

# Lecture 1: Introductory Topics

Ethan Ebbighausen

May 12, 2025

## 1 References, Goals

The topics of these notes will be based on the contents of Borthwick's *Introduction to Partial Differential Equations*, Evans' *Partial Differential Equations*, and Sung-Jin Oh's Lecture notes on the subject. My goal is to add information to and contextualize the information presented.

Our learning objectives for this lecture are

- Classify General PDEs in ways that describe their behavior.
- Analyze basic structures in differentiability.
- Recall ODE existence theorems and solution methods.
- Apply the change-of-variables to make sense of surface integrals.
- Simplify vector integrals via the Divergence theorem.

## 2 What are Partial Differential Equations?

Welcome to Partial Differential Equations. Throughout the course, we will call such equations PDEs, and call ordinary differential equations ODEs. What are PDEs? You should recall that ODEs take a form such as

$$\begin{cases} x'(t) = f(t, x) \\ x(t_0) = x_0 \end{cases}$$

We call them *ordinary* because they take an ordinary derivative. Example uses cases involve things like modeling population, concentration of a solution, or even temperature as a function of time. We are constrained to one variable, and this limits physical applicability.

Let's focus on the last case. What if we wanted to measure the temperature of a room (in 3D space) as a function of time. We would establish some function  $u(t, x, y, z)$  where  $t$  is our point in time, starting at some arbitrary time we call  $t = 0$ , and  $x, y, z$  give the coordinates of a spot in our room.

If the room is empty and isolated from anything outside (if we had very well-insulated walls), we should expect the temperature to stay the same. However, in the real world, we have some influence via outside phenomena (heat exchange through the walls). Let's say it's some fixed temperature  $T_0$  outside. Let us also say the room is  $U = [0, 1]^3 \subset \mathbb{R}^3$ . We'll explain notation a bit more later, but call  $\partial U$  the walls of the room. We should have something like  $u(t, x, y, z)|_{\partial U} = T_0$ , since changes in our room aren't likely to change the temperature of the whole region. This is called *boundary condition*.

We should also have some model of how heat changes or spreads in the room, and we should take account for sources of heat in the room like electronics or even us. We do this via some some function of  $u$  and its derivatives, and write the PDE in the form

$$\begin{cases} F(t, x, y, z, u, \partial_t u, \partial_x u, \dots, \partial_t^\alpha \partial_x^\alpha \partial_y^b \partial_z^c u) = 0 \\ u|_{\partial U} = T_0 \end{cases}$$

We'll look into the actual function involved later. Since this is a bit cumbersome, we usually denote all spatial variables by a vector  $x$  and use partials  $\partial_i = \partial_{x^i} = \frac{\partial}{\partial x^i}$ . The *order* of a PDE is the order of the highest derivative appearing in the equation, so if we have

$$F(x, u, \partial_j u, \dots, \partial_{j_1} \dots \partial_{j_m} u) = 0$$

we have order  $m$ .

In general, PDEs are very hard and have no unified method for solution. Hence, mathematicians like to classify PDEs to look at when we might apply methods. First, we have classifications based on the operator involved. For simplicity of notation, we will often use multi-index notation. For  $\alpha = (\alpha_1, \dots, \alpha_n)$ , and  $x = (x^1, \dots, x^n) \in \mathbb{R}^n$ , we set  $x^\alpha = \prod_{i=1}^n (x^i)^{\alpha_i}$  and  $\partial^\alpha = \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n}$ . We call  $|\alpha| = \alpha_1 + \dots + \alpha_n$  the order of  $\alpha$ . By  $D^k$ , we will denote the vector of all  $\partial^\alpha$  for  $|\alpha| = k$ .

