intervals may be generalized to apply to integration on a variety of sets, which we are calling "objects." It is useful to have an idea of the general abstract construction of such integrals so that one can apply it to objects such as

1. regions (areas) in $\mathbb{R}^2$,
2. volumes in $\mathbb{R}^3$,
3. regions in $\mathbb{R}^n$ for $n > 3$ (i.e., "hypervolumes" or n-dimensional manifolds),
4. curves in $\mathbb{R}^n$,
5. surfaces in $\mathbb{R}^n$.

Integration on all these "objects" is possible, and we describe such integration here in abstract terms.

4 Abstract Integration

Several ingredients are required. Most are very easy to understand. We need an object, which we will call, generally, $\mathcal{V}$ and a real valued function defined on the object:

$$f : \mathcal{V} \to \mathbb{R}.$$

The function f assigns a real number to each point in $\mathcal{V}$. The basic concept we want to define/discuss is called

Integrating a real valued function f on an object $\mathcal{V}$.

Partition

Next we need a way to cut the object $\mathcal{V}$ up into small pieces. The set of pieces together is called a **partition** and is sometimes denoted by

$$\mathcal{P} \qquad \text{or} \qquad \{\mathcal{V}_j\}.$$

There should be finitely many pieces in the partition, and they are indexed by j. Thus, $\{\mathcal{V}_j\}$ means $\{\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3, \ldots, \mathcal{V}_k\}$ where k is some finite number. An important thing to remember about a partition is that the **union** of the partition pieces gives the entire object. In mathematical notation this is expressed by writing

$$\mathcal{V} = \cup_j \mathcal{V}_j.$$

As a technical point, it's often okay for the partition pieces to overlap a little bit, but you don't want them to overlap too much. We will mention this technical point again below.

Measure

We also need a way to measure the size of the partition pieces. Technically, we often need *two ways* to measure partition pieces, but we'll get to that in a moment. The first way, we'll call a **measure** and denote the size of a piece by

$$\text{meas}(\mathcal{V}_j).$$

If $\mathcal{V} = [a, b]$ is an interval, then we can partition into subintervals $\mathcal{V}_j = [x_{j-1}, x_j]$ where $\mathcal{P} : a = x_0 < x_1 < x_2 < \cdots < x_k = b$, and we can use length for the measure:

$\text{meas}(\mathcal{V}_j) = x_j - x_{j-1}$. If $\mathcal{V}$ is an area in $\mathbb{R}^2$, then we will want $\text{meas}(\mathcal{V}_j) = \text{area}(\mathcal{V}_j)$. You'll note that area and length are very different kinds of measures. We need to choose the measure suitable to the object on which we want to integrate. The measure tells us how much the partiation pieces may overlap: We always want $\text{meas}(\mathcal{V}_i \cap \mathcal{V}_j) = 0$ when $i \neq j$. You can check that this happens with our partition of the interval.

Riemann Sum

Once we have a partition and a measure, we can form a **Riemann sum**:

$$\sum_j f(\mathbf{x}_j^*)\ \text{meas}(\mathcal{V}_j).$$

Actually, we also needed to choose the **evaluation points** $\mathbf{x}_j^*$, but that almost goes without saying. These points are chosen so that $\mathbf{x}_j^* \in \mathcal{V}_j$. There are k of them, and the superscript "$*$" indicates that there might be many alternative choices, i.e., there are probably many points in $\mathcal{V}_j$ from which to choose. Eventually, we want to consider **all possible** choices, but for now we can just imagine having made one such choice.

It will be recalled that the Riemann sum associated with a function defined on an interval $[a, b]$ can be interpreted as an approximation of the (signed) area under the graph of the function. There are various similar interpretations which are, more or less, possible for other Riemann sums.

The Riemann Integral

We are essentially ready to define the integral of f over $\mathcal{V}$. Let us write down the definition and explain it.

$$\int_{\mathcal{V}} f = \lim_{\|\mathcal{P}\| \to 0} \sum_j f(\mathbf{x}_j^*)\ \text{meas}(\mathcal{V}_j). \tag{8}$$

This says that we get the integral by taking a limit of the Riemann sum(s). How many Riemann sums are there? There are usually a lot of them, because there are, first of all, many choices for the partition. Also, there can be many choices for the evaluation points. Having noted and understood all these possible choices, it's relatively easy to explain what is going on in this limit—though this is probably the trickiest part.

Norm of a Partition

Before we give that explanation, we need one more thing. We need to explain $\|\mathcal{P}\|$ which is the **norm of the partition**. The condition $\|\mathcal{P}\| \to 0$ is a (second) way of saying we want all the partiation pieces to get "small." For intervals, we can use the measure to accomplish this:

$$\|\mathcal{P}\| = \max_j (x_j - x_{j-1}).$$

That is, the norm of a partition consisting of intervals is the length of the longest interval in the partition. If we try to use area, on the other hand, for the norm of a partition consisting of areas in the plane, then we run into trouble. Can you see why? The problem is that a very small area in the plane (arbitrarily small actually) can be spread out all over the place.

Exercise 7. *Find a subset $\mathcal{V}_1$ of the unit square $\mathcal{V} = [0,1] \times [0,1]$ with the following properties:*

1. $\text{area}(\mathcal{V}_1) = 1/1000$*, and*
2. *Given* **any** *point $p \in \mathcal{V}$, there is a point $q \in \mathcal{V}_1$ such that $|p - q| < 1/1000$.*

To elaborate on this problem a little more, we'd like (at least when the function f is continuous) to have the individual terms $f(\mathbf{x}_j^*)\,\text{meas}(\mathcal{V}_j)$ in the Riemann sum to have about the same value no matter which evaluation point $\mathbf{x}_j^* \in \mathcal{V}_j$ is chosen. This amounts to having $f(\mathbf{x}_j^*)$ and $f(\mathbf{x}_j^{**})$ close together whenever $\mathbf{x}_j^*$ and $\mathbf{x}_j^{**}$ are any two different evaluation points in $\mathcal{V}_j$. This kind of situation is most easily realized by requiring $\mathcal{V}_j$ does not "spread out" too much. And the measure area does not accomplish this.

In short, we often need another way to measure the size of pieces. Without specifying this measure exactly, let us denote it by diam. Then we can take (at least symbolically)

$$\|\mathcal{P}\| = \max_j \text{diam}(\mathcal{V}_j).$$

Thus, $\|\mathcal{P}\|$ is the "diameter" of the partition piece with the largest "diameter." For many kinds of sets,

$$\text{diam}(\mathcal{V}_j) = \sup\{|\mathbf{x} - \mathbf{y}| : \mathbf{x}, \mathbf{y} \in \mathcal{V}_j\} \tag{9}$$

provides a reasonable notion of diameter. In this expression, the sup (supremum) of a set of real numbers is the *smallest number that is greater than every number in the set.*

That's kind of a mouthful. The supremum is also sometimes called the *least upper bound*, and it's an axiom of the real numbers that any bounded set always has a well-defined supremum or least upper bound. This is basically the way mathematicians try to avoid having gaps in the real number line. For this reason, the axiom requiring the existence of least upper bounds for bounded sets of real numbers is called the **axiom of completeness** as well as the **least upper bound axiom**. But this is taking us on a bit of a tangent. The points are that (1) the existence of the supremum in the definiton of diameter (9) relies on a mathematical axiom about the real numbers—it's not something you can prove and it's something rather complicated—and (2) it is generally possible to make a definition of diameter that makes sense.

The Limit

The existence of the limit in (8) means the following thing:

There is a number L such that for any (small) positive real number $\epsilon > 0$, there is another positive real number $\delta > 0$ having the property that whenever $\|\mathcal{P}\| < \delta$, then

$$\left| L - \sum_j f(\mathbf{x}_j^*)\ \mathrm{meas}(\mathcal{V}_j) \right| < \epsilon.$$

That is, if the largest partition piece has size/diameter smaller than δ, then (no matter how you pick the partition subject to the size requirement and no matter how you pick the evaluation points) then the Riemann sum will be "ϵ close" to the limit L.

If there is such a number L, then we call that number the integral:

$$L = \int_{\mathcal{V}} f.$$

Naturally, this integral on a general object may have various interpretations as the area, volume, or hyper-volume "under" the graph of the function f or some other geometric quantity.

If we were to desire more technical precision, it would be useful to give conditions on the sets $\mathcal{V}$, the partitions $\{\mathcal{V}_j\}$ and the functions $f : \mathcal{V} \to \mathbb{R}$ for which the limit (i.e., the limit of the Riemann sums) actually exists, so that the integral exists and is well-defined. One condition is well-known to imply the existence of Riemann integrals in many cases, so we mention it as a vaguely stated theorem.

Theorem 1. *If $\mathcal{V}$ is a (closed) set on which it makes sense for a real valued function $f : \mathcal{V} \to \mathbb{R}$ to be* **continuous***, and the function f is indeed* **continuous***, then the*

Riemann integral

$$\int_{\mathcal{V}} f$$

is well-defined.

A somewhat more precise statement is given on page 896 of the Thomas Calculus text; see also Theorem 1 in section 5.3, page 319.

5 A Difference

The integration we have introduced, even in the case where $\mathcal{V} = [a, b]$ is an interval, is not exactly the same as the integration on intervals from Calc II. Also, there are various notions of integration on curves and surfaces in calculus that are usually introduced with an analogous difference. To emphasize this difference, the integration on objects we have introduced is sometimes called **integration on sets** as opposed to integration on **oriented sets**. To see how orientation plays a role, let's restrict to intervals. Given an interval $[a, b]$ it's true that

$$\int_a^b f(x)\,dx = \int_{[a,b]} f.$$

These two kinds of integrals are **exactly** the same. However, using the notion of integration from Calc II, it also makes perfectly good sense to write

$$\int_b^a f(x)\,dx \tag{10}$$

and integrate "backwards" on the interval $[a, b]$. You will recall that there is a host of manipulations associated with this kind of backwards integration. For example, we know

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx.$$

Not only is the backwards integral in (10) difficult to express as an integral on a set, but these two kinds of integrals have different change of variables formulas. Let's start with a change of variables in Calc II, or what was called "u-substitution." Quite generally, if we have a change of variables $u = \psi(x)$, then in Calc II we would write

$$\int_a^b f(x)\,dx = \int_{\psi(a)}^{\psi(b)} \frac{f \circ \psi^{-1}(u)}{\psi' \circ \psi^{-1}(u)}\,du. \tag{11}$$

You may not remember the formula looking this complicated. What you may remember is more along the following lines: From $u = \psi(x)$, you have $du = \psi'(x)\,dx$, so

$$\int_a^b g(x)\psi'(x)\,dx = \int_{\psi(a)}^{\psi(b)} g \circ \psi^{-1}(u)\,du. \tag{12}$$

If you compare, you will see that these are saying the same thing, but there is a preconditioning of the integrand $f(x)$ obtained by writing

$$g(x) = \frac{f(x)}{\psi'(x)}.$$

Our point here, however, is that this formula works even when $\psi(a) < \psi(b)$ so that the integral you get when you change variables is a backwards integral.

The corresponding computation for integration on sets is usually expressed in terms of a **scaling factor** σ, and when integrating on (unoriented) sets, the scaling factor is always the **absolute value** of something. In the simple case of a change of variables $\psi : [a, b] \to \mathbb{R}$, the scaling factor is

$$\sigma = |(\psi^{-1})'(u)| = \frac{1}{|\psi' \circ \psi^{-1}(u)|}. \tag{13}$$

Exercise 8. *Differentiate the relation*

$$\psi \circ \psi^{-1}(u) = u$$

to obtain the equivalent expressions for the scaling factor σ in (13). Hint: Be careful with your differentiation and the use of the chain rule.

When $\psi(a) < \psi(b)$, then we can use set integration to express either of the u-substitution rules (11) or (12). For example,

$$\int_{[a,b]} f = \int_{[\psi(a),\psi(b)]} f \circ \psi^{-1}(u)\,\sigma(u).$$

When ψ reverses the direction of the interval and $\psi(b) < \psi(a)$, then the change of variables formula is, superficially, a little different:

$$\int_{[a,b]} f = \int_{[\psi(b),\psi(a)]} f \circ \psi^{-1}(u)\,\sigma(u).$$

This is where the absolute value comes in with regard to the scaling factor, since in this case $\sigma = -1/\psi' \circ \psi^{-1}(u)$.
Warning: You may be used to using the "scaling factor"

$$\frac{1}{\psi' \circ \psi^{-1}(u)}\, du$$

in formula (11) when changing variables in the integral on the oriented interval $[a, b]$, but for integration on an interval as a set, the correct scaling factor is

$$\sigma = \left| \frac{1}{\psi' \circ \psi^{-1}(u)} \right|.$$

6 Calculation/Concrete Integration in one special case

You may leave the discussion above with a feeling of dissatisfaction. In fact, it can be very useful to understand the abstract notion of "integration on objects" we have described above. There are problems that are important for engineers and other people who use mathematics which are very difficult to work without it. On the other hand, if one wants to compute a concrete answer, then the definition

$$\int_{\mathcal{V}} f = \lim_{\|\mathcal{P}\| \to 0} \sum_j f(\mathbf{x}_j^*)\, \mathrm{meas}(\mathcal{V}_j)$$

is of limited use. The same was true in Calc II, and that's why the preponderance of the course was spent on various "techniques" of integration. First you may have used the Riemann sum definition to derive some simple examples like

$$\int_{[a,b]} x^n = \frac{1}{n+1}(b^{n+1} - a^{n+1}),$$

or you may have just memorized the "power rule." At any rate, you eventually memorized some elementary integration formulas for powers, trigonometric functions, and exponential functions. Then you learned other techniques, like u-substitution, various algebraic tricks and partial fractions for rational functions, integration by parts, and others. The definition using Riemann sums was probably lost in the shuffle, but that definition is really what gives meaning to the rest.

The good news is that, for the most part, there is no correspondent torrent of complicated techniques needed to integrate on sets that are more complicated than intervals. Usually, one reduces the calculation of an integral on any set to some equivalent calcuation involving 1-D integrals on intervals—which are then attacked with the techniques from Calc II. As one example of this, let us consider integration on a three-dimensional volume in a couple simple cases.

Integration on a Rectangular Parallelopiped, i.e., A Box

If

$$\mathcal{V} = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3] = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : a_j \leq x_j \leq b_j,\ j = 1, 2, 3\}$$

is a Cartesian product of three intervals and $f : \mathcal{V} \to \mathbb{R}$ is a continuous function defined on $\mathcal{V}$, then

$$\int_{\mathcal{V}} f = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \int_{a_3}^{b_3} f(x_1, x_2, x_3)\, dx_3 dx_2 dx_1.$$

The expression on the right is what's called an **iterated integral**. It just means

> *First evaluate*
>
> $$\int_{a_3}^{b_3} f(x_1, x_2, x_3)\, dx_3$$
>
> *as a 1-D integral, thinking of x_1 and x_2 as constants. Then take the result, which will be a function $g = g(x_1, x_2)$, and evalute the 1-D integral*
>
> $$\int_{a_2}^{b_2} g(x_1, x_2)\, dx_2 = \int_{a_2}^{b_2} \left(\int_{a_3}^{b_3} f(x_1, x_2, x_3)\, dx_3 \right) dx_2$$
>
> *thinking of x_1 as a constant. Finally take the result, which will be a function $h = h(x_1)$, and evalute the 1-D integral*
>
> $$\int_{a_1}^{b_1} h(x_1)\, dx_1 = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \left(\int_{a_3}^{b_3} f(x_1, x_2, x_3)\, dx_3 \right) dx_2 \right) dx_1.$$

The fact that this technique of "iterated integrals" works is called **Fubini's Theorem**.

A Generalization

A version of Fubini's theorem works for a somewhat more general region $\mathcal{V} \subset \mathbb{R}^3$ having the form

$$\mathcal{V} = \{(x, y, z) : \psi_1(x, y) < z < \psi_2(x, y),\ (x, y) \in \mathcal{U}\}$$

where $\mathcal{U} \subset \mathbb{R}^2$ is a region in the plane and ψ_1 and ψ_2 are real valued functions with domain $\mathcal{U}$ satisfying $\psi_1 < \psi_2$. In this case, it is possible to express the integral of a function f over $\mathcal{V}$ as the iterated integral

$$\int_{\mathcal{V}} f = \int_{\mathcal{U}} \left(\int_{\psi_1(x,y)}^{\psi_2(x,y)} f(x, y, z)\, dz \right).$$

This reduction of triple/volume integrals to iterated integrals is discussed in the Thomas text on pages 922-924. Integration on various regions $\mathcal{U} \subset \mathbb{R}^2$ is found in sections 15.2-15.4.

In any case, for our purposes, the main points are:

1. It's useful to know the theory of Riemann sums for integrals on many kinds of objects.

2. There are ways to reduce many integrals on various objects to iterated 1-D integrals.

7 More Concrete Integration

7.1 The Divergence

We have discussed **flux** in one dimension.

Say we want to integrate on a region $U \subset \mathbb{R}^n$ some product expression

$$\int_U \Delta u\, \phi$$

8 Better Integration through measures

The Riemann integral is constructed using a measure, but it is not an integral **based on a measure**, and this makes it limited in certain respects. To understand this

accusation, let's return briefly to one dimension and specifically to the set $\mathbb{R}$ or an interval, say $[a, b] \subset \mathbb{R}$. If we want to integrate discontinuous functions, then we can sometimes use the Riemann integral. In fact, **step functions** were very central to the construction of the Riemann integral, and those are, typically, discontinuous. A **step function** is a function of the form

$$f(x) = \sum_{j=0}^{k-1} f_j \chi_{I_j}(x)$$

where $a = x_0 < x_1 < x_2 < \cdots < x_k = b$ is a partition of $[a, b]$ and I_j is one of the four intervals (x_j, x_{j+1}), $[x_j, x_{j+1})$, $(x_j, x_{j+1}]$, or $[x_j, x_{j+1}]$ and $\chi_{I_j} : \mathbb{R} \to \mathbb{R}$ is the **characteristic function** given by

$$\chi_{I_j}(x) = \begin{cases} 1, & x \in I_j \\ 0, & x \notin I_j. \end{cases}$$

The **integral of a characteristic function** is simply[5]

$$\sum_{j=0}^{k-1} f_j (x_{j+1} - x_j)$$

which you should recognize as essentially the same as a Riemann sum. The only difference is that for a Riemann sum, we are looking at step functions associated with a specific given function $u : [a, b] \to \mathbb{R}$, so that we take $f_j = u(x_j^*)$.

The number $x_{j+1} - x_j$ will be recognized as the **length** of the interval I_j. It may come as a surprise, but length itself is not a measure for various technical reasons we will discuss soon, but length is the restriction of various measures to the intervals. Thus, we can say that the Riemann integral is based on the integration of step functions, and the integration of step functions is defined using a measure, though step functions themselves and hence Riemann integration are based on intervals and length, so Riemann integration is not *based* on a measure.

We now ask a question:

What is a measure?

And, more precisely, we ask another question: What is an **ideal measure**? Before we give the answer, let's settle some generalities: A measure μ should be a function

[5]You may imagine, as you read this, that we have not yet defined the Riemann integral.

whose argument is a set. The domain is a set of sets. And, for our purposes, we should look for values in $[0, \infty)$ the set of non-negative real numbers, or possibly in the non-negative **extended real numbers** $[0, \infty]$. Thus, we are going to talk a bit about functions

$$\mu : \Sigma \to [0, \infty) \tag{14}$$

where Σ is some appropriate collection of subsets of, say, the real line $\mathbb{R}$ or the interval $[a, b]$. If we want a measure on $\mathbb{R}$, then we probably want values in $[0, \infty]$, since there are many sets whose measure we expect to be infinite. For example, the real line is infinitely long. On the other hand, perhaps things will be simpler if we restrict attention to an interval $[a, b]$ of finite length. Then (14) should be adequate.

