Shanghai Lectures on Multivariable Analysis

William G. Faris

October 18, 2016

Contents

Preface

The book

This book originated in lectures given in Fall 2014 at NYU Shanghai for an advanced undergraduate course in multivariable analysis. There are chapters on Differentiation, Integration, Differential Forms, The Metric Tensor, together with an optional chapter on Measure Zero. The topics are standard, but the attempt is to present ideas that are often overlooked in this context. The following chapter by chapter summary sketches the approach. The explanations in the summary are far from complete; they are only intended to highlight points that are explained in the body of the book.

Differentiation

This main themes of this chapter are standard.

- It begins with *fixed point iteration*, the basis of the subsequent proofs.
- The central object in this chapter is a smooth (that is, sufficiently differentiable) *numerical function* $\mathbf{f}$ defined on an open subset U of $\mathbf{R}^k$ with values in $\mathbf{R}^n$. Here k is the domain dimension and n is the target dimension. The function is called "numerical" to emphasize that both inputs and outputs involve numbers. (See below for other kinds of functions.) Sometimes a numerical function is denoted by an expression like $\mathbf{y} \mapsto \mathbf{f}(\mathbf{y})$. The choice of the variable name $\mathbf{y}$ is arbitrary.
- When $k < n$ a numerical function $\mathbf{f}$ from an open subset U of $\mathbf{R}^k$ to $\mathbf{R}^n$ can give a *explicit* (parametric) representation of a k-dimensional surface in $\mathbf{R}^n$.
- When $k < n$ a numerical function $\mathbf{g}$ from an open subset W of $\mathbf{R}^n$ to $\mathbf{R}^{n-k}$ can given an *implicit* representation of a family of k-dimensional surfaces in W.
- The *implicit function theorem* gives conditions for when an implicit representation gives rise to an explicit representation.

- When $k = n$ the function $\mathbf{f}$ defines a transformation from an open subset U of $\mathbf{R}^n$ to an open set V in $\mathbf{R}^n$.
- The *inverse function theorem* gives conditions that ensure that the transformation has an inverse transformation.
- In the *passive interpretation* the smooth transformation $\mathbf{f} : U \to V$ has a smooth inverse: the points in U and the points in V give alternative numerical descriptions of the same situation.
- In the *active interpretation* the smooth transformation $\mathbf{f} : U \to U$ describes a change of state: the state described by $\mathbf{y}$ in U is mapped into the new state described by $\mathbf{f}(\mathbf{y})$ in U. The transformation may be iterated.

Integration

This chapter is about the Riemann integral for functions of several variables. There are several interesting results.

- The *Fubini theorem* for Riemann integrals deals with iterated integrals.
- The *dominated convergence theorem* for Riemann integrals is a result about pointwise convergence. The setting is a sequence of Riemann integrable functions defined on a fixed bounded set with a common bound on their values. The sequence of functions is assumed to converge (in some sense) to another Riemann integrable function. It is elementary to prove that if the functions converge uniformly, then the integrals converge. The dominated convergence theorem says that if the functions converge pointwise, then the integrals converge. The direct proof of the dominated convergence involves a somewhat complicated construction, but the result is spectacularly simple and useful.
- There is a treatment of *approximate delta functions*.
- The *change of variables formula* has an elegant proof via approximate delta functions and the dominated convergence theorem.

Differential Forms

This chapter and the next are the heart of the book. The central idea is geometrical: differential forms are intrinsic expressions of change, and thus the basic mechanism of calculation with differential forms is equally simple for every possible choice of coordinate system. (Of course for modeling a particular system one coordinate system may be more convenient than another.) The fundamental result is Stokes' theorem, which is the natural generalization of the fundamental theorem of calculus. Both the fundamental theorem in one dimension and Stokes' theorem in higher dimensions make no reference to notions of length and area; they simply describe the cumulative effect of small changes.

Because of this intrinsic nature, the expression of Stoke's theorem is the same in every coordinate system.

- Example: Here is a simple example from physics that illustrates how natural it is to have a free choice of coordinate system. Consider an idea gas with pressure P, volume V, and temperature T. The number of gas particles N is assumed constant. Each of the quantities P, V, T is a function of the state of the gas; these functions are related by the ideal gas law

$$PV = NkT.$$

(The constant k transforms temperature units to energy units.) The differential form of this relation is

$$P\,dV + V\,dP = Nk\,dT.$$

This equation is a precise description of the how these variables change for a small change in the state of the gas. A typical use of the fundamental theorem of calculus is the calculation of work done by the system during a change of state where the temperature is constant. This is obtained by integrating $-P\,dV$ along states of constant temperature $T = T_0$. Suppose that in this process the volume changes from V_0 to V_1 and the pressure changes from P_0 to P_1. Then

$$-P\,dV = -NkT_0\,dV/V = -NkT_0\,d\log(V)$$

has integral $NkT_0 \log(V_0/V_1)$. But nothing depends on using volume as an independent variable. On the curve where $T = T_0$ there is a relation $P\,dV + V\,dP = 0$. So on this curve the form is also equal to

$$V\,dP = NkT_0\,dP/P = NkT_0\,\log(P)$$

with integral $NkT_0 \log(P_1/P_0)$. Since $P_0V_0 = P_1V_1 = NkT_0$, this is the same result.

- While many applications of differential forms have nothing to do with length or area, others make essential use of these ideas. For instance, the length of a curve in the plane is obtained by integrating

$$\sqrt{(dx/dt)^2 + (dy/dt)^2}\,dt = \sqrt{(dx/du)^2 + (dy/du)^2}\,du.$$

The motion along the curve may be described either by the t coordinate or the u coordinate, but the length of the curve between two points is a number that does not depend on this choice. In practice such integrals involving square roots of sums of squares can be awkward. Sometimes it helps to use a different coordinate system in the plane. For instance, in polar coordinates the same form has the expression $\sqrt{(dr/dt)^2 + r^2(d\theta/dt)^2}\,dt$. The theory of differential forms gives a systematic way of dealing with all such coordinate changes.

- The chapter introduces a fundamental organizing principle of mathematical modeling. It applies in geometry and in most applications of mathematics, and it deserves a serious discussion. In the approach of this book it is based on the elementary concept of an n-dimensional *manifold patch*. This term is used here for a differentiable manifold M modeled on some open subset of $\mathbf{R}^n$. In more detail, M is a set, and there are certain one-to-one functions from M onto open subsets of $\mathbf{R}^n$ called *coordinate systems*. Suppose $\mathbf{u} = (u_1, \ldots, u_n) : M \to U$ is a coordinate system, and suppose that $\mathbf{w} = (w_1, \ldots, w_n) : M \to W$ is another coordinate system, with $\mathbf{u} = \mathbf{g}(\mathbf{w})$. It is required that $\mathbf{g} : W \to U$ be a smooth numerical function with smooth inverse. There is no preferred coordinate system. The modeling process may employ several coordinate systems in the same discussion.

- A *scalar field* is a smooth function $s : M \to \mathbf{R}$ from the manifold patch to the reals. If $\mathbf{u} = (u_1, \ldots, u_n) : U \to \mathbf{R}^n$ is a coordinate system, then $s = f(\mathbf{u}) = f(u_1, \ldots, u_n)$ for some smooth numerical function $f : U \to \mathbf{R}$. A scalar field is usually pictured in terms of its contour curves (or contour surfaces).

- Example: There are many examples of the modeling process from subjects as varied as physics and economics. There are also examples internal to mathematics, especially in geometry. Here is one that might occur in elementary geometry. Let M be the set of geometric rectangles in the plane. (The rectangles have a common corner and common alignment.) Each rectangle has a length ℓ and a width w. So $\ell : M \to \mathbf{R}$ and $w : M \to \mathbf{R}$ are both scalar fields. Together (ℓ, w) form a coordinate system that maps M onto an open quadrant in $\mathbf{R}^2$. This coordinate system is useful for describing the problem of painting a rectangular surface with given dimensions. The product $A = \ell w$ is also a scalar, the area of the rectangle. The pair (A, w) forms another coordinate system. The new system is useful for painting a rectangular surface with a given width using a fixed amount of paint. The numerical function that relates these two coordinate systems is $(x, y) \mapsto (xy, y)$ with inverse $(s, t) \mapsto (s/t, t)$.

- The concept of scalar field leads to the concept of *exact differential 1-form*. If s is a scalar field, the corresponding exact differential 1-form is written ds. A general *differential 1-form* is a linear combination of products of scalar fields with exact differential 1-forms. These concepts have precise definitions that are given in the book.

- Example (continued): In the example the differential of area A is the exact differential 1-form dA given by

$$dA = w\, d\ell + \ell\, dw.$$

  It is the sum of two differential forms neither of which is exact.

- There is another mathematical object that is easy to confuse with a differential 1-form. This is a *vector field*, often pictured as a field of arrows. The book includes a full discussion of vector fields. For the present discussion it is only necessary to point out that a vector field is a very different object from a differential 1-form. For instance, near every point for which the vector field is non-zero it is possible to find a new coordinate system with respect to which the vector field is constant. This is not true for differential forms: the form $\ell\,dw$ is already an example. It makes sense to say that a differential form is exact or not exact. For vector fields any such notion depends on a choice of metric. Vector fields are easy to picture with arrows. Differential forms are pictured in a quite different way, as explained in the book.

- How does one picture a differential 1-form? For an exact form ds there is a straightforward method: For each point, zoom in on contour curves (or contour surfaces) of s until they look linear near the point. In the example this gives nice pictures for $d\,\ell$ and $d\,w$. What about a form like $\ell\,dw$ that is not exact? The book describes how to draw the picture in this case.

- Example (continued): Scalar fields are fundamental to *mathematical modeling*, both in geometry and in applied mathematics. The letters that are used for the scalars are meaningful, as they describe the actual objects that are being modeled. Thus in the example A is "area" and w is "width". Let $c > 0$, and let M_c be the subset of rectangles for which $\ell + w = 2c$. On M_c there is a relation $d\ell + dw = 0$. A standard problem is to solve for the point in M_c where $dA = \ell\,dw + w\,d\ell = 0$. The answer is obtained by eliminating $d\ell$ and dw from these two equations. The solution to the problem is not numerical; it is the square where $\ell = w = c$, that is, a point in the manifold patch M.

- Here is another important distinction. A *passive transformation* is a change of coordinates. It gives an alternative description of the same situation. In the example M may be equally by well described the coordinate system ℓ, w or in the coordinate system A, w. A passive transformation is always invertible. An *active transformation* is a change in the object being modeled. A simple example is doubling the width of the rectangle and then flipping it across the diagonal. This transformation is described by $\ell \leftarrow 2w, w \leftarrow \ell$. The effect on the area is to double it:

$$A(\ell \leftarrow 2w, w \leftarrow \ell) = 2A.$$

  This is a rigorous equality of scalars: the left hand side is the composition of the scalar function $A : M \to \mathbf{R}$ with the active transformation $\ell \leftarrow 2w, w \leftarrow \ell : M \to M$. An active transformation need not have an inverse: the transformation $\ell \leftarrow \sqrt{\ell w}, w \leftarrow \sqrt{\ell w}$ maps rectangles to squares.

  The distinction between passive and active transformations is widely recognized but not always made explicit. (Some authors use terminology in

which "passive" and "active" are replaced by the awkward terms "alias" and "alibi".)

- Given two manifold patches K and M there is a concept of *manifold map* from K to M. The treatment in the book gives a practical notation that may be used in computations. Suppose K has coordinate system $\mathbf{t} = (t_1, \ldots, t_k)$ and M has coordinate system $\mathbf{u} = (u_1, \ldots, u_n)$. Then there is a smooth numerical function $\mathbf{f}$ such that the manifold map may be represented in the form $\mathbf{u} \leftarrow \mathbf{f}(\mathbf{t})$. What this means is: take the input point in K, find its coordinate values using $\mathbf{t}$, apply the function $\mathbf{f}$ to these values to get new values, and finally define the output point in M to be the point that has these values as $\mathbf{u}$ coordinates. The notation for an active transformation is the special case when $K = M$ and there is only one coordinate system, so the transformation is $\mathbf{u} \leftarrow \mathbf{f}(\mathbf{u})$.

- The left arrow notation for an active transformation is in perfect analogy with the assignment notation used in computer science. In that case, a (simultaneous) assignment means to start with the machine state, find the values of the $\mathbf{u}$ variables, do the computation indicated by $\mathbf{f}(\mathbf{u})$ and store the result in the locations indicated by the $\mathbf{u}$ variables, thus producing a new machine state. (In computer science and in mathematics the equal sign is often used to indicate assignment. Since assignment is not symmetric, this is a clash of notation.)