#### Behavior as an operator:

*Linear PDEs* are PDEs linear in the input function. In other words, we can write a partial differential operator  $L$  and have  $L(u + v) = Lu + Lv$ ,  $L(cu) = cLu$  for  $c \in \mathbb{R}$ . These are important because, if we have a PDE  $Lu = 0$ , two solutions  $u_1$  and  $u_2$  give a new solution  $c_1 u_1 + c_2 u_2$  by linearity. This is called the property of superposition. For example, the general second-order linear operator is

$$L = - \sum_{i,j=1}^n a_{ij} \partial_i \partial_j + \sum_{j=1}^n b_j \partial_j + c$$

where the  $a_{ij}, b_j, c$  are all functions. We will study this exact operator later.

We also have weaker versions of linearity. A *semilinear* PDE is a PDE linear in its highest order derivative with coefficients that do not depend on  $u$ , so

$$\sum_{|\alpha|=k} a_\alpha(x) \partial^\alpha u + b(D^{k-1}u, \dots, Du, u, x) = 0$$

A *quasilinear* PDE is linear in the highest-order derivative with coefficients depending on at most the lower-order derivatives of  $u$ , so

$$\sum_{|\alpha|=k} a_\alpha(D^{k-1}u, \dots, du, u, x) \partial^\alpha u + b(D^{k-1}u, \dots, Du, u, x) = 0$$

#### Behavior as a PDE:

We also like to classify PDEs by the way the operator looks. Roughly speaking, we form a symbolic polynomial from the PDE operator by replacing the derivatives with variables. For example, given  $\partial_1 \partial_2 + \partial_2$ , we would have a polynomial  $\xi^1 \xi^2 + \xi^2$ . The level sets of the polynomial gives some geometric shape, which actually tells us a lot about PDE solutions.

For example, *elliptic* equations act a lot like a specific equation called the Laplace equation  $-\Delta u = 0$  where  $\Delta = \sum \partial_i^2$ . For the general second-order operator above, these correspond to the cases where the matrix  $[a_{ij}]$  is symmetric with strictly positive eigenvalues.

We will also look at *hyperbolic* equations, which look like the wave equation  $(\partial_t^2 - \Delta)u = 0$ , and *parabolic* equations that look like the heat equation  $(\partial_t - \Delta)u = 0$

## 2.1 Posing Problems, Well-Posedness

As in our example at the beginning, we usually pose the PDE on some sort of domain  $U \subset \mathbb{R}^n$ , with a function  $u : U \rightarrow \mathbb{R}$ . We will often have some sort of information about how  $u$  behaves on the boundary of  $U$ , which we call boundary data or boundary conditions. We focus on two common types. *Dirichlet* boundary

conditions specify the value of  $u$  on the boundary, usually like  $u|_{\partial U} = g$ . *Neumann* boundary conditions specify the normal derivatives of  $u$  at the boundary, in the form  $\partial_\eta u = g$ .

Given some sort of data and a PDE, we care about whether the equation is solvable, whether this is stable in our data, and whether the solution is "nice". We define a concept to encapsulate this idea of a "good problem". A PDE is called *well-posed* if 1.) a solution exists, 2.) the solution is uniquely determined by the data and 3.) the solution depends continuously on the data. The third condition is necessary because our input data is usually flawed in practice, so we want to guarantee that small errors in the data mean only small errors in the solution.

## 2.2 Questions:

- 1.) What type of PDE is  $2x^1 \partial_1^2 u = 3\partial_2^2 u + 5u$ ?
- 2.) What about  $\partial_1^5 u + \partial_1^3 \partial_2^2 u + 3u \partial_2^4 u + 16u^2 = 0$ ?

# 3 A Crash-Course in Analysis

## 3.1 Real Numbers, Euclidean Space

Any reader has likely used the real numbers  $\mathbb{R}$  extensively and will have an intuitive idea of their structure. Strictly speaking, one constructs them from the rational numbers using a process like Dedekind cuts to take a sort of completion. We only need this in as much it means the real numbers and  $\mathbb{R}^n$  are complete, or that limits behave nicely. This manifests itself in  $\mathbb{R}$  by the completeness axiom. Given a set  $A \subset \mathbb{R}$ , an upper bound for  $A$  is a number  $M$  such that for all  $x \in A$ ,  $x \leq M$ . The completeness axiom says that any nonempty subset bounded above has a *supremum*, denoted  $\sup(A)$ , which is the least upper bound in the sense that given any other upper bound  $M$  for  $A$ ,  $\sup(A) \leq M$ . A related concept is the *infimum*, or greatest lower bound. If  $A$  is not bounded above, we may say that  $A$  has an infinite supremum.