8.1 Ideal Measure

With some experience, the following list was compiled: Ideally, a measure should have the following properties:

1. Σ should be the set of all subsets (of $[a, b]$).
2. $\mu(\phi) = 0$. The measure of the empty set should be zero.
3. $\mu(I) = \beta - \alpha$ where $I \subset [a, b]$ is any one of the four types of intervals: (α, β), $[\alpha, \beta)$, $(\alpha, \beta]$, or $[\alpha, \beta]$. We say, *the measure of an interval is its length.*
4. $\mu\{x + t : x \in A\} = \mu A$ for every $A \in \Sigma$ and $t \in \mathbb{R}$. This is called *translation invariance.*[6]
5. The following property is called **countable additivity**:

$$\mu\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu E_j$$

whevever $E_1, E_2, E_3, \ldots$ is a sequence of mutually disjoint subsets of $[a, b]$.

Thus, an **ideal measure** on an interval $[a, b]$ of real numbers having finite length is *a countably additive, non-negative, translation invariant, set function for which the measure of an interval is its length.*

[6]You will note, there is an apparent problem with this formulation of translation invariance because generally, the translation $\{x + t : x \in A\}$ cannot be assumed to be a subset of $[a, b]$. The formulation given works for a measure on $\mathbb{R}$. It turns out that it's adequate to assume the desired property whenever $A \subset [a, b]$ and $\{x + t : x \in A\} \subset [a, b]$.

Theorem 2. *(Lebesgue) There is no ideal measure on an interval of positive length. Such a thing does not exist.*

In view of this nominally tragic result, mathematicians have scrambled and scattered in all directions of compromise. Before I attempt to describe some of those, let me give some brief comments concerning the list of properties of the (nonexistent) ideal measure.

The set of all subsets of a given set, say $[a, b]$ is called the **power set** of the given set and is denoted by $\mathcal{P}(A)$ or 2^A. The latter notation is sort of a curious one, and you may be amused to know how it arises: In set theory, numbers and arithmetic are defined solely in terms of sets. In particular, the number 0 is the empty set, the number 1 is the set containing the empty set

$$1 = \{\phi\} = \{0\}$$

and $2 = \{0, 1\}$. This is how the "2" in 2^A should be interpreted. Exponentiation of sets (which is where exponentiation of numbers finds its origin in set theory) is defined in terms of sets of functions (which are also sets). To make a long story short, 2^A is the set of all functions $\phi : A \to 2$. That is, each function assigns to each element of A an element of $2 = \{0, 1\}$. So think of it this way: If $E \subset A$, then there is a function $\phi : A \to \{0, 1\}$ which assigns to $x \in A$ the value 1 if $x \in E$ and the value 0 if $x \notin E$. That is ϕ is the characteristic function χ_E on E. Conversely, given a function $\phi : A \to \{0, 1\}$, it's obvious that such a function is designating a specific subset $E = \phi^{-1}(\{1\}) = \{x \in A : \phi(x) = 1\}$. This identification is clearly a one-to-one correspondence between the subsets of A and the functions from A to $\{0, 1\}$. That is, $\mathcal{P}(A) = 2^A$.

In the special case of an interval $[a, b]$ of zero length, i.e., $[a, a] = \{a\}$, we can write down $\mathcal{P}([a, a]) = \{\phi, \{a\}\}$. And we note that the zero measure is an **ideal measure**. The function $\mu : \mathcal{P}([a, b]) \to \{0\}$ by $\mu A \equiv 0$ satisfies all the properties of an ideal measure listed above. Whenever the length of an interval $[a, b]$ is positive, however, there are obviously many more subsets, and it took the genius (or audacity) of Lebesgue to suggest throwing some of them out for the sake of having a respectable measure. Thus, Lebesgue questioned the necessity of requiring $\Sigma = \mathcal{P}([a, b])$. This meant that he needed to designate which subsets should be in the set of sets Σ and which should not. Then he needed to reformulate the remaining properties and see if they survived in tact. His formulation of the requirements on Σ are what are called the defining properties of a **sigma algebra**. We will discuss some possibilities for choices of the sigma algebra Σ below.

The next property almost requires no mention. For one thing, as one might guess (and as one can show using countable additivity for example) if you have a measure, then the measure should be **monotone** in the sense that

$$\mu A \leq \mu B \qquad \text{whenever} \qquad A \subset B.$$

In particular, if the measure of an interval is its length, then

$$0 \leq \mu(\phi) \leq \mu([a,a]) = \lim_{\epsilon \searrow 0} \mu([a-\epsilon, a+\epsilon]) = 0.$$

So you can prove this "null property" from the others...if you have them. The real reason the property $\mu(\phi) = 0$ is included is because when one starts monkeying around with the other properties, this is one property that can be kept. It's a sort of silly property, but **all measures satisfy** $\mu(\phi) = 0$, and it's the only one of these properties with that property. Everybody agrees on this.

The next two properties, translation invariance and the requirement that the measure of an interval is its length, may be called the **Euclidean properties**, and measures that preserve these two properties are called **Euclidean measures**. Perhaps a short note on translation invariance for finite intervals is in order. First of all, if one is worried about translating a set $A \subset [a,b]$ out of $[a,b]$, then one possibility is **modular arithmetic** or **modular translation**. This works roughly like this: Consider the disjoint translated intervals

$$I_n = (a + n(b-a), b + n(b-a)) \qquad \text{for} \qquad n \in \mathbb{Z}.$$

We say μ is **translation invariant mod** $(b-a)$ if

$$\sum_{n \in \mathbb{Z}} \mu\{p - n(b-a) : p \in T \cap I_n\} = \mu A \qquad \text{for every translate } T = \{x + t : x \in A\}. \tag{15}$$

Note that given $t \in \mathbb{R}$ determining a translate $T = \{x + t : x \in A\}$, the translate T can intersect at most two of the open intervals I_n, and if there are two, then those two must be consecutive. Therefore, at most two of the terms in the sum on the left in (15) are nonzero. Furthermore, if we assume those (at most) two intervals are I_n and I_{n+1}, then

$$\phi : (T \cap I_n) \cup (T \cap I_{n+1}) \to T \qquad \text{by} \qquad \phi(p) = p$$

is an injection onto all but (possibly) three points of T, and the sets

$$\{p - n(b-a) : p \in T \cap I_n\} \qquad \text{and} \qquad \{p - n(b-a) : p \in T \cap I_{n+1}\}$$

whose measures appear on the left in (15) are translates of these sets. Finally, adding in one, two, or three endpoints would not change the measure in view of additivity and the fact that the measure of an interval (of length zero) is it's length. We end these considerations with the following lemma (claimed above in a footnote):

Lemma 1. *Let μ be a (nonexistent) ideal measure. If $\mu\{x+t : x \in A\} = \mu(A)$ whenever $A \subset [a,b]$ and $t \in \mathbb{R}$ with $T = \{x+t : x \in A\} \subset [a,b]$, then μ is translation invariant, say mod $(b-a)$.*

Proof: As mentioned above, given $A \subset [a,b]$ and $t \in \mathbb{R}$, there is some n for which the condition (15) becomes

$$\mu\{p - n(b-a) : p \in T \cap I_n\} + \mu\{p - n(b-a) : p \in T \cap I_{n+1}\} = \mu A.$$

The construction we now carry out is partially illustrated in Figure 5.

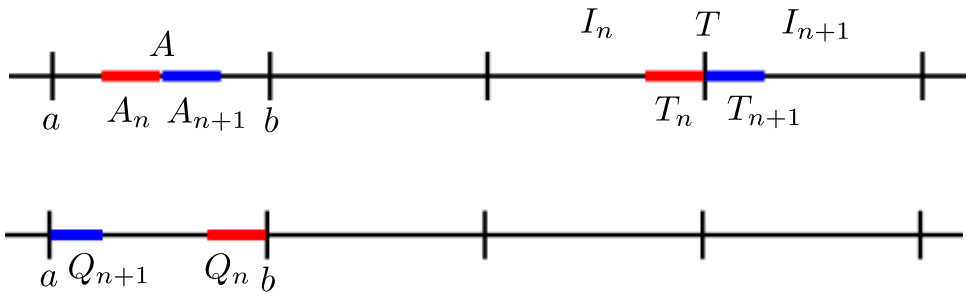

Figure 5: Translating sets and translation invariance for a measure. The red and blue points are initially in a set $A \subset [a,b]$. These are translated to $T = \{x+t : x \in A\} \subset I_n \cup I_{n+1}$ and finally translated back into $[a,b]$ for comparison.

Notice that the endpoint between I_n and I_{n+1} is $m = b + n(b-a)$.
Let $T_n = \{x+t : x \in A\} \cap I_n = T \cap I_n$ and $T_{n+1} = T \cap I_{n+1}$. Furthermore, let

$$A_n = \{p - t : p \in T_n\} \qquad \text{and} \qquad A_{n+1} = \{p - t : p \in T_{n+1}\}.$$

We claim that

$$Q_n = \{p - n(b-a) : p \in T_n\} \subset [a,b].$$

is a translate of A_n, namely,

$$Q_n = \{x + t - n(b-a) : x \in A_n\}. \tag{16}$$

The inclusion is clear since

$$a + n(b-a) < p < b + n(b-a) \qquad \Longrightarrow \qquad a < p - n(b-a) < b.$$

To show (16) take $x + t - n(b+a)$ with $x \in A_n$. Because $x \in A_n$, we have $x = p - t$ for some $p \in T \cap I_n$. Thus we can write

$$x + t - n(b-a) = p - n(b-a) \in Q_n.$$

We have shown $\{x + t - n(b-a) : x \in A_n\} \subset Q_n$. Now take $p - n(b-a) \in Q_n$ with $p \in T_n = T \cap I_n$. Then we know $p = x + t$ for some $x \in A$. Therefore,

$$p - n(b-a) = x + t - n(b-a).$$

It remains to show $x \in A_n$. But this is clear since $x = p - t$ and $p \in T_n$. Thus, we have established (16).

Similarly, it can be shown that

$$Q_{n+1} = \{p - n(b-a) : p \in T_{n+1}\} \subset [a, b].$$

is a translate of A_{n+1}, namely,

$$Q_{n+1} = \{x + t - n(b-a) : x \in A_{n+1}\}.$$

Finally, $A \backslash (A_n \cup A_{n+1}) \subset \{a, b, m\}$. Therefore,

$$\mu Q_n + \mu Q_{n+1} = \mu A_n + \mu A_{n+1} = \mu(A_n \cup A_{n+1} \cup \{a, b, m\}) \geq \mu A$$

because μ is Euclidean and $\mu\{a, b, m\} = 0$. On the other hand,

$$\mu Q_n + \mu Q_{n+1} = \mu A_n + \mu A_{n+1} \leq \mu A$$

by monotonicity. $\square$

The last property of an ideal measure is **countable additivity**. I don't have too much to say, particularly, about countable additivity, except that this condition is often motivated by the weaker property of **finite additivity**:

$$\mu(E_1 \cup E_2) = \mu E_1 + \mu E_2 \qquad \text{when} \qquad E_1 \cap E_2 = \phi$$

which naturally extends by induction to finite collections of mutually disjoint sets. Countable additivity is also important as a property that is relaxed to **countable subadditivity**:

$$\mu\left(\bigcup_{j=1}^{\infty} E_j\right) \leq \sum_{j=1}^{\infty} \mu(E_j). \tag{17}$$

Naturally, the condition of subadditivity does not require the sets in the union to be disjoint. Consequently, it turns out to be natural to consider set functions $\mu^* : \mathcal{P}([a,b]) \to [0,\infty)$ which are (countably) subadditive and are called **outer measures**. The two most famous outer measures Lebesgue outer measure and Hausdorff (outer) measure will be discussed below. Before we get to that, however, let us return to our main course of discussion.

8.2 Euclidean Measures

We have discussed what it would be like to have an ideal measure, and we have confessed up front that there is no such thing. As mentioned above, the suggestion of Lebesgue was to give up the hope of being able to measure all sets and restrict attention to a subcollection $\Sigma \subset \mathcal{P}([a,b])$ which is a **sigma algebra** or more commonly a σ-algebra. A σ-algebra on a set X, for example $X = [a,b]$ has some pretty simple properties:

1. $\phi \in \Sigma$ and $X \in \Sigma$.
2. If $E \in \Sigma$, then the complement of E given by $X \backslash E \in \Sigma$.
3. If $E_1, E_2, E_3, \ldots \in \Sigma$, then $\cup E_j \in \Sigma$.

We say Σ is **closed under taking complements and countable unions**. In view of countable additivity, the closure under countable unions is quite natural. With the crucial definition of a σ-algebra, we can easily recast our five desired properties to obtain what is called a **Euclidean measure** on an interval:

Definition 1. *Given a σ-algebra Σ of subsets of the real interval $[a,b] \subset \mathbb{R}$, a function $\mu : \Sigma \to [0,\infty)$ is said to be a* **Euclidean measure** *on $[a,b]$ if*

1. $\mu(\phi) = 0$.
2. *For every subinterval $I \subset [a,b]$ we have $I \in \Sigma$,*
$$\mu(I) = \text{length}(I),$$
i.e., the measure of an interval is its length.
3. *For every $A \in \Sigma$ and every $t \in \mathbb{R}$ such that $T = \{x + t : x \in A\} \in \Sigma$, we have*
$$\mu\{x + t : x \in A\} = \mu A.$$
This is the property of translation invariance.

4. *If $E_1, E_2, E_3, \ldots$ is a (countable) sequence of sets in Σ which are mutually, i.e., pairwise, disjoint, then*

$$\mu\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu E_j.$$

This is, of course, countable additivity.

With this definition, the nonexistence result Theorem 2 can (and usually is) restated as an existence theorem:

Theorem 3. *(existence of a non-measurable set) If $\mu : \Sigma \to [0, \infty)$ is a Euclidean measure on an interval $[a, b]$ with $b - a > 0$, then there exists a subset $\mathcal{N}$ of $[a, b]$ with $\mathcal{N} \notin \Sigma$.*

Thanks to Lebesgue, we have reduced the requirements on the desired measure from five to four, but we still have not identified such a measure. It is perhaps natural to try to obtain a σ-albebra of subsets which is as large as possible. It seems the credit for identifying this collection, or at least giving the most user friendly characterization of this collection, goes to Constantin Carathéodory. Still, it's not so user friendly, but let me try to describe what happens briefly. For any set $A \subset [a, b]$ one sets

$$\mu^*(A) = \inf_{A \subset \cup I_j} \sum_j \mu(I_j)$$

where $I_1, I_2, I_3, \ldots$ is a countable collection of intervals with $A \subset \cup I_j$. This is called **Lebesgue outer measure.** It satisfies most of our properties for an ideal measure. In fact, it satisfies all of them except countable additivity, and (as you might guess) μ^* is countably subadditive. So this is a good start. Now here is Carathéodory's trick: $\Sigma = \mathfrak{M}$ is the set of all sets E for which

$$\mu^* A = \mu^*(A \cap E) + \mu^*(A \backslash E) \qquad \text{for every } A \subset [a, b], \tag{18}$$

that is for every $A \in \mathcal{P}([a, b])$. Note that A is not the **measureable set** we get out of this, but rather E. The condition is called the **Carathéodory measurability criterion**, and the set of all sets E satisfying (18) is a σ-algebra called either the Carathéodory measurable sets or the Lebesgue measurable sets. This σ-algebra is often denoted by something like $\mathfrak{M}$. According to the Carathéodory criterion, the measurable sets are those which "cut well" an arbitrary set A. It can be shown that if you have any more, then you've included non-measurable sets.

It's a good bit of work, but you can show that taking $\mu : \mathfrak{M} \to [0, \infty)$ as the restriction of Lebesgue outer measure to $\mathfrak{M}$, i.e., $\mu A = \mu^* A$ for all $A \in \mathfrak{M}$, defines a Euclidean measure on $[a, b]$: Of course, the intervals are in $\mathfrak{M}$. The measure is translation invariant and the measure of an interval is its length, i.e., the measure is Euclidean. Moreover, you get countable additivity. Finally one can show the following:

Theorem 4. *Any Euclidean measure $\mu : \Sigma \to [0, \infty)$ satisfies $\Sigma \subset \mathfrak{M}$ and*

$$\mu A = \mu^* A \qquad \text{for every } A \in \Sigma.$$

So you have a rather good measure, and now you can have rather good integration to go along with it. That takes at least a bit more work, so let's pause to set up a framework for what is required and describe some perspective for other matters as well.

8.3 Directions of Generalization

One obvious generalization we need is a Euclidean measure on all of $\mathbb{R}$. This will be an extended real valued function $\mu : \mathfrak{M} \to [0, \infty]$, and you can almost guess the nature of both μ and the σ-algebra $\mathfrak{M}$. They should be some straightforward generalizations of μ^* and the Lebesgue measurable sets on finite intervals $[a, b]$. I won't go through the details. It can be done.

Here is a kind of "to do" list of other directions which are important:

1. Extend Lebesgue measure to all of $\mathbb{R}$. (mentioned above)

2. Find an analogous measure for n-dimensional volume on $\mathbb{R}^n$. That is, generalize to higher dimensions.

3. Consider **non-Euclidean measures** for which we relax, i.e., throw out, the Euclidean requirements of translation invariance and that the measure of an interval is given by length.

4. Consider some **outer measures** for which we relax countable additivity to countable subadditivity.

8.4 Lebesgue Integration

Lebesgue integration is based on the Lebesgue measure as follows: A **simple function** is a function of the form

$$f(x) = \sum_{j=0}^{k} f_j \chi_{E_j}(x)$$

where $E_1, E_2, E_3, \ldots, E_k$ are Lebesgue measurable sets. To integrate a non-negative function $u : [a, b] \to [0, \infty)$, you can take

$$\int_{[a,b]} u = \sup_{f \leq u} \sum_{j=0}^{k} f_j \mu(E_j). \tag{19}$$

It turns out that you won't have a very nice integral here unless u is a decent function that respects measurability. Such a function is called a **measurable function**. There are various ways to define what it means for a function to be measurable. Here is a general definition:

Definition 2. *Given (X, Σ, μ) and (Y, T, ν) two measure spaces, we say $u : X \to Y$ is a* **measurable function** *if*

$$\{x : u(x) \in E\} \in \Sigma \qquad \text{whenever} \qquad E \in T. \tag{20}$$

The set $\{x : u(x) \in E\}$ appearing in (20) is called the **inverse image of** E and is denoted by

$$u^{-1}(E) = \{x : u(x) \in E\}.$$

Notice that this set is always well-defined. In particular, the use of the notation $u^{-1}(E)$ with the set E as an "argument" is not intended to suggest that the function u has an inverse as a function $u : X \to Y$. In particular, one does not mean by this notation that the function u is one-to-one (injective) or onto (surjective).