- There is a general concept of *pullback* of a scalar field s or a differential form by a manifold map. In the case of a scalar field this is just composition. That is, if the scalar field

$$s = g(\mathbf{u}) : M \to \mathbf{R}$$

is composed with the manifold map $\mathbf{u} \leftarrow \mathbf{f}(\mathbf{t}) : K \to M$, then the result is

$$s(\mathbf{u} \leftarrow \mathbf{f}(\mathbf{t})) = g(\mathbf{f}(\mathbf{t})) : K \to \mathbf{R}.$$

The notational device in the above equation is precise and convenient. It is also important as substantive mathematics: pullback together with change of variable is the key to integration of differential forms.

- There is an *exterior product* for differential forms, and this leads to the general notion of *differential k-form*. Furthermore, the *differential* of a $(k-1)$-form ω is a k-form $d\omega$. The main result of the chapter is *Stokes' theorem*. The simplest case of Stokes's theorem is for $k = 2$; in this case it is usually known as Green's theorem. In that special case the 1-form

$$\omega = p\,du + q\,dv$$

has a differential that is the 2-form

$$d\omega = \left(\frac{\partial q}{\partial u} - \frac{\partial p}{\partial v} \right) du\,dv.$$

If R is a two-dimensional region with boundary ∂R, Green's theorem states that

$$\int_R \left(\frac{\partial q}{\partial u} - \frac{\partial p}{\partial v} \right) du\, dv = \int_{\partial R} (p\, du + q\, dv).$$

The intuitive meaning is that circulation within the region) produces change along the boundary. Stokes' theorem is the corresponding general result that works in any number of dimensions, again for arbitrary coordinate systems.

- What is the geometrical picture for Stokes' theorem? It should not be in terms of a vector field, because that is a very different kind of object from a differential form. The treatment in the book presents reasonable pictures, at least in low dimensions.

- The ideas are illustrated by simple computations with static electric and magnetic fields. Physics texts often describe fields E, D, H, B. The electric field E may be thought of as a 1-form, while the electric flux density D is a 2-form. Similarly, the magnetic field intensity H can be a 1-form, while the magnetic flux density B is a 2-form. The forms E and B are exact and can be written in terms of potential by $E = -d\phi$ and $B = dA$. In the presence of charges and currents the forms D and H fail to be exact.

The metric tensor

Some texts on multivariable analysis treat differential forms, but do not explain how they related to notions of length, area, and volume. Or sometimes they make the connection only in Cartesian coordinates, or only in the three-dimensional situation. But such metric notions are unavoidable; the Laplace operator is a central object that makes full use of them.

The treatment in this book gives an account that describes the situation in any number of dimensions. The central object is the *metric tensor*. This may be regarded as a differential 1-form with very special properties that allow it to define the notion of distance. In the simplest case (two-dimensional flat space) it is the usual $\sqrt{(dx/dt)^2 + (dy/dt)^2}\, dt$ in Cartesian coordinates or $\sqrt{(dr/dt)^2 + r^2(d\theta/dt)^2}\, dt$ in polar coordinates.

- Sometimes all that is needed is to specify a *volume form*, which is an n-form. In two dimensional flat space this is the area 2-form $dx\, dy = r\, dr\, d\theta$. For uniformity of terminology this is called volume, more precisely, 2-dimensional volume. The volume form in arbitrary coordinates is conventionally written in a notation such as $\sqrt{g}\, du\, dv$, where the scalar factor $\sqrt{g}$ depends on the coordinate system. Thus in Cartesian coordinates $\sqrt{g} = 1$, while in polar coordinates $\sqrt{g} = r$.

- The *divergence theorem* is a result about vector fields that depends on knowing the volume form. A vector field, perhaps representing fluid flow, determines a *flux form*, which is an $(n-1)$-form that describes how the

flow crosses a surface. In the case of dimension $n = 2$ the vector field has components a and b (that depend on u and v), and the corresponding flux form is the 1-form $a\sqrt{g}\,dv - b\sqrt{g}\,du$. The divergence theorem then states that

$$\int_R \frac{1}{\sqrt{g}} \left(\frac{\partial a\sqrt{g}}{\partial u} + \frac{\partial b\sqrt{g}}{\partial v} \right) \sqrt{g}\,du\,dv = \int_{\partial R} (a\sqrt{g}\,dv - b\sqrt{g}\,du)\,.$$

The intuitive meaning is that mass production within the region) produces flow across the boundary. The flow is described by a vector field, an object that is easier to picture than a differential 1-form. Once the appropriate volume form is given, the divergence theorem may be formulated in any number of dimensions and in arbitrary coordinates. The interpretation and use of the divergence theorem are illustrated in the book by the solution of a *conservation law*.

- The definitions of *gradient* and *Laplace operator* require use of the metric tensor. This leads to a choice of whether to use *coordinate bases* or *orthonormal bases*. The advantages of both appear when it is possible to use *orthogonal coordinates*. The book explains how these ideas are related. In particular, it makes the connection with treatments found in elementary textbooks.

- The chapter concludes with formulas for *surface area*. A central idea is an amazing generalization of the theorem of Pythagoras. The classical theorem says that the length of a vector is the square root of the sum of squares of the lengths of its projections onto the coordinate axes. One case of the generalization says that the area of a parallelogram is the square root of the sum of the squares of the areas of its projections onto coordinate planes. Unfortunately, surface area calculations are not easy.

Measure zero

The notions of Riemann integral and Jordan content are almost obsolete; they have been surpassed by the far more powerful and flexible theories of Lebegue integral and Lebesgue measure. The study of the Riemann integral is still useful, if only to see how a subtle change in a definition can lead to a new theory that improves it in every way. Roughly speaking, the Riemann integral or Jordan content is defined by a one-step limiting process, while the Lebesgue integral or Lebesgue measure uses a two-step limiting process. This makes all the difference.

- The first part of the chapter is a tiny portion of the Lebesgue theory that fits in a nice way with the material presented earlier. This is the theory of sets of *Lebesgue measure zero*. These sets have nice mapping properties, and they also are fundamental to the characterization of Riemann integrable functions.

- The second topic is different. The usual surface area formula works for an explicitly defined surface (that is, a parameterized surface). There is another approach that works for a family of implicitly defined surfaces. For such a family there are two associated objects. The *fiber form* describes how to integrate over such a surface, and the *co-area factor* is a metric quantity associated with the surface. The *co-area formula* states that the surface area form is equal to the co-area factor times the fiber form. This gives a direct way to do surface area calculations for implicitly defined surfaces.

General references

For mathematics at this level it is helpful to see multiple approaches. The course used Rudin [17] as an alternative text; it focused on the two chapters on Functions of Several Variables and on Integration of Differential Forms. The book by Spivak [18] gives a treatment at about the same level as Rudin. Flanders [6] presents differential forms along with many applications. For a more advanced version of the story, the reader may consult Barden and Thomas [2]. Perhaps the relatively technical books by Morita [12] and by Agricola and Friedrich [1] could be useful. The subject matter considered here overlaps with tensor analysis. The book by Lovelock and Rund [9] has a relatively traditional approach; as a consequence one can find many useful formulas. The notes by Nelson [13] are more abstract and sophisticated, but there is good information there too.

Acknowledgements

The students in this course were young, enthusiastic, and able; meeting them over the course of a semester was a delight. Some of them were able to find time to read the manuscript and make valuable comments. In particular, the author is happy to thank Hong-Bin Chen and Xiaoyue Gong for their work on a previous draft.

Chapter 1

Differentiation

1.1 Fixed point iteration (single variable)

The foundational results for the subject of these lectures are the implicit function theorem and the inverse function theorem. Each of these depends on the fact that a certain equation has a unique solution. The simplest and most general technique for proving that an equation has a solution is fixed point iteration. That is the topic that begins these notes.

Let X be a metric space. Let $g : X \to X$ be a function from the space to itself. A *fixed point* is a point x^* in X with $g(x^*) = x^*$.

Take x_0 in X. Define the sequence of points by *iteration*. That is, define x_n in X by

$$x_{n+1} = g(x_n) \tag{1.1}$$

for $n \geq 0$. The set of iterates starting at x_0 is called the *orbit* of x_0.

Proposition 1.1 *Suppose that $g : X \to X$ is continuous. Let x_n be the sequence of iterates starting at x_0. Suppose that $x_n \to x^*$ as $n \to \infty$. Then x^* is a fixed point.*

Example: As an example, take X to be the real line with $g(x) = \cos(x)$. Start with an arbitrary real number. Then the iterates converge to a fixed point x^* that is equal to 0.739 in the first three decimal places. This is an experiment that is easy to do with a scientific calculator. |

When X is the real number line there is a lovely way of picturing the iteration process. Consider the graph of the linear function $y = x$. Each iterate is represented by a point (x, x) on this graph. Consider the graph of the given function $y = g(x)$. The next iterate is obtained by drawing the vertical line from (x, x) to $(x, g(x))$ and then the horizontal line from $(x, g(x))$ to $(g(x), g(x))$. The process is repeated as many times as is needed to show what is going on.

In the following we need stronger notions of continuity. A function $g : X \to X$ is *Lipschitz* if there is a constant $c \geq 0$ such that for all x and y in X we have

$$d(g(x), g(y)) \leq cd(x, y). \tag{1.2}$$

A Lipschitz function is automatically continuous, in fact, it is even uniformly continuous. If the constant $c \leq 1$ then the function is called a *contraction*. If the constant $c < 1$, then the function is called a *strict contraction*.

Theorem 1.2 (Contraction mapping theorem) *Let X be a complete metric space. Let $g : X \to X$ be a function. Suppose that g is a strict contraction. Then g has a unique fixed point x^*. Furthermore, if x_0 is in X, then the corresponding sequence of iterates satisfies $x_n \to x^*$ as $n \to \infty$.*

Proof: There can only be one fixed point. Suppose $x = g(x)$ and $y = g(y)$ are fixed points. Then $d(x, y) \leq d(x, g(x)) + d(g(x), g(y)) + d(g(y), y) = d(g(x), g(y)) \leq cd(x, y)$. Since $c < 1$ we must have $d(x, y) = 0$, so $x = y$. This proves the uniqueness of the fixed point.

The existence of the fixed point is shown via iteration. Start with x_0 and define the corresponding iterates x_n. Then $d(x_{n+1}, x_n) \leq c^n d(x_1, x_0)$. Hence for $m > n$

$$d(x_m, x_n) \leq d(x_m, x_{m-1}) + \cdots + d(x_{n+1}, x_n) \leq [c^m + \cdots + c^n] d(x_1, x_0). \quad (1.3)$$

Hence

$$d(x_m, x_n) \leq [c^{m-n} + \cdots + 1] c^n d(x_1, x_0) \leq \frac{1}{1-c} c^n d(x_1, x_0). \quad (1.4)$$

This shows that the x_n form a Cauchy sequence. Since the metric space is complete, this sequence must converge to some x^*. This is the desired fixed point. □

Example: Take again the example $g(x) = \cos(x)$. Let $1 \leq r < \pi/2$. Take the metric space to be the closed interval $X = [-r, r]$. Since $1 \leq r$ the function $g(x) = \cos(x)$ maps X into itself. Furthermore, since $r < \pi/2$ there is a $c < 1$ such that the derivative $g'(x) = -\sin(x)$ satisfied $|g'(x)| \leq c$. This is enough to show that g is a strict contraction. So the theorem guarantees the existence and uniqueness of the fixed point. Although there is no explicit formula for the solution of $\cos(x) = x$, the theorem defines this number with no ambiguity. |

The reasoning in this one-dimensional case may be formulated in a general way. The following result gives a way of rigorously proving the existence of fixed points. The hypothesis requires a bound on the derivative and the existence of an approximate fixed point p.

Proposition 1.3 *Let p be a real number, and consider the closed interval $[p - r, p + r]$ with $r > 0$. Suppose that $|g'(x)| \leq c < 1$ for x in this interval. Furthermore, suppose that $|g(p) - p| \leq (1 - c)r$. Then g maps the interval into itself and is a strict contraction, so it has a unique fixed point in this interval. Furthermore, iteration starting in this interval converges to the fixed point.*

Proof: It is clear that $|g(x) - g(y)| \leq c|x - y|$. In order to show that g maps the interval into itself, suppose $|x - p| \leq r$. Then $|g(x) - p| \leq |g(x) - g(p)| + |g(p) - p| \leq c|x - p| + (1 - c)r \leq r$. □

Sometimes it is helpful to have a result where one knows there is a fixed point, but wants to show that it is stable, in the sense that fixed point iteration starting near the fixed point converges to it. The following proposition is a variant that captures this idea.

Proposition 1.4 *Let p be a fixed point of g. Suppose that g' is continuous and that $|g'(p)| < 1$. Then for c with $|g'(p)| < c < 1$ there is an $r > 0$ such that $|g'(x)| \leq c < 1$ for x in the closed interval $[p - r, p + r]$. Then g maps the interval into itself and is a strict contraction. Furthermore, iterates starting in this interval converge to the fixed point.*

Fixed point iteration gives a rather general way of solving equations $f(x) = 0$. If a is an arbitrary non-zero constant, then the fixed points of

$$g(x) = x - \frac{1}{a} f(x) \quad (1.5)$$

are the solutions of the equation. The trick is to pick a such that g is a strict contraction. However,

$$g'(x) = 1 - \frac{1}{a} f'(x), \tag{1.6}$$

so the strategy is to take a close to the values of $f'(x)$ at points where $f(x)$ is close to zero. This is illustrated in the following result, which is a reformulation of a previous result.

