Fall 2016 Notes

Carlos Salinas

November 5, 2016

Contents

Chapter 1

Probability

We will devote this chapter to the material that is covered in MA 51900 (discrete probability) as it was covered in DasGupta's class. We will, for the most part, reference Feller's *An introduction to probability theory and its applications, Volume 1* [3] (especially for the discrete noncalculus portion of the class) and DasGupta's own book *Fundamentals of Probability: A First Course* [1].

1.1 Counting

Some counting

Chapter 2

Introduction to Partial Differential Equations

Here we summarize some important points about PDEs. The material is mostly taken from Evans's *Partial Differential Equations* [2] with occasional detours to Strauss's *Partial Differential Equations: An Introduction* [8]. We will be following Dr. Petrosyan's Course Log which can be found here `https://www.math.purdue.edu/~arshak/F16/MA523/courselog/`, i.e., summarizing the appropriate chapters from [2].

2.1 First-Order PDEs

The transport equation

In this section, we consider the simplest first-order PDE, the *transport equation** with constant coefficients, i.e., the PDE

$$u_t + b \cdot Du = 0 \qquad \text{in } \mathbb{R} \times (0, \infty), \tag{2.1}$$

where b is a fixed vector in $\mathbb{R}^n$, and $u\colon \mathbb{R}^n \times [0, \infty) \to \mathbb{R}$ is the solution to the PDE. Our task is to find solutions u which satisfy the equation (2.1).

To address this task, let us suppose for a moment that we have a (smooth) solution u and try to compute it using the PDE (2.1). First, note that (2.1) asserts that the directional derivative $D_{(b,1)}u = 0$. Fix a point $(x, t) \in \mathbb{R}^n \times (0, \infty)$ and define

$$z(s) := u(x + sb, t + s)$$

for $s \in \mathbb{R}$. Then

$$\dot{z}(s) = Du(x + sb, t + s) \cdot b + u_t(x + sb, t + s) = 0.$$

Thus, z is a constant function of s and, consequently for each (x, t), u is constant on the line through (x, t) with direction $(b, 1) \in \mathbb{R}^{n+1}$. Hence if we know the value of u at any point on each such line, we know its value everywhere in $\mathbb{R}^n \times (0, \infty)$.

*For more details, refer to [2, §2.1].

Initial-value problem

Let's now look at the transport equation with *initial conditions*

$$\begin{cases} u_t + b \cdot Du = 0 & \text{in } \mathbb{R}^n \times (0, \infty), \\ u = g & \text{on } \mathbb{R}^n \times \{\, t = 0 \,\}. \end{cases} \tag{2.2}$$

Here $b \in \mathbb{R}^n$ and $g \colon \mathbb{R}^n \to \mathbb{R}$ are known, and u is the unknown. Given (x, t), the line through (x, t) with direction $(b, 1)$ is represented parametrically by $(x + sb, t + s)$ for $s \in \mathbb{R}$. This line hits the plane $\Gamma := \mathbb{R}^n \times \{\, t = 0 \,\}$ when $s = -t$, at the point $(x - tb, 0)$. Since u is constant on the line and $u(x - tb, 0) = g(x - tb)$, we deduce

$$u(x, t) = g(x - tb) \tag{2.3}$$

for $x \in \mathbb{R}^n$, $t \geq 0$. So if (2.2) has a sufficiently regular solution u (at least C^1), it must certainly be given by (2.3).

2.2 Characteristics

We now turn our attention to a very important method for solving first-order PDEs, the method of *characteristics*. This section is paraphrased if copied not verbatim from [2, §3.2].

Derivation of characteristic ODEs

Consider the first-order (possibly non-linear) PDE

$$F(Du, u, x) = 0 \quad \text{in } U, \tag{2.4}$$

subject to the *boundary condition*

$$u = g \quad \text{on } \Gamma, \tag{2.5}$$

where $\Gamma \subset \partial U$ and $g \colon \Gamma \to \mathbb{R}$ are known. We shall assume, for simplicity, that F and g are smooth.

We now develop the *method of characteristics* to solve (2.4), (2.5) by converting the PDE into a system of ODEs. We proceed as follows: Suppose u solves (2.4), (2.5) and fix a point $x \in U$. We would like to calculate $u(x)$ by finding some curve lying within U, connecting x with a point $x^0 \in \Gamma$ and along which we can compute u. Since (2.5) says $u = g$ on Γ, we know the value of u at one end x^0. We hope to be able to calculate u all along the curve, and so in particular at x.

Finding the characteristic curve

But how do we choose a path in U so all of this will work? Suppose the curve is described parametrically by the function $\mathbf{x}(x) = (x^1(s), \ldots, x^n(s))$, the parameter s lying in some subinterval $I \subset \mathbb{R}$. Assuming u is a C^2 solution of (2.4), we define also

$$z(s) := u(\mathbf{x}(s)). \tag{2.6}$$

In addition, set

$$\mathbf{p}(s) := Du(\mathbf{x}(s)); \tag{2.7}$$

that is, $\mathbf{p}(s) = (p^1(s), \ldots, p^n(s))$, where

$$p^i(s) = u_{xi}(\mathbf{x}(s)), \qquad 1 \leq i \leq n.. \tag{2.8}$$

So $z(\,\cdot\,)$ gives us the values of u along the curve and $\mathbf{p}(\,\cdot\,)$ records the values of gradient Du. We must choose a function $\mathbf{x}(\,\cdot\,)$ that will allow us to compute $z(\,\cdot\,)$ and $\mathbf{p}(\,\cdot\,)$.

Differentiating (2.8), we have

$$\dot{p}^i(s) = \sum_{j=1}^{n} u_{x_i x_j}(\mathbf{x}(s))\dot{x}^j(s). \tag{2.9}$$

But this expression is not too promising since it involves second order derivatives of u which we do not know (in fact, our solution need not be so regular as to have second order derivatives). On the other hand, if we differentiate (2.4) with respect to x_i, we have

$$\sum_{j=1}^{n} F_{p_j}(Du,u,x)u_{x_j x_i} + F_z(Du,u,x)u_{x_i} + F_{x_i}(Du,u,x) = 0. \tag{2.10}$$

We can use this identity to get rid of the second order derivatives in (2.12), provided we first set

$$\dot{x}^j(s) = F_{p_j}(\mathbf{p}(s), z(s), \mathbf{x}(s)), \qquad 1 \le j \le n. \tag{2.11}$$

Assuming (2.11) holds, we evaluate (2.10) at $x = \mathbf{x}(s)$, thereby obtaining from (2.6) and (2.7) the identity

$$\sum_{j=1}^{n} F_{p_j}(\mathbf{p}(s), z(s), \mathbf{x}(s))u_{x_i x_j}(\mathbf{x}(s)) + F_z(\mathbf{p}(s), z(s), \mathbf{x}(s))p^i(s) + F_{x_i}(\mathbf{p}(s), z(s), \mathbf{x}(s)) = 0$$

Finally, we differentiate (2.6) to give us

$$\begin{aligned} \dot{z}(s) &= \sum_{j=1}^{n} u_{x_j}(\mathbf{x}(s))\dot{x}^j(s) \\ &= \sum_{j=1}^{n} p^j(s)F_{p_j}(\mathbf{p}(s), z(s), \mathbf{x}(s)), \end{aligned} \tag{2.12}$$

the second equality holding by (2.8) and (2.9).

We summarize our results by rewriting equations (2.11), (2.11), and (2.12) as

$$\begin{cases} \text{(a) } \dot{\mathbf{p}}(s) = -D_x F(\mathbf{p}(s), z(s), \mathbf{x}(s)) - D_z F(\mathbf{p}(s), z(s), \mathbf{x}(s))\mathbf{p}(s), \\ \text{(b) } \dot{z}(s) = D_p F(\mathbf{p}(s), z(s), \mathbf{x}(s)) \cdot \mathbf{p}(s), \\ \text{(c) } \dot{\mathbf{x}}(s) = D_p F(\mathbf{p}(s), z(s), \mathbf{x}(s)). \end{cases} \tag{2.13}$$

Furthermore,

$$F(\mathbf{p}(s), z(s), \mathbf{x}(s)) = 0. \tag{2.14}$$

These identities hold for $s \in I$.

The system (2.13) of $2n+1$ first order ODEs comprises the *characteristic equotions/ODEs* of the nonlinear first-order PDE (2.4). The functions $\mathbf{p}(\,\cdot\,)$, $\mathbf{x}(\,\cdot\,)$ are called *characteristics* and $\mathbf{x}(\,\cdot\,)$ is called the *projected characteristic* (it is the projection of the full characteristics $(\mathbf{p}, z, \mathbf{x}) \subset \mathbb{R}^{2n+1}$ onto the physical region $U \subset \mathbb{R}^n$).

Theorem 2.1 (Structure of characteristic ODEs)**.** *Let $u \in C^2(U)$ solve the nonlinear, first-order partial differential equation* (2.4) *in U. Assume $\mathbf{x}(\cdot)$ solves the ODE* (2.13)(c)*, where $\mathbf{p}(\cdot) = Du(\mathbf{x}(\cdot))$, $z(\,\cdot\,) = u(\mathbf{x}(\,\cdot\,))$. Then $\mathbf{p}(\,\cdot\,)$ solves the ODE* (2.13)(a) *and $z(\,\cdot\,)$ solves the ODE* (2.13)(b)*, for those s such that $\mathbf{x}(s) \in U$.*

We still need to discover appropriate initial conditions for (2.13) to be useful. We do that in the following section.