Applying this general definition to our proposed definition of integration (19) given above, we have $u : \mathbb{R} \to \mathbb{R}$, and we need to choose σ-algebras and measures for $\mathbb{R}$ as the domain and codomain of u. Now, it might be expected that we would take the extension of Lebesgue measure to $\mathbb{R}$ in both cases. That is, it turns out, not what is usually done. *Why does everything have to be so complicated!* Well, the answer in this case is the following:

If you think about the measurability of functions given generally by (20), and you want to have as many measurable functions as possible, then you want as many measurable sets Σ as possible and **as few measurable sets T as possible.**

Consequently, for $u : [a, b] \to \mathbb{R}$ we take $\mathfrak{M}$ on the domain $[a, b]$ but the **smallest σ-algebra containing the open sets** on $\mathbb{R}$. This latter σ-algebra is called the **Borel σ-algebra** or more often the **Borel sets** and is denoted by $\mathfrak{B}$. Thus, for integrability, a function u is said to be Lebesgue measurable if

$$\{x \in [a, b] : u(x) \in A\} \in \mathfrak{M} \qquad \text{whenever } A \in \mathfrak{B}.$$

This condition turns out to be fairly easy to check:

Lemma 2. *$u : [a, b] \to \mathbb{R}$ is Lebesgue measurable if and only if*

$$\{x \in [a, b] : u(x) > c\} \in \mathfrak{M} \qquad \textit{for every } c \in \mathbb{R}.$$

For these functions the integral (19) has all the properties one would like to see in an integral. Furthermore, you can also break any measurable function up as

$$u = u^+ + u^-$$

where u^+ is a non-negative function and u^- is a non-positive function. Then, since $-u^-$ is a non-positive function (and we know how to integrate those by (19)) we can set

$$\int u = \int u^+ - \int (-u^-).$$

There you have it. Oh, incidentally, this integral agrees with the Riemann integral whenever the Riemann integral is defined. This is the integration used for the L^p spaces at least as far as $L^p(a, b)$.

Exercise 9. *Consider the function $u : [0, 1] \to \mathbb{R}$ given by*

$$u(x) = \begin{cases} 0, & 1 \in \mathbb{Q} \cap [0, 1] \\ 0, & x \in [0, 1] \backslash \mathbb{Q} \end{cases}$$

where $\mathbb{Q} = \{p/q : p \in \mathbb{Z},\ q \in \mathbb{N}\}$ is the rational numbers.

1. *Show that u is not Riemann integrable. Hint: Every* **upper sum** *based on a* **step function** *f (based on intervals) with $f \geq u$ has*

$$\int f = \sum u(x_j^*)(x_{j+1} - x_j) = \sum (x_{j+1} - x_j) = 1$$

because there are rational numbers in every interval of positive length. But Every **lower sum** *for u is zero because there are irrationals in every interval of positive length too.*

2. *Show that $\mathbb{Q}$ has outer measure $\mu^*\mathbb{Q} = 0$. Hint(s): $\mathbb{Q}$ is* **countable** *which means you can find a sequence $q_1, q_2, q_3, \ldots$ containing every rational number:*

$$\mathbb{Q} = \{q_1, q_2, q_3, \ldots\}.$$

Now, given any $\epsilon > 0$ take an open interval $(q_j - \epsilon/2^{j+1}, q_j + \epsilon/2^{j+1})$ containing q_j. Conclude that

$$\mu\mathbb{Q} \leq \epsilon. \tag{21}$$

What non-negative numbers $\mu\mathbb{Q}$ satisfy (21) for every $\epsilon > 0$?

3. *Show that every subset $A \subset [0,1]$ with $\mu^* A = 0$ is measurable, i.e., the Carathéodory condition holds, so $A \in \mathfrak{M}$. This is a fairly difficult task for most people, but it is representative of all the hard work that goes into constructing Lebesgue measure.*

4. *Compute the Lebesgue integral of u. Hint(s): Remember how the Lebesgue integral of a non-negative function like u is defined as the supremum of integrals of* **simple functions** *f with $f \leq u$ (based on measurable sets). Note that the integral of a simple function*

$$f(x) = \sum f_j \chi_{E_j} \qquad \text{is given by} \qquad \int f = \sum f_j \mu E_j.$$

u itself is a simple function.

Higher Dimensional Lebesgue Measure

There is also Lebesgue measure on $\mathbb{R}^n$ (and other sets—curves, surfaces, etc.). The resulting integration generalizes Riemann integration on subsets of $\mathbb{R}^n$ (and other sets). This is also the integration used for and the source of the functions in the L^p space defined on, for example, an open set $U \subset \mathbb{R}^n$.

Borel Measure

In some instances, rather than the absolutely largest σ-algebra $\mathfrak{M}$ obtained by only throwing the non-measurable sets out of $\mathcal{P}([a,b])$, one wants a small (or smaller) σ-algebra. The smallest σ-algebra containing all the open sets is called the **Borel σ-algebra** or more often the **Borel sets** often denoted by $\mathfrak{B}$. Restricting Lebesgue measure (or Lebesgue outer measure) to the Borel sets, one obtains Borel measure. Again, this is a Euclidean measure. All the intervals are in $\mathfrak{B}$ and μ is translation invariant and countably additive.

8.5 Other Measures

It turns out that one can relax the other two major categories of what we have called an ideal measure, namely the Euclidean properties and the countable additivity. Relaxing the Euclidean properties, we get the general definition of a measure:

Definition 3. *A* **finite measure** *on a set X with respect to the σ-algebra Σ of subsets of X is a function $\mu : \Sigma \to [0,\infty)$ satisfying*

1. $\mu(\phi) = 0$.
2. $\mu(\cup E_j) = \sum \mu(E_j)$ *when $\{E_j\}$ is a countable collection of disjoint sets in Σ.*

Obviously, we would like to extend Lebesgue measure to all of $\mathbb{R}$ or all of $\mathbb{R}^n$, and this would not be a finite measure. There are also **infinite measures** and even **signed measures**, but let's set those aside for the moment.

Perhaps the simplest way to view non-Euclidean finite measures is in two groups: The **absolutely continuous measures** and the **singular measures**. The absolutely continuous measures preserve the Euclidean aspect of Lebesgue measure to the extent that sets consisting of single points $\{p\}$ have measure zero. In fact, the definition is that a measure ν is **absolutely continuous** if

$$\mu E = 0 \qquad \Longrightarrow \qquad \nu E = 0.$$

In this case, one writes $\nu \ll \mu$. **Singular measures** by contrast have single point sets (singletons) with positive measure. This is especially useful in discrete probability. Say you have a measure space $X = \{0,1\}$ which corresponds to flipping a coin with 0 for heads and 1 for tails. Then you can have a measure with $\mu\{0\} = 1/2$ and $\mu\{1\} = 1/2$.

You can integrate with respect to any measure in much the same way we did with Lebesgue measure. Notice that if you have a singular measure with

$$\mu E = \begin{cases} 1, & \text{if } p \in E, \\ 0, & \text{if } p \notin E, \end{cases}$$

then

$$\int f \, d\mu = f(p).$$

You can't have this kind of thing happen with Lebesgue measure (or with any absolutely continuous measure). On the other hand, whenever you can integrate you can define a functional

$$f \mapsto \int_E f \, d\mu,$$

and assuming f is taken in an appropriate vector space, this will be a linear functional. On the third hand, when you have a fixed (measurable) function f, then

$$\nu E = \int_E f \, d\mu$$

will define a new measure. Technically, we would want f to be non-negative here, but if not, we get one of those **signed measures** mentioned above.

8.6 Hausdorff Measure

As a final comment on relaxing the properties of an ideal measure, aside from the Lebesgue outer measure, there is another very interesting geometric outer measure called Hausdorff measure. The idea for it came from Felix Hausdorff who was inspired by thinking about how bricks fit together on the surface of a wall.

Consider $\mathcal{H}^d : \mathcal{P}(\mathbb{R}^n) \to [0, \infty]$ which is given for any subset A of $\mathbb{R}^n$ and any $d > 0$ by

$$\mathcal{H}^d(A) = \sup_{r>0} \inf \left\{ \alpha_d \sum_{j=1}^{\infty} r^d : A \subset \bigcup_{j=1}^{\infty} B_r(\mathbf{x}_j) \right\}$$

where α_d is a suitable positive constant. In fact, we take

$$\alpha_d = \frac{\pi^{d/2}}{\Gamma(d/2+1)}$$

where Γ is the Gamma function. The constant α_d is the volume of a unit ball in $\mathbb{R}^d$ when d is an integer, i.e., the d-dimensional Lebesgue measure of the ball.

It turns out that for a given set A there is a uniqe d_0 for which $\mathcal{H}^d(A) = \infty$ when $d < d_0$ and $\mathcal{H}^d(A) = 0$ when $d > d_0$. This number d_0 is called the **Hausdorff dimension** of A. If the Hausdorff dimension of a set $A \subset \mathbb{R}^n$ is $d_0 = n$, then

$$\mathcal{H}^n(A) = \mu(A) \qquad \text{where } \mu \text{ is } n\text{-dimensional Lebesgue measure on } \mathbb{R}^n.$$

For a surface $\mathcal{S}$ in $\mathbb{R}^3$, furthermore, one can show that the Hausdorff dimension is $d_0 = 2$ and $\mathcal{H}^2(\mathcal{S})$ is the area of the surface.

This would be considered a Euclidean measure in the sense of being translation invariant and the measure of simple sets (intervals, boxes, balls, etc.) always agrees with Lebesgue measure. On the other hand, it is not a measure technically but rather an outer measure. But most mathematicians call it Hausdorff measure instead of Hausdorff outer measure. In fact, mathematicians who work in geometric measure theory advocate for calling all countably subadditive measures, like Hausdorff measure, simply "measures." They still use the Carathéodory criterion to determine measurable sets (and hence measurable functions) however.

8.7 Radon Measures

9 Fundamental Theorem of Calculus for Measurable Functions

I'm going to outline a couple basic results concerning measurable functions to give you some idea of how one works with such functions. Then I will discuss the fundamental theorem of calculus in this context.

9.1 Integration

When one has a measure, say Lebesgue measure μ on the line, then one defines the integral of a measurable function in terms of the measure. There are various ways to do this, but one of the simplest is to start with non-negative functions and define the integral of a non-negative function $f : [a, b] \to [0, \infty)$ to be

$$\int f = \sup_{\psi \le f} \int \psi$$

where the supremum is taken over **simple functions**. Simple functions are functions $\psi : [a, b] \to \mathbb{R}$ which can be written in the form

$$\psi(x) = \sum_{j=1}^{k} v_j \chi_{E_j}(x)$$

where $v_1, v_2, \ldots, v_k$ are finitely many **values** and $E_1, E_2, \ldots, E_k$ are finitely many measurable **base sets**. Generally, the sets $E_1, E_2, \ldots, E_k$ may overlap, and there is some work required to show things like (1) there is a canonical expression for a simple function where the base sets are pairwise disjoint and (2) the expression for the integral of a simple function:

$$\int \psi = \sum_{j=1}^{k} v_j \mu E_j$$

does not depend on the particular values $v_1, v_2, \ldots, v_k$ or the base sets $E_1, E_2, \ldots, E_k$ used to write down a formula for the function. In any case, this gives one a definition for the integral of a simple function and for the integral of a non-negative function. Then any (measurable) function can be written uniquely as a difference of non-negative measurable functions $f = f_+ - f_-$ where $f_+(x) = \max\{f(x), 0\}$. Then one defines

$$\int f = \int f_+ - \int f_-.$$

A remark on notation: Generally, if a function $f : [a, b] \to \mathbb{R}$, then sets of measure zero will not effect the value of an integral. This is also true for Reimann integrals. When we have written $\int f$ above, this meant/means the integral over the **entire domain** of f. Keeping in mind that sets of measure zero, like single points for Lebesgue measure, make no difference, if we restrict the discussion to Lebesgue measure, then $\int f$ for $f : [a, b] \to \mathbb{R}$ may be written as

$$\int f = \int_{[a,b]} f = \int_{(a,b)} f.$$

More generally, if $E \subset [a, b]$ is a measureable subset of $\mathbb{R}$ (having finite measure) and $f : E \to \mathbb{R}$ is a measurable function, we can write (thinking of E as the **entire domain** of f)

$$\int f = \int_E f = \int f\chi_E = \int_{(a,b)} f\chi_E$$

where in the last two integrals we consider $f\chi_E : [a,b] \to \mathbb{R}$ by

$$f\chi_E(x) = \begin{cases} f(x), & x \in E \\ 0, & x \in [a,b] \backslash E. \end{cases}$$

Even more generally, these considerations allow an extremely convenient notion and notation which is essentially unavailable for the Riemann integral: If $f : E \to \mathbb{R}$ is a measurable function and A is a measurable set with $A \subset E$, then

$$\int_A f = \int f\chi_A.$$

Another remark involving the relation of functions to sets of measure zero: If $f : E \to \mathbb{R}$ is a measurable function and $\tilde{f} : E \to \mathbb{R}$ satisfies

$$\mu\{x \in E : \tilde{f}(x) \neq f(x)\} = 0,$$

then we say $\tilde{f}$ is **a version** of f. Note that in this case,

$$\int_A \tilde{f} = \int_A f \qquad \text{for every measurable set } A \subset E.$$

Here, and in the discussion below, we are primarily restricting attention to measurability and integration of functions with respect to Lebesgue measure μ, but most of these constructions and comments have some analogue for any measure. In particular, we will discuss below in the context of Lebesgue measure a **precise version** of a function $f \in L^1(a,b)$ which, as far as I know, is a notion introduced in the book *Measure Theory and Fine Properties of Functions* by Evans and Geriepy.

In some sense, the next step is to consider functions for which $\int |f| < \infty$. These are sometimes called **integrable** or **summable**. Sometimes the term **integrable** applies also to measurable functions even if $\int |f| = +\infty$. In this case, **integrable** and **measurable** mean the same thing. The set of summable functions is sometimes referred to as $L^1(a,b)$. Technically, it is more common to think of $L^1(a,b)$ as a set consisting of **equivalence classes** of functions

$$[f] = \left\{ g : \int |g - f| = 0 \right\}.$$

9.2 Basic Techniques in $L^1(a,b)$

Here is a basic result about summable functions:

Theorem 5. *If $g \in L^1(a,b)$, then $v : (a,b) \to \mathbb{R}$ by*

$$v(x) = \int_{(a,x)} g \tag{22}$$

satisfies $v \in C^0(a,b)$.

If we had and extension $\overline{g}$ of g with $\overline{g} \in C^0[a,b]$, then the Riemann integral

$$\int_a^x g(t)\,dt$$

would be well-defined and we would have agreement with the Lebesgue integral:

$$\int_{(a,x)} g = \int_a^x g(t)\,dt.$$

Furthermore, the fundamental theorem of calculus for Riemann integrable functions would hold so that v would not only be continuous, but also satisfy $v \in C^1(a,b)$ with

$$v'(x) = g(x).$$

So our theorem is not really too surprising and, in fact, more is true. These functions, however, are a good deal more complicated than continuous functions, and the result we have stated will provide a good illustration of how one can (and needs to) deal with them.

Recall the definition of continuity as applied to v: The function $v : (a,b) \to \mathbb{R}$ is continuous at $x_0 \in (a,b)$ if for each $\epsilon > 0$, there is some $\delta > 0$ such that

$$|x - x_0| < \delta \qquad \Longrightarrow \qquad |v(x) - v(x_0)| < \epsilon. \tag{23}$$

The condition on the right in (23) is

$$\left| \int_{(a,x)} g - \int_{(a,x_0)} g \right| < \epsilon. \tag{24}$$

The quantity appearing on the left in (24) is either

$$\left| \int_{(x,x_0)} g \right| \qquad \text{or} \qquad \left| \int_{(x_0,x)} g \right| \tag{25}$$

depending on whether $x < x_0$ or $x > x_0$. If $x = x_0$, then the difference of integrals in (24) is zero and (23) is clearly satisfied. Assuming $|x - x_0| < \delta$ and $x \in (a, b)$ as stipulated in (23), both quantities in (25) are bounded above by

$$\int_{(x_0-\delta, x_0+\delta)\cap(a,b)} |g|.$$

Therefore, continuity of v clearly follows from the following result:

Lemma 3 (Lebesgue continuity lemma)**.** *If $g : (a, b) \to \mathbb{R}$ with $g \in L^1(a, b)$, then given any $\epsilon > 0$, there is some $\delta > 0$ such that for every measurable set $E \subset [a, b]$*

$$\mu E < \delta \qquad \Longrightarrow \qquad \int_E |g| < \epsilon.$$

As we motivated/introduced the Lebesgue continunity lemma as a means to prove the continuity of the measurable version of an indefinite integral (22), we will use another fundamental result to outline the proof of Lebesgue's lemma, so we proceed with that discussion.

Recall that $g \in L^1(a, b)$ means (g) is measurable and) $\int |g| < \infty$. Consider for $j = 1, 2, 3, \ldots$ the function $h_j : (a, b) \to \mathbb{R}$ by

$$h_j(x) = \max\{|g(x)|, j\}. \tag{26}$$

The function h_j is measurable for each j and

$$\int |h_j| \leq j(b - a),$$

so $h_j \in L^1(a, b)$. Furthermore, $0 \leq h_j(x) \leq |g(x)|$ and the **pointwise limit**

$$\lim_{j\nearrow\infty} h_j(x) = |g(x)| \qquad \text{for every } x \in (a, b).$$

In fact, $h_1(x) \leq h_2(x) \leq h_3(x) \leq \cdots h_j(x) \nearrow |g(x)|$. We denote this condition simply by $h_j \nearrow |g|$. It is precisely under these circumstances that a result which has been called "among the most important assertions in all of analysis" applies.

Theorem 6 (Lebesgue monotone convergence theorem)**.** *If $h_j \in L^1(a, b)$ for $j = 1, 2, 3, \ldots$ and $h \in L^1(a, b)$ satisfy*

$$0 \leq h_j(x) \nearrow h(x) \qquad \textit{for every fixed } x \in (a, b),$$

then

$$\lim_{j\nearrow\infty} \int h_j = \int h.$$

The result can be nominally strengthened:

Theorem 7 (monotone convergence theorem II)**.** *If h_j is measurable for $j = 1, 2, 3, \ldots$ with*

$$0 \leq h_j(x) \nearrow h(x) \qquad \text{for every fixed } x \in (a, b),$$

then

$$\int h_j \nearrow \int h.$$

The important part is that the second version allows the possibility $\int h = \infty$, but in our case we only need the first version.

Proof of the Lebesgue continuity lemma: There is some $N \in \mathbb{N}$ such that $j > N$ implies

$$\left| \int h_j - \int |g| \right| = \int |g| - h_j < \frac{\epsilon}{2}.$$

Fix $j > N$. Then there is some δ such that $\mu E < \delta$ implies

$$\int_E h_j \leq j\mu E < \frac{\epsilon}{2}.$$

In fact, we can take $\delta = \epsilon/(2j)$. Therefore, if $E \subset (a, b)$ with $\mu E < \delta$, then

$$\int_E |g| = \int_E |g| - h_j + \int_E h_j < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \qquad \square$$

Proof of Theorem 5: Let $v(x) = \int_{(a,x)} g$. Take δ according to Lebesgue's lemma for $|g| \in L^1(a, b)$. Then for $x, x_0 \in (a, b)$ with $|x - x_0| < \delta$ we have

$$|v(x) - v(x_0)| \leq \int_{(x_0 - \delta, x_0 + \delta) \cap (a,b)} |g| < \epsilon. \qquad \square$$

It will be noted that we have actually shown v is **uniformly continuous** because δ does not depend in any way on the particular points $x, x_0 \in (a, b)$ to which we apply the lemma. In fact more is true.

