A Practical Introduction to Differential Forms

Alexia E. Schulz

and

William C. Schulz

March 17, 2016

Transgalactic Publishing Company
Flagstaff, Vienna, Cosmopolis

Every creator painfully experiences the chasm between his inner vision and its ultimate expression. Isaac Bashevis Singer

Dedicated to our parents, children, and cats

Contents

Chapter 1

DeRham Cohomology; an introduction

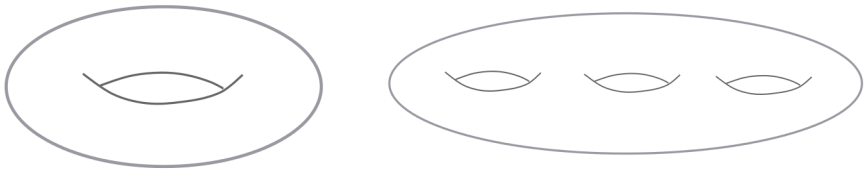
1.1 Introduction

We want to give a short introduction to de Rham cohomology and the Hodge theorem, which says that there is a unique harmonic form in each cohomology class. Everything up to this point has been mostly local in character and the reader may have the feeling that differential forms are not useful in understanding the global shape of a manifold. This is far from the truth. A great deal of global information is indeed accessible through differential forms and this is addressed through de Rham Cohomology.

This is a vast subject and we will be able to present little more than an introduction. Readers whose fancy is caught by our story will want to consult the beautiful book *Differential Forms in Algebraic Topology* by Raul Bott and Loring Tu, where systematic methods are developed. We do not have space to develop these systematic methods and will be satisfied with some ad hoc methods just sufficient to make the subject intelligible.

1.2 A Little History

When we look at a Donut (technically referred to as a *torus*, and we are considering only the skin of the donut, not the meat) we see first the hole in it. The fact that the donut has this hole is really the essential part of its global structure. This is easy to generalize as shown in the illustration.



One hole and three hole tori

The number of holes in the 2-manifold is called its *genus*. A two manifold is characterized by by its genus and whether it is or is not orientable. We are going to confine our work in this chapter to orientable manifolds. This means that any compact orientable differentiable 2-manifold (which of course has no boundary) is diffeomorphic to a surface of genus 0,1,2,.... The sphere of course has genus 0.

Much of this material was worked out for 2-manifolds by Riemann in the 1850's and 60's, but Riemann's methods were not easy to generalize to higher dimensions.

Around the beginning of the 20^{th} Century Henri Poincaré (a name we have previously encountered) devised two methods, Homotopy and Homology, to as-

sign numbers to differentiable manifolds of any dimension (and the same methods worked for topological manifolds) that allowed them to be classified into types. Since Poincaré built upon previous work of Betti, the numbers for Homology are called Betti numbers. Somewhat later, at a conference, Emmy Noether remarked to Pavel Alexandrov that, since the Betti numbers were the ranks of Abelian Groups it might be better to consider the groups themselves. This was a pregnant remark which eventually restructured the entire subject of Homology. The next big change came when mathematicians began to look at the maps of Homology Groups into certain Rings, $\mathbb{Z}$ or $\mathbb{R}$ being popular choices, and these were called *Cohomology Rings*, and this made the theory more powerful.

Originally Homology and Cohomology Groups were difficult to compute but by using algebraic methods and by building up spaces from simpler spaces computation became much easier. We mention the Mayer-Vietoris sequence as the easiest of these methods. In time, many different methods for constructing cohomology groups and rings were discovered. These methods all agree on differentiable manifolds, but may not agree for more general spaces. In this chapter we are going to consider the method pioneered by de Rahm, which uses differential forms, and is called *de Rham Cohomology*. Other methods are simplicial cohomology, singular cohomology, Čech cohomology, etale cohomology, etc.

1.3 De Rham Cohomology

1.3.1 Definitions and Examples

In this section we will deal with differential forms with C^∞ coefficients. This is for convenience and is not essential to the theory. Our basic reference for this material is [2]. Recall that the notation for k-forms on the manifold M is $\Lambda^k(T^*(M))$ which we are going to abbreviate as $\Lambda^k(M)$. Recall that the operator d then gives the following chain of maps, (called a *complex* because $d^2 = 0$).

$$\Lambda^0(M) \xrightarrow{d} \Lambda^1(M) \xrightarrow{d} \Lambda^2(M) \xrightarrow{d} \Lambda^3(M) \xrightarrow{d} \cdots$$

We abbreviate the entire sequence by $\Lambda^*(M)$. The elements of $\Lambda^*(M)$ are called cochains. We have the following basic definitions:

Def A cochain $\omega \in \Lambda^k(M)$ is a *cocycle* if and only if $d\omega = 0$.
The vector space of cocycles is denoted by $Z^k(M) \subseteq \Lambda^k(M)$.

Def A cochain $\omega \in \Lambda^k(M)$ is a *coboundary* if and only if there exists an $\eta \in \Lambda^{k-1}(M)$ for which $d\eta = \omega$.
The vector space of coboundaries is denoted by $B^k(M) \subseteq \Lambda^k(M)$.

Since $dd = 0$, we have $B^k(M) \subseteq Z^k(M)$ and we can form the factor vector spaces

$$H^k_{dR}(M) = Z^k(M)/B^k(M)$$

which are called the (de Rahm[1]) *Cohomology Vector Spaces*. Because we will

[1] Georges de Rahm, 1903-1990, Swiss mathematician (Canton Vaud). He switched from Literary to Scientific studies to avoid Latin.

not be working with any other kind of cohomology group, we will omit the dR subscript and write merely $H^k(M)$.

It is important to notice that while $Z^k(M)$ and $B^k(M)$ are often infinite dimensional vector spaces, the $H^k(M)$ are usually finite dimensional. The dimensions, often called *rank* in these circumstances, are the Betti numbers of M.