We will often use sequences to capture ideas of limits like this. A *sequence*  $\{x_k\}_{k=1}^\infty$  is just a collection of numbers  $x_i \in \mathbb{R}^n$  indexed by the natural numbers  $\mathbb{N} = \{1, 2, \dots\}$ . A sequence is said to converge to a limit  $x$  if, for all  $\epsilon > 0$ , there is some  $N$  such that if  $k > N$ ,  $|x - x_k| < \epsilon$ . In simple terms, the numbers get closer and closer to  $x$  as  $k$  gets large. A function that respects limits is called *continuous*. More rigorously, a function  $f$  is continuous at  $x$  if whenever  $\{x_k\} \rightarrow x$  (converges to  $x$ ),  $\{f(x_k)\} \rightarrow f(x)$ . Equivalently, for all  $\epsilon > 0$  there exists some  $\delta > 0$  so that if  $|x - y| < \delta$ ,  $|f(x) - f(y)| < \epsilon$ . In simple terms,  $f$  takes nearby points to nearby points.

Now, given  $A$  bounded above, we can extract an interesting property. Consider any number  $\epsilon > 0$ . We may find  $x \in A$  so  $x > \sup(A) - \epsilon$ . This is simple to prove by contradiction: if this were not true, then  $\sup(A) - \epsilon \geq x$  for all  $x \in A$ , and so this is an upper bound, which contradicts that the supremum is the least upper bound. If we do this for all the numbers  $\frac{1}{n}$  for  $n \in \mathbb{N}$ , we prove the following

**[3.1]** For a nonempty set  $A \subset \mathbb{R}$ , there exists a sequence of points  $x_k \in A$  such that

$$\lim_{k \rightarrow \infty} x_k = \sup(A)$$

and a similar statement is true for the infimum.

In  $\mathbb{R}^n$ , a supremum doesn't make sense, but we can get a similarly useful concept by actually defining completeness. Recall that a sequence is *Cauchy* if, for any  $\epsilon > 0$  there exists some  $N$  so  $|x_n - x_m| < \epsilon$  when  $n, m > N$ . A space is called *complete* if every Cauchy sequence has a limit in the space. We will not need the details of proving the following:

**[3.2]**  $\mathbb{R}^n$  is a complete metric space.

If you haven't seen metrics before, don't worry. They are a way of measuring distance. In  $\mathbb{R}^n$ , we do this by defining the norm or Euclidean length  $|x| = \sqrt{x \cdot x}$  from the dot product. Then, we define the distance  $d(x, y) = |x - y|$ . In particular,  $|x - y| \leq |x - z| + |z - y|$  is the useful property called the *triangle inequality*

which we will use from this distance (the length of the side of a triangle is less than the sum of the lengths of the other sides). We also have that

$$\lim_{k \rightarrow \infty} x_k = w \Leftrightarrow \lim_{k \rightarrow \infty} |x_k - w| = 0$$

On many occasions, we will use balls in  $\mathbb{R}^n$ , denoted  $B(x, r) = \{y \in \mathbb{R}^n : |y - x| < r\}$ , because they capture the concept of "nearby". If  $x \in A$  has  $B(x, \epsilon) \subset A$  for a small  $\epsilon$ , then  $x$  is called an *interior point* of  $A$  and  $A$  is called a *neighborhood* for  $x$ . A set  $A$  is *open* if it is a neighborhood of all of its points. By contrast, a set  $B$  is called closed if its complement  $B^C = \mathbb{R}^n \setminus B$  (everything outside of  $B$ ) is open. An easier way to look at this in the real numbers is to define the *boundary*  $\partial A$ , which is the set of points  $x$  so every neighborhood intersects both  $A$  and its complement. This is exactly "the edge" in simple cases. For example,  $\partial B(x, r) = \{y : |x - y| = r\}$  is the sphere. A set  $B$  is closed if and only if  $\partial B \subset B$ . For this reason, we also often define the closure of a set  $A$  to be  $\bar{A} = A \cup \partial A$ .

