Partial Differential Equations Notes 0

by
Jesse Ratzkin

In these pages we collect some relevant terminology and background material from analysis and linear algebra.

Some terminology and notation: We'll be integrating a lot of functions, and so we'll need a decent notation for integration. First of all, we call an open connected set $D \subset \mathbf{R}^n$ a **domain**. We will pay particular attention to a ball of radius R centered at x_0, which we denote as

$$B_R(x_0) = \{x \in \mathbf{R}^n : |x - x_0| < R\}.$$

We use the number ω_n to denote the volume of the unit ball in $\mathbf{R}^n$, so that

$$\mathrm{Vol}(B_R(x_0)) = \omega_n R^n, \qquad \mathrm{Area}(\partial B_R(x_0)) = n\omega_n R^{n-1}.$$

If $D \subset \mathbf{R}^n$ is a domain and $u : D \to \mathbf{R}$ is a continuous (or, more generally, measurable) function we write the volume integral of u over D as $\int_D u(x)dV(x)$. That is, we denote the volume element as $dV(x)$, with the x indicating the variable of integration. If it is clear from the context, we will leave out the x, just writing the volume element as dV. Sometimes we will also write the volume element as $d\mu$, so that a volume integral is $\int_D u(x)d\mu(x)$. If D has at least a Lipschitz boundary ∂D (see below) then we can integrate functions on the boundary, and we write these boundary integrals as $\int_{\partial D} u(x)dA(x)$, or sometimes $\int_{\partial D} u(x)d\sigma(x)$. Again, if it is clear we just write the area element as dA or $d\sigma$.

It will also be useful to define the support of a function.

Definition 1. *If $D \subset \mathbf{R}^n$ is a domain, and $u : D \to \mathbf{R}^n$ then the support of u is the set* $\mathrm{spt}(u)$, *which is the closure of the set $\{x \in D : u(x) \neq 0\}$. That is,* $\mathrm{spt}(u)$ *is the smallest closed set on which u is nonzero.*

Linear algebra: Recall that a linear map $T : \mathbf{R}^n \to \mathbf{R}^m$ satisfies

$$T(\alpha v + \beta w) = \alpha T(v) + \beta T(w)$$

for all $\alpha, \beta \in \mathbf{R}$ and $v, w \in \mathbf{R}^n$. If we choose bases $\{e_1, \dots, e_n\}$ for $\mathbf{R}^n$ and $\{f_1, \dots, f_m\}$ for $\mathbf{R}^m$, we can represent T with an $n \times m$ matrix $[T]$ with entries T_{ij}, where

$$T(e_i) = \sum_{j=1}^{m} T_{ij} f_j.$$

In this way we can identify the linear map T with left-multiplication by the matrix $[T]$. There are an assortment of operations you can do on the matrix $[T]$ which you should recall, including multiplication (which corresponds to the composition of linear maps), and in the case $n = m$ you can do things like take the determinant and trace of $[T]$. It is a good exercise to check that if $n = m$ then $\det([T])$ and $\mathrm{tr}([T])$ are actually independent on the choice of bases.

An eigenvector $v \in \mathbf{R}^n \backslash \{0\}$ with eigenvalue $\lambda \in \mathbf{R}$ satisfies $T(v) = \lambda v$. In particular, this means $(T - \lambda Id)(v) = 0$, and so $\det(T - \lambda Id) = 0$. This gives the eigenvalues as the roots of an nth degree polynomial, and so an $n \times n$ matrix has precisely n eigenvalues (counting multiplicity). Some particular cases warrant mention. T is called **positive definite** if all its eigenvalues are positive, and **positive semi-definite** if all its eigenvalues are non-negative. Similarly, a **negative definite** matrix has negative eigenvalues, while a **negative semi-definite** matrix has non-positive eigenvalues. If A and B are both $n \times n$ matrices, we say $A > B$ if $A - B$ is positive definite, and we say $A \geq B$ if $A - B$ is positive semi-definite.

It is also useful to recall the spectral theorem:

Theorem 1. *If A is a positive semi-definite, symmetric matrix then there exists an orthogonal matrix O such that $O^t AO = O^{-1}AO = D$ is diagonal, with the diagonal entries of D being the eigenvalues of A. In fact, on can take the columns of O to be the eigenvectors of A.*

A version of the spectral theorem remains true for a positive semi-definite, symmetric, bilinear form on a Hilbert space.