Proposition 1.5 *Let p be a real number, and consider the closed interval $[p - r, p + r]$ with $r > 0$. Suppose that for some c with $0 < c < 1$ we have $1 - c < \frac{1}{a} f'(x) < 1 + c$ for x in this interval. Furthermore, suppose that $|f(p)| \leq |a|(1 - c)r$. Then the corresponding fixed point iteration starting in this interval converges to the solution of $f(x) = 0$ in this interval.*

Example: Suppose we want to solve the equation $x = 2\cos(x)$. Since $2\cos(\pi/3) = 1$, this is a fixed point equation with a solution near $\pi/3$. However, since $-2\sin(\pi/3) = -\sqrt{3}$ there is not much hope for stability. So we may write the equation as $f(x) = x - 2\cos(x) = 0$. The derivative is $f'(x) = 1 + 2\sin(x)$. The value of this derivative at $\pi/3$ is $1 + \sqrt{3}$ which is fairly close to 3. So we take $a = 3$, and the iteration function is $g(x) = x - \frac{1}{3}(x - 2\cos(x))$ with derivative $g'(x) = 1 - \frac{1}{3}(1 + 2\sin(x))$. This derivative decreases from $1 - \frac{1}{3}(1 + 1) = 1/3$ at $\pi/6$ to $1 - \frac{1}{3}(1 + 2) = 0$ at $\pi/2$. So we may take $c = 1/3$. Furthermore, $\pi/3 - g(\pi/3) = \frac{1}{3}(\pi/3 - 1)$. It is easy to see that this is bounded above by $(1 - c)r = \frac{2}{3}\pi/6$. So $\pi/3$ is a sufficiently good approximation to the fixed point. Iteration shows that the fixed point is close to 1.03. |

Example: Even with simple examples like the one above, it is convenient to use a computer program to do the calculations. Here, for example, is a complete program written in the computer language R.

```
f <- function (x) x - 2 * cos(x)
g <- function (x) x - f(x)/3
x <- 1
for (i in 1:10) x <- g(x)
x
[1] 1.029867
```

|

Example: Fix z with $0 < z < 1$. Suppose we want to solve the equation $w = ze^w$ for w. We can plot $z = we^{-w}$ and see that there are always two positive solutions, one less than 1 and the other greater than 1. The smaller solution may be computed by fixed point iteration. On the other hand, the greater solution is an unstable fixed point. If we take the smaller solution, then we have a well-defined function $L(z)$ defined for $0 < z < 1$ with values $0 < L(z) < 1$ that satisfies $L(z) = ze^{L(z)}$. There is no obvious explicit formula for this function. |

1.2 The implicit function theorem (single variable)

If one wants to solve an equation $f(x, y) = 0$ for x as a function of y (or for y as a function of x), there can be problems. Consider, for example, the equation $x - y^2 = 0$. There is no problem in solving for x as a function of y. On the other hand, near the origin there is a real issue of how to solve for y as a function of x. The obvious attempts are $y = \sqrt{x}$ for $x \geq 0$ and $y = -\sqrt{x}$ for $x \geq 0$. In either case the solution is non-differentiable at $x = 0$. The implicit function theorem gives insight into this sort of behavior.

Theorem 1.6 (Implicit function theorem) *Let $f(x, y)$ be continuous with continuous partial derivatives near some point $x = a, y = b$ with $f(a, b) = 0$. The goal is to solve $x = h(y)$, so x is to be the dependent variable. Suppose that the partial derivative with respect to the dependent variable $\partial f(x, y)/\partial x \neq 0$ at $x = a, y = b$. Then there is a continuous function $h(y)$ defined for y near b with $f(h(y), y) = 0$. In fact, the function $h(y)$ has a continuous derivative, and*

$$0 = \frac{df(h(y), y)}{dy} = \frac{\partial f(x, y)}{\partial x}|_{x=h(y)} \frac{dh(y)}{dy} + \frac{\partial f(x, y)}{\partial y}|_{x=h(y)}. \tag{1.7}$$

In the following we shall sometimes write the partial derivatives in the form

$$\frac{\partial f(x, y)}{\partial x} = f'_{,1}(x, y) \tag{1.8}$$

and

$$\frac{\partial f(x, y)}{\partial y} = f'_{,2}(x, y). \tag{1.9}$$

Then the equation has the form

$$0 = f'_{,1}(h(y), y)h'(y) + f'_{,2}(h(y), y). \tag{1.10}$$

This may be solved to give

$$h'(y) = -\frac{f'_{,2}(h(y), y)}{f'_{,1}(h(y), y)}. \tag{1.11}$$

Proof: Let $A = \partial f(x, y)/\partial x$ at $x = a, y = b$, that is $A = f'_{,1}(a, b)$. Consider the iteration function

$$g(x, y) = x - \frac{1}{A} f(x, y). \tag{1.12}$$

This has partial derivative

$$\frac{\partial g(x, y)}{\partial x} = 1 - \frac{1}{A} \frac{\partial f(x, y)}{\partial x}. \tag{1.13}$$

At $x = a, y = b$ this has the value 0. Pick some convenient value of c with $0 < c < 1$. Then for x, y sufficiently close to a, b this partial derivative has

absolute value bounded by c. In particular, this is true in some box $|x-a| \le r$, $|y-b| < s$. Fix $r > 0$. We know that $g(a,b)-a = 0$. Hence, if $s > 0$ is sufficiently small, $|g(a,y) - a| \le (1-c)r$. This shows that for all y with $|y-b| < s$ the contraction mapping result works for the function $x \mapsto g(x,y)$ on the interval of x with $|x-a| \le r$. The resulting fixed point is given by some function $x = h(y)$.

It is not too difficult to show that $h(y)$ is a continuous function of y. Consider y' near y. Then $g(h(y),y) = h(y)$ and $g(h(y'),y') = h(y')$. So $h(y') - h(y) = g(h(y'),y') - g(h(y),y) = g(h(y'),y') - g(h(y),y') + g(h(y),y') - g(h(y),y)$. This gives $|h(y') - h(y)| \le c|h(y') - h(y)| + |g(h(y),y') - g(h(y),y)|$. Write this as $(1-c)|h(y') - h(y)| \le |g(h(y),y') - g(h(y),y)|$. Then as $y' \to y$ we have $g(h(y),y') \to g(h(y),y)$, and hence $h(y') \to h(y)$.

It remains to show that h has a continuous derivative. The derivative of h at y is computed as the limit of the difference quotient $(h(y+k) - h(y))/k$ as $k \to 0$. In order to get a handle on this, compute

$$0 = 0 + 0 = f(h(y+k), y+k) - f(h(y), y). \tag{1.14}$$

Expand

$$\begin{aligned} 0 &= f(h(y+k), y+k) - f(h(y), y) \qquad (1.15)\\ &= \int_0^1 \frac{d}{dt} f(th(y+k) + (1-t)h(y), t(y+k) + (1-t)y)\, dt. \end{aligned}$$

By the chain rule for partial derivatives

$$\begin{aligned} 0 &= f(h(y+k), y+k) - f(h(y), y) \qquad (1.16)\\ &= \int_0^1 [f'_{,1}(th(y+k) + (1-t)h(y), t(y+k) + (1-t)y)(h(y+k) - h(y))\\ &\quad + f'_{,2}(th(y+k) + (1-t)h(y), t(y+k) + (1-t)y)k]\, dt. \end{aligned}$$

This has solution

$$\frac{h(y+k) - h(y)}{k} = -\frac{\int_0^1 f'_{,2}(th(y+k) + (1-t)h(y), t(y+k) + (1-t)y)\, dt}{\int_0^1 f'_{,1}(th(y+k) + (1-t)h(y), t(y+k) + (1-t)y)\, dt}. \tag{1.17}$$

Now let $k \to 0$ in each integral on the right. For each fixed t between 0 and 1 the integrand converge to a limit as $k \to 0$. Furthermore, the integrands are bounded uniformly in k. In this circumstance the dominated convergence theorem (considered later in these notes) shows that the limit of the integral is the integral of the limit. In fact, the integrands become independent of t, and we obtain the desired formula for $h'(y)$, that is,

$$h'(y) = -\frac{f'_{,2}(h(y), y)}{f'_{,1}(h(y), y)}. \tag{1.18}$$

The right hand side of this formula is continuous in y, so the left hand is also continuous in y. This shows that the function h has a continuous derivative. □

Example: As a very simple example, consider the equation $f(x, y) = x^2 + y^2 - 1 = 0$ of a circle. The partial derivatives are $\partial f(x, y)/\partial x = 2x$ and $\partial f(x, y)/\partial y = 2y$. Thus for every point on the circle where $x \neq 0$ it is possible to solve for x in terms of y near this point. However, if the point is on the right half of the circle, then $x = \sqrt{1 - y^2}$, while if the point is on the left half of the circle, then $x = -\sqrt{1 - y^2}$. In a completely symmetrical way, for every point on the circle where $y \neq 0$ it is possible to solve for y in terms of x near this point. If the point is on the upper half of the circle, then $y = \sqrt{1 - x^2}$, while if the point is on the lower half of the circle, then $y = -\sqrt{1 - x^2}$. |

Example: For a less trivial example, take $f(x, y) = \cos(x) + \cos(xy) + \cos(y) - 2$. Then $f(\pi/2, 0) = 0$. In order to solve $x = h(y)$ near $y = 0$ we can use fixed point iteration. Say that, for instance, we want the value $h(1/10)$. Notice that $\partial f(x, y)/\partial y = -\sin(x) - y\sin(xy)$. At $x = \pi/2, y = 0$ this has the value -1. So we take the iteration function $g(x) = x + f(x, 1/10)$. There is a fixed point at $x = 1.553753$ which is near $\pi/2 = 1.570796$. This is the value of $h(1/10)$. |

Example: Here is an R program to compute $h(1/10)$ from the previous example.

```
f <- function (x,y) cos(x) + cos(x * y) + cos(y) - 2
g <- function (x) x + f(x,1/10)
x <- pi/2
for (i in 1:30) x <- g(x)
x
[1] 1.553753 |
```

1.3 Linear algebra review (norms)

This section assumes some acquaintance with linear algebra and matrix theory. If A is an m by n real matrix, and $\mathbf{x}$ is an n-component column vector, then $A\mathbf{x}$ is an m-component column vector. Thus matrix defines a linear function that sends n-component vectors to m-component vectors. If B is a k by m matrix, then the matrix product BA is a k by n matrix. The product of BA with $\mathbf{x}$ is the same as the product of B with $A\mathbf{x}$. Thus matrix multiplication coincides with the composition of the functions.

One can think of an n-component *column vector* $\mathbf{x}$ as a n by 1 matrix. Often we shall refer to this simply as a *vector*. An n-component *row vector* ω is a 1 by n matrix. This will be called a *covector* or *linear form*. The product $\omega\mathbf{x}$ is then a scalar.

Let A be an m by n matrix. Then its *transpose* A^T is the n by m matrix with the columns and rows reversed. There is a fundamental identity

$$(BA)^T = A^T B^T. \tag{1.19}$$

between n by k matrices.

An n by n square matrix Q has an *inverse* R if $QR = I$ and $RQ = I$. The inverse of Q is denoted Q^{-1}. It is well-known linear algebra fact that if the inverse exists on one side, for instance $QR = I$, then also $RQ = I$, and the

inverse exists on both sides. For matrices there are natural notions of addition and subtraction and of (non-commutative) multiplication. However some care must be taken with division, since PQ^{-1} is in general different from $Q^{-1}P$. Note the important identity

$$(AB)^{-1} = B^{-1}A^{-1}. \tag{1.20}$$

Furthermore, suppose both A and B have inverses. Then we may take $B = A + B - A$, multiply to get $A^{-1}B = I + A^{-1}(B - A)$, multiply again to get $A^{-1} = B^{-1} + A^{-1}(B - A)B^{-1}$. A square matrix with an inverse is often called *invertible* or *non-singular*. Of course it is said to be *singular* if it does not have an inverse.

An m by n matrix A cannot have an inverse matrix unless $m = n$. However there are important related concepts. Such a matrix has a *rank* r, which is the dimension of the range, that is, of the space spanned by the columns. It is always true that $r \leq n$ and $r \leq m$. If $r = n \leq m$ or $r = m \leq n$, then the matrix is said to be of *maximal rank*. In the first case this is equivalent to the n by n matrix $A^T A$ being invertible, while in the second case it is equivalent to the m by m matrix AA^T being invertible. A matrix like $A^T A$ is sometimes called a *Gram matrix*.