Now as an example we will look at $H^0(M)$ which is very simple. We are looking at manifolds M which include open subsets of $\mathbb{R}^n$. Consider $\omega \in \Lambda^0(M)$. Then ω is just a function on M and it will be a cocycle just when $d\omega = 0$, which means it will be a constant function. There are no coboundaries in this circumstances. If M is connected, then the cocycles form a one dimensional vector space which is then $\mathbb{R}$ and we have $H^0(M) = \mathbb{R}$ and Betti number $h^0 = 1$. On the other hand, if M is *not* connected, then say it has connected components $M_1, M_2, \ldots, M_r$. Then ω may be specified by giving a constant function $f_1, f_2, \ldots, f_r$ on each of the components independently, and thus we have, as above, $H^0(M) = \mathbb{R} \oplus \mathbb{R} \oplus \cdots \oplus \mathbb{R}$ with r summands, and $h^0 = r$. Thus essentially H^0 counts the number of connected components. Notice there is nothing in the game that keeps r from being infinite; if we want finite dimensional H^0 we must specify finitely many components.

Since this is just an introduction, we will now specify that in our future examples M will be connected so that $H^0(M) = \mathbb{R}$ and $h^0 = 1$.

Before progressing futher, let us introduce the normal notation to handle the situation. Let $\omega \in Z^k(M)$. Then ω determines a cohomology class in the factor vector space $H^k(M) = Z^k(M)/B^k(M)$ and we denote this cohomology class by $[\omega]$. Then, by definition, $[\omega] = [\omega + d\eta]$ for any $\eta \in \Lambda^{k-1}(M)$ and if $[\omega_1] = [\omega_2]$ then there exists and $\eta \in \Lambda^{k-1}(M)$ so that $\omega_2 - \omega_1 = d\eta$.

For our next example let us look at $M = \mathbb{R}$. We know $H^0(\mathbb{R}) = \mathbb{R}$. What about $H^1(\mathbb{R})$? Any $\omega \in \Lambda^1(\mathbb{R})$ will have $d\omega = 0$ since $\Lambda^2(\mathbb{R}) = 0$. Thus $\omega \in \Lambda^2(\mathbb{R})$ will look like $\omega = f\,dt$. We form

$$g(t) = \int_0^t f(u)\,du$$

Then clearly $dg = f(t)\,dt = \omega$ so that ω is always a coboundary, and we have $H^1(\mathbb{R}) = 0$. For $k > 1$ all $\Lambda^k(\mathbb{R}) = 0$, so we have

$$H^k(\mathbb{R}) = \begin{cases} \mathbb{R} & \text{if } k = 0 \\ 0 & \text{if } k > 0 \end{cases}$$

Now we look at the circle S^1. We take a unit circle and θ as the parameter. Notice, and this is critically important, that θ is not a continuous function on S^1, and ultimately this will lead us to $H^1(S^1)$. We take for our $\omega \in \Lambda^1(S^1)$ the element $\omega = d\theta$. $\omega \in Z^1(S^1)$ since $\Lambda^2(S^1) = 0$. Divide the circle into two arcs, the upper half circle and the lower half circle. On the upper half circle the parameter θ takes the values $0 \le \theta \le \pi$ and on the lower half circle θ takes the

values $\pi \leq \theta \leq 2\pi$. These are the distinct θ's that the expression $d\theta$ refers to. We use these parameters to compute

$$\begin{aligned}\int_{S^1} \omega &= \int_0^{\pi} d\theta + \int_{\pi}^{2\pi} d\theta \\ &= \theta\big|_0^{\pi} + \theta\big|_{\pi}^{2\pi} \\ &= \pi + \pi = 2\pi\end{aligned}$$

This verifies that $\omega \notin B^1(S^1)$ since if $\omega = df$ then

$$\int_{S^1} \omega = \int_{S^1} df = \int_{\partial S^1} f = 0$$

since $\partial S^1 = 0$. We have now shown that $Z^1(S^1) \neq 0$. To complete the calculation of $H^1(S^1)$ we need the following lemma.

Lemma Let $\omega = f\,d\theta$ be in $Z^1(S^1)$. Then ω is in $B^1(S^1)$ if and only if $\int_{S^1} \omega = 0$.

Proof $\Rightarrow$: If $\omega = df$ where f is a (single valued) function on S^1 then

$$\int_{S^1} \omega = \int_{S^1} df = f\big|_0^{2\pi} = f(2\pi) - f(0) = 0$$

since $f(2\pi) = f(0)$ since f is a function on S^1.

$\Leftarrow$: Suppose $\int_{S^1} \omega = 0$. Set

$$f(\theta) = \int_0^{\theta} \omega$$

Then since $\int_{S^1} \omega = 0$, f is well defined for all θ and satisfies $f(\theta + 2\pi) = f(\theta)$ making it a well defined function on S^1. Clearly $df = \omega$ and $\omega \in B^1(S^1)$.

Now let $\omega = f\,d\theta$ be in $Z^2(S^1)$ and let $\alpha = \int_{S^1} \omega$. Then $\omega - \frac{\alpha}{2\pi} d\theta$ has integral around S^1 equal to 0. Indeed

$$\int_{S^1} (\omega - \frac{\alpha}{2\pi} d\theta) = \alpha - \frac{\alpha}{2\pi} 2\pi = 0$$

and by the lemma we have $\omega - \frac{\alpha}{2\pi} d\theta = df$ for some f on S^1. Hence we have $[\omega - \frac{\alpha}{2\pi} d\theta] = [df] = 0$ in ${}^1(S^1)$ and thus

$$[\omega] = [\frac{\alpha}{2\pi}\, d\theta + \omega - \frac{\alpha}{2\pi} d\theta] = [\frac{\alpha}{2\pi}\, d\theta + df] = \frac{\alpha}{2\pi}[d\theta]$$

Thus $H^1(S^1)$ consists of just the multiples of $[d\theta]$ and so $H^1(S^1) = \mathbb{R}$.