A set  $A$  is *bounded* if  $A \subset B(0, R)$  for some  $R > 0$ . A set which is closed and bounded (in  $\mathbb{R}^n$ ) is called *compact*.

A set  $A \subset \mathbb{R}^n$  is called path connected if any two points in the set may be joined by a continuous path in the set. In other words, for  $a, b \in A$ , there is some  $\sigma : [0, 1] \rightarrow A$  so  $\sigma(0) = a$ ,  $\sigma(1) = b$ , and  $\sigma$  is continuous. Lastly, we call an open, path connected set  $\Omega \subset \mathbb{R}^n$  a *domain*. (Note, we really only care about this set being *connected*, not path connected, but for an open set in  $\mathbb{R}^n$ , the concepts are equivalent)

#### Examples:

- 1.) the set  $(0, 1)$  is open, and has boundary  $\{0, 1\}$ . Then,  $[0, 1]$  is closed (and compact).
- 2.)  $B(x, r)$  is open
- 3.) Any interval  $[a, b]$  is connected, as is any ball.

### 3.2 Questions:

- 1.) What is  $\lim_{n \rightarrow \infty} \frac{n!}{n^n}$ ?
- 2.) Is  $(0, 2) \cup (3, 4)$  connected? Is it complete? Is it bounded?
- 3.) Let  $A = \{(x, \sin(x)) \mid x \in (0, 1]\} \cup \{(0, 0)\}$ . Is this set path connected? Can you find open  $C, D$  disjoint so for  $C_A = C \cap A$ ,  $D_A = D \cap A$ ,  $C_A \cup D_A = A$ ?

### 3.3 Differentiability, ODEs and Solutions

To understand PDEs, we must first understand derivatives and ODEs. Recall that a function  $f : \mathbb{R} \rightarrow \mathbb{R}$  is called differentiable at  $x \in \Omega$  if

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

exists, in which case we denote it  $f'(x)$ . Given a domain  $\Omega \subset \mathbb{R}^n$ ,  $f : \Omega \rightarrow \mathbb{R}$  is differentiable if the limits

$$\lim_{h \rightarrow 0} \frac{f(x+he) - f(x)}{h}$$

exist for all unit vectors  $e$ , in which case we denote this  $\partial_e f$ . Recall that the partial derivatives are simply setting these to be the standard basis vectors. We denote the space of functions which admit continuous partial derivatives up to order  $m$  by  $C^m(\Omega)$ . A *smooth* function has continuous partial derivatives of all orders, and we denote it  $C^\infty(\Omega)$ . We need an open set for this definition because these limits don't make sense at boundary points. Therefore, we define by  $C^m(\bar{\Omega})$  the set of functions which admit  $C^m$  extensions on some open set containing  $\bar{\Omega}$ . We sometimes denote continuous functions by  $C(\Omega)$  or  $C^0(\Omega)$ .

#### Examples:

- 1.) Any polynomial  $p(x) = a_m x^m + \dots + a_1 x + a_0$  is smooth on  $\mathbb{R}$ .
- 2.)  $f(x) = |x|$  is not differentiable on  $\mathbb{R}$ , but it is on  $(1, 2)$ .  $f(x) = \frac{1}{1+x^2}$  is smooth on  $\mathbb{R}$ .
- 3.)  $f(x) = \frac{1}{x}$  is smooth on  $(0, 1)$ , but is not in  $C^m([0, 1])$  for any  $m$

One other tool we will need in the future is the Leibniz Integral Rule, which is a way of differentiating under an integral. The standard, general statement is the following