If A is an n by n square matrix, then there are several interesting numbers that one can associate to it. The *trace* of A is the number $\mathrm{tr}(A)$ obtained by taking the sum of the diagonal entries. If A and B are such matrices, then

$$\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B) \tag{1.21}$$

and

$$\mathrm{tr}(AB) = \mathrm{tr}(BA). \tag{1.22}$$

The most interesting number associated to a square matrix A is its *determinant* $\det(A)$. It has a relatively complicated definition, but there is one property that is particularly striking. If A and B are such matrices, then

$$\det(AB) = \det(BA) = \det(A)\det(B). \tag{1.23}$$

Also $\det(I) = 1$. (Here I is the n by n identity matrix, with 1s on the main diagonal and 0s elsewhere.) We always have $\det(A) = \det(A^T)$. A matrix A has an inverse A^{-1} if and only if $\det(A) \neq 0$. In that case $\det(A^{-1}) = 1/\det(A)$.

An n by n real matrix H is *symmetric* if $A = A^T$. An n by n real matrix P is *orthogonal* if $P^T P = PP^T = I$, that is, $P^{-1} = P^T$. Notice that if P is orthogonal, then $P^T = P^{-1}$ is also orthogonal. An n by n real matrix Λ is *diagonal* if all entries off the main diagonal are 0.

Theorem 1.7 (Spectral theorem) *Suppose H is real and symmetric. Then there is a real orthogonal matrix P and a real diagonal matrix Λ such that*

$$HP = P\Lambda. \tag{1.24}$$

The columns of P form an orthonormal basis, and these are the eigenvectors of H. The diagonal entries in Λ are the eigenvalues of H. In general it is quite difficult to compute such eigenvalues and eigenvectors.

The equation $HP = P\Lambda$ may be written in various ways. Multiplication on the right by P^T gives $H = P\Lambda P^T$. If λ_j are the eigenvalues, and $\mathbf{p}_j$ are the columns of P, then this gives the *spectral representation* of the matrix:

$$H = \sum_j \lambda_j \mathbf{p}_j \mathbf{p}_j^T . \tag{1.25}$$

Example: Here is a simple example where one can do the computation. The symmetric matrix is

$$H = \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix} . \tag{1.26}$$

It is easy to see that this matrix has dependent rows, so the determinant is zero. As a consequence at least one eigenvalue is zero. In this situation it is easy to find the other two eigenvalues λ_1, λ_2. Use $\lambda_1 + \lambda_2 = \mathrm{tr}(A) = 34$ and $\lambda_1^2 + \lambda_2^2 = \mathrm{tr}(A^2) = 706$. (For a symmetric matrix $\mathrm{tr}(A^2) = \mathrm{tr}(A^T A)$ is the 2-norm.) All that remains is to solve a quadratic equation to get the non-zero eigenvalues 25, 9. The corresponding eigenvectors are found by solving linear systems. The eigenvectors form the column matrix

$$R = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & -2 \\ 0 & 4 & -1 \end{bmatrix} . \tag{1.27}$$

Since the eigenvalues are distinct, the eigenvectors are automatically orthogonal. This says that $R^T R$ is a diagonal matrix. If we normalize each column to be a vector of length one, then we get a new matrix P such that $P^T P$ is the identity matrix. In particular, we get the representation

$$H = 25\frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} + 9\frac{1}{18}\begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & -1 & 4 \end{bmatrix} . \tag{1.28}$$

|

Our main goal is to get a number to measure the size of a matrix. There is a particularly simple definition of the size of a real column vector, namely

$$|\mathbf{x}| = \sqrt{\mathbf{x}^T \mathbf{x}}. \tag{1.29}$$

This will be called the *Euclidean norm* of the vector. There is a corresponding notion of inner product of two real column vectors:

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y}. \tag{1.30}$$

The famous Schwarz inequality says that $|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}||\mathbf{y}|$.

If A is a real m by n matrix, then we may define the size of the matrix A to be the least number $\|A\|$ such that for all real n component column vectors we have

$$|A\mathbf{x}| \leq \|A\||\mathbf{x}|. \tag{1.31}$$

It follows that

$$|A\mathbf{x} - A\mathbf{y}| = |A(\mathbf{x} - \mathbf{y})| \leq \|A\||\mathbf{x} - \mathbf{y}|, \tag{1.32}$$

so $\|A\|$ is the best Lipschitz constant for the linear transformation A. In these lectures it will be called the *Lipschitz norm* of A (corresponding to the Euclidean norm for the vectors).

It is not difficult to show that the norm behaves reasonably under sum and product. In fact, $\|A + B\| \leq \|A\| + \|B\|$ and $\|AB\| \leq \|A\|\|B\|$. The norm also preserves the transpose: $\|A^T\| = \|A\|$. This has an easy proof that begins with $\|A^T\mathbf{x}\|^2 = A^T\mathbf{x} \cdot A^T\mathbf{x} = \mathbf{x} \cdot AA^T\mathbf{x}$. The reader will have no difficulty using the Schwarz inequality to complete the proof that $\|A^T\| \leq \|A\|$. This is half the proof. But then one can use $A = A^{TT}$ to prove that $\|A\| \leq \|A^T\|$.

For the inverse it is easy to see that $\|A^{-1}\| \geq 1/\|A\|$. If B is close enough to A and A^{-1} exists, then B^{-1} also exists. All that is required is that $B - A$ be small enough so that $\|A^{-1}\|\|(B - A)\| < 1$. In fact, there is an even better result, given in the following proposition.

Proposition 1.8 *Suppose that A^{-1} exists and that $\|A^{-1}(B - A)\| < 1$. Then B^{-1} exists, and*

$$B^{-1} = A^{-1} - A^{-1}(B - A)B^{-1}. \tag{1.33}$$

Proof: The space of n by n matrices with the Lipschitz norm is a complete metric space. Define the map $g(C) = A^{-1} - A^{-1}(B - A)C$. By hypothesis this is a contraction mapping. Hence it has a fixed point satisfying $C = A^{-1} - A^{-1}(B - A)C$. Then $AC = I - (B - A)C = I - BC + AC$, and hence $BC = I$. So C is the inverse of B. □

This norm $\|A\|$ is quite difficult to compute. In fact, write $H = A^TA$ and notice that it is symmetric. So $HP = P\Lambda$, or, taking $Q = P^{-1}$, $QH = \Lambda Q$. Thus

$$|A\mathbf{x}|^2 = (A\mathbf{x})^T(A\mathbf{x}) = \mathbf{x}^TA^TA\mathbf{x} = \mathbf{x}^TH\mathbf{x} = \mathbf{x}^TQ^T\Lambda Q\mathbf{x} = (Q\mathbf{x})^T\Lambda(Q\mathbf{x}). \tag{1.34}$$

Note that this expression easily leads to the conclusion that all the eigenvalues $\lambda_i \geq 0$. It can be compared to

$$|\mathbf{x}|^2 = \mathbf{x}^T\mathbf{x} = \mathbf{x}^TQ^TQ\mathbf{x} = (Q\mathbf{x})^T(Q\mathbf{x}). \tag{1.35}$$

Clearly $|A\mathbf{x}|^2 \leq \lambda_{\max}|\mathbf{x}|^2$, and this is the least such bound. This is summarized in the following result.

Theorem 1.9 *The Lipschitz norm of A has the explicit expression*

$$\|A\| = \sqrt{\lambda_{\max}}, \tag{1.36}$$

where $\lambda_{\max}$ is the largest eigenvalue of A^TA.

For a real symmetric matrix $A^T A$ the norm is the largest eigenvalue. So another consequence of this reasoning is

$$\|A\| = \sqrt{\|A^T A\|}. \tag{1.37}$$

In other words, to compute the norm of A one computes the norm of the symmetric matrix $A^T A$ and takes the square root.
Example: Consider the matrix

$$A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}. \tag{1.38}$$

It is easy to see that $\|A\|_2 = \sqrt{34}$. On the other hand, to compute $\|A\|$ takes some work. But the matrix $A^T A$ is the matrix H of the previous example, which has eigenvalues 25, 9, 0. So $\|A\| = \sqrt{25} = 5$.

There is a considerably easier way to do this computation, namely compute $\|A^T\|$, which is square root of the largest eigenvalue of AA^T. The pleasure of this is left to the reader. |

Since eigenvalues are difficult to compute, this is bad news. However, there is another norm of A that is easier to compute. This is the *Euclidean norm*

$$\|A\|_2 = \sqrt{\mathrm{tr} A^T A}. \tag{1.39}$$

It has the happy feature that it may be obtained from the matrix entries. The j,k entry of $A^T A$ is $\sum_i a_{ij} a_{ik}$. Thus the diagonal j,j entry is $\sum_i a_{ij} a_{ij} = \sum_i a_{ij}^2$ Thus

$$\|A\|_2^2 = \mathrm{tr} A^T A = \sum_j \sum_i a_{ij}^2 \tag{1.40}$$

Even better news is that the Euclidean norm is an upper bound for the Lipschitz norm. In fact,

$$|A\mathbf{x}|^2 = \sum_i (\sum_j a_{ij} x_j)^2 \le \sum_i (\sum_j a_{ij}^2)(\sum_k x_k^2) = \sum_i \sum_j a_{ij}^2 (\sum_k x_k^2) = \|A\|_2^2 |x|^2. \tag{1.41}$$

This is summarized in the following proposition.

Proposition 1.10 *The Lipshitz norm of a matrix is related to the Euclidean norm of the matrix by*

$$\|A\| \le \|A\|_2. \tag{1.42}$$

The other direction is not so precise. There is a bound, but it depends on the dimension of the m by n matrix. One result is the following.

Proposition 1.11 *The Euclidean norm of a matrix is related to the Lipschitz norm of the matrix by*

$$\|A\|_2 \le \sqrt{n} \|A\|. \tag{1.43}$$

The proof is easy. Write $a_{ij} = \sum_p a_{ip}\delta_{pj}$. This is the i component of A applied to the vector δ_j that is 0 except for 1 in the jth place. Then

$$\|A\|_2^2 = \sum_j \sum_i a_{ij}^2 = \sum_j \sum_i (\sum_p a_{ip}\delta_{pj})^2 = \sum_j |A\delta_j|^2 \leq \sum_j \|A\|^2 |\delta_j|^2 = n\|A\|^2. \tag{1.44}$$

1.4 Linear algebra review (eigenvalues)

This section is devoted to a deeper analysis of the structure of a square matrix A with real entries. A complex number λ is an *eigenvalue* if and only if $A - \lambda I$ is a singular matrix. The set of eigenvalues is sometimes called the *spectrum* of A. A real matrix A has either real eigenvalues λ_k or complex eigenvalue pairs $\lambda_k = |\lambda_k|(\cos(\phi_k) + i\sin(\phi_k))$, $\bar{\lambda}_k = |\lambda_k|(\cos(\phi_k) - i\sin(\phi_k))$. In the complex case $|\lambda_k| > 0$ and ϕ_k is not a multiple of π. The *spectral radius* $\rho(A)$ is defined to be the largest of the $|\lambda_k|$.

A complex eigenvalue pair manifests itself in the form of a real rotation matrix. Recall that a *rotation matrix* is a matrix of the form

$$R(\phi) = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix}. \tag{1.45}$$

Again we shall mainly be interested in the cases when R is not diagonal, that is, ϕ is not a multiple of π.

One more ingredient is necessary. A square matrix N is said to be *nilpotent* if some power $N^p = 0$ for $p \geq 1$. Such a matrix is small in the sense that every eigenvalue of N must be zero. In other words, the spectral radius $\rho(N) = 0$. However the norm $\|N\|$ can be quite large.

Much can be learned from the special case of 2 by 2 matrices. A 2 by 2 matrix A defines a linear function from $\mathbf{R}^2$ to $\mathbf{R}^2$. For each $\mathbf{x}$ in $\mathbf{R}^2$ there is a corresponding $A\mathbf{x}$ in $\mathbf{R}^2$. It is difficult to imagine the graph of such a function. However there is a nice pictorial representation that is very helpful. One picks several values of $\mathbf{x}$ and draws corresponding vectors $A\mathbf{x} - \mathbf{x}$ from $\mathbf{x}$ to $A\mathbf{x}$.

In the 2 by 2 case it is not difficult to compute the eigenvalues. The sum of the eigenvalues is $\mathrm{tr}(A)$ and the product of the eigenvalues is $\det(A)$. From this it is easy to see that the eigenvalues are the solutions of the quadratic equation $\lambda^2 - \mathrm{tr}(A)\lambda + \det(A) = 0$.

If A has positive eigenvalues less than one, then the vectors point in the general direction of the origin. This is a stable case when the orbits go to the fixed point at zero. If A has positive eigenvalues greater than one, then they move away from the origin. If A has a positive eigenvalue greater than one and another less than one, then the orbits move toward the eigenspace corresponding to the larger of the two eigenvalues and then outward. If A has negative eigenvalues, then there is some overshoot. However one can always look at A^2 and get pictures with just positive eigenvalues.