It is worth noting that in the proof it was very important to find a path in S^1 (which was S^1 itself) and a differential $\omega = d\theta$ for which $\int_{\text{path}} \omega \neq 0$. The non-zero output of the integral is called a *period* of the differential form and periods have many uses. Digesting the result

For the Unit Circle S^1

$$H^k(S^1) = \begin{cases} \mathbb{R} & \text{if } k = 0, 1 \\ 0 & \text{if } k > 1 \end{cases}$$

If we use more powerful methods we can get more powerful results. If you think about the converse of the Poincarè Lemma you will see that it immediately gets the following result since it says that, for positive k and an open set U of $\mathbb{R}^n$ all the cocycles in $Z^k(U)$ are coboundaries; $Z^k(U) = B^k(U)$ and thus we have

For an open subset U of $\mathbb{R}^n$

$$H^k(U) = \begin{cases} \mathbb{R} & \text{if } k = 0 \\ 0 & \text{if } k > 0 \end{cases}$$

We will present without proof the results for the n-sphere S^n. It is not so much a question of not being able to prove this; we could probably fight our way through it, but we would have to come up with ad hoc methods which would be of little value in general. So we think it is better to refer to [2] for this material where more powerful methods are developed which have general applicability. We do not want to develop these methods ourselves because they are a little advanced for this book and would require rather a lot of effort. The formulas for S^n are

For the n-sphere S^n

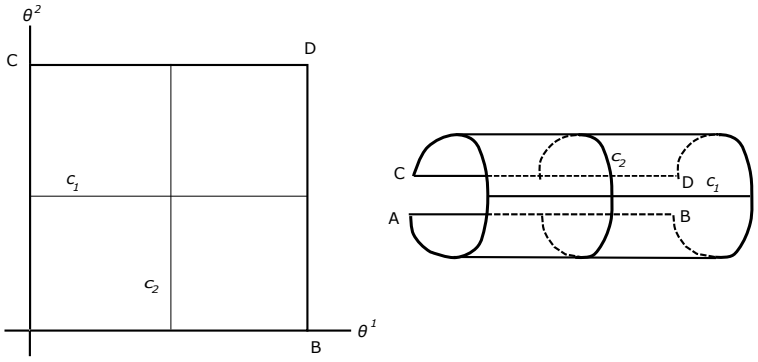
$$H^k(S^n) = \begin{cases} \mathbb{R} & \text{if } k = 0, n \\ 0 & \text{if } k \text{ is not 0 or } n \end{cases}$$

1.3.2 Tori

The *torus* (plural tori) is one of the most important mathematical shapes and we now attempt to explain why. There is a torus for each dimension. It is made by identifying the opposite sides of the unit cube. What makes tori important is that they are simultaneously differentiable manifolds and Abelian Groups. Different group structures may be put on the same torus but this has no effect on the cohomology so we can ignore it. This is closely connected with the ideas of Complex Torus and Abelian Variety but it would require uncomfortable degressions to explain this further and we don't need it for what we are doing.

The simplest torus (aside from the dot or 0-torus) is the 1-torus which is a circle. You can conveniently construct a 1-torus by taking a segment of length 2π and bending it into a circle. Essentially this identifies the two ends of the segment at 0 and 2π. A convenient coordinate is θ, $0 \leq \theta \leq 2\pi$ and we can recover the group structure by identifying the point of the circle with $e^{i\theta}$. Notice the coordinate θ does not give a quite proper coordinate system for the circle. In fact, θ is not even a function on S^1. By now you should see why. However, we can ignore this for integration purposes. We worked out the cohomology of the circle in the previous subsection so we can now move on from 1-tori. The group, as we have defined it, is isomorphic to the group $\mathbb{R}/2\pi\mathbb{Z}$ where the group operation is addition of real numbers but two real numbers r and s are identified if they differ by a multiple of 2π; that is $s - r \in 2\pi\mathbb{Z}$. This is called *reduction modulo* 2π.

The 2-torus can be constructed by identifying the opposite sides of a square of side 2π. (The size is for convenience and doesn't really matter.) We have attempted to draw an illustration showing part of the process.

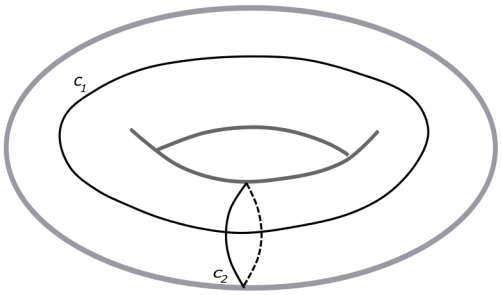


Constructing the 2-torus

This shows the first part of the process. To complete it one glues the identified sides AB and CD together and then bends the resulting tube around and glues the ends together. Note the coordinates θ^1 and θ^2 shown on the original square. These become coordinates on the torus. To add the points (θ_1^1, θ_1^2) and (θ_2^1, θ_2^2) on the torus, form $(\theta_1^1 + \theta_2^1, \theta_1^2 + \theta_2^2)$ and then reduce modulo 2π to get the coordinates of the sum.

Note the two curves c_1 and c_2. These curves are important for many purposes. Note θ^2 is constant on c_1 and θ^1 is constant on c_2. This is the important property of c_1 and c_2. The precise value of the constants is not important.

The final result of all the gluing, on which we have also drawn the curves c_1 and c_2, is



The 2-torus

It is critical to realize that the dimension 2 had no important role to play in what we have done. We could just have easily constructed the n-torus from a hypercube with coordinates $\theta^1, \theta^2, \ldots, \theta^n$ with $0 \leq \theta^i \leq 2\pi$ and curves

$c_1, c_2, \ldots, c_n$. Nothing would be different. (We are using 2π only for convenience; any positive number would work.) Also, for convenience

Notation We will denote the n-torus by T^n.

The first thing we will do is to find the cohomology of T^2. We know $H^0(T^2) = \mathbb{R}$ because T^2 is connected and we strongly suspect that $H^2(T^2) = \mathbb{R}$ because we know the volume topform is going to be a nonzero element in $H^2(T^2)$ since its integral over T^2 is the area and hence non-zero.