Examples

But before we move on, we look at some examples to show you how to use (2.13) to find solutions to (2.4).

The linear case

Suppose (2.4) is linear, i.e., has the form

$$F(D, u, x) = \mathbf{x} \cdot Du(x) + c(x)u(x) = 0, \qquad x \in U. \tag{2.15}$$

Then, rewriting (2.15) in terms of p, z, and x, we have$F(p, z, x) = \mathbf{b}(x) \cdot p + c(x)z$, so

$$D_pF = \mathbf{b}(x)$$

so (2.13)(c) becomes

$$\dot{\mathbf{x}}(s) = \mathbf{x}(\mathbf{x}(s)), \tag{2.16}$$

an ODE involving only the function $\mathbf{x}(\,\cdot\,)$. Furthermore (2.13)(b) becomes

$$\dot{z}(s) = \mathbf{b}(\mathbf{x}(s)) \cdot \mathbf{p}(s). \tag{2.17}$$

Then equation (2.14) simplifies (2.17), yielding

$$\dot{z}(s) = -c(\mathbf{x}(s))z(s).$$

This ODE is linear in $z(\,\cdot\,)$, noce we know the function $\mathbf{x}(\,\cdot\,)$ by solving (2.16). In summary, we have

$$\begin{cases} \text{(a) } \dot{\mathbf{x}}(s) = \mathbf{b}(\mathbf{x}(s)), \\ \text{(b) } \dot{z}(s) = -c(\mathbf{x}(s))z(s). \end{cases} \tag{2.18}$$

Example 2.2. Let's now look at a simple example to see how to use (2.24) to solve a PDE. Consider the PDE

$$\begin{cases} x_1 u_{x_2} - x_2 u_{x_1} = u & \text{in } U, \\ \hfill u = g & \text{on } \Gamma, \end{cases} \tag{$*$}$$

where $U = \{\, x_1 > 0, x_2 > 0 \,\}$ and $\Gamma = \{\, x_1 > 0, x_2 = 0 \,\} \subset \partial U$. The PDE $(*)$ is of the form (2.15) with $\mathbf{b} = (-x_2, x_1)$ and $c = -1$. Thus the equations (2.24) read

$$\begin{cases} \dot{x}^1 = -x^2, \quad \dot{x}^2 = x^1, \\ \dot{z} = z. \end{cases} \tag{$**$}$$

Solving this system of ODEs we have

$$\begin{cases} x^1(s) = x^0 \cos s, \quad x^2(s) = x^0 \sin x, \\ \ \ z(s) = z^0 \mathrm{e}^s \\ \qquad\ \ = g(x^0)\mathrm{e}^s, \end{cases}$$

where $x^0 \geq 0$, $0 \leq s \leq \frac{\pi}{2}$. Now, fix $(x_1, x_2) \in U$. Select $s > 0$, $x^0 > 0$ so that $(x_1, x_2) = (x^1(s), x^2(s)) = (x^0 \cos s, x^0 \sin s)$ and solve for x^0, in this case, $x^0 = \sqrt{x_1^2 + x_2^2}$, $s = \arctan(\frac{x_2}{x_1})$, and therefore

$$\begin{aligned} u(x) &= u(x^1(s), x^2(s)) \\ &= z(s) \\ &= g(x^0)\mathrm{e}^s \\ &= g\Big(\sqrt{x_1^2 + x_2^2}\Big)\mathrm{e}^{\arctan(\frac{x_2}{x_1})}. \end{aligned}$$

The quasilinear case

Let's look at the quasilinear case now, i.e., (2.4) with the form

$$F(Du, u, x) = \mathbf{b}(x, u(x)) \cdot Du(x) + c(x, u(x)) = 0. \tag{2.19}$$

In this circumstance $F(p, z, x) = \mathbf{b}(x, z) \cdot p + c(x, z)$, whence

$$D_p F = \mathbf{b}(x, z).$$

Hence equation (2.13)(c) reads

$$\dot{\mathbf{x}}(s) = \mathbf{b}(\mathbf{x}(s), z(s)), \tag{2.20}$$

an ODE involving only the function $\mathbf{x}$. Furthermore, (2.13)(c) becomes

$$\dot{z}(s) = \mathbf{b}(\mathbf{x}(s)) \cdot \mathbf{p}(s), \tag{2.21}$$

which, after applying (2.14), turns into

$$\dot{z}(s) = -c(\mathbf{x}(s))z(s). \tag{2.22}$$

In summary, we have

$$\begin{cases} \text{(a) } \dot{\mathbf{x}}(s) = \mathbf{b}(\mathbf{x}(s)), \\ \text{(b) } \dot{z}(s) = -c(\mathbf{x}(s))z(s). \end{cases} \tag{2.23}$$

We will see later that the equation for $\mathbf{p}(\cdot)$ is in fact not needed (at least in the linear and quasilinear cases).

Example 2.3. Let's look at an example of a quasilinear PDE. Consider the PDE

$$\begin{cases} u_{x_2} + u_{x_1} = u & \text{in } U, \\ \qquad\quad\ \ u = g & \text{on } \Gamma. \end{cases} \tag{$*$}$$

Here $U = \{ x_2 > 0 \}$ and $\Gamma = \{ x_2 = 0 \} = \partial U$ with $\mathbf{b} = (1,1)$ and $c = -z^2$. Thus, the equations (2.23) yield

$$\begin{cases} \dot{x}^1 = 1, \quad \dot{x}^2 = 1, \\ \dot{z} = z^2. \end{cases}$$

Consequently

$$\begin{cases} x^1(s) = x^0 + s, & x^2(s) = s, \\ z(s) = \dfrac{z^0}{1 - sz^0} = \dfrac{g(x^0)}{1 - sg(x^0)}, \end{cases}$$

where $x^0 \in \mathbb{R}$, $s \geq 0$, provided the denominator is not zero.

Now, fix a point $(x_1, x_2) \in U$ and select $s > 0$ and $x^0 \in \mathbb{R}$ so $(x_1, x_2) = (x^1(s), x^2(s)) = (x^0 + s, s)$, i.e., $x^0 = x_1 - x_2$, $s = x_2$. Then

$$\begin{aligned} u(x) &= u(x^1(s), x^2(s)) \\ &= z(s) \\ &= \frac{g(x^0)}{1 - sg(x^0)} \\ &= \frac{g(x_1 - x_2)}{1 - x_2 g(x_1 - x_2)}. \end{aligned}$$

This solution of course only makes sense if $1 - x_2 g(x_1 - x_2) \neq 0$.

The fully nonlinear case

In the general case, we must integrate the full characteristics (2.13) if possible. In this case, we cannot generally reduce (2.13) and we must look at the PDE on a case-by-case basis.

Example 2.4. Let's look at an example. Consider the fully nonlinear PDE

$$\begin{cases} u_{x_1} u_{x_2} = u & \text{in } U, \\ u = x_2^2 & \text{on } \Gamma, \end{cases} \tag{$*$}$$

where $U = \{ x_2 > 0 \}$, $\Gamma = \partial U = \{ x_1 = 0 \}$. Here, $F(p, z, x) = p_1 p_2 - z$, and hence (2.13) yield

$$\begin{cases} \dot{p}^1 = p^1, & \dot{p}^2 = p^2, \\ \dot{z} = 2p^1 p^2, \\ \dot{x}^1 = p^2, & \dot{x}^2 = p^1. \end{cases}$$

After integrating these equations, we have

$$\begin{cases} x^1(s) = p_2^0(\mathrm{e}^s - 1), & x^2(s) = x^0 + p_1^0(\mathrm{e}^s - 1), \\ z(s) = z^0 + p_1^0 p_2^0(\mathrm{e}^{2s} - 1), \\ p^1(s) = p_1^0 \mathrm{e}^s, & p^2(s) = p_2^0 \mathrm{e}^s, \end{cases}$$

where $x^0 \in \mathbb{R}$, $s \in \mathbb{R}$, and $z^0 = (x^0)^2$.