Another situation is the complex eigenvalue case, when A is a multiple of a rotation. If the angle ϕ is reasonably small, then one gets a picture of a spiral around the origin, moving inward if $|\lambda| < 1$ and outward if $|\lambda| > 1$.

The case of a nilpotent 2 by 2 matrix A is quite interesting. There is an eigenspace corresponding to eigenvalue 0. The arrows all land in that eigenspace, and the ones starting from the eigenspace lead to the origin.
Example: Consider the matrix

$$A = \frac{2}{5} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}. \tag{1.46}$$

This has eigenvalues equal to 2/5 times $2 \pm i$. The maginitude of the eigenvalues is 2/5 times $\sqrt{5}$ which is about 0.8944272. This means the behavior is stable. The orbit should spiral in to zero. |
Example: Here is a computation for the previous example.

```
x <- vector()
y <- vector()
u <- 100
v <- 0
for (i in 1:100) {
u <- (2*u - v) * 2 / 5
v <- (u + 2*v) * 2 / 5
x[i] <- u
y[i] <- v }
frame()
plot(x,y)
lines(x,y)  |
```

Theorem 1.12 *Let A be an n by n matrix with real entries. Suppose that it has n distinct eigenvalues. Then there exists a real matrix Λ with the same (possibly complex) eigenvalues λ_k as A. Here Λ is a real matrix whose non-zero entries are either real diagonal entries λ_k or two-by-two blocks of the form $|\lambda_k| R(\phi_k)$, where $|\lambda_k| > 0$ and $R(\phi_k)$ is a rotation matrix. Furthermore, there exists a real invertible matrix P such that*

$$AP = P\Lambda. \tag{1.47}$$

Suppose for simplicity that the eigenvalues are distinct and real, so Λ is a real diagonal matrix. In this case there is a spectral theorem that directly generalizes the result for the symmetric case. Let P have column vector $\mathbf{p}_j$ and let P^{-1} have row vectors χ_k. Then we have matrices that are *similar matrices*

$$A = P\Lambda P^{-1} = \sum_j \lambda_j \mathbf{p}_j \chi_k. \tag{1.48}$$

Example: Consider the matrix

$$A = \begin{bmatrix} 11 & -18 \\ 6 & -10 \end{bmatrix}. \tag{1.49}$$

Then

$$A = P\Lambda Q = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}. \tag{1.50}$$

Here $Q = P^-$ is the inverse of P. The spectral representation is

$$A = 2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{bmatrix} 2 & -3 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} \begin{bmatrix} -1 & 2 \end{bmatrix}. \tag{1.51}$$

In this situation there is usually nothing gained by normalizing the vectors. |

Theorem 1.13 *Let A be an n by n matrix with real entries. Then there exists a real matrix Λ with the same (possibly complex) eigenvalues λ_k as A. Here Λ is a real matrix whose non-zero entries are either real diagonal entries λ_k or two-by-two blocks of the form $|\lambda_k|R(\phi_k)$, where $|\lambda_k| > 0$ and $R(\phi_k)$ is a rotation matrix. Furthermore, for every $\delta > 0$ there exists a real invertible matrix P and a real nilpotent matrix N, such that*

$$AP = PJ, \tag{1.52}$$

where

$$J = \Lambda + N \tag{1.53}$$

and $\|N\| \leq \delta$.

This theorem says that A may be represented as $A = PJP^{-1}$, where we know a lot about the size of $J = \Lambda + N$. In fact, the norm of Λ is the spectral radius $\rho(A)$ of A. The norm of N is less that δ. So the norm of J is bounded

$$\|J\| \leq \|\Lambda\| + \|N\| \leq \rho(A) + \delta. \tag{1.54}$$

Unfortunately, this does not mean that the norm of $\|A\|$ is particularly small. In fact, all we know is that

$$\|A\| = \|PJP^{-1}\| \leq \|P\|\|J\|\|P^{-1}\|. \tag{1.55}$$

Unfortunately, $\|P\|\|P^{-1}\|$ can be much larger than one.

There is an artificial trick for getting around this difficulty. This trick is to find a new norm that is adapted to the particular matrix A. It is convenient to write $Q = P^{-1}$, so that Q is an invertible matrix with

$$A = Q^{-1}JQ. \tag{1.56}$$

The matrix $G = Q^TQ$ is symmetric with eigenvalues > 0. Define a new norm on vectors by

$$|\mathbf{x}|_G = \sqrt{\mathbf{x}^T G \mathbf{x}} = \sqrt{\mathbf{x}^T Q^T Q \mathbf{x}} = \sqrt{(Q\mathbf{x})^T (Q\mathbf{x})} = |Q\mathbf{x}|. \tag{1.57}$$

This defines a new Lipschitz norm on matrices. We can compute this norm in the case of A (for which it was designed). We have

$$|A\mathbf{x}|_G = |QA\mathbf{x}| = |JQ\mathbf{x}| \leq \|J\||Q\mathbf{x}| = \|J\||\mathbf{x}|_G. \tag{1.58}$$

With respect to this new Lipschitz norm we have

$$\|A\|_G \leq \|J\| \leq \rho(A) + \delta. \tag{1.59}$$

This discussion is summarized in the following theorem. The theorem says that the size of the matrix A is in some more profound sense determined by the spectral radius $\rho(A)$.

Theorem 1.14 *For every real square matrix A with spectral radius $\rho(A)$ and for every $\delta > 0$ there is a new Lipschitz norm defined with respect to a symmetric matrix G with strictly positive eigenvalues such that*

$$\|A\|_G \leq \rho(A) + \delta. \tag{1.60}$$

One can wonder whether one has the right to change norm in this way. The first observation is that notions of continuity do not change. This is because if $G = Q^T Q$, then

$$\begin{aligned} |\mathbf{x}|_G &\leq \|Q\||\mathbf{x}| \\ |\mathbf{x}| &\leq \|Q^{-1}\||\mathbf{x}|_Q \end{aligned} \tag{1.61}$$

Furthermore, the geometry only changes in a rather gentle way. Thus, a ball $|\mathbf{x}|_G^2 < r^2$ in the new norm is actually an ellipsoid $|Q\mathbf{x}|^2 = \mathbf{x}^T G \mathbf{x} < r^2$ in the original picture.

The weakness of this idea is that the new norm is specially adapted to the matrix A. If one is dealing with more than one matrix, then the new norm that is good for one may not be the new norm that is good for the other.

Example: Consider the previous example with distinct eigenvalues $2, -1$. The new norm is $|Q\mathbf{x}| = \sqrt{(2x - 3y)^2 + (-x + 2y)^2} = \sqrt{5x^2 - 16xy + 13y^2}$. With respect to this norm on vectors the matrix A has norm 2 (the spectral radius). |

In the case of a 2 by 2 matrix A with a repeated eigenvalue λ it is easy to find the decomposition. Let $\mathbf{u}$ be an eigenvector, so that $(A - \lambda I)\mathbf{u} = 0$. Find some other vector such that $(A - \lambda I)\mathbf{v} = \delta \mathbf{u}$. Take $P = [\mathbf{u}, \mathbf{v}]$ to be the matrix with these vectors as columns. Then

$$AP = P \begin{bmatrix} \lambda & \delta \\ 0 & \lambda \end{bmatrix}. \tag{1.62}$$

Example: Let

$$A = \begin{bmatrix} -4 & 4 \\ -9 & 8 \end{bmatrix} \tag{1.63}$$

with repeated eigenvalue 2. Then

$$A = PJQ = \begin{bmatrix} 2 & -\frac{1}{3}\delta \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 2 & \delta \\ 0 & 2 \end{bmatrix} \frac{1}{\delta} \begin{bmatrix} 0 & \frac{1}{3}\delta \\ -3 & 2 \end{bmatrix}. \tag{1.64}$$

So the new norm $|Q\mathbf{x}|$ is determined by

$$\delta^2|Q\mathbf{x}| = \sqrt{(-3x+3y)^2 + \frac{1}{9}\delta^2 y^2} = \sqrt{9x^2 - 12xy + (4+\frac{1}{9}\delta^2)y^2}. \quad (1.65)$$

As long as $\delta \neq 0$ this expression is only zero at the origin, and so it defines a legitimate norm. |

We conclude this discussion of linear algebra with one more digression. We have been considering a matrix as a linear transformation, in which case the relevant classification is given by *matrix similarity* transformations. Thus if $\mathbf{y} = A\mathbf{x}$, and we write $\mathbf{x} = P\mathbf{x}'$ and $\mathbf{y} = P\mathbf{y}'$, we get $\mathbf{y}' = P^{-1}AP\mathbf{x}'$. This means that we transform the matrix A to $P^{-1}AP$ in an attempt to simplify it. However it is also possible to regard a symmetric matrix G as a *quadratic form*. This means that we associate with G the quadratic form

$$q(\mathbf{x}) = \mathbf{x}^T G\mathbf{x}. \quad (1.66)$$

If we make a transformation $\mathbf{x} = P\mathbf{y}$, then the quadratic form becomes

$$q(\mathbf{x}) = \mathbf{x}^T G\mathbf{x} = (P\mathbf{y})^T gP\mathbf{y} = \mathbf{y}^T P^T GP\mathbf{y}. \quad (1.67)$$

In other words, the relevant transformation is a *matrix congruence* transformation that sends G to $P^T GP$. The matrix P is required to be non-singular, but is not required to be orthogonal. So the eigenvalues are not preserved. Remarkably, the signs of the eigenvalues are preserved. In fact, the matrix may be brought to the diagonal form $P^T GP = E$, where the diagonal entries are ± 1 or 0. The quadratic form is *non-degenerate* precisely when all these entries are ± 1. It is *positive definite* when all these entries are 1.

These quadratic form ideas will occur in the context of the Hessian matrix of second partial derivatives. An even more important application is the metric tensor that provides the notions of length, area, and volume.

1.5 Differentiation (multivariable)

It is convenient to deal with a rather general class of functions, those defined on some domain $E \subseteq \mathbf{R}^n$ with values in $\mathbf{R}^m$. If $\mathbf{x}$ is in E, then there is a corresponding value

$$\mathbf{u} = \mathbf{f}(\mathbf{x}) \quad (1.68)$$

in $\mathbf{R}^m$. This may be written somewhat more explicitly as $u_i = f_i(\mathbf{x}) = f_i(x_1, \ldots, x_n)$.

It may help to see an example. Take $n = 3$ and $m = 2$. Then the equation says

$$\begin{aligned} u_1 &= f_1(x_1, x_2, x_3) \\ u_2 &= f_2(x_1, x_2, x_3). \end{aligned} \quad (1.69)$$

In many practical situations this is written without subscripts, for example

$$\begin{aligned} u &= f(x,y,z) \\ v &= g(x,y,z). \end{aligned} \tag{1.70}$$

Differential calculus has an important generalization to this situation. The starting point is the matrix of partial derivatives $\mathbf{f}'(\mathbf{x})$. We take E to be an open subset of $\mathbf{R}^n$. For each $\mathbf{x}$ in E we have the matrix of partial derivatives $\mathbf{f}'(\mathbf{x})$ with components

$$f'(\mathbf{x})_{i,j} = \frac{\partial u_i}{\partial x_j} = \frac{\partial f_i(\mathbf{x})}{\partial x_j}. \tag{1.71}$$

Each *partial derivative* is just an ordinary derivative in the situation when all variables but one are held constant during the limiting process. The matrix of partial derivatives is also called the *Jacobian matrix*.

Again it may help to look at an example. We have

$$\mathbf{f}'(\mathbf{x}) = \begin{bmatrix} f'_{1,1}(\mathbf{x}) & f'_{1,2}(\mathbf{x}) & f'_{1,3}(\mathbf{x}) \\ f'_{2,1}(\mathbf{x}) & f'_{2,2}(\mathbf{x}) & f'_{2,3}(\mathbf{x}) \end{bmatrix}. \tag{1.72}$$

An alternate notation using variables might be

$$\frac{\partial u_1, \partial u_2}{\partial x_1, \partial x_2, \partial x_3} = \begin{bmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} & \frac{\partial u_1}{\partial x_3} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} & \frac{\partial u_2}{\partial x_3} \end{bmatrix}. \tag{1.73}$$

The reader should be warned that a notation like the one on the left is often used in situations where the matrix is a square matrix; in that case some authors use it to denote the matrix, others use it to denote the determinant of the matrix.

If $u = f(x,y,z)$ and $v = g(x,y,z)$ we can also write

$$\begin{bmatrix} du \\ dv \end{bmatrix} = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \end{bmatrix} \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix}. \tag{1.74}$$

One way to interpret this equation is to think of x, y, z as functions of a parameter t. Then replace dx, dy, dz by $dx/dt, dy/dt, dz/dt$ and du, dv by $du/dt, dv/dt$.