Two candidates for $H^1(T^2)$ are $[d\theta^1]$ and $[d\theta^2]$. These cannot be coboundaries, since the integrals $\int_{c_i} d\theta^j = 2\pi\delta_i^j$. On the other hand suppose $\omega \in Z^1(T^2)$. We can form $\alpha_i = \int_{c_i} \omega$. Then, as we did with the circle, we can show

$$\int_{c_i} (\omega - \frac{\alpha_i}{2\pi}\, d\theta^i) = 0$$

and from this we could show, with a little more effort than before, that we could manufacture a function

$$f(\theta_1, \theta_2) = \int_{(0,0)}^{(\theta^1,\theta^2)} (\omega - \frac{\alpha_1}{2\pi}\, d\theta^1 - \frac{\alpha_2}{2\pi}\, d\theta^2)$$

which is well defined since its integral over any closed curve is 0. (This is not quite obvious, but not too difficult to prove). Then, as before

$$[\omega] = \frac{\alpha_1}{2\pi}[d\theta^1] + \frac{\alpha_2}{2\pi}[d\theta^2]$$

so that

$$H^1(S^2) = \mathbb{R} \otimes \mathbb{R}$$

Now we can make an educated guess about $H^k(T^n)$. Based on what we have seen we can form the vector space $Z^k(T^n)$ by using

$$d\theta^{i_1} \wedge d\theta^{i_2} \wedge \cdots \wedge d\theta^{i_k} \qquad \text{where} \qquad i_1 < i_2, \cdots, i_k$$

as a basis. We can then show in a similar but more complicated way that if we use I to index the increasing sequences $i_1 < i_2, \cdots, i_k$, and $\omega \in Z^k(T^n)$ then

$$[\omega] = \sum_{j=1}^{k} \frac{\alpha_I}{2\pi}[d\theta^{i_1} \wedge d\theta^{i_2} \wedge \cdots \wedge d\theta^{i_k}]$$

Since there are $\binom{n}{k}$ possible ways to form such increasing sequences, we see that

$$H^k(T^n) = \bigoplus_{j=1}^{\binom{n}{k}} \mathbb{R}$$

1.4 The Hodge Decomposition and the Hodge Theorem

1.4.1 Orientation

In this section we will give a sort of derivation of the Hodge Decomposition and use it to derive the Hodge theorem that there is a unique Harmonic Form in each cohomology class of a compact manifold. We cannot give complete proofs of these results because it would almost double the size of the book, but we can give an argument that makes it very likely the results are true. The missing piece is called *regularity theory* and the idea is that if you find a solution to your problem in a large space which contains many more functions than you want to deal with, there is a possibility that the solutions you found are actually inside a much smaller space consisting of much nicer functions. For example, pretty much no matter what space you find a solution to $\triangle u = 0$, the solutions will actually be C^∞. This is the prototype result of regularity theory and was proved by Hermann Weyl long ago. It turns out that if L is an elliptic differential operator and f is C^∞ then the solutions u to $Lu = f$ will also be C^∞. This is a vast, subtle, and difficult theory and we cannot go into it in this book, but we will give references for those who wish to pursue it. Most users of regularity theory simply use the theorems and do not concern themselves with inner workings of the regularity machine. Vast areas of modern mathematics are dependent on regularity results, as for instance the material which we are going to develop. It is sometimes possible to avoid regularity by means of ad hoc methods, which may be of great interest in their own right. We will discuss more details and give references at the appropriate place in our development.

So what we will give here is the abstract outline of the results without emphasizing the regularity theory. Needless to say this cuts out about four fifths of the difficulty.

One further warning. Our idiosyncratic method of derivation of the Hodge decomposition theorem is not necessarily the optimal way to attack the problem when giving a complete proof.

In order to even present the material we need to take a detour and learn a little functional analyis. Since we are using Real Hilbert Space this is mostly pretty easy. We will learn a little functional analysis now, use that to get our main results, and then enthusiasts can learn a little more functional analysis so we can present how the regularity theory works. Naturally you may omit that part if you are not curious about it. However, we warn that the tiny amount we present would pay big dividends in understanding how much of modern analysis works, so try to crank up the curiosity.

1.4.2 Functional Analysis I

Orientation and Definitions

Linear Algebra, as met early in ones mathematical career, is the theory of finite dimensional vector spaces. These are not adequate for many applications and generalization is necessary. The direct generalization does not work well; infinite dimensional spaces without additional structure are not suitable for most applications. For example, a basis (called a Hamel basis) contains an uncountable number of elements, which makes calculation difficult. (In this situation each element of the infinite dimensional vector space is a *finite* linear combination of basis elements, and this is also too restrictive)

The solution to this problem is to add additional structure to the infinite dimensional vector space, usually a topology, which may appear as a result of other additions. The most direct method is perhaps to add a size function in the guise of *norm* which satisfies the following axioms

$$\begin{aligned} ||v|| &> 0 \text{ for all } u \neq 0 \\ ||\lambda u|| &= |\lambda|\, ||u|| \\ ||u+v|| &\leq ||u|| + ||v|| \end{aligned}$$

A vector space with a norm is called a *normed linear space*. A normed linear space is a metric space with distance function $\rho(u,v) = ||u-v||$. One usually wants to work in a complete space[2] and so if the space is not complete one usually completes it by the standard method.

Def A *Banach space* is a complete normed linear space.

Banach spaces are the simplest of the infinite dimensional spaces in some sense, but they do not necessarily have bases, which is just one of their inconveniences.

For applications it is often better to generalize not the naked vector space but the vector space with an inner product, to get an infinite dimensional inner product space. This does not work well unless the inner product is positive definite and we restrict our attention to this case.

Def A Real Hilbert Space is a space with and inner product which is *complete* with respect to the norm $||v|| = (v,v)^{\frac{1}{2}}$. The inner product satisfies

1. $(u,u) > 0$ for $u \neq 0$

2. $(u,v) = (v,u)$

3. $(u+v,w) = (u,w) + (v,w)$

4. $(\lambda u, v) = \lambda\,(u,v)$ for $\lambda \in \mathbb{R}$

(For a complex Hilbert Space replace 2. by $(u,v) = \overline{(v,u)}$ and 4. by $(\lambda u, v) = \overline{\lambda}\,(u,v)$ for $\lambda \in \mathbb{C}$, but we do not need this.) Hilbert spaces are special cases

[2] See Topology section

of Banach spaces (with the norm $||v|| = (v, v)^{\frac{1}{2}}$) and are much more convenient to work in. For example, the ones usually used in applications, the separable ones, all have bases.