**[3.3]** Let  $f(x, t)$  be a function such that both  $f$  and  $f_x$  are continuous in  $t$  and  $x$  in some region of the  $x, t$  plane including  $a(x) \leq t \leq b(x)$ ,  $x_0 \leq x \leq x_1$ . Suppose both  $a(x)$  and  $b(x)$  are  $C^1([x_0, x_1])$ . Then,

$$\partial_x \int_{a(x)}^{b(x)} f(x, t) dt = f(x, b(x))b'(x) - f(x, a(x))a'(x) + \int_{a(x)}^{b(x)} \partial_x f(x, t) dt$$

### 3.3.1 Supports and Smooth Bumps

The *support* of a function is the closure of place where it is nonzero,  $\text{supp}(f) = \overline{\{x \in \Omega : f(x) \neq 0\}}$ . We use this to locate where the function behaves meaningfully. We denote by  $C_c^m(\Omega)$  the set of  $C^m$  functions with compact support in  $\Omega$ . We will often use a special type of function with compact support called a smooth bump  $\psi(x)$ . The goal of this function is that, when we have a compact set  $C$  and open subset  $U$  so  $C \subset U \subset \Omega$ , we may find  $\psi \in C_c^\infty$  so  $\psi = 1$  on  $C$  and  $\text{supp}(\psi) \subset U$ .

#### Constructing $\psi$ :

Set  $d(x, C) = \inf\{d(x, y) : y \in C\}$ . Since  $C$  is compact and  $U$  is open, we may use that  $B(C, \epsilon) \subset U$  for some  $\epsilon > 0$  (if you haven't seen compactness before, just take this as given).

We may start with a function  $h : \mathbb{R} \rightarrow \mathbb{R}$  given by

$$h(x) = \begin{cases} e^{-\frac{1}{1-x^2}} & |x| < 1 \\ 0 & |x| \geq 1 \end{cases}$$

This gives one example of a nonzero smooth compactly supported function.

You may check that  $h$  is smooth using L'Hopital's rule. Then,  $\phi(x) = \frac{\int_{-\infty}^x h(t) dt}{\int_{-\infty}^{\infty} h(t) dt}$  is smooth and has  $\phi(x) = 0$  for  $x \leq -1$ ,  $\phi(x) = 1$  for  $x \geq 1$ . We sometimes call this a *smooth transition function*.

Lastly,

$$\psi(x) = \phi\left(1 - \left(\frac{2d(C, x)}{\epsilon}\right)^2\right)$$

has the desired properties

### 3.3.2 Picard-Lindelöf

Along with defining derivatives, we will also commonly use a result from ODEs called the Picard-Lindelöf theorem. We won't prove it fully here (see the ODEs Existence Document on my website for an in-depth treatment and proof). We call a function  $F : \Omega \rightarrow \mathbb{R}$  Lipschitz continuous if there exists some  $M$  so  $|F(x) - F(y)| \leq M|x - y|$ . We also consider the standard form for an ODE:

$$\begin{cases} x'(t) = f(t, x) \\ x(t_0) = x_0 \end{cases}$$

**[3.4]** Picard-Lindelöf: Consider the ODE above, where  $f : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  is continuous in  $t$  and Lipschitz continuous in  $x$ . Then, there exists some  $\epsilon > 0$  admitting a  $C^1$  solution to the ODE  $X : (t_0 - \epsilon, t_0 + \epsilon) \rightarrow \mathbb{R}^n$ , and this is the unique solution to the ODE in this time interval.

### 3.3.3 Solving ODEs

Recall that we only need to consider a first-order ODE since a higher order ODE may be reduced to a first-order system of ODEs. We will use Picard-Lindelöf several times throughout the course to claim that ODE solutions exist. Occasionally, we will also need to solve ODEs. There are a few easy cases to remember. The first is that of separable equations

$$x' = \frac{f(t)}{g(x)}$$

in this case, we may integrate to find  $\int g(x)dx = \int f(t)dt$ . Presuming we may invert  $\int g(x)dx$  (solve for  $x$ , perhaps nonconstructively using something like the inverse function theorem), then we obtain a solution  $x = G^{-1}(F(t))$ . For example,

$$x' = ax$$

has the general family of solutions  $x(t) = x_0 e^{at}$ .