Of course one can dispense with matrix notation altogether and write

$$\begin{aligned} du &= \frac{\partial u}{\partial x}\,dx + \frac{\partial u}{\partial y}\,dy + \frac{\partial u}{\partial z}\,dz \\ dv &= \frac{\partial v}{\partial x}\,dx + \frac{\partial v}{\partial y}\,dz + \frac{\partial v}{\partial z}\,dz. \end{aligned} \tag{1.75}$$

The existence of the matrix of partial derivatives does not fully capture the notion of differentiability. Here is the actual definition of the *derivative* $\mathbf{f}'(\mathbf{x})$. Write

$$\mathbf{f}(\mathbf{x}+\mathbf{h}) = \mathbf{f}(\mathbf{x}) + \mathbf{f}'(\mathbf{x})\mathbf{h} + \mathbf{r}(\mathbf{x},\mathbf{h}). \tag{1.76}$$

Thus the function is written as the value at $\mathbf{x}$ plus a linear term (given by multiplying a matrix times a vector) plus a remainder term. The requirement on the remainder term is that

$$\frac{|\mathbf{r}(\mathbf{x},\mathbf{h})|}{|\mathbf{h}|} \to 0 \tag{1.77}$$

as $\mathbf{h} \to 0$. Sometimes it is easier to deal with this with an equivalent definition that does not involve fractions:

$$|\mathbf{r}(\mathbf{x},\mathbf{h})| \le \epsilon(\mathbf{x},\mathbf{h})|\mathbf{h}|, \tag{1.78}$$

where $\epsilon(\mathbf{x},\mathbf{h}) \to 0$ as $\mathbf{h} \to 0$. In other words, the remainder term has to be of higher order, in this precise sense of being bounded by the norm of $\mathbf{h}$ with a coefficient that goes to zero. From this definition one can conclude that if the derivative $\mathbf{f}'(\mathbf{x})$ exists in this sense, then function is continuous and the matrix of partial derivatives exists and coincides with $\mathbf{f}'(\mathbf{x})$. Since the derivative provides a linear transformation that approximates changes in the function, it is often called the *linearization* of the function.

Example: There are examples where the matrix of partial derivatives exists and yet the derivative does not exist. Take $f(x,y) = xy/(x^2+y^2)$ away from the origin, $f(0,0) = 0$. Away from the origin it is easy to compute the partial derivatives $\partial f(x,y)/\partial x = xy(y-x)/(x^2+y^2)^2$ and $\partial f(x,y)/\partial y = xy(x-y)/(x^2+y^2)^2$. Since $f(x,0) = 0$ and $f(0,y) = 0$, it follows that the two partial derivatives exist also at the origin. However the function is discontinuous at the origin. |

There is a theorem that says that if the matrix of partial derivatives exists and is continuous, then the derivative exists. This theorem is proved below. In most cases we shall assume that the conditions of the theorem are satisfied. There is a very convenient notation that is used in such contexts. A function defined on an open set is called a C^k *function* if all partial derivatives up to order k exist and are continuous. Thus a continuous function is a C^0 function, and a differentiable function with continuous derivative is a C^1 function. Often in theoretical work it is convenient to assume that one is dealing with a C^1 function, or even a C^2 function. So as not to worry too much about technicalities, it is quite common to assume that the function under consideration is a C^∞ *function*. This just means that it is in C^k for every $k = 0, 1, 2, 3, \ldots$. Often a C^∞ function is called a *smooth function*. In theoretical discussions one often has various things to worry about other than exactly how differentiable a function has to be. So it is common to limit the discussion to smooth functions. Sometimes one instead talks about "sufficiently smooth" function, which just means C^k with k large enough for the purpose at hand.

The notion of differentiable function is very general, and it is revealing to see how various cases apply to various important situations. First, consider the case when $n < m$. In that case it is common to think of the set of $\mathbf{u} = \mathbf{f}(\mathbf{x})$ for $\mathbf{x}$ in the domain to be a *parameterized surface* in $\mathbf{R}^m$ of dimension n. Suppose that the rank of $\mathbf{f}'(\mathbf{x})$ is n. (The rank is the dimension of the range, spanned

by the columns.) For each $\mathbf{x}$ one can look at the *tangent space* to the surface at $\mathbf{u} = \mathbf{f}(\mathbf{x})$. The tangent space is n dimensional and has the parametric form $\mathbf{f}(\mathbf{x}) + \mathbf{f}'(\mathbf{x})\mathbf{h}$, where $\mathbf{h}$ is in $\mathbf{R}^n$ and parameterizes the tangent space.

In the special case when $n = 1$ the surface is actually just a curve. The derivative at a point is the tangent vector. The tangent space at a point on the curve is a line tangent to the point. For example, take the case when $m = 2$ and the curve is given by

$$\begin{aligned} u &= \cos(t) \\ v &= \sin(t) \end{aligned} \tag{1.79}$$

This is a parametric representation of a circle. The tangent to the circle at a particular point is given by

$$\begin{aligned} \bar{u} &= \cos(t) - \sin(t)k \\ \bar{v} &= \sin(t) + \cos(t)k, \end{aligned} \tag{1.80}$$

where k is an arbitrary real constant.

When $n = 2$ one gets a parameterized surface. Here the classic example with $m = 3$ is

$$\begin{aligned} u &= \sin(s)\cos(t) \\ v &= \sin(s)\sin(t) \\ w &= \cos(s) \end{aligned} \tag{1.81}$$

This describes a spherical surface. The angle s is the co-latidude and the angle t is the longitude. The tangent plane is given by two parameters h, k as

$$\begin{aligned} \bar{u} &= \sin(s)\cos(t) + \cos(s)\cos(t)h - \sin(s)\sin(t)k \\ \bar{v} &= \sin(s)\sin(t) + \cos(s)\sin(t)h + \sin(s)\cos(t)k \\ \bar{w} &= \cos(s) - \sin(s)h. \end{aligned} \tag{1.82}$$

Next, consider the case when $m < n$. In this case the set of $\mathbf{x}$ with $\mathbf{f}(\mathbf{x}) = \mathbf{c}$ can be the *implicit* definition of a surface of dimension $n - m$ in $\mathbf{R}^n$. Suppose that $\mathbf{f}'(\mathbf{x})$ has rank m. Then the tangent space to the surface at a point $\mathbf{x}$ should have dimension $n - m$. It should consist of the points $\bar{\mathbf{x}}$ such that $\mathbf{f}'(\mathbf{x})(\bar{\mathbf{x}} - \mathbf{x}) = 0$.

When $m = 1$ the surface has dimension $n - 1$ and is called a *hypersurface*. If $u = f(x_1, \ldots, x_n)$, then the derivative $\partial u/\partial x_i = f'_{,i}(x_1, \ldots, x_n)$ is a covector. We often write this in the differential notation as

$$du = \frac{\partial u}{\partial x_1}\, dx_1 + \cdots + \frac{\partial u}{\partial x_n}\, dx_n. \tag{1.83}$$

A simple example is a sphere given by

$$x^2 + y^2 + z^2 = 1. \tag{1.84}$$

The differential of the left hand side is $2x\,dx + 2y\,dy + 2z\,dz$. The tangent plane at a point is found by solving the equation

$$2x(\bar{x} - x) + 2y(\bar{y} - y) + 2z(\bar{z} - z) = 0. \tag{1.85}$$

When $m < n$ it turns out that it is often possible to convert an implicit representation $\mathbf{f}(\mathbf{x}) = \mathbf{c}$ into a parametric representation. Information on how to do this is given by the implicit function theorem.

Yet another story is when $m = n$. In applications the equation $\mathbf{u} = \mathbf{f}(\mathbf{x})$ can have several interpretations. In the *passive* interpretation some object is described by coordinates $\mathbf{x}$, and it is found convenient to have an alternative description in terms of new coordinates $\mathbf{y}$. In this case one would like to be able to solve for the $\mathbf{x}$ in terms of the $\mathbf{u}$, so that the two descriptions are equivalent. In particular, one would like each matrix $\mathbf{f}'(\mathbf{x})$ to have an inverse matrix. We shall see that the inverse function theorem gives information that applies to this situation.

Example: A typical passive operation is the change from polar coordinates to cartesian coordinates. This is just

$$\begin{aligned} x &= r\cos(\phi) \\ y &= r\sin(\phi). \end{aligned} \tag{1.86}$$

A given point has two descriptions. |

In the *active* interpretation there is only one coordinate system. The function $\mathbf{f}(\mathbf{x})$ then describes how the state of the object is changed. This process can be iterated, so that $\mathbf{x}$ at the next stage becomes $\mathbf{f}(\mathbf{x})$ from the previous stage. In this situation fixed points play an important role. If $\mathbf{x}^*$ is a fixed point with $\mathbf{f}(\mathbf{x}^*) = \mathbf{x}^*$, then the eigenvalues of $\mathbf{f}'(\mathbf{x}^*)$ given important information about the stability of the fixed point.

Example: Here is an example called the Hénon map. It occurred in a study of an extremely simplified model of atmospheric motion. It may be viewed as the composition of two simpler maps. The first is

$$\begin{aligned} p &= x \\ q &= 1 - ax^2 + y. \end{aligned} \tag{1.87}$$

Here $a > 0$ is a parameter. This particular transformation preserves area. Combine this with the transformation

$$\begin{aligned} u &= q \\ v &= bp. \end{aligned} \tag{1.88}$$

Here b is a parameter with $0 < b < 1$, representing some kind of contraction or dissipation. It decreases area, but in a simple way. The combined transformation is the Hénon map

$$\begin{aligned} u &= 1 - ax^2 + y \\ v &= bx. \end{aligned} \tag{1.89}$$

It may be thought of as a prediction of the future state of the system from the present state. Notice that this is an active operation; the state changes. It is possible to iterate this map many times, in an attempt to predict the state far into the future. This example has been the subject of much research. It turns out that reliable prediction far into the future is quite difficult. |

The *composition* of two functions $\mathbf{g}$ and $\mathbf{f}$ is the function $(\mathbf{g} \circ \mathbf{f})$ defined by $(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x}))$. The *chain rule* describes the derivative of such a function.

Theorem 1.15 (Chain rule) *Suppose that the derivatives* $\mathbf{f}'(\mathbf{x})$ *and* $\mathbf{g}'(\mathbf{f}(\mathbf{x}))$ *exist. Then*

$$(\mathbf{g} \circ \mathbf{f})'(\mathbf{x}) = \mathbf{g}'(\mathbf{f}(\mathbf{x}))\mathbf{f}'(\mathbf{x}). \tag{1.90}$$

The left hand is the derivative of the composition of the two functions, while the right hand side is the matrix product representing the composition of their derivatives, evaluated at the appropriate points.

Proof: Suppose $\mathbf{f}(\mathbf{x}+\mathbf{h}) = \mathbf{f}(\mathbf{x})+\mathbf{f}'(\mathbf{x})\mathbf{h}+\mathbf{r}(\mathbf{x},\mathbf{h})$ with $|\mathbf{r}(\mathbf{x},\mathbf{h})| \leq \epsilon(\mathbf{x},\mathbf{h})|\mathbf{h}|$. Similarly, suppose $\mathbf{g}(\mathbf{u}+\mathbf{k}) = \mathbf{g}(\mathbf{u})+\mathbf{g}'(\mathbf{u})\mathbf{k}+\mathbf{s}(\mathbf{u},\mathbf{k})$ with $|\mathbf{s}(\mathbf{u},\mathbf{k})| \leq \eta(\mathbf{u},\mathbf{k})|\mathbf{k}|$. Take $\mathbf{u} = \mathbf{f}(\mathbf{x})$ and $\mathbf{k} = \mathbf{f}'(\mathbf{x})\mathbf{h}$. Then

$$\mathbf{g}(\mathbf{f}(\mathbf{x}+\mathbf{h})) = \mathbf{g}(\mathbf{u}+\mathbf{k}+\mathbf{r}(\mathbf{x},\mathbf{h})) = \mathbf{g}(\mathbf{u})+\mathbf{g}'(\mathbf{u})\mathbf{k}+\mathbf{g}'(\mathbf{u})\mathbf{r}(\mathbf{x},\mathbf{h})+\mathbf{s}(\mathbf{u},\mathbf{k}+\mathbf{r}(\mathbf{x},\mathbf{h})). \tag{1.91}$$

We need to show that the two remainder terms are appropriately small. First,

$$|\mathbf{g}'(\mathbf{u})\mathbf{r}(\mathbf{x},\mathbf{h})| \leq \|\mathbf{g}'(\mathbf{u})\|\epsilon(\mathbf{x},\mathbf{h})|\mathbf{h}|. \tag{1.92}$$

Second,

$$s(\mathbf{u},\mathbf{k}+\mathbf{r}(\mathbf{x},\mathbf{h})) \leq \eta(\mathbf{u},\mathbf{k}+\mathbf{r}(\mathbf{x},\mathbf{h}))|\mathbf{k}+\mathbf{r}(\mathbf{x},\mathbf{h})| \leq \eta(\mathbf{u},\mathbf{f}'(\mathbf{x})\mathbf{h}+\mathbf{r}(\mathbf{x},\mathbf{h}))(\|\mathbf{f}'(\mathbf{x})\|+\epsilon(\mathbf{x},\mathbf{h}))|\mathbf{h}|. \tag{1.93}$$

□

The chain rule has various important consequences. For instance, in the case when $m = n$ it is possible that $\mathbf{f}$ has an inverse function $\mathbf{g}$ such that $\mathbf{f}(\mathbf{g}(\mathbf{y})) = \mathbf{y}$. It follows from the chain rule that

$$\mathbf{g}'(\mathbf{y}) = \mathbf{f}'(\mathbf{g}(\mathbf{y}))^{-1}. \tag{1.94}$$

In other words, the derivative of an inverse function is obtained by taking the inverse of a matrix. In this case when the matrix is square, the Jacobian matrix has a determinant, called the *Jacobian determinant*. It follows that when the function has an inverse, the Jacobian determinant is non-zero.