Some function spaces: In these paragraphs we quickly recall some notions of regularity for functions. We consider a domain (*i.e.* open, connected subset) $D \subset \mathbf{R}^n$ and a function $u : D \to \mathbf{R}$. For simplicity, most of the time we will take D to be bounded, and for most examples it will suffice to take D to be the unit ball centered at the origin 0.

Definition 2. *The space of continuous functions $u : D \to \mathbf{R}$ is written $C^0(D)$, and it has the norm*

$$\|u\|_{C^0(D)} = \sup_{x\in D} |u(x)|.$$

For $k \geq 1$, we say $u \in C^k(D)$ if u has all partial derivatives up to (and including) order k, and they are all continuous. We can place the norm

$$\|u\|_{C^k(D)} = \|u\|_{C^0(D)} + \sum_{l=1}^{k} \sup_{x\in D} \sqrt{\sum_{i_1<i_2<\cdots i_l} \left(\frac{\partial^l u}{\partial x_{i_1}\cdots\partial x_{i_l}}\right)^2}.$$

For instance, we $u \in C^1(D)$ if its gradient exists and is continuous, and it has $C^1(D)$-norm

$$\|u\|_{C^1(D)} = \sup_{x\in D} |u(x)| + \sup_{x\in D} \sqrt{\sum_{i=1}^{n} \left(\frac{\partial u}{\partial x_i}\right)^2} = \|u\|_{C^0(D)} + \||\nabla u|\|_{C^0(D)}.$$

We'll also use certain spaces of integrable functions. A functions $u : \mathbf{R}^n \to \mathbf{R}$ is integrable if $\int_{\mathbf{R}^n} |u(x)| dV(x) < \infty$, where the integral is with respect to the Lebesque measure, and we denote the space of integrable functions as $L^1(\mathbf{R}^n)$. More generally, if $p > 0$ we denote

$$L^p(\mathbf{R}^n) = \{u : \mathbf{R}^n \to \mathbf{R} : \int_{\mathbf{R}^n} |u(x)|^p dV(x) < \infty\},$$

and we endow $L^p(\mathbf{R}^n)$ with the norm

$$\|u\|_{L^p(\mathbf{R}^n)} = \left(\int_{\mathbf{R}^n} |u(x)|^p dV(x)\right)^{1/p}.$$

It is a fact that this norm makes $L^p(\mathbf{R}^n)$ into a Banach space; in particular it is a complete metric space. One can make similar definitions for funtions defined on a domain $D \subset \mathbf{R}^n$, obtaining the spaces $L^p(D)$ for $p > 0$. We'll usually restrict our attention to the case $p \geq 1$.

It follows from the definitions of Lebesque measure that the space of simple functions

$$f = \sum_{i=1}^{k} \alpha_i \chi_{U_i}, \quad \alpha_i \in \mathbf{R}, \quad U_i \text{ measurable}$$

are dense in L^p. (Here χ_{U_i} is the indicator function of U_i; it is 1 for points in U_i and 0 otherwise.) Two important convergence theorems you will see in the Measure Theory course are the monotone convergence theorem and the dominated convergence theorem.

Theorem 2. *Let f_n be a sequence of measurable functions such that*

- *$0 \leq f_n(x) \leq f_{n+1}(x)$ for almost every x and*

- *there is a measurable function f such that $f_n(x) \to f(x)$ for almost every x.*

Then

$$\lim_{n\to\infty} \int f_n = \int f,$$

where we interpret the limit as going towards ∞ if the integral of f is infinite.

Theorem 3. *Let f_n, g be measurable functions with $|f_n(x)| \leq g(x)$ for almost every x, and suppose $f_n(x) \to f(x)$ for almost every x. Then*

$$\lim_{n\to\infty} \int f_n = \int f.$$

Exercise: Show that if $D \subset \mathbf{R}^n$ is bounded and $1 \leq p < q$ then $L^q(D) \subset L^p(D)$. (Hint: Hölder's inequality.) Why do you need D to be bounded? Can you weaken that assumption at all?

It turns out that, at least for bounded domains, if $u \in L^p(D)$ for large enough p then it is also continuous. This is one part of the Sobolev embedding theorem, and you'd see it in an advanced functional analysis course. (Or you can look in chapter 7 of Gilbarg and Trudinger's book on second order, elliptic PDE.)