We must determine $p^0 = (p_1^0, p_2^0)$. Since $u = x_2^2$ on Γ, $p_2^0 = u_{x_2}(0, x^0) = 2x^0$. Furthermore the PDE $u_{x_1x_2} = u$ itself implies $p_1^0 p_2^0 = z^0 = (x^0)^2$, and so $p_1^0 = \frac{x^0}{2}$. Consequently the formulas above become

$$\begin{cases} x^1(s) = 2x^0(\mathrm{e}^s - 1), \quad x^2(s) = \dfrac{x^0}{2}(\mathrm{e}^s + 1), \\ z(s) = (x^0)^2 \mathrm{e}^{2s}, \\ p^1(s) = \dfrac{x^0}{2}\mathrm{e}^s, \qquad p^2(s) = 2x^0\mathrm{e}^s. \end{cases}$$

Fix a point $(x_1, x_2) \in U$. Select s and x^0 so that $(x_1, x_2) = (x^1(s), x^2(s)) = (2x^0(\mathrm{e}^s - 1), \frac{x^0}{2}(\mathrm{e}^s + 1))$. This equality implies

$$x^0 = \frac{4x_2 - x_1}{4}; \qquad \mathrm{e}^s = \frac{x_1 + 4x_2}{4x_2 - x_1};$$

so

$$\begin{aligned} u(x) &= u(x^1(s), x^2(s)) \\ &= z(s) \\ &= (x^0)^2 \mathrm{e}^{2s} \\ &= \frac{(x_1 + 4x_2)^2}{16}. \end{aligned}$$

Compatibility conditions on boundary data

Let $x^0 \in \Gamma$. We intend to use the characteristic ODEs (2.13) to construct a solution u to (2.4), (2.5), at least near x^0, and for this, we must determine appropriate initial conditions

$$\mathbf{p}(0) = p^0, \qquad z(0) = z^0, \qquad \mathbf{x}(0) = x^0. \tag{2.24}$$

We will assume throughout that Γ is flat (i.e., isometric to $\{\, x_n = 0 \,\}$) at least near x^0.†

Clearly if the curve $\mathbf{x}(\,\cdot\,)$ passes through x^0, we should insist that

$$z^0 = g(x^0). \tag{2.25}$$

What should we require concerning $\mathbf{p}(0) = p^0$? Since (2.5) implies

$$u(x_1, \dots, x_{n-1}, 0) = g(x_1, \dots, x_{n-1})$$

near x^0, we may differentiate this to find

$$u_{x_i}(x^0) = g_{x_i}(x^0).$$

This, along with the PDE (2.4), forces p^0 to satisfy

$$\begin{cases} p_i^0 = g_{x_i}(x^0) \quad 1 \le i \le n-1, \\ F(p^0, z^0, x^0) = 0. \end{cases} \tag{2.26}$$

†This can always be achieved, assuming some regularity of Γ.

These identities provide n equations for the n quantities $p^0 = (p_1^0, \dots, p_n^0)$.

We call (2.25) and (2.26) *compatibility conditions.* A triple $p^0, z^0, x^0 \in \mathbb{R}^{2n+1}$ satisfying (2.25), (2.26), is called *admissible.* Note z^0 is uniquely determined by the boundary condition and our choice of x^0, but p^0 satisfying (2.26) may not exist or be unique.

Having ascertained what are appropriate boundary conditions for the characteristic ODEs with $\mathbf{x}(\,\cdot\,)$ intersecting Γ at x^0, we proceed to construct a solution to (2.4), (2.5), near x^0. We now ask, can we somehow appropriately perturb (p^0, z^0, x^0), keeping the compatibility conditions?

In other words, given a point $y = (y_1, \dots, y_{n-1}, 0) \in \Gamma$, with y close enough to x^0, we intend to solve the ODEs (2.13) with initial conditions

$$\mathbf{p}(0) = \mathbf{q}(y), \qquad z(0) = g(y), \qquad \mathbf{x}(0) = y. \tag{2.27}$$

Our task is now to find a function $\mathbf{q}(\,\cdot\,) = (q^1(\,\cdot\,), \dots, q^n(\,\cdot\,))$, so that

$$\mathbf{q}(x^0) = p^0 \tag{2.28}$$

and $\mathbf{q}(y), g(y), y$ is admissible; i.e., so

$$\begin{cases} q^i(y) = g_{x_i}(y) \quad 1 \le i \le n-1, \\ F(\mathbf{q}(y), g(y), y) = 0, \end{cases} \tag{2.29}$$

hold for all $y \in \Gamma$ close to x^0.

Lemma 2.5 (Noncharacteristic boundary conditions)**.** *There exists a unique solution* $\mathbf{q}(\,\cdot\,)$ *of* (2.28), (2.29)*, for all* $y \in \Gamma$ *sufficiently close to* x^0*, provided*

$$F_{p_n}(p^0, z^0, x^0) \neq 0.$$

See [2, §3.2 c] for proof.

More generally,

$$D_p F(p^0, z^0, x^0) \cdot \boldsymbol{\nu}(x^0) \neq 0,$$

$\boldsymbol{\nu}(x^0)$ denoting the outward unit normal to ∂U at x^0.

Now, given a point $y = (y_1, \dots, y_{n-1}, 0)$ sufficiently close to x^0 we solve the characteristic ODEs (2.28) subject to (2.29).

Write

$$\begin{cases} \mathbf{p}(s) = \mathbf{p}(y, s) = \mathbf{p}(y_1, \dots, y_{n-1}, s), \\ z(s) = z(y, s) = z(y_1, \dots, y_{n-1}, s), \\ \mathbf{x}(s) = \mathbf{x}(y, s) = \mathbf{x}(y_1, \dots, y_{n-1}, s), \end{cases}$$

displaying the dependence of the solution of (2.28), (2.29), with respect to s and y.

Lemma 2.6 (Local invertibility)**.** *Assume we have the noncharacteristic condition* $F_{p_n}(p^0, z^0, x^0) \neq 0$*. Then there exists an open interval* $I \subset \mathbb{R}$ *containing* 0*, a neighborhood* W *of* x^0 *in* $\Gamma \subset \mathbb{R}^{n-1}$*, and a neighborhood* V *of* x^0 *in* $\mathbb{R}^n$*, such that for each* $x \in V$ *there exists a unique* $s \in I$*,* $y \in W$ *such that*

$$x = \mathbf{x}(y, s).$$

The mappings $x \mapsto s, y$ *are* C^2*.*

In view of this lemma, for each $x \in V$, we can uniquely solve (at least locally) the equation

$$\begin{cases} x = \mathbf{x}(y, s), \\ y = \mathbf{y}(x), \qquad s = s(x). \end{cases} \tag{2.30}$$

Define

$$\begin{cases} u(x) := z(\mathbf{y}(x), s(x)), \\ \mathbf{p}(x) = \mathbf{p}(\mathbf{y}(x), s(x)), \end{cases} \tag{2.31}$$

for $x \in V$, s and y as in (2.30).

Putting all of this together, we have the following result:

Theorem 2.7 (Local existence theorem)**.** *The function u defined above is C^2 and solves the PDE*

$$F(Du(x), u(x), x) = 0 \qquad x \in V,$$

with the boundary condition

$$u(x) = g(x) \qquad x \in \Gamma \cap V.$$

See [2, §3.2.4] for more details.

2.3 Power Series

We now turn our attention to an important class of solutions to PDEs, analytic solutions. We begin this section with a brief discussion on noncharacteristic surfaces.

Noncharacteristic surfaces

Consider the k^{th}-order quasilinear PDE

$$\sum_{|\alpha|=k} a_\alpha(D^{k-1}u, \dots, u, x) D^\alpha u + a_0(D^{k-1}u, \dots, u, x) = 0, \tag{2.32}$$

in some region in $U \subset \mathbb{R}^n$. Let us assume for simplicity that Γ is a smooth $(n-1)$-dimensional hypersurface in U. For any $x^0 \in \Gamma$ let $\boldsymbol{\nu}(x^0) = \boldsymbol{\nu} = (\nu_1, \dots, \nu_n)$ denote the unit normal to Γ at x^0. Furthermore, we define the j^{th} *normal derivative* of u at $x^0 \in \Gamma$ by

$$\frac{\partial^j u}{\partial \nu^j} := \sum_{|\alpha|=j} \binom{j}{\alpha} D^\alpha u \boldsymbol{\nu}^\alpha = \sum_{\alpha_1 + \dots + \alpha_n = j} \binom{j}{\alpha} \frac{\partial^j u}{\partial^{\alpha_1} x_1 \cdots \partial^{\alpha_n} x_n} \nu_1^{\alpha_1} \cdots \nu_n^{\alpha_n}.$$

Let $g_0, \dots, g_{k-1} \colon \Gamma \to \mathbb{R}$ be k given functions. The *Cauchy problem* is to find a function u solving the PDE (2.32), subject to the boundary conditions

$$\begin{cases} u = g_0, \\ \dfrac{\partial u}{\partial \nu} = g_1, \\ \quad \vdots \\ \dfrac{\partial^{k-1} u}{\partial \nu^{k-1}} = g_{k-1} \end{cases} \qquad \text{on } \Gamma. \tag{2.33}$$

We say that (2.33) prescribe the *Cauchy data* on Γ.

Evans demonstrates how to extend the notion of noncharacteristic curve/surface to the PDE (2.32), (2.33), first in the case that Γ is flat and then makes some remarks about the case that Γ is nonflat. The most important result from this section is the notion of noncharacteristic surface for (2.32), (2.33) which is summarized in the following definition:

Definition 2.8. We say the surface Γ is noncharacteristic for the PDE (2.32) provided

$$\sum_{|\alpha|=k} a_\alpha \boldsymbol{\nu}^\alpha \neq 0. \tag{2.34}$$

for all values of the arguments of the coefficients $a_\alpha (|\alpha| = k)$.