Another nice consequence of the chain rule is the mean value theorem. There is a problem with the mean value theorem for vector-valued functions. One would like a theorem of the following form: There exists a t with $0 \leq t \leq 1$ and

$$\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) = \mathbf{f}'((1-t)\mathbf{x} + t\mathbf{y})\,(\mathbf{y} - \mathbf{x}). \tag{1.95}$$

This would say that there is a point on the segment between $\mathbf{x}$ and $\mathbf{y}$ where the derivative accurately predicts the change. But this can be false!

The following is a statement of a true version of the *mean value theorem*. The idea is to average over the segment. (The hypothesis of this particular version is that $\mathbf{f}'(\mathbf{x})$ not only exists but is continuous in $\mathbf{x}$. A version requiring differentiability but not continuous differentiability may be found in Rudin.)

Theorem 1.16 (Mean value theorem) *Let E be an open convex set in $\mathbf{R}^n$, and let $\mathbf{f}(\mathbf{x})$ be differentiable with derivative $\mathbf{f}'(\mathbf{x})$ continuous in $\mathbf{x}$ in E. Then the difference in function values is predicted by the average of the derivative over the segment:*

$$\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) = \int_0^1 \mathbf{f}'((1-t)\mathbf{x} + t\mathbf{y})\, dt\, (\mathbf{y} - \mathbf{x}). \tag{1.96}$$

Furthermore, suppose that for all $\mathbf{z}$ in E we have $\|\mathbf{f}'(\mathbf{z})\| \leq M$. Then it follows that

$$|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})| \leq M|\mathbf{y} - \mathbf{x}|. \tag{1.97}$$

Proof: Use the fundamental theorem of calculus and the chain rule to compute

$$\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) = \int_0^1 \frac{d}{dt}\mathbf{f}((1-t)\mathbf{x} + t\mathbf{y})\, dt = \int_0^1 \mathbf{f}'((1-t)\mathbf{x} + t\mathbf{y})(\mathbf{y} - \mathbf{x})\, dt. \tag{1.98}$$

The continuous differentiability assumption guarantees that the integrand is continuous in t. This gives the identity.

From the identity we get

$$|f(\mathbf{y}) - f(\mathbf{x})| \leq \int_0^1 \|\mathbf{f}'((1-t)\mathbf{x} + t\mathbf{y})\||\mathbf{y} - \mathbf{x}|\, dt \leq \int_0^1 M|\mathbf{y} - \mathbf{x}|\, dt. \tag{1.99}$$

But the integrand in the last integral does not depend on t. So this is just $M|\mathbf{y} - \mathbf{x}|$, as in the theorem. □

The mean value theorem idea also works to prove the result about continuous partial derivatives.

Theorem 1.17 *Consider a function $\mathbf{f}(\mathbf{x})$ that is defined and continuous on some open set. Then $\mathbf{f}'(\mathbf{x})$ exists and is continuous in $\mathbf{x}$ if and only if the partial derivatives exist and are continuous in $\mathbf{x}$.*

Proof: It is evident that if the derivative exists and is continuous, then the partial derivatives exist and are continuous. All the work is to go the other way.

The existence and continuity of the partial derivatives implies the following statement. Let $\mathbf{z}$ be in the open set on which the partial derivatives exist and are continuous. Let $\mathbf{h}$ be a vector in one of the coordinate directions. Then $d\mathbf{f}(\mathbf{z} + t\mathbf{h})/dt = \mathbf{f}'(\mathbf{z} + t\mathbf{h})\mathbf{h}$ exists for sufficiently small t, and the matrix of

partial derivatives $\mathbf{f}'(\mathbf{z})$ is continuous in $\mathbf{z}$. From here on it is sufficient to work with this hypothesis.

We need to examine $\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x})$ when $\mathbf{h}$ is not in one of the coordinate directions. Denote the orthogonal projection of $\mathbf{h}$ on the ith coordinate direction by $\mathbf{h}_{(i)}$. Write $\mathbf{h}_{[i]} = \sum_{j=1}^{i} \mathbf{h}_{(j)}$. Then

$$\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) = \sum_{i=1}^{n} [\mathbf{f}(\mathbf{x}+\mathbf{h}_{[i-1]}+\mathbf{h}_{(i)}) - \mathbf{f}(\mathbf{x}+\mathbf{h}_{[i-1]})] \tag{1.100}$$

This represents the total change as the sum of changes resulting from incrementing one coordinate at a time. We can use the fundamental theorem of calculus to write this as

$$\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) = \sum_{i=1}^{n} \int_0^1 \mathbf{f}'(\mathbf{x}+\mathbf{h}_{[i-1]}+t\mathbf{h}_{(i)})\mathbf{h}_{(i)}\, dt. \tag{1.101}$$

Notice that each term in the sum only involves one coordinate direction. Furthermore each integrand is continuous in t. It follows that

$$\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h} = \sum_{i=1}^{n} \int_0^1 [\mathbf{f}'(\mathbf{x}+\mathbf{h}_{[i-1]}+t\mathbf{h}_{(i)}) - \mathbf{f}'(\mathbf{x})]\mathbf{h}_{(i)}\, dt. \tag{1.102}$$

Hence

$$|\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h}| \le \sum_{i=1}^{n} \int_0^1 \|\mathbf{f}'(\mathbf{x}+\mathbf{h}_{[i-1]}+t\mathbf{h}_{(i)}) - \mathbf{f}'(\mathbf{x})\|\, dt |\mathbf{h}|. \tag{1.103}$$

Then it is not difficult to show that

$$\epsilon(\mathbf{x},\mathbf{h}) = \sum_{i=1}^{n} \int_0^1 \|\mathbf{f}'(\mathbf{x}+\mathbf{h}_{[i-1]}+t\mathbf{h}_{(i)}) - \mathbf{f}'(\mathbf{x})\|\, dt \to 0 \tag{1.104}$$

as $\mathbf{h} \to 0$. This is seen as follows. For each t between 0 and 1 we have the estimate $|\mathbf{h}_{[i-1]} + t\mathbf{h}_{(i)}| \le |\mathbf{h}|$. If we take $|\mathbf{h}|$ small enough, then from the continuity assumption each integrand $\|\mathbf{f}'(\mathbf{x}+\mathbf{h}_{[i-1]}+t\mathbf{h}_{(i)}) - \mathbf{f}'(\mathbf{x})\|$ is small uniformly in t. So each integral is small. □

Here are some chain rule examples. Let $\mathbf{u} = \mathbf{f}(\mathbf{x})$ and $\mathbf{p} = \mathbf{g}(\mathbf{u})$, so the composition is $\mathbf{p} = \mathbf{g}(\mathbf{f}(\mathbf{x}))$. The chain rule says

$$(\mathbf{g}\circ\mathbf{f})'(\mathbf{x}) = \begin{bmatrix} g'_{1,1}(\mathbf{u}) & g'_{1,2}(\mathbf{u}) \\ g'_{2,1}(\mathbf{u}) & g'_{2,2}(\mathbf{u}) \end{bmatrix} \begin{bmatrix} f'_{1,1}(\mathbf{x}) & f'_{1,2}(\mathbf{x}) & f'_{1,3}(\mathbf{x}) \\ f'_{2,1}(\mathbf{x}) & f'_{2,2}(\mathbf{x}) & f'_{2,3}(\mathbf{x}) \end{bmatrix}. \tag{1.105}$$

We can also write

$$\begin{bmatrix} \frac{\partial p}{\partial x} & \frac{\partial p}{\partial y} & \frac{\partial p}{\partial z} \\ \frac{\partial q}{\partial x} & \frac{\partial q}{\partial y} & \frac{\partial q}{\partial z} \end{bmatrix} = \begin{bmatrix} \frac{\partial p}{\partial u} & \frac{\partial p}{\partial v} \\ \frac{\partial q}{\partial v} & \frac{\partial q}{\partial v} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z}. \end{bmatrix}. \tag{1.106}$$

This matrix notation is just another way of writing six equations. For example, one of these equations is

$$\frac{\partial q}{\partial y} = \frac{\partial q}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial q}{\partial v}\frac{\partial v}{\partial y}. \tag{1.107}$$

There are several ambiguities in such an expression. The first is that if q is a quantity that depends on u and some other variables, then $\partial q/\partial u$ depends on the choice of variables. If we specify $q = g_2(u, v)$, and define $\partial q/\partial u = g'_{2,1}(u, v)$, then there is no problem. But if there were some other coordinate system u, w for which u is one coordinate, then the function relating q to u, w might be quite different from the function relating q to u, v. Thus $\partial q/\partial u$ holdig w constant is quite different from $\partial q/\partial u$ holding v constant. For this reason many scientists use a notation such as $\partial q/\partial u|v = \text{const}$ or just $\partial q/\partial u|v$.
Example: Consider the situation where

$$p = h(x, g(x, y)). \tag{1.108}$$

Then

$$\frac{\partial p}{\partial x} = h'_{,1}(x, g(x, y)) + h'_{,2}(x, g(x, y))g'_{,1}(x, y). \tag{1.109}$$

When the functional relationships are specified there is no ambiguity. However this could also be written with $p = h(x, v)$, $v = g(x, y)$ and hence $p = h(x, g(x, y))$. Then

$$\frac{\partial p}{\partial x} = \frac{\partial p}{\partial x} + \frac{\partial p}{\partial v}\frac{\partial v}{\partial x}. \tag{1.110}$$

Now the problem is evident: the expression $\partial p/\partial x$ is ambiguous, at least until it is made clear what other variable or variables are held constant. If we indicate the variable that is held constant with a subscript, we get a more informative equation

$$\frac{\partial p}{\partial x}|_y = \frac{\partial p}{\partial x}|_v + \frac{\partial p}{\partial v}|_x \frac{\partial v}{\partial x}|_y. \tag{1.111}$$

|

In general, if $p = h(u, v)$, $u = f(x, y)$, $v = g(x, y)$, a more precise notation for partial derivatives should be

$$\frac{\partial p}{\partial x}|_y = \frac{\partial p}{\partial u}|_v \frac{\partial u}{\partial x}|_y + \frac{\partial p}{\partial v}|_u \frac{\partial v}{\partial x}|_y. \tag{1.112}$$

In practice one usually does not indicate which variables are held constant unless there is risk of confusion. But one should be clear that the partial derivative with respect to a variable depends on the entire coordinate system.

The second ambiguity is special to the chain rule. Say that $p = g_1(u, v)$, $q = g_2(u, v)$ and $u = f_1(x, y)$, $v = f_2(x, y)$. Then

$$\frac{\partial q}{\partial y} = g'_{2,1}(f_1(x, y), f_2(x, y))f'_{1,2}(x, y) + g'_{2,2}(f_1(x, y), f_2(x, y))f'_{2,2}(x, y). \tag{1.113}$$

This is usually written

$$\frac{\partial q}{\partial y} = \frac{\partial q}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial q}{\partial v}\frac{\partial v}{\partial y}. \tag{1.114}$$

The right hand side first differentiates q regarded as a function of u (holding v constant), and only after that one replaces u by $f_1(x, y)$ and v by $f_2(x, y)$. So a more precise notation might indicate that q is defined with the replacement $u \leftarrow f_1(x, y), v \leftarrow f_2(x, y)$ before the differentiation $\partial q/\partial y$. On the other hand, $\partial q/\partial u$ is defined with the same replacement, but after the differentiation $\partial q/\partial u$. It is uncommon to indicate this sort of replacement explicitly, but it should be kept in mind. In the passive interpretation it is hardly necessary, since in that case one regards $u = f_1(x, y), v = f_2(x, y)$. There may be other situations where it is necessary for clarity.

Remark: In virtually every application of calculus, the variables that are employed have specific meanings, and it is natural to write the formulas in terms of these variables. Even as strict an author as Rudin eventually introduces notations that employ variables. In multivariable calculus these notations are not entirely standard. One wants to use notations that will seem familiar to the mathematical reader, but that work together in a consistent way. Here are some suggestions.