A topological space is *separable*[3] if and only if it has a countable dense subset. For our purposes we quote the theorem that a Hilbert Space has a countable (or finite) orthonormal basis if and only if it is separable. We will work only with separable Hilbert spaces. We explicitly *do not* exclude finite dimensional spaces from being Hilbert spaces, because to do so is very inconvenient for Quantum Mechanics.

The norm topology of a Hilbert Space can be described in terms of limits. We define $\lim_{n\to\infty} x_n = x$ or more tersely $x_n \to x$ by $\lim_{n\to\infty} ||x - x_n|| = 0$

For separable Hilbert spaces there is a good analogue of expansion in terms of an orthonormal basis. We write

$$f = \sum_{i=1}^{\infty} a^i e_i$$

if and only if for any $\epsilon > 0$ there exists and N so that for any $n \geq N$

$$||f - \sum_{i=1}^{n} a^i e_i|| < \epsilon$$

and the usual formulas continue to work:

$$\begin{aligned} a^i &= (f, e_i) \\ \sum_{i=1}^{n} (a^i)^2 &\leq ||f||^2 \quad \text{Parseval's Inequality} \\ \sum_{i=1}^{\infty} (a^i)^2 &= ||f||^2 \quad \text{Parseval's Equality} \end{aligned}$$

All the usual relationships involving perpendicularity work in a Hilbert Space just as they do in a finite dimensional Euclidean space with a positive definite inner product and one can define angles between vectors the same way.

Linear Manifolds and Subspaces

Def Let H be a Hilbert Space and $V \subseteq H$ be a subset of H. Then V is a linear manifold of H if and only if V is closed under addition and scalar multiplication.

Linear manifolds are nice, but even nicer are the *closed* linear manifolds.

Def A linear manifold V of a Hilbert Space H is a *subspace* if and only if V is closed as a subset of H.

We remind the reader that this means if $x_n \to x$ and $x_n \in V$ then $x \in V$. This of course need not be true for linear manifolds.

[3]There is no clear reason for this name.

The distinction is important. The domains of differential operators are not naturally closed, so they are linear manifolds but not subspaces. We mention that if a linear manifold happens to be finite dimensional, then it is automatically closed.

Linear Operators

We will be dealing with two kinds of operators. First there are continuous or bounded operators and second there are closed operators.

The first kind of operator $T : H^1 \to H_2$ we will describe is the continuous or bounded operator. These kind of operators are always defined on the whole space H_1. It is simpler in this context to use the bounded description, which is equivalent to the continuous description. Continuous is of course the usual concept:

$$\text{If } x_n \to x \text{ then } T(x_n) \to T(x)$$

This is equivalent to the concept of *bounded* operator which can be defined in several ways. Perhaps the easiest is to say that the image of the unit ball $\overline{B}_1(0)$ is bounded. That is, there exists a C so that

$$\text{If } ||x|| \leq 1 \text{ then } ||T(x)|| \leq C$$

By considering $x/||x||$ we then see that

$$||T(x)|| \leq C||x||$$

The infimum of all C for which this is true is called the norm of T and denoted by $||T||$. This then gives

$$||T(x)|| \leq ||T||\,||x||$$

It is obvious that if T is bounded then it is continuous and the other direction is only a little harder.

The second type of operator important for our purposes is the *closed* operator. For many readers this will be a new concept. It is a kind of weakening of continuity suitable for differential operators. *It is most important to realize that the normal mode of discussion of functions is suspended in one detail for closed operators.* When we write

$$T : H_1 \to H_2$$

we **do not** assume that $\mathrm{D}(T)$, the domain of T, is *all* of H_1. We **do** assume that $\mathrm{D}(T)$ is *dense* in H_1. Notice this is contrary to normal usage but it is standard in the field and one must get used to it. The definition of closed operator is then

$$\begin{aligned} T \text{ is closed} \quad \Longleftrightarrow \quad & \text{if } x_n \to x \text{ and } T(x_n) \to y \\ & \text{then } x \in \mathrm{D}(T) \text{ and } T(x) = y \end{aligned}$$

Let us put this into words hoping to clarify it a bit. If $x_n \to x$ and $T(x_n)$ just happens to approach something y, (and nothing guarantees this will happen), then x *must* be in the domain of T and $T(x) = y$.

This condition may be described in another way. We may form the Hilbert space $H_1 \times H_2$ and we can make this into a Hilbert space by setting

$$(\,(x_1, y_1), (x_2, y_2)\,) = (x_1, x_2) + (y_1, y_2) \qquad \text{for} \qquad x_i \in H_1 \text{ and } y_i \in H_2$$

and then the graph of T consists of the points $(x, T(x)) \in H_1 \times H_2$ for all $x \in \mathrm{D}(T)$. Then we have the easily proved result

$$T \text{ is a closed operator} \Longleftrightarrow \text{The graph of } T \text{ is closed in } H_1 \times H_2$$

It is important to emphasize that in contrast to bounded operators where the domain of the operator is always H_1, with closed operators the domain is extremely important. For example, with differential operators in $L^2(-\pi, \pi)$ the boundary conditions select the domain, and different boundary condition can make for wildly different operators. It is common to have two differential operators with the same formula but different domains which then are to be considered different operators. It is also possible for the domain of one operator, S, to be a subset of the domain of a second operator T *with the same formula,* and we then write $S \subseteq T$. More formally:

Def Let S and T be two operators, $S, T : H_1 \to H_2$, and $\mathrm{D}(S) \subseteq \mathrm{D}(T)$ and also for $u \in \mathrm{D}(S)$ we have $S(u) = T(u)$. Then we write $S \subseteq T$.

For example let $H = L^2(-\pi, \pi) = \{u(t) | \int_{-\pi}^{\pi} |u|^2 dt < \infty\}$ and $S, T : H \to H$ and $S(u) = T(u) = \frac{d}{dt}\, u$ and $\mathrm{D}(S) = \{u | u(-\pi) = u(\pi) = 0\}$ and $\mathrm{D}(T) = \{u | u(-\pi) = u(\pi)\}$[4]. Then $S \subseteq T$

Everything up to this point has been very easy, but we now come to one of the four great theorems of Functional Analysis, and this one is not so easy to prove. We refer to [13] for the proof. For more information on closed operators for the enthusiast we also cite [11].