It turns out that usual continuity is often not enough to prove the results we want, and we will need something slightly stronger, called Hölder continuity.

Definition 3. *For $0 < \alpha < 1$ and a bounded domain D, we say $u \in C^{0,\alpha}(D)$ if*

$$\|u\|_{C^{0,\alpha}(D)} = \|u\|_{C^0(D)} + \sup_{x\neq y} \frac{|u(x) - u(y)|}{|x-y|^\alpha} < \infty.$$

There is a corresponding notion of Hölder continuous for unbounded domains, but you need to phrase the definition a little more carefully. Basically, the idea is the same, but you only want to compare two points which are close together.

We can allow the limit case of $\alpha \to 1$, and in this case we say $u \in C^{0,1}(D)$ is Lipschitz continuous. You might recall the following theorem from your course in measure theory:

Theorem 4. *If $u \in C^{0,1}(D)$ then u is differentiable almost everywhere.*

Exercise: Show that $u(x) = x^{1/3} \in C^{0,\alpha}(-1,1)$ precisely for $0 < \alpha \leq 1/3$. Is there anything special about the exponent $1/3$, or can you make a similar statement about the function $u(x) = |x|^p$ for any exponent $0 < p < 1$?

Exercise: If D is bounded, show that any $u \in C^1(D)$ is also Lipschitz continuous: $C^1(D) \subset C^{0,1}(D)$ if D is bounded.

Exercise: Show that if $0 < \alpha < \beta < 1$ and $u \in C^{0,\beta}(D)$ then we also have $u \in C^{0,\alpha}(D)$. We can write this containment as

$$C^{0,1}(D) \subset C^{0,\beta}(D) \subset C^{0,\alpha}(D) \subset C^0(D).$$

Exercise: Show that if $\alpha > 1$ and $u \in C^0(0,1)$ satisfies

$$\sup_{x\neq y} \frac{|u(x) - u(y)|}{|x-y|^\alpha} < \infty$$

then u is constant.

Finally, we say a domain D has a C^k boundary, for $k \geq 0$, if one can locally write ∂D as the graph of a C^k function. This means, for any $x_0 \in \partial D$ there is a neighborhood U of x_0, and a coordinate system $\{x_1, \ldots, x_n\}$ in U such that $\partial D \cap U = \{(x_1, x_2, \ldots, x_{n-1}, f(x_1, \ldots, x_{n-1}))\}$, and f is a C^k function. Similarly, D has a Lipschitz boundary, written $\partial D \in C^{0,1}$, if we can locally write ∂D as the graph of a Lipschitz function.

The Arzela-Ascoli theorem: The Arzela-Ascoli theorem (or Ascoli-Arzela theorem, or Montel's theorem, or ...) is one of the most useful theorems in analysis, because it's one of the easiest ways to produce a convergent sequence of functions. The proof is also an instuctive example of the Cantor diagonalization trick. To prove this theorem, we will first need to recall some definitions.

Definition 4. *Let X be a metric space. A subset A is dense in X if its closure $\bar{A}$ is all of X. A metric space X is called separable if it contains a countable, dense subset.*

For example, the real line is separable, because the rational numbers for a countable, dense subset. In the same way, $\mathbf{R}^n$ is separable for any n, which in particular means any domain $D \subset \mathbf{R}^n$ is separable.

Definition 5. *Let X be a metric space and let $\mathcal{F}$ be a family of functions from X to $\mathbf{R}$. The family $\mathcal{F}$ is equicontinuous if for every $\epsilon > 0$ there is a $\delta > 0$ such that*

$$\operatorname{dist}(x_1, x_2) < \delta \Rightarrow |f(x_1) - f(x_2)| < \epsilon$$

for all $f \in \mathcal{F}$.

Notice that in the definition we choose the **same** δ for all the functions $f \in \mathcal{F}$, and so this is a very strong condition. In particular, each $f \in \mathcal{F}$ is uniformly continuous. For instance, the family $\mathcal{F}_1 = \{f_n(x) = x^n : n = 1, 2, 3, \dots, 0 \leq x \leq 1\}$ is not equicontinuous, but the family $\mathcal{F}_2 = \{x^n : n = 1, 2, 3, \dots, 0 \leq x \leq 1/2\}$ is.