9.3 Absolute Continuity and the Fundamental Theorem

Recall that the **fundamental theorem of calculus**, based on Riemann integration, is basically about **continuous** functions. More precisely, the theorem gives relations between a continuous function and a special kind of continuous function, a C^1 function. The theorem may be stated in two, essentially equivalent, ways:

1. If $f \in C^0[a,b]$ and $F \in C^1[a,b]$ with $F' = f$, then

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

2. If $f \in C^0[a,b]$, then $F : [a,b] \to \mathbb{R}$ by

$$F(x) = \int_a^x f(t)\,dt$$

satisfies $F \in C^1[a,b]$, and $F'(x) = f(x)$.

We want to give here, to the extent reasonable and possible, a version of this theorem for measurable functions. As one might expect, there will be two special kinds of measurable functions, and we will seek analogues of 1 and 2. The two special kinds of measurable functions are **summable functions** $u \in L^1(a,b)$, which we discussed above, and **absolutely continuous functions** $v \in \mathrm{AC}[a,b]$, which we will discuss below. The condition $u \in L^1(a,b)$ takes the place of continuity; these are the functions we can integrate with the Lebesgue integral. The condition $v \in \mathrm{AC}[a,b]$ takes the place of differentiability or $F \in C^1[a,b]$. The analogues are the following:

1. If $u \in L^1(a,b)$ and $v \in \mathrm{AC}[a,b]$ with **derivative**[7] $v' = u \in L^1(a,b)$, then

$$\int_{(a,b)} f(x)\,dx = v(b) - v(a).$$

2. If $u \in L^1(a,b)$, then $v : [a,b] \to \mathbb{R}$ by

$$v(x) = \int_{[a,x)} u$$

satisfies $v \in \mathrm{AC}[a,b]$, and $v' = u \in L^1(a,b)$.

The condition for **absolute continuity** looks, more or less, like a straightforward modification of continuity:

[7]This is, of course, not a classical derivative, but a measurable derivative defined almost everywhere. We will elaborate below.

Definition 4. *A function $v : [a, b] \to \mathbb{R}$ is* **absolutely continuous** *if for any $\epsilon > 0$, there is some $\delta > 0$ such that whenever*

$$a \leq a_1 \leq b_1 \leq a_2 \leq b_2 \leq \cdots \leq a_k \leq b_k \leq b \qquad \text{with} \qquad \sum_{j=1}^{k} b_j - a_j < \delta,$$

then

$$\sum_{j=1}^{k} |v(b_j) - v(a_j)| < \epsilon.$$

It's obvious that an absolutely continuous function is uniformly continuous. It's less obvious that an absolutely continuous function is differentiable at almost every point and has a summable derivative. It's even less obvious that having a derivative almost everywhere which is summable implies absolute continuity, but all these things are true, along with the fundamental theorem stated above. Thus, absolute continuity is essentially the optimal condition to go along with summability in the fundamental theorem.

Exercise 10. *We can also say $v : (a, b) \to \mathbb{R}$ is absolutely continuous if for any $\epsilon > 0$, there is some $\delta > 0$ such that whenever*

$$a < a_1 \leq b_1 \leq a_2 \leq b_2 \leq \cdots \leq a_k \leq b_k < b \qquad \text{with} \qquad \sum_{j=1}^{k} b_j - a_j < \delta,$$

then

$$\sum_{j=1}^{k} |v(b_j) - v(a_j)| < \epsilon.$$

Show that if v is absolutely continuous in this sense on an interval (a, b) with finite endpoints, then there is a unique extension overline$v \in \mathrm{AC}[a, b]$.

9.4 Lebesgue Points

Another useful thing to know about summable functions concerns **averages** of the values. Let us consider here $u : U \to \mathbb{R}$ where U is an open subset of $\mathbb{R}^n$, the function u is measurable, and we consider integration with respect to Lebesgue measure. The **average** value of u over a measurable set $E \subset U$ with $0 < \mu E < \infty$ is

$$\frac{1}{\mu E} \int_E u.$$

If $\int_K |u| < \infty$ whenever $K \subset U$ is compact, we say $u \in L^1_{loc}(U)$.

Theorem 8 (Lebesgue's differentiation theorem)**.** *If $u \in L^1_{loc}(U)$, then for almost every $x_0 \in U$*

$$\lim_{r \searrow 0} \frac{1}{\mu B_r(x_0)} \int_{B_r(x_0)} u = u(x_0). \tag{27}$$

Definition 5. *Given a function $u \in L^1_{loc}(U)$, a point $x_0 \in U$ is called a* **Lebesgue point** *of u if*

$$\lim_{r \searrow 0} \frac{1}{\mu B_r(x_0)} \int_{B_r(x_0)} |u - u(x_0)| = 0.$$

Exercise 11. *Show that (27) holds at every Lebesgue point x_0, but that (27) can also hold at a point $x_0 \in U$ which is not a Lebesgue point. And of course it is possible that*

$$\lim_{r \searrow 0} \frac{1}{\mu B_r(x_0)} \int_{B_r(x_0)} u$$

exists but

$$\lim_{r \searrow 0} \frac{1}{\mu B_r(x_0)} \int_{B_r(x_0)} u \neq u(x_0).$$

Nevertheless, the following is true:

Theorem 9. *If $u \in L^1_{loc}(U)$, then almost every point of U is a Lebesgue point.*

Definition 6. *Given $u \in L^1_{loc}(U)$, we say the function*

$$u_1(x) = \begin{cases} \lim_{r \searrow 0} \frac{1}{\mu B_r(x_0)} \int_{B_r(x_0)} u, & \text{if this limit exists} \\ 0, & \text{otherwise.} \end{cases}$$

is the **precise pointwise version** *of u.*

Exercise 12. *We say $\tilde{u} = u \in L^1_{loc}(U)$ if*

$$\int_K |\tilde{u} - u| = 0 \qquad \text{for every compact set } K \subset U.$$

Show that if $\tilde{u} = u \in L^1_{loc}(U)$, then one **cannot** *say $\tilde{u}(x) = u(x)$ at any particular single point $x \in U$. Show, however, that if $\tilde{u} = u \in L^1_{loc}(U)$, then the precise pointwise versions are equal:*

$$\tilde{u}_1(x) \equiv u_1(x) \qquad \text{for every } x \in U.$$

Note: Theorem 1 of §4.9 of Evans and Gariepy states that[8] if $u \in W^1(\mathbb{R})$, then the precise pointwise version of u satisfies $u_1 \in \mathrm{AC}[a,b]$ for every $a, b \in \mathbb{R}$ with $a < b$ and $u_1 \in L^1_{loc}(\mathbb{R})$. He shows (absolute) continuity of u_1 using a mollification argument I do not completely understand, but it should provide an alternative to the argument we used above to show continuity in Theorem 5. This is also relevant to Problem 1 on Assignment 9 (MATH 6702 Spring semester 2020) and Problems 1 and 2 on Assignment 11 (of the same course). In particular, one should be able to use some version of the mollification approach of Evans and Gariepy to modify the discussion of Problems 1 and 2 of Assignment 11 resulting in a proof of continuity.

10 Integration Spaces

10.1 The L^p spaces

Given an open set $U \subset \mathbb{R}^n$, we consider the following sets of real valued measurable functions:

$$L^p(U) = \left\{ u : \int |u|^p < \infty \right\} \qquad \text{for } 0 < p < \infty.$$

$$L^\infty(U) = \left\{ u : \operatorname{essup} |u| < \infty \right\}.$$

It may be important to note that the "functions" in these classes are ususally considered as **equivalence classes** of functions rather than functions that are traditionally pointwise defined. This means for example that as functions in $L^1(0,1)$ the functions $f, g, h : (0,1) \to \mathbb{R}$ defined pointwise by

$$f(x) = 0, \qquad g(x) = \begin{cases} 0, & x \neq 0 \\ 1, & x = 0, \end{cases} \qquad \text{and} \qquad h(x) = \begin{cases} 0, & x \notin \mathbb{Q} \\ 1, & x \in \mathbb{Q} \end{cases}$$

respectively, are considered **essentially indistinguishable** from one another. The point can be expressed in many ways. One can observe, for example, that

$$\int_{(0,1)} f\phi = \int_{(0,1)} g\phi = \int_{(0,1)} h\phi \qquad \text{for all } \phi \in C_c^\infty(0,1).$$

Alternatively,

$$\int_{(0,1)} |f - g| = \int_{(0,1)} |f - h| = \int_{(0,1)} |g - h| = 0.$$

[8] See the discussion of weak derivatives below.

For any function $u \in L^p(U)$, we set

$$\|u\|_{L^p} = \left(\int_U |f|^p \right)^{1/p}. \tag{28}$$

This is called the L^p **norm** on $L^p(U)$. When $1 \leq p < \infty$, the expression (28) defines a norm. The non-negative homogeneity and positive definiteness or essentially obvious. The triangle inequality

$$\|u + v\|_{L^p} \leq \|u\|_{L^p} + \|v\|_{L^p} \tag{29}$$

is called the **Minkowski inequality**. This is used, for example, to show L^p is closed under addition, and it follows that $L^p(U)$ is a **normed vector space**.

Even for $0 < p < 1$, the space $L^p(U)$ is a vector space, but the L^p norm fails to be a norm in these cases. It is still non-negative homogeneous and positive definite, but the trangle inequality must be relaxed to

$$\|u + v\|_{L^p} \leq M(\|u\|_{L^p} + \|v\|_{L^p})$$

for some constant M. Such a function is called a **quasinorm** and so, while $L^p(U)$ for $0 < p < 1$ is not a normed space, setting

$$d_p(u, v) = \|u - v\|_{L^p}^p$$

still makes $L^p(U)$ for $0 < p < 1$ a metric space.[9]

The norm on $L^\infty(U)$ is a little more complicated but rather instructive with regard to the nature of these "functions."

$$\|u\|_{L^\infty} = \sup\{m : \mu\{x : |u(x)| \geq m\} > 0\}.$$

The expression on the right here is called the **essential sup**, i.e., essential supremum, of $|u|$. In addition to the L^∞ norm, $\| \cdot \|_{L^\infty}$ is also called the **essential sup norm**. If $\|f - g\|_{L^p} = 0$, then $\{x : f(x) \neq g(x)\}$ has measure zero. Thus, f and g are in the same equivalence class as L^p or measurable functions.

The L^∞ norm is also a norm. The triangle inequality here is also considered a form of the Minkowski inequality.

When U is a bounded open subset of $\mathbb{R}^n$, then the L^p spaces are nested with

$$L^q(U) \subset L^p(U) \qquad \text{for } 1 \leq p < q \leq \infty. \tag{30}$$

[9]Recall the definition of **metric space** from the notes on differentiation. See also further discussion of metric spaces below.

One approach to showing this relation is to use **Hölder's inequality**: If $1 < p \leq q < \infty$ and

$$\frac{1}{p} + \frac{1}{q} = 1,$$

then p and q are called **Hölder conjugate exponents**, and

$$\|uv\|_{L^1} \leq \|u\|_{L^p}\|v\|_{L^q} \qquad \text{whenever } u \in L^p(U) \text{ and } v \in L^q(U).$$

It may be viewed as a consequence of this inequality that if $u \in L^p(U)$ and $v \in L^q(U)$, then the product $uv \in L^1(U)$. The exponents $p = 1$ and $q = \infty$ are also considered to be **Hölder conjugates**, and the inequality

$$\|uv\|_{L^1} \leq \|u\|_{L^1}\|v\|_{L^\infty} \qquad \text{whenever } u \in L^1(U) \text{ and } v \in L^\infty(U)$$

also holds and is called a Hölder inequaility.

Exercise 13. *Use the Hölder inequalities to show (30) holds for* $1 \leq p \leq \infty$.

Exercise 14. *Prove the following* **generalized Hölder inequality** *by induction: If* $1 \leq p_1 \leq p_2 \leq \cdots \leq p_k \leq \infty$ *with*

$$\sum_{p_j \neq \infty} \frac{1}{p_j} = 1$$

and $u_j \in L^{p_j}(U)$ *for* $j = 1, 2, \ldots, k$, *then* $u_1 u_2 \cdots u_k \in L^1(U)$ *and*

$$\|u_1 u_2 \cdots u_k\|_{L^1(U)} \leq \|u_1\|_{L^{p_1}(U)}\|u_2\|_{L^{p_2}(U)} \cdots \|u_k\|_{L^{p_k}(U)}.$$

10.2 Metric Completeness

Recall the definition of a **metric space** from the notes on differentiation as a set X with a notion of distance, or a **distance function** defined on $X \times X$. A sequence $x_1, x_2, x_3, \ldots$ in a metric space is said to be **Cauchy** if for any $\epsilon > 0$ there is some $N \in \mathbb{N}$ with

$$j, k > N \qquad \Longrightarrow \qquad d(x_j, x_k) < \epsilon.$$

This condition says that the elements in the seauence "bunch up" at the "end" with respect to the distance. You may recall that a sequence, like the one above, is said to **converge** to an element $x \in X$ if for any $\epsilon > 0$, there is some $N \in \mathbb{N}$ such that

$$j > N \qquad \Longrightarrow \qquad d(x_j, x) < \epsilon.$$

In this case we say the sequence **has limit** x, and we write

$$\lim_{j\to\infty} x_j = x.$$

Notice, first of all, that the definition of convergence has integral to it, the point x which is the **limit** while the Cauchy condition does not even mention any specific limit. This is no accident.

Exercise 15. *Show that if a sequence $\{x_j\}_{j=1}^{\infty}$ in a metric space X converges to a limit $x \in X$, then the sequence is Cauchy.*

Exercise 16. *Recall that any subset of a metric space is a metric space itself using the restriction of the metric on the larger space. (This sort of thing will not work* **at all** *for a normed space, since general subsets will fail to be vector subspaces.) In particular, the interval $X = (0, 1)$ is a perfectly good metric space. Show that $\{1/j\}_{j=1}^{\infty}$ is Cauchy but does not converge in X. Hint: This sequence converges in the larger metric space $\mathbb{R}^1$.*

The previous two exercises are, in some sense, the key to understanding **metric completeness**.

Definition 7. *A metric space is said to be* **complete** *if every Cauchy sequence converges to some limit (in the space).*

Thus, $X = (0, 1)$ is not complete as a metric space, but $X = [0, 1]$ is complete. The vector space $C^0[a, b]$ is not complete under the norm induced metric

$$\|f - g\|_{L^1} = \int_a^b |f(x) - g(x)|\, dx,$$

but $C^0[a, b]$ is complete under the C^0 norm. (Can you prove that?) Also, the space $L^p(U)$ is complete under (the distance induced by) the L^p norm. With this distinction of metric completeness in mind, here are some common disignations you may be glad to know about:

> A normed vector space which is metrically complete with respect to the distance induced by the norm is called a **Banach space**.
>
> An inner product space which is metrically complete with respect to the distance induced by the inner product is called a **Hilbert space**.

See below for the definition of an **inner product space**.

10.3 $L^2(U)$ as an Inner Product Space

We have already seen that when we have a norm $\| \cdot \| : V \to [0, \infty)$ on a vector space, then we can define a distance function or metric $d : V \times V \to [0, \infty)$. The metric does not need the vector space structure of V, and in this way were were motivated to consider a less structured kind of "space" called a metric space. We repeat: Every normed space is a metric space with distance function $d(v, w) = \|v - w\|$. There are axioms for a metric space which can even be just a set X with no vector space structure.

It is also possible to have a vector space with **more structure** than a normed vector space.

Definition 8. *An* **inner product space** *is a vector space V together with a bilinear form $\langle \cdot\ ,\ \cdot \rangle : V \times V \to \mathbb{R}$ satisfying the following conditions:*

(i symmetric) *$\langle v, w \rangle = \langle w, v \rangle$ for all $v, w \in V$.*

(ii bilinear) *$\langle av + bw, z \rangle = a\langle v, z \rangle + b\langle w, z \rangle$ for all $a, b \in \mathbb{R}$ and $v, w, z \in V$. By symmetry one also has*

$$\langle z, av + bw \rangle = a\langle z, v \rangle + b\langle z, w \rangle \qquad \text{for all } a, b \in \mathbb{R} \text{ and } v, w, z \in V.$$

(iii positive definite) *$\langle v, v \rangle \geq 0$ with $\langle v, v \rangle = 0$ if and only if $v = \mathbf{0}$.*

Any time one has an inner product, one has that

$$\|v\| = \sqrt{\langle v, v \rangle} \qquad \text{is a norm.} \tag{31}$$

This is called the **norm induced by the inner product**. Thus, every inner product space is a normed space. The usual approach to showing the norm given in (31) satisfies the triangle inequality for norms is to use the following inequality which is called the **Cauchy-Schwarz inequality** and holds in any inner product space:

$$|\langle u, v \rangle| \leq \|u\|\|v\| \qquad \text{for all } u, v \in V.$$

In the Cauchy-Schwarz inequality, the norm used is the one induced by the inner product. (There could be other norms on the same space.)

In particular, $L^2(U)$ is an inner product space with

$$\langle u, v \rangle = \int_U uv.$$

Note that one needs to know $uv \in L^1$ when $u, v \in L^2(U)$. This follows from the Hölder inequality since $p = 2$ is the Hölder conjugate exponent of itself. Thus, the Cauchy-Schwarz inequality in $L^2(U)$ tells us

$$\int_U |uv| \leq \left(\int_U u^2 \right)^{1/2} \left(\int_U v^2 \right)^{1/2}.$$

Recall that a normed space which is metrically complete with respect to the distance induced by the norm is called a Banach space. Similarly, an inner product space which is metrically complete as a metric space (i.e., a Banach space with respect to the norm) is called a **Hilbert space**.

10.4 Weak Derivatives and Sobolev Spaces

Given a function $u \in L^1_{loc}(U)$ with U a bounded open subset of $\mathbb{R}^n$, we say $g_j \in L^1_{loc}(U)$ is **a weak j-th partial derivative of** u if

$$\int_U u \frac{\partial \phi}{\partial x_j} = -\int_U g_j \phi \qquad \text{for every } \phi \in C_c^\infty(U).$$

The vector space of functions $u \in L^1_{loc}(U)$ having first weak partial derivatives $g_j \in L^1_{loc}(U)$ for each index $j = 1, 2, \ldots, n$ is denoted by $W^1_{loc}(U)$. If u and the weak partial derivatives g_j for $j = 1, 2, \ldots, n$ of a function $u \in W^1_{loc}(U)$ are actually in $L^1(U)$, then we write $u \in W^1(U)$. Either of the spaces $W^1_{loc}(U)$ or $W^1(U)$ may be called **the space of weakly differentiable functions**. Both are vector spaces. The latter is a normed space which, when equipped with the norm

$$\|u\|_{W^1} = \int_U |u| + \sum_{j=1}^n \int_U |g_j| = \|u\|_{L^1(U)} + \sum_{j=1}^n \|g_j\|_{L^1(U)},$$

is also denoted by $W^{1,1}(U)$ and is called **the space of weakly differentiable functions with one derivative in** L^1. The terms in the norm involving the derivatives comprise a **natural seminorm** on $W^{1,1}(U)$ in a manner very similar to the C^α Hölder seminorm. This natural seminorm is denoted by

$$[u]_{W^1} = \sum_{j=1}^n \int_U |g_j|.$$

As usual, a **seminorm** is non-negative valued, non-negative homogeneous, and satisfies the triangle inequality for norms, but may be only non-negative semi-definite

(and fail to be positive definite). In particular, this seminorm vanishes on all constant functions just like the C^α seminorm. We will see an example of an *unnatural seminorm* below which has the positive definiteness fail in a different way.