Theorem 2.9 (Cauchy data and noncharacteristic surfaces)**.** *Assume that Γ is noncharacteristic for the PDE* (2.32)*. Then if u is a smooth solution of* (2.32) *satisfying the Cauchy data* (2.33)*, we can uniquely compute all of the partial derivatives of u along Γ in terms of Γ, the functions $g_0, \ldots, g_{k-1}$, and the coefficients $a_\alpha (|\alpha| = k)$, a_0.*

See [2, §4.6.b] for a proof.

Real analytic functions

We now briefly review the concept of analytic functions.

Definition 2.10. A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is said to be *(real) analytic near* x_0 if there exists $r > 0$ and constants $\{f_\alpha\}$ such that

$$f(x) = \sum_\alpha f_\alpha (x - x_0)^\alpha \qquad |x - x_0| < r,$$

the sum taken over all multiindices α.

Remarks 2.11.

(i) Remember that when we write x^α we mean $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$.
(ii) If f is analytic near x_0, then f is C^∞ near x_0. Furthermore the coefficients f_α are computed as $f_\alpha = \frac{D^\alpha f(x_0)}{\alpha!}$, where $\alpha! = \alpha_1! \cdots \alpha_n!$. Thus f equals its Taylor formula about x_0

$$f(x) = \sum_\alpha \frac{1}{\alpha!} D^\alpha f(x_0)(x - x_0)^\alpha \qquad |x - x_0| < r.$$

To simplify, we hereafter take x_0 to be 0.

Example 2.12. Let us look at an example of an analytic function. For $r > 0$, set

$$f(x) := \frac{r}{r - (x_1 + \cdots + x_n)} \qquad \text{for } |x| < \frac{r}{\sqrt{n}}.$$

Then

$$\begin{aligned} f(x) &= \frac{1}{1 - \dfrac{x_1 + \cdots + x_n}{r}} \\ &= \sum_{k=0}^{\infty} \left(\frac{x_1 + \cdots + x_n}{r} \right)^k \\ &= \sum_{k=0}^{\infty} \frac{1}{r^k} \sum_{|\alpha|=k} \binom{|\alpha|}{\alpha} x^\alpha \\ &= \sum_{\alpha} \frac{|\alpha|!}{r^{|\alpha|} \alpha!} x^\alpha. \end{aligned}$$

We employed the multinomial theorem for the third equality above and recalled that $\binom{|\alpha|}{\alpha} = |\alpha|!/\alpha!$. This power series is absolutely convergent for $|x| < r/\sqrt{n}$. Indeed,

$$\sum_{\alpha} \frac{|\alpha|!}{r^{|\alpha|} \alpha!} |x^\alpha| = \sum_{k=0}^{\infty} \left(\frac{|x_1| + \cdots + |x_n|}{r} \right)^k < \infty,$$

since $|x_1| + \cdots + |x_n| \leq |x|\sqrt{n} < r$.

The analytic equation we have just considered is quite important as it allows us to *majorize*, and therefore confirm the convergence of, other power series.

Definition 2.13. Let

$$f = \sum_{\alpha} f_\alpha x^\alpha, \qquad g = \sum_{\alpha} g_\alpha x^\alpha$$

be two power series. We say that g *majorizes* f, written $g \gg f$, if

$$g_\alpha \geq |f_\alpha| \qquad \text{for all } \alpha.$$

Here is an important result:

Lemma 2.14 (Majorants)**.**

(i) If $g \gg f$ and g converges for $|x| < r$, then f also converges for $|x| < r$.
(ii) If $f = \sum_\alpha f_\alpha x^\alpha$ converges for $|x| < r$ and $0 < s\sqrt{n} < r$, then f has a majorant for $|x| < s/\sqrt{n}$.

Cauchy–Kovalevskaya theorem

We now turn our attention to a very important theorem regarding analytic solutions to (2.32) with analytic Cauchy data (2.33) specified on an analytic, noncharacteristic hypersurface Γ.

Reduction to a first-order system

We intend to construct a solution u as a power series, but we must first transform (2.32), (2.33) into a more convenient form.

First of all, upon flattening out the boundary by an analytic mapping, we can reduce to the situation that $\Gamma \subset \{ x_n = 0 \}$. Additionally, by subtracting off appropriate analytic functions, we may assume that the Cauchy data are identically zero. Consequently, we may assume, without loss of generality, that our problem reads:

$$\begin{cases} \displaystyle\sum_{|\alpha|=k} a_\alpha(D^{k-1}u, \dots, u, x)D^\alpha u + a_0(D^{k-1}u, \dots, u, x) = 0 & \text{for } |x| < r, \\ u = 0, & \\ \dfrac{\partial u}{\partial x_n} = 0, & \text{for } |x'| < r,\ x_n = 0, \\ \vdots & \\ \dfrac{\partial^{k-1} u}{\partial x_n^{k-1}} = 0, & \end{cases} \tag{2.35}$$

$r > 0$ to be found. Here $a_\alpha(|\alpha| = k)$ and a_0 are analytic are analytic, and as usual we write $x' = (x_1, \dots, x_{n-1})$.

Finally, we this to a first-order *system*. To do so, we set

$$\mathbf{u} := \left(u, \tfrac{\partial}{\partial x_1}u, \dots, \tfrac{\partial^{k-1}}{\partial x_n^{k-1}}u\right),$$

the components of which are all the partial derivatives of u of order less than k. Let m hereafter denote the number of components of $\mathbf{u}$, so $\mathbf{u}\colon \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{u} = (u^1, \dots, u^m)$. Observe from the boundary condition in (2.35) that $\mathbf{u} = 0$ for $|x'| < r$, $x_n = 0$.

For $1 \le k \le m-1$, we can compute $u^k_{x_n}$ in terms o $\{\mathbf{u}_{x_j}\}_{j=1}^{n-1}$. Furthermore in view of the noncharacteristic condition $a_{(0,\dots,0,k)} \neq 0$ near 0, we can utilize the PDE in (2.35) to solve for $u^m_{x_n}$ in terms of $\mathbf{u}$ and $\{\mathbf{u}_{x_j}\}_{j=1}^{n-1}$.

Employing these relations, we can consequently transform (2.35) into a boundary-value problem for a first-order system $\mathbf{u}$, the coefficients of which are analytic functions. This system is of the form

$$\begin{cases} \mathbf{u}_{x_n} = \displaystyle\sum_{j=1}^{n-1} \mathbf{B}_j(\mathbf{u}, x')\mathbf{u}_{x_j} + \mathbf{c}(\mathbf{u}, x') & \text{for } |x| < r, \\ \mathbf{u} = 0 & \text{for } |x'| < r,\ x_n = 0, \end{cases} \tag{2.36}$$

where we are given the analytic function $\mathbf{B}_j\colon \mathbb{R}^m \times \mathbb{R}^{n-1} \to \mathrm{Mat}(m, \mathbb{R})$, $1 \le j \le n-1$, and $\mathbf{c}\colon \mathbb{R}^m \times \mathbb{R}^{n-1} \to \mathbb{R}^m$. We will write $\mathbf{B}_j = \big((b^{k\ell}_j)\big)$ and $\mathbf{c} = (c^1, \dots, c^m)$. Carefully note that we have assumed $\{\mathbf{B}_j\}_{j=1}^{n-1}$ and $\mathbf{c}$ do not depend on x_n> We can always reduce to this situation by introducing if necessary a new component u^{m+1} of the unknown $\mathbf{u}$, with $u^{m+1} \equiv x_n$.

In particular, the components of the system of partial differential equations in (2.36) read

$$u^k_{x_n} = \sum_{j=1}^{n-1}\sum_{\ell=1}^{m} b^{k\ell}_j(\mathbf{u}, x')u^\ell_{x_j} + c^k(\mathbf{u}, x'), \qquad 1 \le k \le m. \tag{2.37}$$

Having reduced to the special form (2.36), we can now expand $\mathbf{u}$ into a power series near 0.

Theorem 2.15 (Cauchy–Kovalevskaya theorem)**.** *Assume* $\{B_j\}_{j=1}^{n-1}$ *and* $\mathbf{c}$ *are real analytic functions. Then there exists* $r > 0$ *and a real analytic function*

$$\mathbf{u} = \sum_{\alpha} \mathbf{u}_\alpha x^\alpha \tag{2.38}$$

solving the boundary-value problem (2.36)*.*

2.4 Second-order PDEs

We now turn our attention to second-order partial differential equations. For this section, we will mainly refer to [5, §3] starting with a minor detour to [2, §7.2.5].

Equations in two variables

In this section, we consider second-order hyperbolic PDEs involving only two variables. The rough idea is that since a function in two variables has only three second order partial derivatives, algebraic and analytic simplifications in the structure of the PDE may be possible, which are unavailable for more than two variables.

We begin by considering a general linear second-order PDE in two variables

$$\sum_{i,j=1}^{2} a^{ij} u_{x_i x_j} + \sum_{i=1}^{2} b^i u_{x_i} + cu = 0, \tag{2.39}$$

where the coefficients a^{ij}, b^i, c, $1 \leq i, j \leq 2$, with $a^{ij} = a^{ji}$, and the unknown u are functions of the two variables x_1 and x_2 in some region $U \subset \mathbb{R}^2$. Note that for the moment, and in contrast to the theory developed above, we do *not* identify ether x_1 or x_2 wit the variable t denoting time.