Another common case is to use an integrating factor. An ODE of the form

$$y' + p(t)y = q(t)$$

looks a bit like a chain rule. To realize this, consider that

$$(ye^{a(t)})' = y'e^{a(t)} + a'(t)ye^{a(t)}$$

so setting  $a(t) = \int p(t)dt$  gives

$$(ye^{\int p})' = qe^{\int p}$$

a separable equation we may solve as

$$y(t) = e^{-\int p} \left( \int qe^{\int p} dt \right)$$

Our last ODE case to consider is a specific family of ODEs,

$$x'' = -k^2 x$$

for  $k > 0$ . To solve this, we convert to the system  $y = (x, x')$  and

$$y' = \begin{bmatrix} 0 & 1 \\ -k^2 & 0 \end{bmatrix} y$$

we solve such an equation by diagonalizing the matrix. Its eigenvalues are  $\pm ik$  with eigenvectors  $\begin{bmatrix} 1 \\ \pm ik \end{bmatrix}$ . Hence, we see solutions  $x = c_1 e^{ikt} + c_2 e^{-ikt}$ . Reducing this to real numbers makes life a bit easier, and becomes  $x = a_1 \sin(kt) + a_2 \cos(kt)$ .

## 3.4 Vector Calculus, Divergence Theorem, Greens' Identities

We next treat vector calculus as a way to approach manipulating PDEs. One of the most useful tools will be the Divergence theorem, which gives a way to "integrate by parts". For this, we recall a couple concepts.

First, recall that the *gradient* of  $f \in C^1(\Omega)$  is the vector-valued function  $\nabla f = (\partial_1 f, \dots, \partial_n f)$  (the symbol is called nabla). We define the *divergence* of a vector-valued  $F \in C^1(\Omega; \mathbb{R}^n)$  with components  $v_i$  to be

$$\nabla \cdot v = \sum \partial_i v_i$$

In this language, the Laplacian is  $\Delta u = \nabla \cdot (\nabla u)$ , and is sometimes written as  $\nabla^2$  in differential geometry (where directional derivatives are hard to define).

We also note that there are multi-dimensional versions of the Leibniz rule, often under other names like the Reynolds' Transport Theorem. We shall tend to use the multi-dimensional version

$$\partial_t \int_{\Omega} u(t, x) dx = \int_{\Omega} \partial_t u(t, x) dx$$

where  $u$  and  $u_t$  are assumed continuous on  $(a, b) \times \bar{\Omega}$  and  $\Omega$  is a bounded domain.

### 3.4.1 Boundary Regularity

We need the notion of a "nice" boundary for our domain, in essence so we can differentiate with respect to the normal vector to the boundary. This is a bit more complicated. We say that a domain  $\Omega$  has  $C^k$  boundary if, for every  $x \in \partial\Omega$ , there exists some  $\epsilon > 0$  and some  $C^k$  function  $\sigma : B(0, \delta) \subset \mathbb{R}^{n-1} \rightarrow B(x, \epsilon)$  such that  $\partial\Omega \cap B(x, \epsilon) = \sigma(B(0, \delta))$ . In simpler terms, the boundary is locally the graph of a  $C^k$  function on  $\mathbb{R}^{n-1}$ . The map  $\sigma$  is called a *coordinate parameterization*.

Using some higher theory (the implicit function theorem), it turns out that we can write this same definition in a "backward" way. We may say that the boundary is  $C^k$  at  $x$  if there exists a function  $\gamma : \mathbb{R}^{n-1} \rightarrow \mathbb{R}$  such that, up to relabeling coordinate axes,  $\Omega \cap B(x, \epsilon) = \{y \in B(x, \epsilon) : y^n > \gamma(y^1, \dots, y^{n-1})\}$ . This is a bit weird, but it's a way to "flatten" the boundary by writing it in coordinates  $\Phi(x) = (x^1, \dots, x^{n-1}, x^n - \gamma(x^1, \dots, x^{n-1}))$  (notice that  $\Phi^{-1}|_{x_n=0}$  gives a coordinate parameterization). It also gives us a way to compute a normal vector, since the boundary is now a level surface to  $x^n - \gamma(x)$ , so a *normal vector* is

$$(-\partial_1\gamma, \dots, -\partial_{n-1}\gamma, 1)$$

and we may get an outward unit normal vector by dividing by the magnitude and checking whether this points inward or outward.