Recall that $L^1_{loc}(U)$ consists of (equivalence classes of) measurable functions $u : U \to \mathbb{R}$ which have the property

$$\int_K |u| < \infty \qquad \text{for every } K \subset\subset U.$$

The space $L^1_{loc}(U)$, and consequently the space $W^1_{loc}(U)$ is not a normed space; **there is no W^1_{loc} norm**. There is, however, a notion of **convergence** of sequences (of functions) in the space $W^1_{loc}(U)$:

> A sequence $\{u_j\}_{j=1}^\infty \subset W^1_{loc}(U)$ is said to **converge** to a function $u \in W^1_{loc}(U)$ if
>
> $$\lim_{j\to\infty} \|u_j - u\|_{W^{1,1}(K)} = 0$$
>
> for every $K \subset\subset U$.

There is no problem, in this context of considering the norm and the functions u and g_j for $j = 1, 2, \ldots, n$ on a potentially more complicated set $K \subset\subset U$; we are already assuming the functions are defined on U. When this convergence is considered, the space $W^1_{loc}(U)$ is denoted by $W^{1,1}_{loc}(U)$. Just to be clear, there is no fundamental difference between $W^1(U)$ and $W^{1,1}(U)$ (or $W^1_{loc}(U)$ and $W^{1,1}_{loc}(U)$). The former notation is used simply to signal the consideration (or presence or appearance in the discussion) of **weak derivatives**. The latter notation suggests also the consideration of a **topology** on the space.

Remark: The point of considering spaces like $L^1_{loc}(U)$ and $W^1_{loc}(U)$ is to allow important functions which might be otherwise non-integrable. For example, the function $u(x) = 1/x$ is a perfectly reasonable function to consider in $C^0(0,1)$. Note, however, that $u \notin L^1(0,1)$, but $u \in L^1_{loc}(0,1)$. In fact, as has been observed in the notes on continuity and differentiability, u is not in $C^0(a,b)$ when considered as a normed space.[10] Thus, we can say also $u \in C^\infty_{loc}(0,1)$.

[10]Here is an example, where the name of a function space is used to indicate two different spaces and one, generally, needs to know from the context which one is being considered. Specifically, it is commone to write $u \in C^0(a,b)$ to simply mean u is continuous on the open interval (a,b). But when the (norm) topology is under consideration, $u \in C^0(a,b)$ means

$$u \in \{u \in C^0(a,b) : \|u\|_{C^0} = \sup |u| < \infty\}$$

Topology

As a partial aside, let me try to explain the use of the word **topology** here. You may recall that we have considered various notions of what has loosely been called **structure** on sets. More precisely, there is potentially algebraic structure and analytic structure on a set. If the set is a vector space, then that constitutes an algebraic structure on the set. Most sets of functions we have considered are vector spaces. If there is a norm (or inner product) on a vector space, then the norm constitutes an analytic structure on the space. An inner product always gives a norm, so having an inner product constitutes more (analytic) structure on a vector space. Any normed (vector) space has a notion of distance and, consequently, is a metric space. And we have seen that any set (not necessarily) a vector space may be endowed with a distance function (or metric) with certain axiomatic properties. Thus, a metric space can have less structure than a normed space both analytically and algebraically. A metric space (and all the spaces with at least this structure) are said to have **geometric structure**. Note that "geometry" means literally **measurement of the Earth** or more generally "the measurement of the world" or **space**, and note the presence of "metric" in "geo-metric."

It is possible to have a very natural structure on a set which is possessed by every metric space and which can be isolated axiomatically but does not allow measurement (it is not geometric) but it does allow some other notions like convergence, and thus (in some abstract sense) a notion of closeness. This lower level of structure on a set is called a **topology**. A set of subsets $\mathcal{T}$ of a set X is said to be a **topology** on X or **collection of open sets** and X is said to be a **topological space** if the following hold:

1. $\phi, X \in \mathcal{T}$. (The empty set and the entire space are open.)
2. Whenever $\{U_\alpha\}_{\alpha\in\Gamma} \subset \mathcal{T}$, one has
$$\bigcup_{\alpha\in\Gamma} U_\alpha \in \mathcal{T}.$$
(Any arbitrary union of open sets is open; the collection (the topology) $\mathcal{T}$ is closed under arbitrary unions.)
3. Whenever $U_1, U_2, \ldots, U_k \in \mathcal{T}$, then
$$\bigcap_{j=1}^{k} U_j \in \mathcal{T}.$$

(An intersection of finitely many open sets is open. The topology is closed under finite intersections.)

Exercise 17. *Let X be any metric space X with distance function $d : X \times X \to [0, \infty)$. Consider the collection $\mathcal{T}$ of all sets $U \subset X$ with the following property:*

Whenever $p \in X$, then there exists some $r > 0$ for which

$$B_r(p) = \{x \in X : d(x, p) < r\} \subset U.$$

Show that $\mathcal{T}$ is a topology.

A sequence $\{x_j\}_{j=1}^{\infty}$ in a topological space X is said to **converge** to an element $x \in X$ if for any open set U with $x \in U$, there exists some $N \in \mathbb{N}$ for which

$$j > N \qquad \Longrightarrow \qquad x_j \in U.$$

One must be a bit careful (at least sometimes) with this notion of "closeness."

Exercise 18. *Show that there exists a topological space containing a sequence which converges to two different points. That is, limits of sequences in a topological space are not always unique. Hint: Consider a "space" with only two points.*

Most topological spaces considered in analysis have the following property:

A topological space X is said to be **Hausdorff** if given any two (distinct) points x_1 and x_2 in X, there exist disjoint open sets U_1 and U_2 with $x_j \in U_j$ for $j = 1, 2$.

Exercise 19. *Show that the limit of a sequence in a Hausdorff space is unique and every metric space is a Hausdorff space.*

The topology defined in Exercise 17 is called the **metric topology**. And naturally, since every normed (or inner product) space is a metric space, any such space has a metric topology. It will be noted, however, that we did not define a norm on $W^1_{loc}(U) = W^{1,1}_{loc}(U)$. This space does not have a norm, and it doesn't have a natural metric either. Nevertheless, it is a (Hausdorff) topological space. The topology on $W^{1,1}_{loc}(U)$ (or $L^1_{loc}(U)$ for that matter) is called a **topology of local convergence**. Certainly, the easiest way to think about such topologies is in terms of the convergence of sequences described above. The formal definition of open sets in terms of this convergence on compact subsets is a little delicate. Perhaps the easiest way to describe

this topology is in terms of a metric. Though there is no natural metric, the resulting topology is *metrizable* which means there does exist a metric so that the topology is the metric topology with respect to that metric. Such a metric is given as follows: Take a sequence of compactly nested open sets $U_1 \subset\subset U_2 \subset\subset U_3 \subset\subset \cdots \subset U$ with $\cup U_j = U$. (This is called an *exhaustion* of U.) On each such subset U_j, we have the $W^{1,1}(U_j)$ norm. Using this norm (via restriction) we can define a **seminorm**

$$]u[_j = \|u\|_{W^{1,1}(U_j)}$$

on $W^1_{loc}(U)$. Then

$$d(u,v) = \sum_{j=1}^{\infty} \frac{]u-v[_j}{1+]u-v[_j} \tag{32}$$

is a metric inducing the correct topology.

Exercise 20. *Show that $d : W^1_{loc}(U) \times W^1_{loc}(U) \to [0,\infty)$ is a metric and that if a sequence $\{u_j\}_{j=1}^{\infty}$ converges to $u \in W^{1,1}_{loc}(U)$ in the induced metric topology, then the sequence converges to the (restriction of the) same limit in $W^{1,1}(V)$ for every $V \subset\subset U$.*

Metric Completeness

It turns out that the spaces $W^{1,1}(U)$ and $W^{1,1}_{loc}(U)$ are metrically complete. In particular $W^{1,1}(U)$ is a Banach space.

Other Sobolev Spaces

The spaces $W^{1,1}(U)$ and $W^{1,1}_{loc}(U)$ are called Sobolev spaces. More generally, for $p \geq 1$, the spaces of L^p functions with weak (partial) derivatives of order k in L^p are denoted by $W^{k,p}(U)$. The locally integrable counterparts are $W^{k,p}_{loc}(U)$.

These spaces are also all metrically complete. In particular, $W^{k,p}(U)$ is a Banach space for every k and every p, and very significantly $W^{k,2}(U)$ is a Hilbert space with the natural inner product

$$\langle u, v\rangle_{W^{k,2}} = \sum_{|\alpha|\leq k} \int_U D^\alpha u D^\alpha v.$$

Here $D^\alpha u$ denotes the weak partial derivative of multi-index order α defined as a function $g_\alpha \in L^1_{loc}(U)$ satisfying

$$\int_U g_\alpha\, \phi = (-1)^{|\alpha|} \int_U u\, D^\alpha \phi \qquad \text{for each } \phi \in C^\infty_c(U).$$

Of course, on the right

$$D^\alpha \phi = \frac{\partial^{|\alpha|}\phi}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}$$

is a classical derivative. The operator $L^* : C^k(U) \to C^0(U)$ by

$$L^* v = (-1)^{|\alpha|} D^\alpha v$$

is called the **weak adjoint** of D^α. The Hilbert space $W^{k,2}(U)$ is usually called $H^k(U)$. This is the central space under consideration in showing the existence of weak solutions for elliptic PDE.

Additional Properties of Sobolev Spaces

If U is a bounded open set in $\mathbb{R}^n$ and ∂U is C^1, then

$$W^{1,\infty}(U) = \mathrm{Lip}(U).$$

Also, if U is just an open subset of $\mathbb{R}^n$, then

$$W^{1,\infty}_{loc}(U) = \mathrm{Lip}_{loc}(U).$$

Taking $p = 2$, the space $H^1(U) = W^{1,2}$ is a Hilbert space with

$$\langle u, v \rangle_{W^{1,2}} = \langle u, v \rangle_{H^1} = \langle u, v \rangle_{L^2} + \sum_{j=1}^{n} \langle D_j u, D_j v \rangle_{L^2} = \int uv + \int Du \cdot Dv.$$

10.5 Poincaré Inequalities

Here we state two results concerning L^2 bounds for a function in terms of its derivative. These results can and will be used to show that the bilinear form $B : H_0^1(U) \times H_0^1(U) \to \mathbb{R}$ by

$$B[u, v] = \int_U Du \cdot Dv$$

is an **alternative inner product** on $H_0^1(U)$. We will prove the first Poincaré inequality in a nontrivial special case. Recall that in general, say on $H^1(U)$ or $W^{1,p}(U)$, the quantity

$$[u]_{W^{1,p}} = \left(\sum_{j=1}^{n} \|D_j u\|_{L^p}^p \right)^{1/p} \tag{33}$$

is not a norm, though it is non-negative homogeneous and satisfies the triangle inequality for norms. The expression in (33) is even non-negative semi-definite, but it defines only a **seminorm** because it fails to be positive definite. It vanishes on a nonzero constant for example. This same defect keeps $B[u, v]$ from being an inner product: B is symmetric, bilinear, and non-negative semi-definite, but in general B is not positive definite for exactly the same reason $[\,\cdot\,]_{W^{1,2}}$ is not positive definite.

Notice, however, that if we restrict to functions $\phi \in C_c^\infty(U)$, then by forcing the boundary values to be zero, if we have

$$[\phi]_{W^{1,p}} = 0,$$

then we do have that the constant function ϕ is $\phi \equiv 0$. Thus, we seem to have gone some distance in making the defect in this seminorm go away. The Poincaré inequalities establish that we can actually make B an inner product if we restrict to $p = 2$ and $H_0^1(U) = W_0^{1,2}(U)$.

Theorem 10. *(smooth Poincaré inequality) If U is a bounded open subset of $\mathbb{R}^n$, then there is a constant C depending only on the dimension n and the domain U such that for any $\phi \in C_c^\infty(U)$ one has*

$$\|\phi\|_{L^2(U)} \leq C\|D\phi\|_{L^2(U)}$$

where

$$\|D\phi\|_{L^2(U)} = \| \, |D\phi| \, \|_{L^2(U)} = \left(\int_U |D\phi|^2 \right)^{1/2}$$

and

$$|D\phi| = \sqrt{\sum_{j=1}^{n} \left(\frac{\partial \phi}{\partial x_j} \right)^2}. \tag{34}$$

The constant we will obtain in the case $n = 2$ is $C = \sqrt{\mu(U)}$ where $\mu(U)$ is the n-dimensional Lebesgue measure of U. Note that in (34) these are classical derivatives of ϕ.

Proof in the case $n = 2$ when $U \subset \mathbb{R}^2$: Naturally, we may assume ϕ is defined on all of $\mathbb{R}^2$ by setting $\phi(\mathbf{x}) \equiv 0$ for $\mathbf{x} \in \mathbb{R}^2 \setminus \operatorname{supp}(u)$. If $\mathbf{x} = (x_1, x_2) \in U$, then by the fundamental theorem of calculus

$$\phi(\mathbf{x}) = \int_{-\infty}^{x_1} D_1\phi(\xi_1, x_2)\, d\xi_1$$

and

$$\phi(\mathbf{x}) = \int_{-\infty}^{x_2} D_2\phi(x_1, \xi_2)\, d\xi_2.$$

Consequently,

$$|\phi(\mathbf{x})|^2 \leq \left(\int_{\xi_1\in\mathbb{R}} |D\phi(\xi_1, x_2)|\right)\left(\int_{\xi_2\in\mathbb{R}} |D\phi(x_1, \xi_2)|\right).$$

Therefore,

$$\begin{aligned}\int_{x_1\in\mathbb{R}} |\phi(\mathbf{x})|^2 &\leq \int_{x_1\in\mathbb{R}} \left(\int_{\xi_1\in\mathbb{R}} |D\phi(\xi_1, x_2)|\right)\left(\int_{\xi_2\in\mathbb{R}} |D\phi(x_1, \xi_2)|\right) \\ &= \left(\int_{\xi_1\in\mathbb{R}} |D\phi(\xi_1, x_2)|\right)\int_U |D\phi|\end{aligned}$$

and

$$\int_U |\phi|^2 \leq \int_{x_2\in\mathbb{R}} \left(\int_{\xi_1\in\mathbb{R}} |D\phi(\xi_1, x_2)|\right)\int_U |D\phi| = \left(\int_U |D\phi|\right)^2.$$

That is,

$$\|\phi\|_{L^2(U)} \leq \| \, |D\phi| \, \|_{L^1(U)} = \|D\phi\|_{L^1(U)}.$$

Finally, because U is a bounded domain, we know the constant $\chi_U \equiv 1 \in L^2(U)$, so by the Cauchy-Schwarz inequality

$$\|D\phi\|_{L^1(U)} = \langle |D\phi|, \chi_U\rangle_{L^2(U)} \leq \|\chi_U\|_{L^2(U)}\|D\phi\|_{L^2(U)} = \sqrt{\mu(U)}\|D\phi\|_{L^2(U)}.$$

Therefore,

$$\|\phi\|_{L^2(U)} \leq \sqrt{\mu(U)}\|D\phi\|_{L^2(U)}$$

as was to be shown in the special case. □

A stronger form of the Poincaré inequality above can be obtained in the case $n > 2$ which is called the Gagliardo-Nirenberg-Sobolev inequality. The following exercise outlines the first few steps in obtaining the Gagliardo-Nirenberg-Sobolev inequality (which are a relatively straightforward generalization of the argument given above). See Theorem 1 of section 5.6 in Craig Evans' book *Partial Differential Equations* for details.

Exercise 21. *Given p and p^* with $1 \leq p < n$ and*

$$\frac{1}{p} - \frac{1}{p^*} = \frac{1}{n},$$

we say p and p^* are **Sobolev conjugate exponents**.

1. *Show/observe that $p^* > p$.*

2. *Show that*
$$p^* = \frac{np}{n-p}.$$
In particular, if $p = 2$ (and $n \geq 3$), then $p^* = 2n/(n-2) > 2$.

3. *Let $\phi \in C_c^\infty(\mathbb{R}^n)$ and consider the case $p = 1$ and $p^* = n/(n-1)$. Let $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Use the fundamental thoerem of calculus to write*
$$\phi(\mathbf{x}) = \int_{-\infty}^{x^j} D_j\phi(x_1, \ldots, \xi_j, \ldots, x_n)\, d\xi_j.$$
Notice that this can be done for $j = 1, 2, \ldots, n$, and it follows that
$$|\phi(\mathbf{x})| \leq \int_{\xi_j \in \mathbb{R}} |D\phi(x_1, \ldots, \xi_j, \ldots, x_n)|$$
and
$$|\phi(\mathbf{x})|^{1/(n-1)} \leq \left(\int_{\xi_j \in \mathbb{R}} |D\phi(x_1, \ldots, \xi_j, \ldots, x_n)| \right)^{1/(n-1)}.$$
Conclude that
$$|\phi(\mathbf{x})|^{n/(n-1)} \leq \prod_{j=1}^{n} \left(\int_{\xi_j \in \mathbb{R}} |D\phi(x_1, \ldots, \xi_j, \ldots, x_n)| \right)^{1/(n-1)}.$$
Now, integrate interatively with respect to x_1, x_2, …, x_n, simplifying after each integration to conclude
$$\int_U |\phi|^{n/(n-1)} \leq \left(\int_U |Du| \right)^{n/(n-1)}. \tag{35}$$
Hint: Note that one of the factors at each stage is a constant with respect to the integration on the right. Then use the general Hölder inequality on the other integral.

4. *The inequality (35) is the Gagliardo-Nirenberg-Sobolev inequality when $p = 1$. (Note that when $p = 1$ we have $C = 1$.) The general Gagliardo-Nirenberg-Sobolev inequality is*
$$\|\phi\|_{L^{p^*}(\mathbb{R}^n)} \leq C\|D\phi\|_{L^p(\mathbb{R}^n)} \qquad \text{for } \phi \in C_c^\infty(\mathbb{R}^n).$$

Use this inequality with $p = 2$ and the Hölder inequality to show the smooth Poincaré inequality stated above.

Using the fact that $H_0^1(U)$ is the closure of $C_c^\infty(U)$, we can get the following result:

Theorem 11. *(Poincaré Sobolev inequality) If U is a bounded open subset of $\mathbb{R}^n$, then there is a constant C depending only on the dimension n and the domain U such that for any $v \in H_0^1(U)$ one has*

$$\|v\|_{L^2(U)} \leq C\|Dv\|_{L^2(U)}$$

where

$$\|Dv\|_{L^2(U)} = \| \, |Dv| \, \|_{L^2(U)} = \left(\int_U |Dv|^2 \right)^{1/2}$$

and

$$|Dv| = \sqrt{\sum_{j=1}^{n} (D_j v)^2}. \tag{36}$$

In (36) these are weak derivatives of v.

11 Continuous/Bounded Linear Operators

The introduction of a number of spaces of integrable functions, namely the L^p spaces and the Sobolev spaces $W^{k,p}$, provides a natural collection of **normed spaces** on which to apply the following concepts which are properly part of the subject of **functional analysis**. The C^k and $C^{k,\alpha}$ Hölder spaces are also normed spaces and certain aspects of this discussion apply. Of special interest here, however, are the **inner product spaces** the main examples of which are the space L^2 and $H^k = W^{k,2}$. In these spaces we have the **Riesz representation theorem** which will be covered here.

11.1 Examples and Some Basics

Let V and W be normed vector spaces. A function $f : V \to W$ is called an **operator**. Examples are the **differential operators** like

$$\frac{\partial}{\partial x_j} : C^1(U) \to C^0(U) \qquad \text{and} \qquad \Delta : C^2(U) \to C^0(U).$$

Other examples are given by **integral operators** like

$$\mathcal{F} : C_c^\infty(U) \to \mathbb{R} \qquad \text{by} \qquad \mathcal{F}[\phi] = \int_U f\phi.$$

This is the **distributational representation** of a real valued function $f \in L^1_{loc}(U)$. In the special case of an operator $\mathcal{F} : V \to W$ where W is the field $\mathbb{R}$, we call $\mathcal{F}$ a **functional**. Thus, $\mathcal{F}$ here is an integral functional.