Theorem (The Closed Graph Theorem) If $T : H_1 \to H_2$ is a closed operator and $\mathrm{D}(T)$ is all of H_1 then T is bounded.

The reason this theorem is so important is that if T is a closed operator that is one to one, and the range $\mathrm{R}(T)$ of T is all of H_2 then T^{-1} is also a closed operator whose domain is all of H_2, as is very easy to prove, and then the closed graph theorem tells us that T^{-1} is bounded, which is halfway to Nirvana. It tells us, for example, that the original problem is well posed, since if there is some small error in the f of $T(u) = f$ then the effect on u will be small since T^{-1} is continuous.

The reader now of course wishes to see an example of closed operator, but wait; things are not so simple as that. Our interest is in differential operators which as originally defined are not quite closed. Hence we introduce one more type of operator, the *closable* operator, which is almost closed, and just needs a little adjustment.

[4]These operators are closeable but not closed; see below.

Def An operator T is *closable* if and only if the closure of its graph is the graph of an operator, called $\overline{T}$.

How could this fail? We could have $y_n \to y$ and $T(y_n) \to z_1$ so $(y, z_1) \in \overline{\text{graph}(T)}$ and also $\tilde{y}_n \to y$ and $T(\tilde{y}_n) \to z_2$ so $(y, z_2) \in \overline{\text{graph}(T)}$. If $y_1 \neq y_2$ then $\overline{\text{graph}(T)}$ is not the graph of an operator. To prevent this, we have the condition

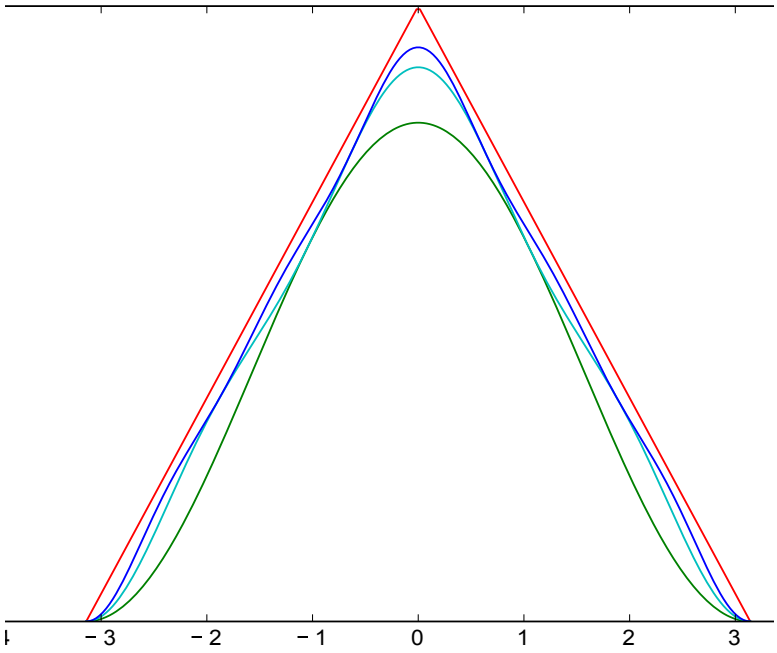
Theorem An operator $T : H_1 \to H_2$ is closable $\Longleftrightarrow$ if $x_n \to 0$ and $T(x_n) \to z$ then $z = 0$.

Now let $H_1 = H_2 = L^2(\Omega)$ where Ω is an open subset of $\mathbb{R}^n$ and $\partial\Omega$ is made up of a finite number of C^∞ manifolds glued together nicely. Let $T = \frac{\partial}{\partial u^i}$ with domain all the functions u on Ω which are C^∞ and which are 0 at the boundary. Then T is a closable operator on the Hilbert Space $L^2(\Omega) = \{f : \int_\omega |f|^2 dx^1 \wedge \ldots \wedge dx^n < \infty\}$.

We want to have a closer look at what happens when an closable operator is closed, so for a simple example we take $H_1 = H_2 = L^2([-\pi, \pi])$ and $T = \frac{d}{du}$. For the domain D(T) we take C^∞ functions which are 0 at the end points; $f(-\pi) = f(\pi) = 0$. Then it is pretty easy to see that T is closable and we want to look at something in the domain of the closure $\overline{T}$ of T. Let $f(t) = 1 - |\frac{t}{\pi}|$ and let us look at the slightly modified Fourier polynomials of $f(t)$. They are

$$
\begin{aligned}
f(t) &= 1 - (1/\pi)|t| && \text{red} \\
a_1 &= (1/2 - (4/\pi^2)) \\
f_1(t) &= 1/2 + (4/\pi^2)\cos(t) - a_1 && \text{green} \\
a_2 &= 1/2 - (4/\pi^2) - (4/(9\pi^2)) \\
f_2(t) &= 1/2 + (4/\pi^2)\cos(t) + (4/(9\pi^2))\cos(3t) - a_2 && \text{cyan} \\
a_3 &= 1/2 - (4/\pi^2) - (4/(9\pi^2)) - (4/(25\pi^2)) \\
f_3(t) &= 1/2 + (4/\pi^2)\cos(x) + (4/(9\pi^2))\cos(3 * x) + (4/(25\pi^2))\cos(5 * x) - a_3 && \text{blue}
\end{aligned}
$$

The a_i are the values of the regular Fourier polynomials at π and $-\pi$ and are subtracted from the regular Fourier polynomials so that $f_i(\pm\pi)$ will be 0 and thus the f_i, all of which are C^∞ are in D(T). Here is the graph of the situation.