**Example:** Consider the simple domain  $\omega = B(0, 1) \subset \mathbb{R}^n$ . Let us consider  $x_0 = (0, \dots, 0, -1)$ . In the neighborhood  $\{x : x^n > 0\}$ , we get a coordinate parameterization

$$\phi(y^1, \dots, y^{n-1}) = (y^1, \dots, y^{n-1}, -\sqrt{1 - \sum_{i=1}^{n-1} (y^i)^2})$$

and we also have  $\gamma(x^1, \dots, x^{n-1}) = -\sqrt{1 - \sum_{i=1}^{n-1} (x^i)^2}$ . Hence, we may find a normal via the gradient

$$\left( \frac{x^1}{\sqrt{1 - \sum_{i=1}^{n-1} (x^i)^2}}, \dots, \frac{x^{n-1}}{\sqrt{1 - \sum_{i=1}^{n-1} (x^i)^2}}, 1 \right)$$

Dividing by the magnitude shows that the normal to the unit sphere at point  $x$  is just the vector  $x$  itself (you should also be able to justify this geometrically)!

### 3.4.2 Surface Integration, Polar Integration

The last piece of theory we need to set up the divergence theorem is the idea of integrating over a surface. The main concept is to write coordinates or a parameterization for the surface like we did for the boundary case above. Since we are interested in these boundaries, we will only consider surfaces of dimension  $n-1$  in  $\mathbb{R}^n$ , but this theory may be easily generalized.

Given some  $C^1$  parameters  $\sigma : V \subset \mathbb{R}^{n-1} \rightarrow U$  where  $U$  is the surface to integrate over, we may consider  $U$  and  $V$  as "the same" for our integration. Integrating over  $V$  using the change of variables tells us how to integrate over  $U$ . Let  $\nu$  be the outward unit normal to  $\sigma(V)$ . If you haven't seen the change-of-variables theorem before, just take the following formula as given

$$\int_{\sigma(V)} f dS = \int_V f(\sigma(w)) |\det(\partial_{w^1}\sigma, \dots, \partial_{w^{n-1}}\sigma, \nu)| dx$$

Recall that the determinant tells us how a linear map affects the volume of a unit cube. Thus, the Jacobian determinant above tells us the instantaneous change in volume caused by applying the transformation  $\sigma$ .

**Example: Polar Integration**

Let us consider the unit sphere  $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$ . Consider  $\omega(y)$  a parameterization of the sphere by coordinates in  $\mathbb{R}^{n-1}$ .

Consider the change-of-variables  $(r, y) \mapsto x = r\omega(y)$ , and the integration formula gives

$$\begin{aligned} dx &= r^{n-1} |\det(\partial_{y^1}\omega, \dots, \partial_{y^{n-1}}\omega, \omega)| dr dy^1 \dots dy^{n-1} \\ &= r^{n-1} dr dS(y) \end{aligned}$$

since the outward unit normal is exactly  $\omega(y)$ .  
for the polar formula

$$\int_{B(0,R)} f dx = \int_{S^{n-1}} \int_0^R f(r\omega(y)) r^{n-1} dr dS(y)$$

### 3.4.3 Divergence Theorem

We now have enough to state (without proof) the divergence theorem of Gauss.

**[3.5]** Suppose  $\Omega \subset \mathbb{R}^n$  is a bounded domain with a piecewise  $C^1$  boundary. For a vector field  $F \in C^1(\overline{\Omega}; \mathbb{R}^n)$ ,

$$\int_{\Omega} \nabla \cdot F dx = \int_{\partial\Omega} F \cdot \nu dS$$

where  $\nu$  is the outward unit normal to  $d\Omega$  and  $dS$  is the surface measure.