As far as the definition is concerned, it is not important that the target space is $\mathbf{R}$, and one can define equicontinuous familes of functions with the target space being any metric space Y at all. However, the main use of equicontinuity is the Arzela-Ascoli theorem below, so most of the time one will want the target to at least be a vector space.

Theorem 5. *(The Arzela-Ascoli Theorem) Let X be a separable metric space, and let $\mathcal{F}$ be an equicontinuous, pointwise bounded family of functions. (That is, there is a function $M : X \to (0, \infty)$ such that for all x and all $f \in \mathcal{F}$ we have $|f(x)| \leq M(x)$.) Then every sequence $\{f_n\}$ of functions in $\mathcal{F}$ has a subsequence $\{f_{n_k}\}$ which converges uniformly on all compact subsets of X.*

Proof. The space X is separable, so it contains a countable dense subset A. Enumerate the points in A as

$$A = \{x_1, x_2, x_3, \dots\}.$$

Now consider our sequence $\{f_n\}$ of functions in the family $\mathcal{F}$. The sequence of numbers

$$\{f_1(x_1), f_2(x_1), f_3(x_1), \dots\}$$

is a bounded seqence of real numbers and so it has a convergent subsequence $\{f_{n,1}(x_1)\} \to y_1$. We proceed to the point x_2, and we see that

$$\{f_{1,1}(x_2), f_{2,1}(x_2), f_{3,1}(x_2), \dots\}$$

is also a bounded sequence of real numbers, and so it also has a convergent subsequence $\{f_{n,2}(x_2)\} \to y_2$. Notice that we automatically have $\{f_{n,2}(x_1)\} \to y_1$, because we selected it as a subseqence of the convergent sequence $\{f_{n,1}(x_1)\}$. We proceed inductively, so that after the kth step we have a convergent sequence $\{f_{n,k}(x_k)\} \to y_k$, where we automatically have $\{f_{n,k}(x_j)\} \to y_j$ for $j \leq l$. Now we are ready to extract our convergent subsequence. We define $g_k = f_{k,k}$ and observe that for every l we have

$$\lim_{k \to \infty} g_k(x_l) = \lim_{k \to \infty} f_{k,k}(x_l) = y_l.$$

This is exactly the Cantor diagonal trick, and it's very useful to remember.

Now choose a compact set $K \subset X$, and let $\epsilon > 0$. The family of functions $\mathcal{F}$ is equicontinuous, so there is a $\delta > 0$ such that for all $p, q \in X$

$$\operatorname{dist}(p, q) < \delta \Rightarrow |g_k(p) - g_k(q)| = |f_{k,k}(p) - f_{k,k}(q)| < \epsilon$$

for all $k = 1, 2, 3 \ldots$. Cover K with open balls of radius $\frac{\delta}{2}$ and extract a finite subcover $B_1, B_2, \ldots, B_m$. The set A is dense in X, so for each $i = 1, \ldots, m$ there is a point $x_i \in A \cap B_i$. However, $\lim_{k\to\infty} g_k(x_i) = y_i$ exists, which means there is an integer N such that

$$|g_k(x_i) - g_l(x_i)| < \epsilon$$

so long as $k, l > N$. (This last line is just the fact that $\{g_k(x_i)\}$ is a Cauchy sequence written out.) Notice that we can choose one N to work for all the points $x_1 \in B_1, x_2 \in B_2, \ldots, x_m \in B_m$, because there are only finitely many of these balls which cover K.

Finally, pick an arbitrary $x \in K$. Then $x \in B_i$ for some i, and so $\operatorname{dist}(x, x_i) < \delta$. We combine the equicontinuity of $\mathcal{F}$ and the convergence $\{g_k(x_i)\} \to y_i$ to get that for $k, l > N$ we have

$$|g_k(x) - g_l(x)| \leq |g_k(x) - g_k(x_i)| + |g_k(x_i) - g_l(x_i)| + |g_l(x_i) - g_l(x)| < \epsilon + \epsilon + \epsilon = 3\epsilon.$$

In other words, we've just shown that when restricted to k the sequence $\{g_k\} = \{f_{k,k}\}$ converges uniformly, because it satisfies the Cauchy criterion. $\square$

We'll use the Arzela-Ascoli theorem several times in class (see the proof of Weyl's lemma, for instance), and it illustrates the power of a priori bounds.