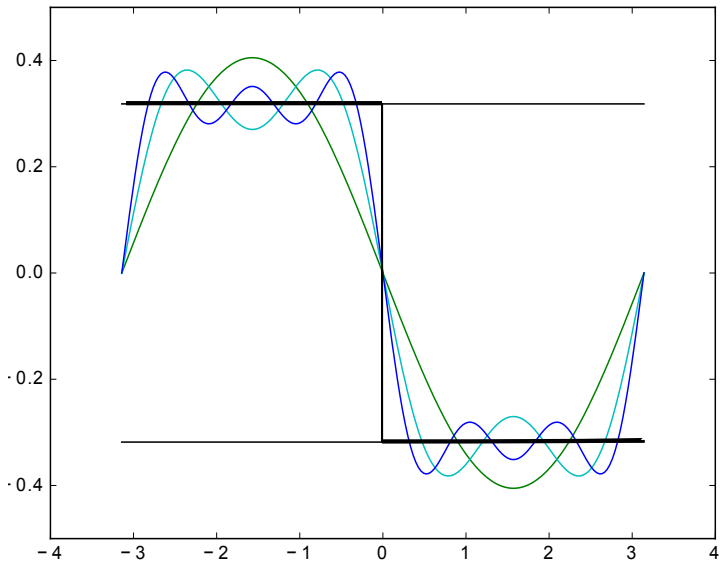


$g(t)$ and its modified Fourier approximations $f_i(t)$

It is clear from the illustration that $f_i \to f$. Now let's look at the derivatives of the $f_i(t)$.

$$
\begin{array}{lcll}
f_1'(t) & = & -(4/\pi^2)\sin(t) & \text{green} \\
f_2'(t) & = & -(4/\pi^2)\sin(t) - (4/(3\pi^2))\sin(3t) & \text{cyan} \\
f_3'(t) & = & -(4/\pi^2)\sin(t) - (4/(3\pi^2))\sin(3t) - (4/(5\pi^2))\sin(5*x) & \text{blue}
\end{array}
$$

Next we look at the graph, with the same color codings as before.



$f'(t)$ and its modified Fourier approximations $f_i'(t)$

Clearly There is a function $g(t)$ that the $f_i'(t)$ are approaching: $f_i'(t) \to g(t)$. We have drawn $g(t)$ (in heavy black) on the graph, and further investigation suggests that

$$g(t) \quad = \quad \begin{cases} 1/\pi & \text{if } t < 0 \qquad \text{red} \\ -1/\pi & \text{if } t > 0 \qquad \text{red} \end{cases}$$

Note we are here using the $L^2(-\pi, \pi)$ norm which for finite intervals essentially measures the area between the curves. This area is decreasing to 0 as $f_i'(t) \to g(t)$. The anomalous situation at $t = 0$ does not effect the area or the L^2 convergence. Thus our situation is exactly that for a closed operator; $f_i \to f$ and $f_i' \to g$ so $f \in \mathrm{D}(\overline{T})$ and $\overline{T}(f) = g$. Notice, and this is the point, that the original operator T had only smooth functions in it's domain, but it's closure $\overline{T}$ has the function $f(t)$ with a corner in its domain. Thus closing an operator may extend its domain to functions we don't especially want to deal with, and it is sometimes extremely difficult to know just which functions we have let in by this process. Regularity theory, which we have mentioned before, deals with the question of exactly how bad the monsters are that we have let in. Closing the second derivative operator $\frac{d^2}{dt^2}$ on the same domain will not let in nearly as bad functions.

Another important point is that this is a method of generalizing operators. Notice that $\overline{T}$ can differentiate f, with its corners. The closure process gives us the proper generalization of T, which is to take the derivative of each of the pieces of the piecewise differentiable function f and ignore the single bad point

$t = 0$. The closure proecess can be regarded as finding the smallest extension of T which is natural for the given Hilbert space.

And finally, note that the series we used were Fourier series, and it is typical for discontinuous functions for the Fourier series $f_i'(t)$ to converge to the average of the left and right limits of $g(t)$ at $t = 0$. There is no reason to think that if we used some other approximation $\tilde{f}_i \to f$ it would give the same output at 0.

This is a simple example of the *application* of functional analysis to the derivative operator. It is quite simple because we worked in one dimension. Things become more complex in higher dimensions. If you enjoyed it, you might wish to learn more functional analysis, and we recommend [13] for the general theory and [1] or [11] for Hilbert Space. These books cover a lot of material, and there are many other fine books on this subject, many of which are easier.

Continuous Linear Functionals

aaaaa

Adjoint Operators

Given an operator $T : H_1 \to H_2$ in Hilbert spaces with the domain $\mathrm{D}(T)$ dense in H_1, it is possible to define another operator $T^* : H_2 \to H_1$ called the Adjoint operator[5]. Here is the definition:

Def Let $T : H_1 \to H_2$ be a closed linear operator in Hilbert spaces having a dense domain. For all $w \in \mathrm{D}(T)$ suppose there is a $v \in H_1$ and a $u \in H_2$ for which $(u, T(w)) = (v, w)$. Then $u \in \mathrm{D}(T^*)$ and $T^*(u) = v$.

From this we get the equation

$$(u, T(w)) = (T^*(u), w) \qquad \text{for all } w \in \mathrm{D}(T) \text{ and } u \in \mathrm{D}(T^*)$$

It is necessary that T have dense domain for $T^*(u)$ to be uniquely defined. Also, we have restricted the adjoint to operators T with dense domain because this guarantees that $\mathrm{D}(T^*)$ will also have a dense domain, see [11].

Now we have

Thm T^* is always closed.

Proof: Suppose $u_i \to u$ and $T(u_i) \to v$. Then we have

$$(u_i, T(w)) = (T^*(u_i), w) \qquad \text{for all } w \in \mathrm{D}(T)$$

Taking limits with respect to i and remembering that the inner product is continuous in each variable, we have

$$(u, T(w)) = (v, w) \qquad \text{for all } w \in \mathrm{D}(T)$$

But this means, by definition of T^*, that $u \in \mathrm{D}(T^*)$ and $T^*(u) = v$, which is what is required for T^* to be closed.

[5] In many circumstances the notation $T^\dagger$ is used but this time we are going to go with the math notation.