We now pose the following question:

> *Is it possible to simplify the structure of the PDE* (2.39) *by introducing new independent variables?*

In other words, can we expect to turn the PDE into a nicer form by rewriting in terms of new variables $y = \mathbf{\Phi}(x)$?

More precisely, set

$$\begin{aligned} y_1 &= \Phi^1(x_1, x_2) \\ y_2 &= \Phi^2(x_1, x_2) \end{aligned} \tag{2.40}$$

2.5 Second-order linear PDEs

That last section was a bit useless, now we'll get down to developing some useful theory.

Characteristics for linear and quasilinear second-order equations

We start with the general quasilinear second-order equation for a function with $u(x, y)$

$$au_{xx} + 2bu_{xy} + cu_{yy} - d = 0, \tag{2.41}$$

where a, b, c, and d depend on x, y, u, u_x, u_y. Here the *Cauchy problem* consists of finding a solution u of (2.41) with given (compatible) values of u, u_x, u_y on a curve Γ in the xy-plane. Thus, for Γ given parametrically by

$$x = f(s), \quad y = g(s), \tag{2.42}$$

we prescribe on Γ

$$u = h(s), u_x \quad = \varphi(s), u_y \quad = \psi(s). \tag{2.43}$$

The values of any function $v(x, y)$ and of its first derivative $v_x(v, y, v_y(x, y))$ along the curve Γ are connected by the compatibility condition

$$\dot{v} = v_x f'(s) + v_y g'(s)$$

which follows by differentiating $v(f(s), g(s))$ with respect to s.

Chapter 3

Algebraic Geometry

A summary to a course on an introduction to sheaf cohomology. We will mostly reference Donu's notes available here `https://www.math.purdue.edu/~dvb/classroom.html`, but also cite Ravi Vakil's *Fundamentals of Algebraic Geometry* [?] available here `https://math216.wordpress.com/`.

3.1 The statement of de Rham's theorem

These are almost verbatim Arapura's notes on the de Rham Complex and cohomology.

Before doing anything fancy, let's start at the beginning. Let $U \subseteq \mathbb{R}^3$ be an open set. In calculus class, we learn about operations

$$\{\,\text{functions}\,\} \xrightarrow{\nabla} \{\,\text{vector fields}\,\} \xrightarrow{\nabla\times} \{\,\text{vector fields}\,\} \xrightarrow{\nabla\cdot} \{\,\text{functions}\,\}$$

such that $(\nabla\times)(\nabla) = 0$ and $(\nabla\cdot)(\nabla\times) = 0$. This is a prototype for a *complex.* An obvious question: does $\nabla \times v = 0$ imply that v is a gradient? Answer: sometimes yes (e.g. if $U = \mathbb{R}^3$) and sometimes no (e.g. if $U = \mathbb{R}^3$ minus a line). To quantify the failure we introduce the first de Rham cohomology

$$H^1_{\mathrm{dR}}(U) = \frac{\{\, v \text{ a vector field on } U : \nabla \times v = 0 \,\}}{\{\, \nabla f \,\}}.$$

Contrary to first appearances, for reasonable U this is finite dimensional and computable. This follows from the de Rham's theorem, which we now explain. First, let's generalize this to an open set $U \subset \mathbb{R}^n$. Once $n > 3$ vector calculus is useless, but there is a good replacement. A differential form of degree p, or p-form, is an expression

$$\alpha = \sum f_{i_1,\ldots,i_p}(x_1,\ldots,x_n)\, dx_{i_1} \wedge \cdots \wedge dx_{i_p}$$

such that the x_i are coordinates, the f are C^∞ functions, $dx_{i_1} \wedge \cdots \wedge dx_{i_p}$ are symbols where $\wedge$ is an anticommutative product. Let $\mathcal{E}^p(U)$ denote the vector space of p-forms. Define the exterior derivative by

$$d\alpha = \sum \sum_j \frac{\partial f_{i_1,\ldots,i_p}}{\partial x_j}\, dx_j \wedge \cdots \wedge dx_{i_p}.$$

This is a $(p+1)$-form.

Lemma 3.1. $d^2 = 0$.

Proof. We prove it for $p = 0$. In this case, we have

$$\begin{aligned} df &= \sum_i \frac{\partial f}{\partial x_i}\, dx_i \\ d(df) &= \sum_{i,j} \sum \frac{\partial^2}{\partial x_j \partial x_i}\, dx_j \wedge dx_i. \end{aligned}$$

Using anticommutativity, we can rewrite this as

$$\sum_{j<i} \left(\frac{\partial^2 f}{\partial x_j \partial x_i} - \frac{\partial^2 f}{\partial x_i \partial x_j} \right) dx_j \wedge dx_i = 0.$$

∎

A cochain complex is a collection of Abelian groups M^i and homomorphisms $d\colon M^i \to M^{i+1}$ such that $d^2 = 0$. We define the p^{th} cohomology of this by

$$H^p_{\mathrm{dR}}(M^\bullet, d) = \frac{\operatorname{Ker} d\colon M^p \to M^{p+1}}{\operatorname{Im} d\colon M^{p-1} \to M^p}.$$

So we have an example of a complex $(\mathcal{E}^\bullet(U), d)$ called the de Rham complex of U. It's cohomology is the de Rham cohomology $H^p_{\mathrm{dR}}(U) = H^p\big(\mathcal{E}^\bullet(U), d\big)$. Here is a basic computation.

Theorem 3.2 (Poincaré's lemma)**.**

$$H^p_{\mathrm{dR}}(\mathbb{R}^n) = \begin{cases} \mathbb{R} & \text{if } p = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. We show this for $n \leq 2$. We first treat the case $n = 1$. Clearly $H^p_{\mathrm{dR}}(\mathbb{R})$ consists of constant functions. If $\alpha = f(x)\, dx$, then

$$d\left(\int_0^x f(t)\, dt \right) = \alpha.$$

There are no p-forms for $p > 1$.

Next, we treat $n = 2$ which contains all of the ideas of the general case. Let x, y be coordinates. We define some operators

$$\begin{tikzcd} \mathcal{E}^\bullet(\mathbb{R}^2) \arrow[r, bend left, "s*"] & \mathcal{E}^\bullet(\mathbb{R}), \arrow[l, bend left, "\pi^*"] \end{tikzcd}$$

where π^* is the pullback along the projection $\mathbb{R}^2 \to \mathbb{R}$. It takes a form in x and treats it as a form in x, y. The pullback along the zero section s^* sets y and dy to zero. Note that $s^* \circ \pi^*$ is the identity. Although $\pi^* \circ s^*$ is not the identity, we will show that it induces the identity on cohomology. This will show that $H^*_{\mathrm{dR}}(\mathbb{R}^2) \cong H^*_{\mathrm{dR}}(\mathbb{R})$, which is all we need. This involves a new concept. We introduce an operator $H\colon \mathcal{E}^p(\mathbb{R}^2) \to \mathcal{E}^{p-1}(\mathbb{R}^2)$ of degree -1 called a *homotopy*. It integrates y as follows:

$$\begin{aligned} H\big(f(x,y)\big) &= 0 \\ H\big(f(x,y)\,dx\big) &= 0 \\ H\big(f(x,y)\,dy\big) &= \int_0^y f(x,t)\,dt \\ H\big(f(x,y)\,dx \wedge dy\big) &= \left[\int_0^y f(x,t)\,dt\right] dx. \end{aligned}$$

A computation using nothing more than the fundamental theorem of calculus shows that

$$1 - \pi^* s^* = \pm(Hd - dH).$$

This implies that the left side induces 0 on $H^*_{\mathrm{dR}}(\mathbb{R}^2)$, or equivalently $\pi^* \circ s^*$ acts like the identity on cohomology. ∎

Before describing de Rham's theorem, we have to say what's happening at the othe rend. The standard n dimensional simplex, or n-simplex, $\Delta^n \subset \mathbb{R}^{n+1}$ is the convex hull of the unit vectors $(1,0,\ldots,0), (0,1,0,\ldots,0), \ldots$. The convex hull of the subset of these is called a face. This is homeomorphic to a simplex of smaller dimension. Omitting all but the i^{th} vertex is called the i^{th} face of Δ^n. We have a standard homeomorphism

$$\delta_i\colon \Delta^{n-1} \to i^{\text{th}} \text{ face of } \Delta^n.$$

A geometric simplicial complex is given by a collection of simplices glued along faces. Historically, the first cohomology theory was defined for simplicial complexes. A bit later singular cohomology was developed, which is an arbitrary topological space X. A (real/complex) singular p-cochain α is an integer (real/complex) valued function on the set of all continuous maps $f\colon \Delta^p \to X$. It might help to think of $\alpha(f)$ as a combinatorial integral $\int_f \alpha$. Let $S^p(X)$ ($S^p(X,\mathbb{R})$, $S^p(X,\mathbb{C})$) denote the group of these cochains. Define $\delta\colon S^p(X) \to S^{p+1}(X)$ by

$$\delta(\alpha)(f) = \sum (-1)^i \alpha(f \circ \delta_i).$$

Lemma 3.3.

$$\delta^2 = 0.$$

For $p = 0$. Let $\alpha \in S^0$. Fix $f\colon \Delta^2 \to X$. Label the restriction of f to the vertices by $0, 1, 2$ and faces $01, 02, 12$. Then

$$\begin{aligned} \delta^2(f) &= \delta\alpha(12) - \delta\alpha(02) + \delta\alpha(01) \\ &= \alpha(1) - \alpha(2) - \alpha(0) + \alpha(2) + \alpha(0) - \alpha(1) \\ &= 0. \end{aligned}$$

∎

Thus we have a complex. Singular cohomology is defined by $H^p(X,\mathbb{Z}) = H^p(S^\bullet(X), \delta)$, and similarly for real or complex valued singular cohomology. These groups are highly computable.