Another integral operator is the **convolution** or **mollification operator**

$$\mathcal{G} : L^1_{loc}(\mathbb{R}^n) \to C_c^\infty(\mathbb{R}^n) \qquad \text{by} \qquad \mathcal{G}[u] = \int_{\xi \in U} u(\xi)\phi(x-\xi).$$

Here ϕ is a fixed function in $C_c^\infty(\mathbb{R}^n)$ and $u^* = \mathcal{G}[u]$ may be shown to approximate u in various ways. Notice that the integral expression is a function of x. For this functional, the domain $L^1_{loc}(\mathbb{R}^n)$ appears as a stand in for a space of functions with minimal regularity. In applications, this space is often replaced with a space having a specific useful norm in which to measure the approximation.

As a third example of an integral operator, we can take the **weak Laplacian adjoint operator** or the **Dirchlet form**:

$$B : H_0^1(U) \times H_0^1(U) \to \mathbb{R} \qquad \text{by} \qquad B[u,v] = \int_U Du \cdot Dv.$$

The functional $\mathcal{D} : H_0^1(U) \to \mathbb{R}$ by $\mathcal{D}[u] = B[u,u] = \int_U |Du|^2$ is called the Dirichlet energy. If u measures the temperature in a region, this is a measure of the potential energy of heat diffusion.

An operator which is neither a differential operator nor an integral operator is the **evaluation functional** or **delta distribution**:

$$\mathcal{E}_p : C^0(U) \to \mathbb{R} \qquad \text{by} \qquad \mathcal{E}_p[u] = u(p).$$

All of the operators above are **linear** (or in the case of the Dirichlet form bilinear).

Exercise 22. *Show the evaluation functional is linear.*

There are also nonlinear operators. For example, the **area functional** from the calculus of variations given by

$$\text{area} : C^1(\overline{U}) \to \mathbb{R} \qquad \text{by} \qquad \text{area}[u] = \int_U \sqrt{1 + |Du|^2}$$

is not linear. The **weak mean curvature adjoint form**

$$\mathcal{M} : H_0^1(U) \times H_0^1(U) \to \mathbb{R} \qquad \text{by} \qquad \mathcal{M}[u, v] = \int_U \frac{Du \cdot Dv}{\sqrt{1 + |Du|^2}}$$

is another example.

We are primarily interested here in linear operators and linear functionals, and particularly continuous linear operators. We denote[11] the space of **continuous linear operators** by

$$\beth^0(V \to W) \subset C^1(V \to W).$$

Recall that **continuity at a point** $v_0 \in V$ for $L : V \to W$ in this context means the following:

> For any $\epsilon > 0$, there is some $\delta > 0$ such that
>
> $$\|v - v_0\| < \delta \qquad \Longrightarrow \qquad \|L(v) - L(v_0)\| < \epsilon.$$

First of all, with the addition of linearity, pointwise continuity has a remarkable property.

Lemma 4. *If $L : V \to W$ is linear and continuous at one point $v_0 \in V$, then $L \in \beth^0(V \to W)$. That is, L is continuous at all points.*

Proof: Say L is continuous at $v_0 \in V$ and we take another point $w_0 \in V$. If $\epsilon > 0$, then we know there is some $\delta > 0$ for which

$$\|Lv - Lv_0\| < \epsilon \qquad \text{whenever} \qquad \|v - v_0\| < \delta. \tag{37}$$

Note that one often omits the parentheses $L(w)$ or $L[w]$ for linear operators and simply writes Lw. Therefore, for any $w \in V$ with $\|w - w_0\| < \delta$, we can take $v = v_0 + w - w_0$ and conclude

$$\|Lw - Lw_0\| = \|Lv_0 + Lw - Lw_0 - Lv_0\| = \|L(v_0 + w - w_0) - Lv_0\| < \epsilon.$$

This means L is continuous at w_0. $\square$

Corollary 5. *If $L : V \to W$ is linear and L is continuous at $\mathbf{0} \in V$, then*

$$L \in \beth^0(V \to W).$$

[11] This notation using the Hebrew letter "beth" is (highly) nonstandard. I'm quite confident you will not find it in any book. Some further discussion is given in the endnotes to this section.

The following result is of central importance:

Theorem 12. *(bounded linear functionals) If $L \in \beth^0(V \to W)$, then there exists some $\Lambda > 0$ such that*

$$\|Lv\| \leq \Lambda \|v\| \qquad \textit{for all } v \in V. \tag{38}$$

Proof: L is continuous at $v_0 = \mathbf{0} \in V$. Therefore, there exists some $\delta > 0$ such that

$$\|v\| < \delta \qquad \Longrightarrow \qquad \|Lv\| < 1.$$

Now, take any $w \in V \backslash \{\mathbf{0}\}$. Notice that

$$\left\| \frac{\delta}{2} \frac{w}{\|w\|} \right\| = \frac{\delta}{2} < \delta.$$

Therefore,

$$\left\| L\left(\frac{\delta}{2} \frac{w}{\|w\|} \right) \right\| < 1.$$

That is,

$$\frac{\delta}{2\|w\|} \|Lw\| < 1 \qquad \text{or} \qquad \|Lw\| < \frac{2}{\delta} \|w\|. \tag{39}$$

This essentially proves the theorem. In fact, setting $\Lambda = \delta/2$, if $v \neq \mathbf{0}$, then we may take $v = w$ and apply (39). If $v = \mathbf{0}$, then we clearly have

$$\|Lv\| \leq \Lambda \|v\|$$

as the theorem asserts. $\square$

Remarks: Notice that the theorem says that every continuous linear operator between normed linear spaces is Lipschitz. That is,

$$\|Lv - Lw\| \leq \Lambda \|v - w\| \qquad \text{for all } v, w \in V.$$

The converse, that every Lipschitz operator between normed linear spaces is continuous, is of course true as well, so

$$\beth^0(V \to W) = \mathrm{Lip}(V \to W).$$

In this case, however, we usually do not refer to the continuity condition as Lipschitz continuity. Notice that (38) is even simpler than the Lipschitz condition (though pretty obviously equivalent in this context). An operator in this context, i.e., a linear operator satisfying (38), is said to be **bounded**. Thus,

$\beth^0(V \to W)$ is called the space of **bounded linear operators**.

Definition 9. *A linear operator $L : V \to W$ of normed spaces V and W is said to be* **bounded** *(or Lipschitz bounded) if there is some Λ for which*

$$\|Lv\| \leq \Lambda \|v\| \qquad \text{for all } v \in V.$$

Exercise 23. *Show that if $L : V \to W$ is bounded, then $L : V \to W$ is continuous.*

In summary, a linear operator $L : V \to W$ of normed spaces V and W is continuous if and only if it is bounded.

We are about to use these remarkable properties of linear continuous functions (between normed spaces) to turn $\beth^0(V \to W)$ into a normed vector space itself. This opens an entirely new arena for investigation which, at some level, is the real starting point for functional analysis. Before me make this construction, let's pause and look back at the ingredients which have brought us to this point and note that what we are doing constitutes a fundamentally new level of structure. We started with **functions**, say real valued functions with domains U which were open sets in $\mathbb{R}^n$. We classified these functions according to their properties relating to differentation and integration (which we studied) into various **normed function spaces**. These were, for example, $C^k(U)$, $C^{k,\alpha}(U)$, $L^p(U)$ and $W^{k,p}(U)$. Then we considered **operators** from between normed function and **functionals** from normed function spaces to their fields. Finally, we have identified the **continuous linear operators**

$$\beth^0(V \to W)$$

where we are especially interested in cases where V and W are among our normed function spaces. For a bounded linear operator $L \in \beth^0(V \to W)$, we set

$$\|L\| = \sup_{v \neq 0} \frac{\|Lv\|}{\|v\|}.$$

This is called the **operator norm** and it makes $\beth^0(V \to W)$ a normed vector space. The norm in the numerator of the quotient

$$\frac{\|Lv\|}{\|v\|} = \frac{\|Lv\|_W}{\|v\|_V}$$

is the norm from W, and the norm in the denominator is the norm from V. That the supremum of such numbers is finite follows immediately from the fact that L is bounded:

$$\frac{\|Lv\|}{\|v\|} \leq \frac{\Lambda \|v\|}{\|v\|} = \Lambda.$$

The properties of a norm may then be verified. The operator norm is non-negative homogeneous, positive definite, and satisfies the triangle inequality for norms.

11.2 Dual Spaces and Adjoints

11.3 Riesz Representation in $\mathbb{R}^n$

We will prove the Riesz representation theorem for bounded linear functionals on a Hilbert space below. It may be helpful to consider the special case of this result in the familiar setting of finite dimensional linear mappings. Say we have a linear function

$$\ell : \mathbb{R}^n \to \mathbb{R},$$

that is, a **linear functional**. Such a function is always continuous and has a number of other properties with which are (or should be) quite familiar from linear algebra. One of these is that there exists a particular vector $\mathbf{u} \in \mathbb{R}^n$ such that

$$\ell(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v} \qquad \text{for all } \mathbf{v} \in \mathbb{R}^n.$$

This observation, and the fact that the vector $\mathbf{u}$ is uniquely determined by the functional ℓ, constitute the essential assertion of the Riesz representation theorem. The linear functional ℓ is said to be "represented" by the vector $\mathbf{u}$. The right side of this relation uses the usual Euclidean dot product, and the Riesz representation theorem essentially replaces this with an arbitrary inner product (and applies to bounded linear functionals on a (potentially) infinite dimensional Hilbert space).

Let's attempt to think a little bit about the vector $\mathbf{u}$ and why this kind of representation occurs. The usual way to do this is to note (or observe) that

$$\mathbf{u} = (\ell(\mathbf{e}_1), \ell(\mathbf{e}_2), \ldots, \ell(\mathbf{e}_n))$$

where $\mathbf{e}_j$ is the **j-th standard unit basis vector** (with 1 in the j-th entry and zeros in all other entries). This choice clearly gives

$$\ell(\mathbf{e}_j) = \mathbf{e}_j \cdot \mathbf{u} \qquad \text{for } j = 1, 2, \ldots, n.$$

The identity

$$\ell(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v} \qquad \text{for every } \mathbf{v} \in \mathbb{R}^n$$

follows by linearity since we can write $\mathbf{v} = \sum (\mathbf{v} \cdot \mathbf{e}_j)\, \mathbf{e}_j$.

In an infinite dimensional inner product space, like $L^2(0, L)$, if you consider an orthonormal basis like $\{\sin j\pi x/L\}_{j=1}^{\infty}$, then you will have to consider series representation, which can be done, but then you will have questions of convergence and other inconveniences to deal with. In some sense, it is the point of **functional analysis** to

avoid such messy details and give a fundamentally different argument avoiding reference to a basis, or limits, or estimates—but sticking to ideas of linear algebra. Let's see if we can suggest how this might be done in the finite dimensional case where we understand everything.

First of all, if we have a representation $\ell(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$, then while the connection between ℓ and $\mathbf{u}$ may not be at all obvious, there is one connection which is easy to make: The subspace

$$\ker(\ell) = \{\mathbf{x} : \ell(\mathbf{x}) = 0\}$$

which is the **kernel** or **null space** of ℓ will consist of the vectors which are perpendicular to $\mathbf{u}$. That is, on the one hand $\ker(\ell)$ is a vector subspace of $\mathbb{R}^n$ and, on the other hand, the **orthogonal complement** of any vector $\mathbf{u}$

$$\mathbf{u}^\perp = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{u} = 0\}$$

is also a subspace. If our representation construction is going to work, **these subspaces must match**. Note that this gives us a place to start without ever mentioning a basis.

If $\ker(\ell) = \mathbb{R}^n$, then we can get a representation using $\mathbf{u} = \mathbf{0}$ (the zero vector). In fact, in this case, $\mathbf{u} = \mathbf{0}$ is the unique choice since

$$\mathbf{u} \cdot \mathbf{v} = \tilde{\mathbf{u}} \cdot \mathbf{v} \quad \text{for all } \mathbf{v} \qquad \Longrightarrow \qquad (\mathbf{u} - \tilde{\mathbf{u}}) \cdot \mathbf{v} = 0.$$

And taking $\mathbf{v} = \mathbf{u} - \tilde{\mathbf{u}}$, we get $|\mathbf{u} - \tilde{\mathbf{u}}| = 0$.

Otherwise, there is some **nonzero** vector $\mathbf{u}_0 \in \ker(\ell)^\perp$. The basic observation of the Riesz theorem is then that **some scaling $\mathbf{u} = \alpha\mathbf{u}_0$ of $\mathbf{u}_0$ will work**. If that is correct, then we need

$$\ell(\alpha\mathbf{u}_0) = \alpha\mathbf{u}_0 \cdot (\alpha\mathbf{u}_0) = \alpha^2|\mathbf{u}_0|^2,$$

so $\alpha = \ell(\mathbf{u}_0)/|\mathbf{u}_0|^2$. Then the question is: Will

$$\mathbf{u} = \frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|^2}\,\mathbf{u}_0 = \frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|}\,\frac{\mathbf{u}_0}{|\mathbf{u}_0|} \tag{40}$$

work?

In fact, we will show that this does work, even in the general infinite dimensional case. Before we do that, however, it may be instructive to consider a very specific example. Let $\ell : \mathbb{R}^3 \to \mathbb{R}^1$ by determined by

$$\begin{cases} \mathbf{e}_1 \mapsto 0 \\ \mathbf{e}_2 \mapsto 1 \\ \mathbf{e}_3 \mapsto 2. \end{cases} \tag{41}$$

The linearity argument above tells us $\mathbf{u} = (0, 1, 2)$ should be the unique vector such that $\ell(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$, and clearly this works.

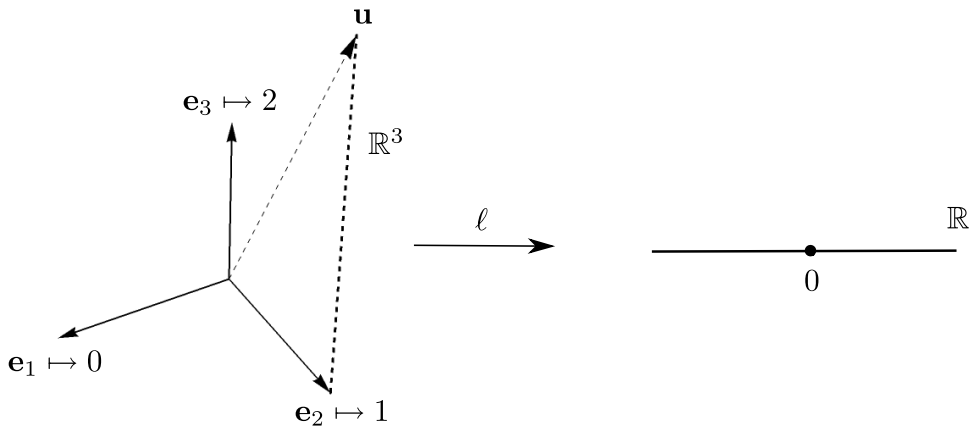

Figure 6: An example of a linear functional $\ell : \mathbb{R}^3 \to \mathbb{R}$.

We are supposed, however, to see/find this vector $\mathbf{u} = (0, 1, 2)$ without using the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. To do so, we look at

$$\ker(\ell) = \{\mathbf{z} \in \mathbb{R}^3 : \ell(\mathbf{z}) = 0\}.$$

We might be inclined, looking at the definition of ℓ in (41), to think

$$\ker(\ell) = \operatorname{span}\{\mathbf{e}_1\}. \tag{42}$$

Let's go with that assumption for a moment. Then our argument says to choose an arbitrary vector $\mathbf{u}_0 \in \ker(\ell)^\perp$. We could take $\mathbf{u}_0 = \mathbf{e}_2$ for example. Then a scaling $\mathbf{u} = \alpha\mathbf{u}_0$ is supposed to work. At this point, clearly something has gone wrong because we will never get $\mathbf{u} = (0, 1, 2)$ as a scaling of $\mathbf{u}_0 = (0, 1, 0)$. What has gone wrong?

What has gone wrong is that our identification of the null space in (42) is incorrect. The null space $\ker(\ell)$ is **larger** than $\operatorname{span}\{\mathbf{e}_1\}$. In fact, if we had thought about it a little bit (and maybe you did) the dimension theorem says

$$\dim \operatorname{Dom}(\ell) = \dim \operatorname{Im}(\ell) + \dim \ker(\ell). \tag{43}$$

In this case, $\dim \ker(\ell) = \dim \operatorname{Dom}(\ell) - \dim \operatorname{Im}(\ell) = 3 - 1 = 2$. And in general, for $\ell : \mathbb{R}^n \to \mathbb{R}^1$, we must have $\dim \ker(\ell) = n - 1$. Thus, in the finite dimensional case,

we must have a rather **large** kernel. In particular, the orthogonal complement of $\ker(\ell)$ must always satisfy

$$\dim \ker(\ell)^{\perp} = 1.$$

Thus, it is no surprise to find the representing vector there. This is the underlying idea also in the case of an infinite dimensional Hilbert space:

> Riesz representation follows because $\ker(L)$ is **large** to the extent that $\ker(L)^{\perp}$ is **one dimensional**.

We will return to this point. For the moment, let us give a more careful account of our example.

We should be able to find another vector in $\ker(\ell)$ for the linear map $\ell : \mathbb{R}^3 \to \mathbb{R}$ determined by (41). A moment's thought tells us $\ell : (0, -2, 1) \mapsto 0 \cdot 0 - 2 \cdot 1 + 1 \cdot 2 = 0$. Thus, $\ker(\ell)$ is a (two dimensional) plane and $\mathbf{u} = (0, 1, 2)$ is clearly a normal to that

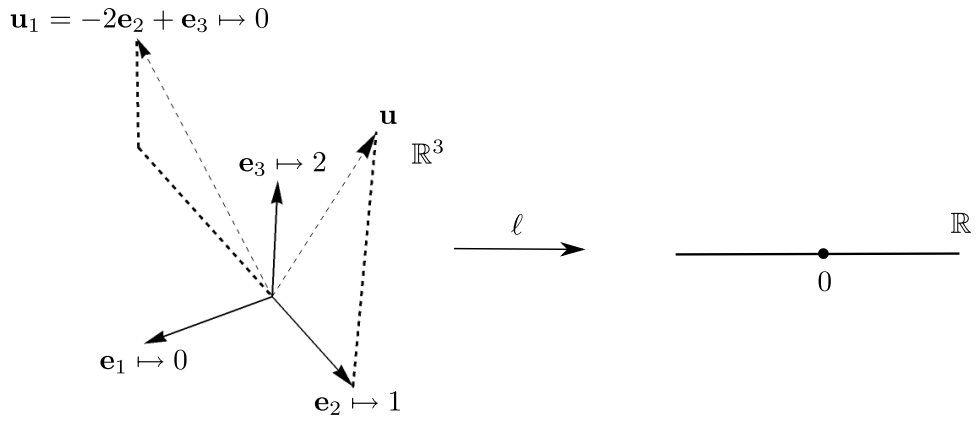

Figure 7: An example of a linear functional $\ell : \mathbb{R}^3 \to \mathbb{R}$; correctly identified null space.

plane.