**Example:** We consider the unit ball again, specifically  $B(0,1) \subset \mathbb{R}^3$ . We will show the divergence theorem via direct integration here. Since the integral is linear and we may decompose vector fields by coordinates, consider  $F = (0, 0, f)$ , whose divergence is simply  $\partial_3 f$ .

Then, using cylindrical coordinates

$$\begin{aligned} \int_{B(0,1)} \nabla \cdot F dx &= \int_0^{2\pi} \int_0^1 \int_{-\sqrt{1-\rho^2}}^{\sqrt{1-\rho^2}} \partial_3 f \rho dz d\rho d\theta \\ &= \int_0^{2\pi} \int_0^1 [f(\rho \cos(\theta), \rho \sin(\theta), \sqrt{1-\rho^2}) - f(\rho \cos(\theta), \rho \sin(\theta), -\sqrt{1-\rho^2})] \rho d\rho d\theta \end{aligned}$$

Notice that the coordinates of the entries of  $f$  are now on the sphere. As with the above examples, we consider the upper and lower hemisphere separately, and obtain volume element  $dS = \frac{\rho}{\sqrt{1-\rho^2}} d\rho d\theta$ . Then,

$$\begin{aligned} \int_0^{2\pi} \int_0^1 [f(\rho \cos(\theta), \rho \sin(\theta), \sqrt{1-\rho^2}) - f(\rho \cos(\theta), \rho \sin(\theta), -\sqrt{1-\rho^2})] \rho d\rho d\theta \\ = \int_{S_+^2} f \sqrt{1-\rho^2} dS + \int_{S_-^2} f(-\sqrt{1-\rho^2}) dS \\ = \int_{S^2} F \cdot \nu dS \end{aligned}$$

as desired.

### 3.4.4 Greens' Identities

Greens' Identities are special cases of the divergence theorem when we consider the Laplacian. If  $F = \nabla u$ , then  $\nabla \cdot F = \Delta u$ . Further,  $\nabla u \cdot \nu = \partial_{\nu} u$  on the boundary. This means that the divergence theorem may be written in this special case as

$$\int_{\Omega} \Delta u dx = \int_{\partial\Omega} \partial_{\nu} u dS$$

Green extended this in the process of solving some PDEs to consider  $F = v\nabla u$  and applying the product rule:

$$\nabla \cdot (v\nabla u) = \nabla v \cdot \nabla u + v\Delta u$$

which gives the following

**[3.6]** Let  $\Omega \subset \mathbb{R}^n$  be a bounded domain with piecewise  $C^1$  boundary.

1.) For  $u \in C^2(\overline{\Omega})$  and  $v \in C^1(\overline{\Omega})$ ,

$$\int_{\Omega} \nabla v \cdot \nabla u + v\Delta u dx = \int_{\partial\Omega} v\partial_{\nu} u dS$$

2.) If  $u, v \in C^2(\overline{\Omega})$ ,

$$\int_{\Omega} v\Delta u - u\Delta v dx = \int_{\partial\Omega} (v\partial_{\nu} u - u\partial_{\nu} v) dS$$

**Example:** We compute the radial Laplacian, which we shall use to simplify some PDEs later. We consider  $\Omega = B(0, R)$  and a radial function  $h(x) = g(|x|)$ . Then, by integrating in polar coordinates,

$$\int_{B(0,a)} \Delta h dx = |S^{n-1}| \int_0^a \Delta g(r) r^{n-1} dr$$

Using the divergence theorem, the left is

$$\int_{\partial B(0,a)} \partial_{\nu} g dS = \int_{\partial B(0,a)} \partial_r g(a) dS = |S^{n-1}| a^{n-1} \partial_r g(a)$$

such that

$$\int_0^a \Delta g(r) r^{n-1} dr = a^{n-1} \partial_r g(a)$$

To recover the Laplacian, differentiate this formula and obtain

$$a^{n-1} \Delta g(a) = \partial_a [a^{n-1} \partial_r g(a)]$$

or that

$$\Delta g = r^{1-n} \partial_r [r^{n-1} \partial_r] g$$