Theorem 3.4 (de Rham)**.** *If $X \subset \mathbb{R}^n$ is open, or more generally a manifold, then $H^p_{\mathrm{dR}}(X, \mathbb{R}) \cong H^p(X, \mathbb{R})$ for all p.*

We will give a proof of this later on as an easy application of sheaf theory. Sheaf methods will help obtain parallel theorems.

Theorem 3.5 (Holomorphic de Rham)**.** *If $X \subset \mathbb{C}^n$ is a complex manifold, then $H^p(X, \mathbb{C})$ can be computed using algebraic differential forms.*

The last theorem is due to Grothendieck. The proof is a lot harder, so we'll try to give the proof by the end of the semester, but there's no guarantee.

3.2 A crash course in homological algebra

By the 1940s techniques from algebraic topology began to be applied to pure algebra, giving rise to a new subject. To begin with, recall that a category $\mathscr{C}$ consists of a set or class of objects (e.g., sets, groups, topological spaces) and morphisms (e.g., functions, homomorphisms, continuous maps) between pairs of objects $\mathrm{Hom}_{\mathscr{C}}(A, B)$. We require an identity $\mathrm{id}_A \in \mathrm{Hom}(A, A)$ for each object A, and associative composition law.

In this section, we will focus on one particular example. Let R be an associative (but possibly noncommutative ring) with identity 1, and let R-**Mod** be the category of left R-modules and homomorphisms. We write $\mathrm{Hom}_R(\,\cdot\,,\,\cdot\,)$ for the morphisms. It is worth noting that $\mathbb{Z}$-**Mod** is the category of Abelian groups. These categories have the following features:

1. $\mathrm{Hom}_R(\,\cdot\,,\,\cdot\,)$ is an Abelian group, and composition is distributive.
2. There is a zero object 0 such that $\mathrm{Hom}_R(0, M) = \mathrm{Hom}_R(M, 0) = 0$.
3. Every pair of objects A, B has a direct sum $A \oplus b$ characterized by certain universal properties.
4. Morphisms have kernels and images, characterized by the appropriate universal properties.

We will encounter other categories satisfying these conditions later on. Such categories are called Abelian. We have been a bit vague about the precise axioms; se Weibel's Homological Algebra for this.

Diagram Chasing

A sequence

$$A \xrightarrow{f} B \xrightarrow{g} C$$

is called *exact* if $\operatorname{Ker} g = \operatorname{Im} f$. A useful skill in this business is to be able to prove things by diagram chasing.

Exercise 3.1. Given a commutative diagram with exact rows

$$\begin{tikzcd}
0 \arrow[r] & A \arrow[r] \arrow[d, "f"] & B \arrow[r] \arrow[d, "g"] & C \arrow[r] \arrow[d, "h"] & 0 \\
0 \arrow[r] & A' \arrow[r] & B' \arrow[r] & C' \arrow[r] & 0,
\end{tikzcd}$$

show that g is an isomorphism if f and h are isomorphisms.

Solution. ■

Theorem 3.6 (Snake lemma)**.** *If*

$$\begin{tikzcd} 0 \arrow[r] & A \arrow[r] \arrow[d, "f"] & B \arrow[r] \arrow[d, "g"] & C \arrow[r] \arrow[d, "h"] & 0 \\ 0 \arrow[r] & A' \arrow[r] & B' \arrow[r] & C' \end{tikzcd}$$

is a commutative diagram with exact rows, then there is an exact sequence

$$0 \longrightarrow \operatorname{Ker} f \longrightarrow \operatorname{Ker} g \longrightarrow \operatorname{Ker} h \xrightarrow{\partial} \operatorname{Coker} f \longrightarrow \operatorname{Coker} g \longrightarrow \operatorname{Coker} h.$$

3.3 Hom Functors

A *(covariant) functor* F from one category to another is a function taking objects to objects and morphisms to morphisms such that if $f\colon A \to B$ then $F(f)\colon F(A) \to F(B)$, $F(\mathrm{id}_A) = \mathrm{id}_{F(A)}$, and $F(f \circ g) = F(f) \circ F(g)$. A contravariant functor reverses direction in the sense that $F(f)\colon F(B) \to F(A)$, $F(\mathrm{id}_A) = \mathrm{id}_{F(A)}$, and $F(f \circ g) = F(g) \circ F(f)$. Here are two basic examples: If $M \in R$-**Mod**, then $F(\,\cdot\,) = \mathrm{Hom}_R(M, \,\cdot\,)$ is a covariant functor from R-**Mod** to $\mathbb{Z}$-**Mod**.

$$\begin{tikzcd} A \arrow[r, "f"] & B \\ M \arrow[u, "g"] \arrow[ur, dashed, "F(f)=f\circ g"'] & \end{tikzcd}$$

When R is commutative, $F(\,\cdot\,)$ is naturally an R-module, but not otherwise. Similarly, $\mathrm{Hom}_R(\,\cdot\,, M)$ is a contravariant functor from R-**Mod** to $\mathbb{Z}$-**Mod** (or R-**Mod**) when R is commutative).

Lemma 3.7. *Suppose that*

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

is exact. Then

(a)
$$0 \longrightarrow \mathrm{Hom}(M, A) \longrightarrow \mathrm{Hom}(M, B) \longrightarrow \mathrm{Hom}(M, C),$$

(b)
$$0 \longrightarrow \mathrm{Hom}(C, M) \longrightarrow \mathrm{Hom}(B, M) \longrightarrow \mathrm{Hom}(A, M)$$

are both exact.

The proof is straight forward and will be omitted.

Exercise 3.2. Prove the lemma.

Exercise 3.3. Prove that

$$0 \longrightarrow \operatorname{Hom}(M, A) \longrightarrow \operatorname{Hom}(M, B) \longrightarrow \operatorname{Hom}(M, C),$$

and

$$0 \longrightarrow \operatorname{Hom}(C, M) \longrightarrow \operatorname{Hom}(B, M) \longrightarrow \operatorname{Hom}(A, M)$$

are exact when the sequence

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

is split exact. This means that there exists a map $s \colon C \to B$, called a splitting, such that $p \circ s = \mathrm{id}_C$.

A (contravariant) functor is called exact if it preserves exact sequences. The lemma says that the Hom functors have the weaker property left exactness. They are not exact, in general:

Example 3.8. Let $R = \mathbb{Z}$, $M = \mathbb{Z}/2$. Note that $\operatorname{Hom}(M, \mathbb{Z}) = 0$ and $\operatorname{Hom}(M, M) = \mathbb{Z}/2$. So $\operatorname{Hom}(M, \,\cdot\,)$ applied to

$$0 \longrightarrow \mathbb{Z} \xrightarrow{2} \mathbb{Z} \longrightarrow \mathbb{Z}/2 \longrightarrow 0$$

yields the sequence

$$0 \longrightarrow 0 \longrightarrow 0 \longrightarrow \mathbb{Z}/2.$$

The last map is certainly not onto.

Exercise 3.4. Find an example for which $\operatorname{Hom}(\,\cdot\,, M)$ isn't exact.

Lemma 3.9. *If M is a free module, then* $\operatorname{Hom}(M, \,\cdot\,)$ *is exact.*

Proof. Let $M = \bigoplus_S R$, where S might be infinite. Given $f \colon B \to C$ surjective, we have

$$\begin{tikzcd} \operatorname{Hom}(M, B) \arrow[r, "f"] \arrow[d, "\cong"'] & \operatorname{Hom}(M, C) \arrow[d, "\cong"] \\ \prod_S B \arrow[r, "\prod f"] & \prod_S C. \end{tikzcd}$$

The horizontal map on the bottom is clearly surjective. ∎

Given a module M, let

$$R^{(M)} = \bigoplus_{m \in M} R.$$

This is a very big free module which maps onto M by sending the 1 in the m^{th} copy of R to m. Let $\operatorname{Ker} M$ be the kernel. We have a *canonical* exact sequence

$$0 \longrightarrow \operatorname{Ker} M \longrightarrow R^{(M)} \longrightarrow M \longrightarrow 0. \tag{3.1}$$

Ext

Inductively, define

$$\begin{aligned}\operatorname{Ext}^1_R(M,N) &= \operatorname{Coker}\big\{\operatorname{Hom}(R^{(M)},N) \to \operatorname{Hom}(\operatorname{Ker} M,N)\big\},\\ \operatorname{Ext}^{i+1}_R(M,N) &= \operatorname{Ext}^i_R(\operatorname{Ker} M,N).\end{aligned}$$

This is not the way these groups are usually defined, but we will get to that later. These are clearly covariant functors in the second variable.