In the infinite dimensional case we do not have recourse to the dimension theorem/relation (43) telling us $\ker(\ell)$ is large, say of dimension $n - 1$, so that $\ker(\ell)^{\perp}$ is small, having dimension 1. These statements, however, do translate into infinite dimensions (for a continuous functional $L \in \mathcal{H}^*$ on a Hilbert space $\mathcal{H}$) in the following form

$\ker(L)$ is **large** to the extent that

$$\dim \ker(L)^\perp = 1 \qquad \text{and} \qquad \mathcal{H} = \ker(L) \oplus \ker(L)^\perp.$$

The crucial smallness of $\ker(L)^\perp$ holds in particular, so that representation is to be expected. I mention these things now, in part, because these facts do not come out in the proof. In the proof, we do something that uses much weaker hypotheses but is, in some sense, much trickier. Nevertheless, it is worth noting that this (true) description is what underlies and drives the result.

As just mentioned, we more or less need to do something rather tricky to show

$$\mathbf{u} \cdot \mathbf{v} = \frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|^2}\mathbf{u}_0 \cdot \mathbf{v} = \ell(\mathbf{v}) \qquad \text{for all } \mathbf{v} \in \mathbb{R}^n$$

as required by the choice (40). Frigyes Riesz seems to have been particularly good at this kind of trickery. Given our previous discussion based on misidentification of $\ker(\ell)$, it is clear that we must use rather strongly that $\mathbf{u}_0 \in \ker(\ell)^\perp$. The intuition is that we must use this fact in some way that takes account of the fact that $\mathbf{u}_0$ is orthogonal to **every** vector in $\ker(\ell)$, and not just some of them. In the finite dimensional case, every vector $\mathbf{v}$ decomposes uniqely as a sum

$$\mathbf{v} = \mathbf{z} + \mathbf{w} \qquad \text{for some } \mathbf{z} \in \ker(\ell) \text{ and some } \mathbf{w} \in \ker(\ell)^\perp.$$

In fact, taking $\ker(\ell)^\perp = \text{span}\{\mathbf{u}_0\}$, we can write

$$\mathbf{v} = \left(\mathbf{v} - \frac{\mathbf{v} \cdot \mathbf{u}_0}{|\mathbf{u}_0|^2}\mathbf{u}_0\right) + \frac{\mathbf{v} \cdot \mathbf{u}_0}{|\mathbf{u}_0|^2}\mathbf{u}_0 \tag{44}$$

where $\mathbf{w} = (\mathbf{v} \cdot \mathbf{u}_0)\mathbf{u}_0/|\mathbf{u}_0|^2$ is the projection of $\mathbf{v}$ onto $\ker(\ell)^\perp$. You may recognize this construction as a part of the Gram-Schmidt orthonormalization procedure, and it is clear that the residual vector

$$\mathbf{z} = \mathbf{v} - (\mathbf{v} \cdot \mathbf{u}_0)\mathbf{u}_0/|\mathbf{u}_0|^2 \tag{45}$$

is in $[\ker(\ell)^\perp]^\perp = \ker(\ell)$. What is immediately clear from the Gram-Schmidt construction is that $\mathbf{z} \in \text{span}\{\mathbf{u}_0\}^\perp = [\ker(\ell)^\perp]^\perp$. The fact that $[\ker(\ell)^\perp]^\perp = \ker(\ell)$ or more generally that the **double orthogonal complement** $V^{\perp\perp} = (V^\perp)^\perp$ satisfies

$$V^{\perp\perp} = V \qquad \text{for any subspace } V$$

requires proof.

Lemma 6. *If V is a subspace of $\mathbb{R}^n$, then $V^{\perp\perp} = V$.*

Proof: One inclusion is easy. Specifically, if $\mathbf{v} \in V$, then clearly $\mathbf{v} \cdot \mathbf{w} = 0$ for all $\mathbf{w} \in V^{\perp}$. This follows from the definition of

$$V^{\perp} = \{\mathbf{w} \in \mathbb{R}^n : \mathbf{w} \cdot \mathbf{v} = 0 \text{ for all } \mathbf{v} \in V\}.$$

But this is also the definition of what it means to have $\mathbf{v} \in (V^{\perp})^{\perp} = V^{\perp\perp}$.

The reverse inclusion is trickier: If $\mathbf{v} \in V^{\perp\perp}$, then there are unique vetors $\mathbf{z} \in V$ and $\mathbf{w} \in V^{\perp}$ with

$$\mathbf{v} = \mathbf{z} + \mathbf{w}. \tag{46}$$

Recalling from the first inclusion that $V \subset V^{\perp\perp}$, we know $\mathbf{z} \in V^{\perp\perp}$ and, of course, $V^{\perp\perp}$ is a subspace. Therefore,

$$\mathbf{w} = \mathbf{v} - \mathbf{z} \in V^{\perp\perp} \cap V^{\perp}.$$

This implies $|\mathbf{w}|^2 = \mathbf{w} \cdot \mathbf{w} = 0$. Hence $\mathbf{w} = 0$ and $\mathbf{v} = \mathbf{z} \in V$. $\square$

It will be noted that we have used the following result to obtain (46).

Lemma 7. *If V is a subspace of $\mathbb{R}^n$, then $\mathbb{R}^n = V \oplus V^{\perp}$. That is, each $\mathbf{v} \in \mathbb{R}^n$ is expressed uniquely as $\mathbf{v} = \mathbf{z} + \mathbf{w}$ with $\mathbf{z} \in V$ and $\mathbf{w} \in V^{\perp}$. The vector $\mathbf{z}$ is called the* **projection** *of $\mathbf{v}$ onto V.*

Proof: If we want to get out of this easily, then we must allow recourse to the fact that $\mathbb{R}^n$ and V are finite dimensional, i.e., these spaces admit bases with finitely many elements. Say $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is a basis for V which we can assume (by Gram-Schmidt orthonormalization) is an orthonormal basis. Then

$$\mathbf{z} = \text{proj}_V(\mathbf{v}) = \textstyle\sum_{j=1}^{k} (\mathbf{v} \cdot \mathbf{u}_j)\, \mathbf{u}_j.$$

As a direct extension of our assertion concerning (44) we see that $\mathbf{v} - \mathbf{z} \in V^{\perp}$. Thus, we have $\mathbf{v} = \mathbf{z} + \mathbf{w}$ with $\mathbf{z} \in V$ and $\mathbf{w} \in V^{\perp}$.

For uniqueness, note that if $\mathbf{v} = \tilde{\mathbf{z}} + \tilde{\mathbf{w}} = \mathbf{z} + \mathbf{w}$ with $\tilde{\mathbf{z}} \in V$ and $\tilde{\mathbf{w}} \in V^{\perp}$, then $\mathbf{z} - \tilde{\mathbf{z}} = \tilde{\mathbf{w}} - \mathbf{w} \in V \cap V^{\perp}$. That is,

$$|\tilde{\mathbf{w}} - \mathbf{w}|^2 = (\tilde{\mathbf{w}} - \mathbf{w}) \cdot (\tilde{\mathbf{w}} - \mathbf{w}) = 0 \qquad \text{and} \qquad |\mathbf{z} - \tilde{\mathbf{z}}|^2 = 0. \qquad \square$$

Returning to our discussion of (44) and (45), if we know the vector $\mathbf{z}$ from (45) satisfies $\mathbf{z} \in \ker(\ell)$, then it is immediate that

$$\ell(\mathbf{v}) = \ell\left(\frac{\mathbf{v} \cdot \mathbf{u}_0}{|\mathbf{u}_0|^2}\, \mathbf{u}_0\right) = \frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|^2}\, \mathbf{v} \cdot \mathbf{u}_0 = \mathbf{v} \cdot \frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|^2}\, \mathbf{u}_0,$$

and we have the Riesz representation proposed in connection with (40).

The argument just given using the finite orthonormal basis for V does not work when V is infinite dimensional. Nevertheless, the assertions of Lemmas 6 and 7 do both hold for any **closed subspace V of a Hilbert space**. In fact, once Lemma 7 is established for a closed subspace V of a Hilbert space, then the argument given for Lemma 6 is valid in the same context.

Summary: Riesz' Trick

Returning to our specific application to the mapping $\ell : \mathbb{R}^3 \to \mathbb{R}$ determined by (41), the argument above should give us that the vector

$$\mathbf{z} = \mathbf{v} - \frac{\mathbf{v} \cdot \mathbf{u}_0}{|\mathbf{u}_0|} \frac{\mathbf{u}_0}{|\mathbf{u}_0|} \tag{47}$$

from the Gram-Schmidt procedure given in (45) satisfies $\mathbf{z} \in \ker(\ell)$. It will be observed, however, that this is hardly the case. In fact, I see no clear and obvious way to see

$$\ell(\mathbf{z}) = \ell(\mathbf{v}) - \frac{\mathbf{v} \cdot \mathbf{u}_0}{|\mathbf{u}_0|} \frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|} = 0.$$

Of course, you can write $\mathbf{z}$ as $\mathbf{z} = \mathbf{z}_0 + \mathbf{w}_0$ with $\mathbf{z}_0 \in \ker(\ell)$ and $\mathbf{w}_0 \in \ker(\ell)^\perp$ and then follow through the argument of Lemma 6. Even if you do that to conclude $\mathbf{z} = \mathbf{z}_0 \in \ker(\ell)$, the relation between the decomposition in (47) involving $\mathbf{v}$ and $\mathbf{u}_0$ and the decomposition $\mathbf{z} = \mathbf{z}_0 + \mathbf{w}_0$ remains (as far as I can see) still rather obscure. In particular, the relationship depends on the use of some basis for $V = \ker(\ell)$ which you don't even have (at least easily) in the infinite dimensional case.

Riesz had the very clever (and elegant) idea of decomposing $\mathbf{v}$ as in (44) and hence $\mathbf{z}$ as in (45) using a *different multiple of* $\mathbf{u}_0$. That is, consider

$$\mathbf{v} = (\mathbf{v} - \beta\mathbf{u}_0) + \beta\mathbf{u}_0 \tag{48}$$

where β is some constant to be determined. This gives the residual vector

$$\mathbf{z} = \mathbf{v} - \beta\mathbf{u}_0 \qquad \text{with} \qquad \ell(\mathbf{z}) = \ell(\mathbf{v}) - \beta\ell(\mathbf{u}_0). \tag{49}$$

Now, all we need to know is that $\mathbf{u}_0 \notin \ker(\ell)$ and we can take $\beta = \ell(\mathbf{v})/\ell(\mathbf{u}_0)$. Then we get $\ell(\mathbf{z}) = 0$, i.e., $\mathbf{z} \in \ker(\ell)$, automatically. In this way we have $\ell(\mathbf{v}) = \beta\ell(\mathbf{u}_0)$.

But, do we get Riesz representation from this and the choice suggested in (40)? In fact,

$$\mathbf{v} \cdot \frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|^2}\,\mathbf{u}_0 = (\mathbf{v} - \beta\mathbf{u}_0) \cdot \frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|^2}\,\mathbf{u}_0 + \beta\mathbf{u}_0 \cdot \frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|^2}\,\mathbf{u}_0 = \beta\,\ell(\mathbf{u}_0) = \ell(\mathbf{v})$$

since $\mathbf{v} - \beta\mathbf{u}_0 \in \ker(\ell)$ and $\mathbf{u}_0 \in \ker(\ell)^\perp$.

11.4 Riesz Representation

Theorem 13. *(Riesz representation theorem) If $\langle\ \cdot\ ,\ \cdot\ \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ is* **any** *inner product on a Hilbert space $\mathcal{H}$ and $L : \mathcal{H} \to \mathbb{R}$ is a bounded linear functional, i.e., $L \in \beth^0(\mathcal{H}) = \mathcal{H}^*$, then there exists a unique $u \in \mathcal{H}$ such that*

$$Lv = \langle u, v\rangle \qquad \textit{for all } v \in \mathcal{H}.$$

Proof: Consider

$$\mathcal{N} = \{z \in \mathcal{H} : Lz = 0\}$$

the **null space** of L. This is a closed vector subspace of $\mathcal{H}$.

If $\mathcal{N} = \mathcal{H}$, then we have representation using the zero vector:

$$Lv = \langle \mathbf{0}, v\rangle,$$

and the representation is unique since

$$\langle u, v\rangle = 0 \quad \text{for all } v \in V \qquad \Longrightarrow \qquad \|u\|^2 = \langle u, u\rangle = 0.$$

So, obviously the more interesting case is when $\mathcal{N} \subsetneq \mathcal{H}$. In this case, we can take a vector $u_0 \in \mathcal{N}^\perp \backslash \{0\}$. Now, normally one would think there are many such vectors u_0, so it's not so obvious that you can count on anything special from this one. Apparently, however, the fact that the image of the functional L is $\mathbb{R}$ (which is conspicuously one dimensional) somehow narrows the possibilities.

Thus, the crucial ansatz is to look for a scaling of u_0 as the choice of u. That is, we look for some $\alpha \in \mathbb{R}$ for which

$$Lv = \langle \alpha u_0, v\rangle \qquad \text{for all } v \in \mathcal{H}.$$

Once this ansatz is written down, then you know the identity of the scalar α because you must have

$$Lu_0 = \alpha\|u_0\|^2.$$

That is,

$$\alpha = \frac{Lu_0}{\|u_0\|^2}.$$

Once that determination is made, one simply needs to see if (or show that)

$$u = \frac{Lu_0}{\|u_0\|^2}\, u_0 \qquad \text{works.}$$

The idea for accomplishing this is somewhat reminiscent of the decomposition of a vector in a direct sum. That is, we take an arbitrary vector v and decompose it in terms of a component along $\mathcal{N}^\perp$ or more precisely along span(u_0):

$$v = (v - \beta u_0) + \beta u_0.$$

In order to have $v - \beta u_0 \in \mathcal{N}$, we take $\beta = Lv/Lu_0$. This is a well-defined vector since $u_0 \notin \mathcal{N}$ and so $Lu_0 \neq 0$. Note then, that

$$L(v - \beta u_0) = 0, \qquad \text{so} \qquad v - \beta u_0 \in \mathcal{N}, \qquad \text{and} \qquad \langle u_0, v - \beta u_0 \rangle = 0.$$

Then we can compute

$$\begin{aligned} \langle \alpha u_0, v \rangle &= \langle \alpha u_0, v - \beta u_0 \rangle + \langle \alpha u_0, \beta u_0 \rangle \\ &= \langle \alpha u_0, \beta u_0 \rangle \\ &= \alpha\beta \|u_0\|^2 \\ &= Lv \end{aligned}$$

since

$$\alpha\beta = \frac{Lu_0}{\|u_0\|^2}\, \frac{Lv}{Lu_0} = \frac{Lv}{\|u_0\|^2}.$$

Thus, we have existence of a vector $u = \alpha u_0$ for which

$$\langle u, v \rangle = Lv \qquad \text{for all } v \in V.$$

Uniqneness is, again rather easy: If $\langle u, v \rangle = \langle \tilde{u}, v \rangle$, then taking $v = u - \tilde{u}$, we get

$$\|u - \tilde{u}\|^2 = \langle u - \tilde{u}, u - \tilde{u} \rangle = 0. \qquad \square$$

One thing to note about our discussion of the Riesz representation theorem: Essentially no inequalities, estimates, questions of convergence, or limits were used. Basically no (hard) analysis was mentioned. There was a good deal of (perhaps tricky)

algebra and especially linear algebra. To be fair, the usual proofs that a **closest vector** in a subspace to a given vector outside that subspace exists involve showing some sequence is Cauchy, and that involves some elementary estimation, but that is about it. Incidentally, this is the point where the completeness of a Hilbert space comes in.

Exercise 24. *Let V be a closed subspace of a Hilbert space $\mathcal{H}$ and let $p \in \mathcal{H} \backslash V$. Consider a sequence of points $x_j \in V$ for $j = 1, 2, 3, \ldots$ with*

$$\lim_{j\to\infty} \|x_j = h\|_{\mathcal{H}} = \operatorname{dist}(p, V) = \inf_{x\in V} \|x - p\|_{\mathcal{H}}. \tag{50}$$

(i) *Show the sequence $\{x_j\}_{j=1}^{\infty}$ is Cauchy.*

(ii) *Use the completeness of $\mathcal{H}$ (and the fact that V is closed) to conclude the limit $\lim_{j\to\infty} x_j = x$ achieves the minimum value in the infemum of (50).*

(iii) *Show the difference $x - p \in V^{\perp}$.*

Exercise 25. *Show that a closed subset of a Hilbert space is metrically complete. In particular a closed subspace of a Hilbert space is a Hilbert space. Give an example of a subspace of a Hilbert space which is neither closed nor complete.*

Exercise 26. *Show the null space* $\ker(L)$ *of a bounded linear functional $L \in \mathcal{H}^*$ is closed.*

Summary: Final Remarks

Once we have proved the Riesz representation theorem in Hilbert space, we can look back and see

$$\ker(L) = \{\mathbf{z} \in \mathcal{H} : L\mathbf{z} = 0\} = \{\mathbf{z} \in \mathcal{H} : \langle \mathbf{u}, \mathbf{z}\rangle = 0\} = \operatorname{span}\{\mathbf{u}\}^{\perp}.$$

Thus, since the span $\operatorname{span}\{\mathbf{u}\}$ of the representing vector is one-dimensional, the space $\ker(L)$ is the (very large) orthogonal complement. Our intuition was correct, even though we didn't use it directly in the proof.

It will noted that all estimates involved here (all analysis) has been swept under the rug in Exercise 24. That analysis, moreover, is aimed at showing the existence of a vector closest to the closed linear subspace which is the null space of the function ℓ. In our application to existence and uniqueness of weak solutions of Poisson's equation, the actual operator from the PDE is actually further hidden away in the inner product,

and we only really see the abstract properties of the inner product in the proof of the Riesz theorem above. In this way, there is a kind of double sweeping of the analysis under the rug. The only means to get back to what is actually happening with the operator, in this case the Laplace operator, is in regard to the Poincaré inequality which essentially renders the weak adjoint

$$B[u,v] = \int Du \cdot Dv \qquad \text{with} \qquad [u]_{W^{1,1}} = B[u,u]$$

an inner product.

Exercise 27 (challenge)**.** *Where is the positive definiteness of the inner product used/required in the proof of* **existence** *in the Riesz representation theorem?*

12 Existence and Uniqueness of Weak Solutions for Laplace's and Poisson's PDE

Here we want to prove the following result:

Theorem 14. *Given any bounded open set $U \subset \mathbb{R}^n$ and $f \in L^2(U)$, there is a unique function $u \in H_0^1(U)$ for which*

$$-\int_U Du \cdot D\phi = \langle f, \phi \rangle_{L^2} \qquad \text{for all } \phi \in C_c^\infty(U). \tag{51}$$

The condition (51) formulates the weak version of the boundary value problem

$$\begin{cases} -\Delta u = f & \text{on } U \\ u\big|_{\partial U} \equiv 0. \end{cases} \tag{52}$$

If fact, if $u \in C^2(U) \cap C^0(\overline{U})$ is a classical solution of (52), then it can be checked that u satisfies (51). The homogeneous boundary condition is captured weakly by assuming $u \in H_0^1(U)$ where $H_0^1(U) = W_0^{1,2}(U)$ is the closed subspace of $H^1(U) = W^{1,2}(U)$ obtained as the **closure** of the subspace

$$C_c^\infty(U) \subset W^{1,2}(U)$$

with respect to the $W^{1,2}$ norm. We recall that $H^1(U) = W^{1,2}(U)$ is a Hilbert space with inner product given by

$$\langle u, v \rangle_{W^{1,2}} = \sum_{j=1}^n \langle D_j u, D_j v \rangle_{L^2} + \langle u, v \rangle_{L^2}. \tag{53}$$

Now, conversely, if $u \in C^2(U) \cap C^0(\overline{U})$ satisfies the condition (51) of the weak formulation, then it can be shown that u is a classical solution of (52).