1.5 Functional Analysis II

Functional Analysis courses often deal with little more than bounded (that is, continuous) operators. In physics and engineering the usual situation is to begin the problem with differential operators. These are *not* bounded operators. However, there are usually *closed* operators, a sort of substitute for bounded. The closed graph theorem states that a closed operator with closed domain is bounded. This is not directly helpful, since differential operators do not have closed domains. But we can still win, sort of, by looking at the inverse operators, since the inverse of a closed operator is also closed. Now should the operator have a range that is a closed set; the closed graph theorem tells us that the inverse if bounded, which in physics terms (after some adjustments) tells us that the problem is well posed, which means the solution exists and depends continuously on the data. This is halfway to Nirvana.

Thus the question of whether an operator has closed range is one of serious importance.

This, however, will not cover the necessary ground for applications because it does not address the important question of eigenvalues and eigenvector expansions. For eigenvalue expansions to be practically useful, we need three things. First, the eigenvalues must exist and be real. Second, each eigenvalue must have a finite dimensional eigenspace. Third, the eigenvalues (of the differential operator) must tend to infinity, which is more easily handled with the inverse operator whose eigenvalues then must tend to 0. If the inverse operator has a zero eigenvalue it may not be necessary to require the dimension of the 0 eigenspace be finite dimensional.

Now it turns out that these properties are generally true of a class of bounded operators called *compact operators*. These operators are better than bounded operators, and their defining property is that they take bounded sets into sets which are precompact, which means the closure of the image of a ball is compact. These are widely studied in first courses in functional analysis under the name of Riess-Schauder theory, although the reason to study them is often a closely guarded secret. Here is the reason; the inverse of a differential operator with closed range is very often a compact operator. And that gives the us the eigenvalues we want.

Originally the approach to the compactness of the inverses of differential operators was a terrible nightmare, depending ultimately on the Arzelà-Ascoli theorm, but in the later 1930s Sergei Lvovich Sobolev in Russia and Charles B. Morrey in America came up with a new way to attack the problem. Unfortunately various hostilities interrupted the smooth flow of mathematical research and after the war Morrey's contributions were neglected but Sobolev's slowly gained ground and eventually became a beautiful and important theory.

Again, because the details here are often hard to ferret out, and because the developments are in widely differing sections of the textbooks, we will give a brief description of how it goes. Ordinary derivatives, as operators, have some serious deficiencies, but these can be fixed by extending the domains of the differential operators somewhat. The extended differential operators are called

(of course) generalized differential operators. Let us work in a domain Ω which is an open subset of $\mathbb{R}^n$ whose closure $\overline{\Omega}$ has a reasonable (say piecewise smooth) boundary $\partial\overline{\Omega}$. The convenient Hilbert Space to start with is $L^2(\overline{\Omega})$ which we also refer to as $H^0(\overline{\Omega})$. This has the usual inner product

$$(f,g) = \int_\Omega f(x)g(x)dx^1 \wedge \cdots \wedge dx^n$$

(The first entry $f(x)$ would normally be conjugated but we are working with *real* Hilbert spaces.) The Sobelev spaces $H^s(\overline{\Omega})$ are now defined, for functions all of whose s^{th} generalized derivatives are defined, by

$$\begin{aligned}(f,g)_s &= \int_\Omega f(x)g(x) + (\text{sum of all the } s^{\text{th}}\text{derivatives of } f \text{ and } g)\, dx^1 \wedge \cdots \wedge dx^n \\ ||f||_s &= (f,f)^{\frac{1}{2}}\end{aligned}$$

where in the sum the number of derivatives of f and the number of derivatives of g are to have total s. The details are of absolutely no importance. Because we are using generalized derivatives it turns out the the $H^s(\overline{\Omega})$ are complete Hilbert Spaces.

There are two important theorems about Sobelev spaces which make everything work. First, if s is sufficiently large, the functions in $H^s(\overline{\Omega})$ are continuous, which is the *Sobelev Embedding Theorem* and says that $H^s(\overline{\Omega})$ may be embedded in $C^k(\overline{\Omega})$ for sufficiently large s. The *Rellich-Kondrachev* theorem says that the for $s > t$ the embedding of $H^s(\overline{\Omega})$ into $H^t(\overline{\Omega})$ is a compact operator.

Bibliography

[1] Akhiezer, N. I. and Glazman, I. M. THEORY OF LINEAR OPERATORS IN HILBERT SPACE, Dover, New York, 1993.

[2] Bott, Raoul and Tu, Loring W., DIFFERENTIAL FORMS IN ALGEBRAIC TOPOLOGY, Springer, New York, 1982

[3] Frankel, Theodore THE GEOMETRY OF PHYSICS, 3rd edition, Cambridge 2012 Cambridge University Press, Cambridge,etc, 2012

[4] Gilkey, Peter B. INVARIANCE THEORY, THE HEAT EQUATION AND THE ATIZYA-SINGER INDEX THEOREM, 2nd Edition Boca Raton 1995

[5] Liusternik L. A. and Sobolev V. J. ELEMENTS OF FUNCTIONAL ANALYSIS New York, 1961

[6] Lovelock, David and Rund, Hanno TENSORS, DIFFERENTIAL FORMS AND VARIATIONAL PRINCIPLES, Wiley Interscience

[7] Misner, Thorne and Wheeler, GRAVITATION, W.H. Freeman, San Francisco, 1973

[8] Nakahara, Mikio GEOMETRY, TOPOLOGY, AND PHYSICS, Institute of Physics Publishing, London, 2003

[9] Roe, John, ELLIPTIC OPERATORS, TOPOLOGY AND ASYMPTOTIC METHODS, 2nd Edition, Boca Raton, 1998

[10] Rosenberg, S. THE LAPLACIAN ON A RIEMANNIAN MANIFOLD, Cambridge, 1997

[11] Schmüdgen, Konrad, UNBOUNDED SELF ADJOINT OPERATORS ON HILBERT SPACE, Springer, Dordrecht, 2012

[12] Schulz, William, THEORY AND APPLICATION OF GRASSMANN ALGEBRA, http://www.cefns.nau.edu/ schulz/grassmann.pdf

[13] Yosida, Kôsaku FUNCTIONAL ANALYSIS, 4th Edition, Springer, New York, 1974.