Theorem 3.10. $\operatorname{Ext}^i(\,\cdot\,,N)$ *is a covariant functor in the first variable. Given a short exact sequence*

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

we have an infinite long exact sequence

$$0 \longrightarrow \operatorname{Hom}(C,N) \longrightarrow \operatorname{Hom}(B,N) \longrightarrow \operatorname{Hom}(N,A) \longrightarrow \operatorname{Ext}^1(C,N) \longrightarrow \operatorname{Ext}^1(B,N) \longrightarrow \cdots .$$

Proof. We prove the first part about the theorem. The second part is similar. One constructs a commutative diagram with exact rows and columns

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & \operatorname{Ker} A \arrow[r] \arrow[d] & \operatorname{Ker} B \arrow[r] \arrow[d] & \operatorname{Ker} C \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & R^{(A)} \arrow[r] \arrow[d] & R^{(B)} \arrow[r] \arrow[d] & R^{(C)} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & A \arrow[r] \arrow[d] & B \arrow[r] \arrow[d] & C \arrow[r] \arrow[d] & 0 \\
 & 0 & 0 & 0. &
\end{tikzcd}$$

Note that the middle column is split exact. Hom the top two rows into N to get

$$\begin{tikzcd}
0 \arrow[r] & \operatorname{Hom}(R^{(C)},N) \arrow[r] \arrow[d] & \operatorname{Hom}(R^{(B)},N) \arrow[r] \arrow[d] & \operatorname{Hom}(R^{(A)},N) \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \operatorname{Hom}(\operatorname{Ker} C,N) \arrow[r] & \operatorname{Hom}(\operatorname{Ker} B,N) \arrow[r] & \operatorname{Hom}(\operatorname{Ker} A,N). &
\end{tikzcd}$$

■

We use the split exactness to see that the top row is exact. Now the snake lemma gives the first 6 terms of the exact sequence. Applying this to

$$0 \longrightarrow \operatorname{Ker} A \longrightarrow \operatorname{Ker} B \longrightarrow \operatorname{Ker} C \longrightarrow 0$$

yields

$$\ldots \longrightarrow \mathrm{Hom}(\mathrm{Ker}\, C, N) \xrightarrow{\delta} \mathrm{Ext}^2(A, N) \longrightarrow \ldots .$$

We need to show that this map factors through $\mathrm{Ext}^1(C, N)$. To see this, use the fact that z below is zero because b is surjective.

$$\begin{tikzcd}
& \mathrm{Hom}(R^{(B)}, N) \arrow[r, "b"] & \mathrm{Hom}(R^{(C)}, N) \arrow[d] \arrow[dr, "z"] & \\
\ldots \arrow[rr] & & \mathrm{Hom}(\mathrm{Ker}\, C, N) \arrow[r, "\delta"] \arrow[d] & \mathrm{Ext}^2(A, N) \\
& & \mathrm{Ext}^1(C, N) \arrow[ur, dashed] &
\end{tikzcd}$$

It follows that the original 6 terms sequence can be continued to a 9 term sequence. This can be continued indefinitely.

Lemma 3.11. *If M is free then* $\mathrm{Ext}^1(M, N) = 0$ *for any N.*

Proof. If M is free, then we can choose a basis m_i. Let $s\colon M \to R^{(M)}$ be the homomorphism which sends m_i to 1 in the ${m_i}^{\text{th}}$ copy of R. This gives a splitting of (3.1). It follows that

$$\mathrm{Hom}(R^{(M)}, N) \longrightarrow \mathrm{Hom}(\mathrm{Ker}\, M, N)$$

is surjective. Therefore Ext^1 vanishes. ■

Lemma 3.12. *If*

$$0 \longrightarrow K \longrightarrow F \longrightarrow M \longrightarrow 0$$

is exact, then

$$\mathrm{Ext}^1(M, N) \cong \mathrm{Coker}\big\{ \mathrm{Hom}(F, N) \to \mathrm{Hom}(K, N) \big\}.$$

Proof. This follows from the previous lemma and theorem 3.10.

We have a free resolution

$$\ldots \longrightarrow F_2 \longrightarrow F_1 \longrightarrow K_0 \longrightarrow 0.$$

Denote the "tail" by $F_{\geq 1}$. The previous case implies that

$$\mathrm{Ext}^2(M, N) \cong \mathrm{Ext}^1(K_0, N) \cong H^1(\mathrm{Hom}(F_{\geq 1}, N)) = H^2(\mathrm{Hom}(F_\bullet, N)).$$

This implies the theorem for $i = 2$, etc. ■

Notethat in the usual approach, it is not a priori clear that $H^i(\mathrm{Hom}(F_\bullet, N))$ is well defined. This would require proof. Here is an example which shows the utility of this description.

Exercise 3.5. Let $R = k[x, y]$ and $M = R/(x, y) \cong k$. Construct a free resolution of M (which is a special case of the Koszul complex)

$$0 \longrightarrow R \xrightarrow{\left[\begin{smallmatrix} -y \\ x \end{smallmatrix}\right]} R^2 \xrightarrow{\left[\begin{smallmatrix} x & y \end{smallmatrix}\right]} M \longrightarrow 0$$

where the maps are given by the indicated matrices. Using this, we can see that

$$\operatorname{Ext}^i(M, M) = \begin{cases} k^2 & \text{if } i = 1, \\ k & \text{if } i = 2, \\ 0 & \text{if } i > 2. \end{cases}$$

This description can also be used

Chapter 4

Algebraic Topology

From my meetings with Mark. We reference Hatcher's *Algebraic Topology* [**?**] freely available here `https://www.math.cornell.edu/~hatcher/#ATI`.

4.1 Cohomology

Let's look at some examples to get an idea of what cohomolgy is all about. Take the simplest case: Let X be a 1-dimensional Δ-complex, i.e., an oriented graph. For a fixed abelian group G, the set of all functions from vertices of X to G forms an abelian group, which we denote by $\Delta^0(X;G)$ in the natural sense, i.e., by point-wise addition. Similarly the set of all functions assigning an element of G to each edge of X forms an abelian group $\Delta^1(X;G)$. We are concerned about homomorphisms $\delta\colon \Delta^0(X;G) \to \Delta^1(X;G)$ sending $\varphi \in \Delta^0$ to the function $\delta\varphi \in \Delta^1(X;G)$ whose value on an oriented edge $[v_0, v_1]$ is the difference $\varphi(v_1) - \varphi(v_0)$. For example, X here might be the graph formed by a system of trails on a mountain, with vertices at the junctions between trails. The function φ could assign to each junction its elevation above sea level, in which case $\delta\varphi$ would measure the net change in elevation along the trail from one junction to the next.

Regarding the map $\delta\colon \Delta^0(X;G) \to \Delta^1(X;G)$ as a chain complex with 0s before and after the two terms, the homology of groups of this chain complex are by definition the simplicial cohomolgy groups of X, namely $H^0(X;G) = \operatorname{Ker}\delta \subset \Delta^0(X;G)$ and $H^1(X;G) = \Delta^1(X;G)/\operatorname{Im}\delta = \operatorname{Coker}\delta$. For simplicity we are using here the same notation as will be used for singular cohomolgy; we later prove that for Δ-complexes, the two theories in fact coincide.

The group $H^0(X;G)$ is easy to describe explicitly. A function $\varphi \in \Delta^0(X;G)$ has $\delta\varphi = 0$ if and only if φ takes the same value at both ends of each edge of X. This is equivalent to saying that φ is constant on each component of X. So $H^0(X;G)$ is the group of all functions from the set of components of X to G. This is a direct product of copies of G, one for each component of X.

The cohomology group $H^1(X : G) = \Delta^1(X;G)/\operatorname{Im}\delta$ will be trivial if and only if $\delta\varphi = \psi$ has a solution $\varphi \in \Delta^0(X;G)$ for each $\varphi \in \Delta^1(X;G)$. Solving this equation means deciding whether specifying the change in φ across each edge of X determines an actual function $\varphi \in \Delta^0(X;G)$. This is rather like the calculus problem of finding a function having a specified derivative, with the difference operator δ playing the role of differentiation. As in calculus, if a solution of $\delta\varphi = \psi$ exists, it will be unique up to adding an element of the kernel of δ, i.e., a f unction constant on each component of X.

The equation $\delta\varphi = \psi$ is always solvable if X is a tree since if we choose arbitrarily a value for φ at a base point vertex v_0, then if the change in φ across each edge of X is specified, this uniquely determines the value of φ at ever other vertex v by induction along the unique path from v_0 to v in a tree. Then, since every vertex lies in one of these maximal trees, the values of ψ on the edges of the maximal trees determine φ uniquely up to a constant on each component of X. But in order for the equation $\delta\varphi = \psi$ to hold, the value of ψ on each edge is not in any of the maximal trees must equal the difference in the already-determined values of φ at the two ends of the edge. This condition need not be satisfied since ψ can have arbitrary values on these edges. Thus we see that the cohomology group $H^1(X;G)$ is a direct product of copies of the group G, one copy for each edge of X not in one of the chosen maximal trees. This can be compared with the homology group $H_1(X;G)$ which consists of a direct *sum* of copies of G, one for each edge of X not in one of the maximal trees. Note that the relation between $H^1(X;G)$ and $H^1(X;G)$ is the same as the relation between $H^0(X;G)$ and $H_0(X;G)$, with $H^0(X;G)$ being a direct product of copies of G and $H_0(X;G)$ a direct sum, with one copy for each component of X in either case.