Our proof relies on the Riesz representation theorem, but we will not apply it using the natural inner product (53) but rather an equivalent inner product adapted to the formulation (51), namely

$$B : H^1(U) \times H^1(U) \to \mathbb{R} \qquad \text{by} \qquad B[u, v] = \int Du \cdot Dv.$$

It has already been discussed why this doesn't look like an inner product (and is not on $H^1(U) = W^{1,2}(U)$) but is an inner product on $H_0^1(U)$ where it is positive definite by the Poincaré inequality. In terms of the bilinear form B, the weak formulation reads

$$B[u, \phi] = -\langle f, \phi \rangle_{L^2} \qquad \text{for all } \phi \in C_c^\infty(U).$$

Before we begin the proof, we may do well to review the situation with solvability in some more general terms. Perhaps the first thing to remember is that the classical problem

$$\begin{cases} \Delta u = 0 \quad \text{on } U \\ u\big|_{\partial U} \equiv g \end{cases} \tag{54}$$

for Laplace's equation considered on a bounded open set U with smooth boundary (say $\partial U \in C^\infty$) and with smooth boundary values $g \in C^\infty$ extending to U so that we have an extension $g \in C^\infty(\overline{U})$ is essentially equivalent to the homogeneous boundary value problem (52) for Poisson's equation. Schematically:

$$\begin{cases} \Delta u = 0 \quad \text{on } U \\ u\big|_{\partial U} \equiv g \end{cases} \qquad \longleftrightarrow \qquad \begin{cases} -\Delta u = f \quad \text{on } U \\ u\big|_{\partial U} \equiv 0. \end{cases}$$

This leads to consideration of the solvability for Poisson's equation with homogeneous boundary values. Thinking in terms of operators classically, we have

$$-\Delta : C^2(\overline{U}) \to C^0(\overline{U}),$$

and we are particularly interested in the image of this operator in relation to the subspace

$$\{f \in C^0(\overline{U}) : f\big|_{\partial U} \equiv 0\}.$$

Thus, this is starting to look rather like a linear algebra problem. From this point of view, the existence problem can be posed by saying: Is it true that

$$\{f \in C^0(\overline{U}) : f\big|_{\partial U} \equiv 0\} \subset -\Delta[C^2(\overline{U})] = \{-\Delta u : u \in C^2(\overline{U})\}?$$

The answer turns out[12] to be "yes," but the proof is somewhat difficult. Also, some proofs are better than others. Some proofs generalize to other PDE. Perhaps one of the simplest "proofs" involves constructing the Green's function for Poisson's equation in terms of the **fundamental solution** for Laplace's equation. This approach, on the one hand, gives a formula for the solution. It is usually the treatment given in engineering courses. On the other hand, it still requires the assumption of solvability for the boundary value problem for Laplace's equation, so without some other existence theorem it is somewhat incomplete.

The existence (and uniqueness) theorem we will actually prove completely has the advantage that it is relatively easy and it turns out that it can be generalized to other equations, like general second order linear uniformly elliptic PDE

$$\sum a_{ij} D_{ij} u + \sum b_j D_j u + cu = f,$$

without too much difficulty. (One needs a version of the Riesz representation theorem called the Lax-Milgram theorem.) The disadvantage is that the unique solution one gets is only a **weak solution**, and then one needs to prove that, under regularity assumptions on ∂U and f that the weak solution u is actually regular and is a classical solution. This turns out to be true as well, at least in the classical framework described above:

Theorem 15. *(regularity) If u is the weak solution of (52) obtained from Thereom 14 above, and $\partial U \in C^\infty$ and $f \in C^\infty(\overline{U})$, then $u \in C^\infty(\overline{U})$ and u is a classical solution.*

Of course, it should also not be overlooked that our result Theorem 14 also gives the uniqueness of a much broader class of classical solutions, say classical solutions in $u \in C^2(U) \cap C^0(\overline{U})$, though that particular result (as we have seen) follows from the weak maximum principle.

Proof of Thereom 14: Again, recall that we may start by considering

$$\begin{array}{rcl} \Delta & : & C^2(\overline{U}) \to C^0(\overline{U}) \\ & & \cup \\ & & C_0^2(\overline{U}) = \{u \in C^2(\overline{U}) : u\big|_{\partial U} \equiv 0\} \\ & & \cup \\ & & C_c^\infty(U). \end{array}$$

[12] And this is something you should take note of even if you can't cross all the t's and dot all the i's to give a complete proof.

Since $B : C_c^\infty(U) \times C_c^\infty(U) \to \mathbb{R}$ by

$$B[u,v] = \int_U Du \cdot Dv = \sum_{j=1}^n \int_U D_j u D_j v \tag{55}$$

is a continuous bilinear form on the subspace $C_c^\infty(U) \times C_c^\infty(U)$ of $H^1(U) \times H^1(U)$ in the $W^{1,2}$ norm, we have that B extends to a continuous bilinear form on the closure $H_0^1(U) \times H_0^1(U)$ by the same formula (55) with the derivatives interpreted as weak derivatives instead of classical derivatives.

This is the bilinear form appearing in the basic formulation of our problem:

$$\text{Given } f \in L^2(U), \text{ find } u \in H_0^1(U) \text{ with } B[u,\phi] = \int_U f\phi \text{ for all } \phi \in C_c^\infty(U).$$

The integral functional associated with the function $f \in L^2$ may be written and extended as follows:

$$\begin{array}{rcl} \mathcal{F} & : & C_c^\infty(U) \to \mathbb{R} \quad \text{by } \mathcal{F}[\phi] = \int_U f\phi \\ & & \downarrow \\ \mathcal{F} & : & H_0^1(U) \to \mathbb{R} \quad \text{by } \mathcal{F}[v] = \int_U fv = \langle f, v\rangle_{L^2}. \end{array}$$

We claim that $\mathcal{F} \in [H_0^1(U)]^* = \beth(H_0^1(U))$ is an element of the Hilbert dual of $H_0^1(U)$, i.e., is a bounded linear functional on $H_0^1(U)$. To see this, observe

$$|\mathcal{F}[v]| = \left| \int_U fv \right| \le \|f\|_{L^2} \|v\|_{L^2} \le \|f\|_{L^2} \|v\|_{W^{1,2}}.$$

Finally, we briefly recall that B is a general inner product on $H_0^1(U)$ by virtue of the fact that B is bilinear, symmetric, non-negative, and

$$\int_U |u|^2 \le C \int_U |Du|^2 \qquad \text{by the Poincaré inequality.} \tag{56}$$

Thus, if $B[u,u] = 0$, then $u = 0$. On this topic we note, finally, that the inequality of (56) holds for $u \in H_0^1(U)$ because for such a function, one can take a sequence of functions $\phi_j \in C_c^\infty(U)$ with $\|\phi_j - u\|_{W^{1,2}} \to 0$ as $j \to \infty$, and the smooth Poincaré inequality gives

$$\int_U |\phi_j|^2 \le C \int_U |D\phi_j|^2.$$

Therefore,

$$\begin{aligned}\left|\int_U |\phi_j|^2 - \int_U |u|^2\right| &\le \int_U ||\phi_j| + |u||||\phi_j| - |u|| \\ &\le \||\phi_j| + |u|\|_{L^2}\||\phi_j| - |u|\|_{L^2} \\ &\le (\|\phi_j\|_{L^2} + \|u\|_{L^2})\|\phi_j - u\|_{L^2} \\ &\le (\|\phi_j\|_{L^2} + \|u\|_{L^2})\|\phi_j - u\|_{W^{1,2}} \\ &\to 0\end{aligned}$$

and

$$\begin{aligned}\left|\int_U |D\phi_j|^2 - \int_U |Du|^2\right| &\le \int_U ||D\phi_j| + |Du||||D\phi_j| - |Du|| \\ &\le \||D\phi_j| + |Du|\|_{L^2}\||D\phi_j| - |Du|\|_{L^2} \\ &\le (\|D\phi_j\|_{L^2} + \|Du\|_{L^2})\|D\phi_j - Du\|_{L^2} \\ &\le M(\|D\phi_j\|_{L^2} + \|Du\|_{L^2})\|D\phi_j - Du\|_{W^{1,2}} \\ &\to 0.\end{aligned}$$

These limits establish (56):

$$\begin{array}{ccc} \int_U |\phi_j|^2 & \le & C\int_U |D\phi_j|^2 \\ \downarrow & & \downarrow \\ \int_U |u|^2 & \le & C\int_U |Du|^2. \end{array}$$

Once we know B is an inner product on $H_0^1(U)$ and $\mathcal{F} \in [H_0^1(U)]^*$, then we have unique representation

$$B[u, v] = \mathcal{F}[v] = \int fv$$

for some unique $u \in H_0^1(U)$. That's the theorem. $\square$

12.1 Endnotes

In spite of its importance, as far as I know, there is no standard notation for the set of bounded linear operators $L : V \to W$ where V and W are normed vector spaces. Two notations which are sometimes used are $B(V, W)$ and $L(V, W)$. I don't

particularly care for these because I use (with many others) B in $B_r(\mathbf{p})$ to denote a ball in Euclidean space. That is relatively standard. And some form of $L(V,W)$ is used more generally (sometimes $\mathcal{L}(V,W)$) to denote linear mappings from a vector space V to a vector space W (without reference to continuity). I have used the notation $\sqsupset^0(V \to W)$ for this important space.[13] This strikes me (at the moment) as a pretty good option. Perhaps a comment or two about why I like it may be amusing. Another nonstandard notation for $\sqsupset^0(V \to W)$ I have considered is

$$\sqsubset^0(V \to W).$$

The use of $\sqsubset$, interpreted as a "square letter C" has the advantage of maintaining a connection with the standard notation C^0 for continuity, and the straight lines of the symbol may be thought suggestive of linearity. I already use this symbol, however, for a different purpose, namely to denote the piecewise continuous functions $\sqsubset^0[a,b]$ and the piecewise C^k functions $\sqsubset^k[a,b]$ on an interval, which are important and used quite frequently in the calculus of variations but also seem to have associated with them no standard notation. In that context one can think of a real valued function on an interval whose graph consists of concatenated straight lines, which is perhaps the same mental image associated with a Lipschitz function on an interval. Thus the straight lines of "$\sqsubset$" seem to serve a good purpose there too. It occured to me that for the bounded linear operators, I could turn "$\sqsubset$" backwards and write $\sqsupset^0(V \to W)$ or simply $\sqsupset(V \to W)$ with the "b" of "backwards," at least, sympathetic with the "b" of "bounded." Of course, that's a bit of a strained connection.

Then, looking through all the available strange symbols, I ran across $\beth$. It looks rather like "$\sqsupset$" or like a backwards "C." It also contains something that looks vaguely like a backwards "L" for "linearity." And finally, $\beth$ is the second letter of the Hebrew alphabet, is pronounced like "b," and may be considered an equivalent or version of "B" for "bounded." Finally, I have retained the superscript zero to emphasize the connection with C^0 and that one has primarily continuity rather than differentiability in mind in this context, though I am aware of the possibility of a theory of differentiation for certain operators. I have never heard of "continuously

[13] I have also replaced (V,W) in the argument of mapping classes with the more suggestive $(V \to W)$ which is also nonstandard, but that is a different eccentricity. Actually, I agree with the criticism that sometimes this use of arrows can become irritating. For example, to denote a bounded linear operator between C^2 mappings of $\mathbb{R}^2$ to itself, one will have $\sqsupset^0(C^2(\mathbb{R}^2 \to \mathbb{R}^2) \to C^2(\mathbb{R}^2 \to \mathbb{R}^2))$. On the other hand, one doesn't encounter/write such a thing often, and I'm not sure this is much worse than $\sqsupset^0(C^2(\mathbb{R}^2,\mathbb{R}^2), C^2(\mathbb{R}^2,\mathbb{R}^2))$. Also, I think for many students, who are still trying to assimulate the idea of a function as a mapping, the more explicit arrows can be, as I have said, more suggestive.

differentiable operators," so I will leave that for later consideration—or more likely the consideration of others. If someone has a need for it, however, I certainly give them my blessing in using $\beth^k(V \to W)$.

13 Notes, Remarks, and Solutions

13.1 Exercise 3

Exercise. *Observe that the integral*

$$\int_I f \, d\nu_k,$$

with respect to the measure ν_k of **any** *pointwise well-defined function $f : I \to \mathbb{R}$ makes sense. Write down a formula for how to compute this integral, and show*

$$N_k : F \to \mathbb{R} \qquad \text{by} \qquad N_k[f] = \int_I f \, d\nu_k$$

defines a **linear functional**[14] *on the vector space F of real valued functions $f : I \to \mathbb{R}$.*

Exercise. *Restricting the functional N_k of the previous exercise to the the subspace $C_c^\infty(I)$, show that given any $\phi \in C_c^\infty(I)$, we have*

$$\lim_{k\to\infty} N_k[\phi] = \int_I \phi \, \frac{m}{\ell}$$

where the integral on the right may be taken as a Riemann integral:

$$\int_I \phi \, \frac{m}{k} = \int_a^b \phi(x) \, \frac{m}{\ell} \, dx \qquad \text{where } I \text{ has endpoints } a \text{ and } b,$$

or as an integral of ϕ times the constant density function $\rho \equiv m/\ell$ with respect to Lebesgue measure, which is the "usual" measure on an interval of the real line discussed in more detail below.

[14] This just means a linear function with codomain the real numbers

Solution:

$$\begin{aligned} N_k[\phi] &= \int_I \phi \, d\nu_k \\ &= \sum_{j=1}^{k} \phi(x_j) \frac{m}{k} \\ &= \sum_{j=1}^{k} \phi(x_j) \frac{m}{\ell} \, \frac{\ell}{k+1} \, \frac{k+1}{k} \\ &= \sum_{j=1}^{k} \phi(x_j) \frac{m}{\ell} \, \frac{k+1}{k} (x_{j+1} - x_j). \end{aligned}$$

This looks rather like the Riemann sum

$$R_k = \phi(x_0) \frac{m}{\ell} (x_1 - x_0) + \sum_{j=1}^{k} \phi(x_j) \frac{m}{\ell} (x_{j+1} - x_j)$$

based on left endpoints except for a couple minor differences. We know since $\phi \in C^0[a, b]$ that

$$\lim_{k \to \infty} R_k = \int_a^b \frac{m}{\ell} \, \rho(x) \, dx.$$

If we can show $|N_k[\phi] - R_k|$ tends to zero as $k \to \infty$, then we can say

$$\left| N_k[\phi] - \int_a^b \frac{m}{\ell} \, \rho(x) \, dx \right| \leq |N_k[\phi] - R_k| + \left| R_k - \int_a^b \frac{m}{\ell} \, \rho(x) \, dx \right|$$

by the triangle inequality, so

$$\lim_{k \to \infty} N_k[\phi] = \int_a^b \frac{m}{\ell} \, \rho(x) \, dx$$

as desired. Thus, we estimate

$$\begin{aligned}
|N_k[\phi] - R_k| &\leq |\phi(x_0)|\frac{m}{\ell}(x_1 - x_0) + \left|\sum_{j=1}^{k} \phi(x_j)\frac{m}{\ell}\left(\frac{k+1}{k} - 1\right)(x_{j+1} - x_j)\right| \\
&\leq \|\phi\|_{C^0[a,b]}\frac{m}{\ell}\left\{\frac{\ell}{k+1} + \sum_{j=1}^{k}\left(\frac{1}{k}\right)(x_{j+1} - x_j)\right\} \\
&\leq \|\phi\|_{C^0[a,b]}\frac{m}{\ell}\left\{\frac{\ell}{k+1} + \frac{\ell}{k+1}\right\} \\
&= \|\phi\|_{C^0[a,b]}\frac{2m}{k+1}
\end{aligned}$$

which clearly tends to zero as $k \to \infty$.

13.2 Exercise 5

For each $k = 1, 2, 3, \ldots$ partition the interval as indicated in Figure 3 with

$$a = \xi_0 < x_1 < \xi_1 < x_2 < \xi_2 < \cdots < x_k < \xi_k = b$$

where

$$\xi_{j+1} - \xi_j = \frac{\ell}{k} \quad \text{for } j = 0, 2, \ldots, k-1$$

and

$$x_{j+1} - x_j = \frac{\ell}{k} \quad \text{for } j = 1, 2, \ldots, k-1.$$

One may also set $x_1 = a + \ell/(2k)$, but this is not necessary. Now at each point x_j, $j = 1, 2, \ldots, k$ assign the point mass

$$g_j = \int_{\xi_{j-1}}^{\xi_j} \rho(x)\,dx = \int_{(\xi_{j-1},\xi_j)} \rho = \int_{(\xi_{j-1},\xi_j)} \rho\,d\mu.$$

Then, setting

$$\nu_k = \sum_{j=1}^{k} g_j \delta_{x_j}$$

we have

$$\nu_k(I) = \sum_{j=1}^{k} g_j = \int_I \rho\,d\mu.$$

Given a continuous function $\phi \in C^0[a,b]$, we have

$$\begin{aligned} M_k &= \int \phi \, d\nu_k \\ &= \sum_{j=1}^{k} g_j \phi(x_j) \\ &= \sum_{j=1}^{k} \phi(x_j) \int_{\xi_{j-1}}^{\xi_j} \rho(x)\, dx. \end{aligned}$$

On the other hand, we may consider the Riemann sum

$$\begin{aligned} R_k &= \sum_{j=1}^{k} \phi(x_j)\rho(x_j)(\xi_j - \xi_{j-1}) \\ &= \sum_{j=1}^{k} \phi(x_j) \int_{\xi_{j-1}}^{\xi_j} \rho(x_j)\, dx \end{aligned}$$

since

$$\int_{\xi_{j-1}}^{\xi_j} \rho(x_j)\, dx = \rho(x_j)(\xi_j - \xi_{j-1}).$$

Since $\rho \in C^0[a,b]$, we know

$$\lim_{k\to\infty} R_k = \int \phi\, \rho.$$

On the other hand,

$$\begin{aligned} |M_k - R_k| &\le \sum_{j=1}^{k} \left| \int_{\xi_{j-1}}^{\xi_j} [\rho(x) - \rho(x_j)]\phi(x_j)\, dx \right| \\ &\le \|\phi\|_{C^0(I)} \sum_{j=1}^{k} \int_{\xi_{j-1}}^{\xi_j} |\rho(x) - \rho(x_j)|\ dx. \end{aligned}$$

Also, since $\rho \in C^0[a,b]$ given any $\epsilon > 0$, there is some $\delta > 0$ such that

$$\left.\begin{array}{l} |x-\xi| < \delta \\ x \in I \end{array}\right\} \qquad \Longrightarrow \qquad \rho(x) - \rho(\xi)| < \frac{\epsilon}{\|\phi\|_{C^0(I)} + 1}.$$

It follows that for large enough k we have $\xi_j - \xi_{j-1} = \ell/k < \delta$ for all $j = 1, \dots k$ and $|M_k - R_k| < \epsilon$. This shows

$$\lim_{k\to\infty} |M_k - R_k| = 0 \qquad \text{and consequently} \qquad \lim_{k\to\infty} M_k = \int \phi\rho.$$