Now let us move up a dimensino, taking X to be a 2-dimensional Δ-complex. Define $\Delta^0(X;G)$ and $\Delta^1(X;G)$ as before, as functions from vertices and edges of X to be Abelian group G, and define $\Delta^2(X;G)$ to be functions from 2-simplices of X to G, and define $\Delta^2(X;G)$ to be functions from 2-simplices of X to G. A homomorphism $\delta\colon \Delta^1(X;G) \to \Delta^2(X;G)$ is defined by $\delta\psi([v_0,v_1,v_2]) = \psi([v_0,v_1]) + \psi([v_1,v_2]) - \psi([v_0,v_2])$, a signed sum of values of ψ on the three edges in the boundary of $[v_0,v_1,v_2]$, just as $\delta\varphi([v_0,v_1])$ for $\varphi \in \Delta^0(X;G)$ we have $\delta\delta\varphi = \big(\varphi(v_1) - \varphi(v_0)\big) + \big(\varphi(v_2) - \varphi(v_2)\big) - \big(\varphi(v_2) - \varphi(v_0)\big) = 0$. Extending this chain complex by 0s on each end, the resulting homology groups are by definition the cohomology groups $H^i(X;G)$.

Chapter 5

Riemannian Geometry

Notes compiled from Do Carmo's *Riemannian Geometry* book.

5.1 Differentiable manifolds

Definition 5.1. A *differentiable manifold* of dimension n is a set M and a family of injective mappings $\mathbf{x}_\alpha \colon U_\alpha \subset \mathbb{R}^n \to M$ of open sets U_α of $\mathbb{R}^n$ into M such that:

(1) $\bigcup_\alpha \mathbf{x}_\alpha(U_\alpha) = M$;
(2) for any pair α, β, with $\mathbf{x}_\alpha(U_\alpha) \cap \mathbf{x}_\beta(U_\beta) = W \neq \emptyset$, the sets $\mathbf{x}_\alpha^{-1}(W)$ and $\mathbf{x}_\beta^{-1}(W)$ are open sets in $\mathbb{R}^n$ and the mappings $\mathbf{x}_\beta^{-1} \circ \mathbf{x}_\alpha$ are differentiable;
(3) the family $\{(U_\alpha, \mathbf{x}_\alpha)\}$ is maximal relative to the conditions (1) and (2).

The pair $(U_\alpha, \mathbf{x}_\alpha)$ with $p \in \mathbf{x}_\alpha(U_\alpha)$ is called a *parametrization* (or *system of coordinates*) of M at p; $\mathbf{x}_\alpha(U_\alpha)$ is then called a *coordinate neighborhood* at p. A family $\{(U_\alpha, \mathbf{x}_\alpha)\}$ satisfying (1) and (2) is called a *differentiable structure* on M.

Remark 5.2. A differentiable structure on a set M induces a natural topology on M. It suffices to define $A \subset M$ to be an *open set* in M if and only if $\mathbf{x}_\alpha^{-1}(A \cap \mathbf{x}_\alpha(U_\alpha))$ is an open set in $\mathbb{R}^n$ for all α. Observe that the topology is defined in such a way that the sets $\mathbf{x}_\alpha(U_\alpha)$ are open and that the mappings $\mathbf{x}_\alpha$ are continuous.

Example 5.3. The *real projective space* $\mathbb{R}\mathrm{P}^n$. Let us denote by $\mathbb{R}\mathrm{P}$ the set of straight lines of $\mathbb{R}^{n+1}$ which pass through the origin $\mathbf{0} = (0, \dots, 0) \in \mathbb{R}^{n+1}$; that is, $\mathbb{R}\mathrm{P}^n$ is the set of "directions" of $\mathbb{R}^{n+1}$.

Let us introduce a differentiable structure on $\mathbb{R}\mathrm{P}^n$. For this, let $(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1}$ and observe, to begin with, that $\mathbb{R}\mathrm{P}^n$ is the quotient space of $\mathbb{R}^{n+1} \setminus \{\mathbf{0}\}$ by the equivalence relation:

$$(x_1, \dots, x_{n+1}) \sim \lambda(x_1, \dots, x_{n+1}), \qquad (\lambda \neq 0).$$

The points of $\mathbb{R}\mathrm{P}^n$ will be denoted by $[x_1, \dots, x_{n+1}]$. Observe that, if $x_i \neq 0$,

$$[x_1, \dots, x_{n+1}] = \left[\frac{x_1}{x_i}, \dots, \frac{x_{i-1}}{x_i}, 1, \frac{x_{i+1}}{x_i}, \dots, \frac{x_{n+1}}{x_i}\right].$$

Define subsets $V_1, \dots, V_{n+1}$, of $\mathbb{R}\mathrm{P}^n$, by:

$$V_i = \left\{\, [x_1, \dots, x_{n+1}] : x_i \neq 0 \,\right\}, \qquad i = 1, \dots, n+1.$$

Geometrically, V_i is the set of straight lines in $\mathbb{R}^{n+1}$ which pass through the origin and do not belong to the hyperplane $x_i = 0$. We are now going to show that we can take the V_i as coordinate neighborhoods, where the coordinates on V_i are

$$y_1 = \frac{x_1}{x_i}, \dots, y_{i-1} = \frac{x_{i-1}}{x_i}, y_i = \frac{x_{i+1}}{x_i}, \dots, y_n = \frac{x_{n+1}}{x_i}.$$

For this, we will define mappings $\mathbf{x}_i \colon \mathbb{R}^n \to V_i$ by

$$\mathbf{x}_i(y_1, \dots, y_n) = [y_1, \dots, y_{i-1}, 1, y_{i+1}, \dots, y_n], \qquad (y_1, \dots, y_n) \in \mathbb{R}^n,$$

and will show that the family $\{(\mathbb{R}^n, \mathbf{x}_i)\}$ is a differentiable structure on $\mathbb{R}\mathrm{P}^n$.

Indeed, any mapping $\mathbf{x}_i$ is clearly bijective while $\bigcap \mathbf{x}_i(\mathbb{R}^n) = \mathbb{R}\mathrm{P}^n$. It remains to show that $\mathbf{x}_i^{-1}(V_i \cap V_j)$ is an open set in $\mathbb{R}^n$ and that $\mathbf{x}_j^{-1} \circ \mathbf{x}_i$, $j = 1, \dots, n+1$, is differentiable there. Now, if $i > j$, the points in $\mathbf{x}_i^{-1}(V_i \cap V_j)$ are of the form:

$$\big\{ (y_1, \dots, y_n) \in \mathbb{R}^n : y_j \neq 0 \big\}.$$

Therefore $\mathbf{x}_i^{-1}(V_i \cap V_j)$ is an open set in $\mathbb{R}^n$, and supposing that $i > j$ (the case $i < j$ is similar),

$$\begin{aligned}
\mathbf{x}_j^{-1} \circ \mathbf{x}_i(y_1, \dots, y_n) &= \mathbf{x}_j^{-1}[y_1, \dots, y_{i-1}, 1, y_i, \dots, y_n] \\
&= \mathbf{x}_j^{-1}\left[\frac{y_1}{y_j}, \dots, \frac{y_{j-1}}{y_j}, 1, \frac{y_{j+1}}{y_j}, \dots, \frac{y_{i-1}}{y_j}, \frac{1}{y_j}, \frac{y_i}{y_j}, \dots, \frac{y_n}{y_j}\right] \\
&= \left(\frac{y_1}{y_j}, \dots, \frac{y_{j-1}}{y_j}, \frac{y_{j+1}}{y_j}, \dots, \frac{y_{i-1}}{y_j}, \frac{1}{y_j}, \frac{y_i}{y_j}, \dots, \frac{y_n}{y_j}\right),
\end{aligned}$$

which is clearly differentiable.

Bibliography

[1] A. DasGupta. *Fundamentals of Probability: A First Course*. Springer Texts in Statistics. Springer New York, 2010.

[2] L.C. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 2010.

[3] W. Feller. *An introduction to probability theory and its applications*. Number v. 1 in Wiley mathematical statistics series. Wiley, 1950.

[4] J. Harris. *Algebraic Geometry: A First Course*. Graduate Texts in Mathematics. Springer New York, 2013.

[5] F. John. *Partial Differential Equations*. Applied Mathematical Sciences. Springer New York, 1991.

[6] D. Mumford. *The Red Book of Varieties and Schemes*. Blurb, Incorporated, 2015.

[7] S.M. Ross. *A First Course in Probability*. Pearson Prentice Hall, 2010.

[8] W.A. Strauss. *Partial Differential Equations: An Introduction*. Wiley, 1992.