If $\mathbf{y} = \mathbf{f}(\mathbf{x})$, then the m by n Jacobian matrix of partial derivatives may be denoted

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \mathbf{f}'(\mathbf{x}). \tag{1.115}$$

This can also be written out even more explicitly in the form

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \frac{\partial (y_1, \ldots, y_m)}{\partial (x_1, \ldots, x_n)}. \tag{1.116}$$

Warning: When $m = n$ some authors (including Rudin) use the notation on the right hand side for the determinant of the Jacobian matrix.

If $\mathbf{p} = \mathbf{g}(\mathbf{y})$, the chain rule says that

$$\frac{\partial \mathbf{p}}{\partial \mathbf{x}} = \frac{\partial \mathbf{p}}{\partial \mathbf{y}}\frac{\partial \mathbf{y}}{\partial \mathbf{x}}, \tag{1.117}$$

The muliplication on the right is matrix multiplication.

If y is a scalar, then we often write

$$dy = \frac{\partial y}{\partial x_1}\, dx_1 + \cdots + \frac{\partial y}{\partial x_n}\, dx_n. \tag{1.118}$$

This will eventually have a rigorous definition in the context of differential forms. One possible meaning of $d\mathbf{x}$ is as a column vector of dx_i. In this setting we can write formulas such as

$$\frac{dy}{dt} = \frac{\partial y}{\partial \mathbf{x}}\frac{d\mathbf{x}}{dt}. \tag{1.119}$$

On the right hand side this is a row covector times a column vector.

There is another meaning for $d\mathbf{x}$. This is as a formula that occurs in integrands:

$$d\mathbf{x} = dx_1 \cdots dx_n = dx_1 \wedge \cdots \wedge dx_n. \tag{1.120}$$

The product $\wedge$ is the exterior product, to be explained in the chapter on differential forms. We shall see that when $m = n$ it is natural to write

$$\frac{d\mathbf{y}}{d\mathbf{x}} = \det \frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \det \frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)}. \tag{1.121}$$

This is the Jacobian determinant that occurs in change of variable formulas. (Experts may note that this is suggestive of the notation for the Radon-Nikodym derivative that occurs in the theory of measure and integration.)

|

1.6 Fixed point iteration (multivariable)

Proposition 1.18 *Let $\mathbf{p}$ be in $\mathbf{R}^n$ and consider the closed ball of radius r about $\mathbf{p}$. Suppose $\mathbf{g}(\mathbf{x})$ is defined and continuous for $\mathbf{x}$ in some open set including this ball and has values in $\mathbf{R}^n$. Suppose that $\|\mathbf{g}'(\mathbf{x})\| \leq c < 1$ for $\mathbf{x}$ in this set. Furthermore, suppose that $|\mathbf{g}(\mathbf{p}) - \mathbf{p}| \leq (1 - c)r$. Then $\mathbf{g}$ maps the ball into itself and is a strict contraction, so it has a unique fixed point in this ball. Furthermore, iteration starting in this ball converges to the fixed point.*

Proof: From the mean value theorem it follows that $|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y})| \leq c|\mathbf{x} - \mathbf{y}|$. In order to show that g maps the ball into itself, suppose $|\mathbf{x} - \mathbf{p}| \leq r$. Then $|\mathbf{g}(\mathbf{x}) - \mathbf{p}| \leq |\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{p})| + |\mathbf{g}(\mathbf{p}) - \mathbf{p}| \leq c|\mathbf{x} - \mathbf{p}| + (1 - c)r \leq r$. □

Sometimes it is helpful have a result where one knows there is a fixed point, but wants to show that it is stable, in the sense that fixed point iteration starting near the fixed point converges to it. The following proposition is a variant that captures this idea.

Proposition 1.19 *Let $\mathbf{p}$ be a fixed point of $\mathbf{g}$. Suppose that $\mathbf{g}'$ is continuous and that $\|\mathbf{g}'(\mathbf{p})\| < 1$. Then for c with $\|\mathbf{g}'(\mathbf{p})\| < c < 1$ there is an $r > 0$ such that $\|\mathbf{g}'(\mathbf{x})\| \leq c < 1$ for $\mathbf{x}$ satisfying $|\mathbf{x} - \mathbf{p}| < 1$. Then $\mathbf{g}$ maps the interval into itself and is a strict contraction. Furthermore, iterates starting in this interval converge to the fixed point.*

In the multidimensional case this result need not give a particularly good account of stability, since the stability should be established by the spectral radius $\rho(\mathbf{g}'(\mathbf{p}))$, and the norm $\|\mathbf{g}'(\mathbf{p})\|$ can be much larger. So the following result is better.

Proposition 1.20 *Let $\mathbf{p}$ be a fixed point of $\mathbf{g}$. Suppose that $\mathbf{g}'$ is continuous and that the spectral radius $\rho(\mathbf{g}'(\mathbf{p})) < 1$. Then there is a closed ellipsoid centered on $\mathbf{r}$ such that $\mathbf{g}$ maps the ellipsoid into itself and iterates starting in this ellipsoid converge to the fixed point.*

Proof: For every $\delta > 0$ there is a new norm $|\mathbf{x}|_G$ so that $\|\mathbf{g}'(\mathbf{p}))\|_G \leq \rho(\mathbf{g}'(\mathbf{p})) + \delta$. Since $\rho(\mathbf{g}'(\mathbf{p})) < 1$, we can pick the norm so that $\|\mathbf{g}'(\mathbf{p}))\|_G < 1$. Then the continuity of $\mathbf{g}'(\mathbf{x})$ in $\mathbf{x}$ shows that for c with $\|\mathbf{g}'(\mathbf{p})\|_G < c < 1$ there is an $r > 0$ such that $|\mathbf{x} - \mathbf{p}|_G \leq r$ implies $\|\mathbf{g}'(\mathbf{x})\|_G \leq c < 1$. Then $\mathbf{g}$ maps this ball into itself and is a strict contraction. Furthermore, iterates starting in this ball converge to the fixed point. Of course with respect to the original Euclidean norm this ball is an ellipsoid. □

When $n = 2$ there is a lovely way of picturing the function and the iteration process. The idea is to plot vectors $\mathbf{g}(\mathbf{x}) - \mathbf{x}$. A sequence of such vectors with the tail of the next one equal to the tip of the previous one indicates the orbit. The function itself may be pictured by drawing representative orbits. Near a fixed point $\mathbf{p}$ the function $\mathbf{g}(\mathbf{x})$ is close to $\mathbf{g}(\mathbf{p})+\mathbf{g}'(\mathbf{p})(\mathbf{x}-\mathbf{p}) = \mathbf{p}+\mathbf{g}'(\mathbf{p})(\mathbf{x}-\mathbf{p})$. Thus $\mathbf{g}(\mathbf{x}) - \mathbf{p}$ is close to $\mathbf{g}'(\mathbf{p})(\mathbf{x} - \mathbf{p})$, and so the picture resembles the picture for the linear transformation $\mathbf{g}'(\mathbf{p})$. In particular, the eigenvalues give insight into the behavior that is expected.

Example: Define a function by

$$\begin{aligned} u &= f(x,y) = \frac{1}{2}(x^2 - y^2) + \frac{1}{2} \\ v &= g(x,y) = xy + \frac{1}{4} \end{aligned} \tag{1.122}$$

This has a fixed point where x and y are both equal to $1/2$. The linearization at the fixed point is

$$\begin{bmatrix} x & -y \\ y & x \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \tag{1.123}$$

This has eigenvalues $\frac{1}{2} \pm \frac{1}{2}i$. The 2 norm is 1, which is at first alarming, but this is an overestimate. The Lipschitz norm is $\frac{\sqrt{2}}{2} < 1$, so the fixed point is certainly stable. In this special case the spectral radius is equal to the Lipschitz norm, so it gives no additional information. |

Example: Here again is a simple computer program in R.

```
f <- function (x,y) (x^2 - y^2)/2 + 1/2
g <- function (x,y) x * y + 1/4
x <- 1
y <- 0
for (i in 1:20) {
u <- f(x,y)
v <- g(x,y)
x <- u
y <- v }
x
[1] 0.4998379
y
[1] 0.5015273 |
```

Fixed point iteration gives a rather general way of solving equations $\mathbf{f}(\mathbf{x}) = 0$. If A is an arbitrary non-singular matrix, then the fixed points of

$$\mathbf{g}(\mathbf{x}) = \mathbf{x} - A^{-1}\mathbf{f}(\mathbf{x}) \tag{1.124}$$

are the solutions of the equation. The trick is to pick A such that $\mathbf{g}$ is a strict contraction. However,

$$\mathbf{g}'(\mathbf{x}) = I - A^{-1}\mathbf{f}'(\mathbf{x}), \tag{1.125}$$

so the strategy is to take A close to values of $\mathbf{f}'(\mathbf{x})$ near a point where $\mathbf{f}(\mathbf{x})$ is close to zero. This is illustrated in the following result, which is a reformulation of a previous result.

Proposition 1.21 *Let $\mathbf{p}$ be a real number, and consider the closed ball $|\mathbf{x}-\mathbf{p}| \leq r$ with $r > 0$. Suppose that for some c with $0 < c < 1$ we have $\|I - A^{-1}\mathbf{f}'(\mathbf{x})\| \leq c$ for $\mathbf{x}$ in this interval. Furthermore, suppose that $|A^{-1}\mathbf{f}(\mathbf{p})| \leq (1-c)r$. Then the corresponding fixed point iteration starting in this interval converges to the solution of $\mathbf{f}(\mathbf{x}) = 0$ in this interval.*

1.7 The implicit function theorem (multivariable)

Theorem 1.22 (Implicit function theorem) *Let $m \leq n$ and let $\mathbf{f}(\mathbf{x}, \mathbf{y})$ be a function from $\mathbf{R}^n$ to $\mathbf{R}^m$. Here $\mathbf{x}$ is in $\mathbf{R}^m$ and $\mathbf{y}$ is in $\mathbf{R}^{n-m}$. Suppose $\mathbf{f}$ is continuous with continuous derivative near some point $\mathbf{x} = \mathbf{a}, \mathbf{y} = \mathbf{b}$. Let $I = \{1, \ldots, m\}$ and $J = \{m+1, \ldots, n\}$ be the indices corresponding to $\mathbf{x}$ and $\mathbf{y}$. Suppose that the m by m matrix $\mathbf{f}'_I(\mathbf{a}, \mathbf{b})$ has an inverse matrix. Then there is a continuous function $\mathbf{h}(\mathbf{y})$ defined for $\mathbf{y}$ near $\mathbf{b}$ with $\mathbf{f}(\mathbf{h}(\mathbf{y}), \mathbf{y}) = 0$. In fact, the function $\mathbf{h}(\mathbf{y})$ has a continuous derivative, and*

$$0 = \mathbf{f}'_I(\mathbf{h}(\mathbf{y}), \mathbf{y})\mathbf{h}'(\mathbf{y}) + \mathbf{f}'_J(\mathbf{h}(\mathbf{y}), \mathbf{y}). \tag{1.126}$$

The equation in the statement of the theorem may be solved to give

$$\mathbf{h}'(\mathbf{y}) = -\mathbf{f}'_I(\mathbf{h}(\mathbf{y}), \mathbf{y})^{-1}\mathbf{f}'_J(\mathbf{h}(\mathbf{y}), \mathbf{y}). \tag{1.127}$$

Proof: Let $A = \mathbf{f}'_I(\mathbf{a}, \mathbf{b})$. Consider the iteration function

$$\mathbf{g}(\mathbf{x}, \mathbf{y}) = \mathbf{x} - A^{-1}\mathbf{f}(\mathbf{x}, \mathbf{y}). \tag{1.128}$$

This has partial derivative

$$\mathbf{g}'_I(\mathbf{x}, \mathbf{y}) = I - A^{-1}\mathbf{f}'_I(\mathbf{x}, \mathbf{y}). \tag{1.129}$$

At $\mathbf{x} = \mathbf{a}, \mathbf{y} = \mathbf{b}$ this is the zero matrix. Pick some convenient value of c with $0 < c < 1$. Then for $\mathbf{x}, \mathbf{y}$ sufficiently close to $\mathbf{a}, \mathbf{b}$ this partial derivative has absolute value bounded by c. In particular, this is true in some box $|\mathbf{x} - \mathbf{a}| \leq r$, $|\mathbf{y} - \mathbf{b}| < s$. In this box $\|\mathbf{g}'_I(\mathbf{x}, \mathbf{y})\| \leq c$, so by the mean value theorem

$|\mathbf{g}(\mathbf{x}'', \mathbf{y}) - g(\mathbf{x}', \mathbf{y})| \le c|\mathbf{x}'' - \mathbf{x}'|$. Fix $r > 0$. We know that $\mathbf{g}(\mathbf{a}, \mathbf{b}) - \mathbf{a} = 0$. Hence, if $s > 0$ is sufficiently small, $|\mathbf{g}(\mathbf{a}, \mathbf{y}) - \mathbf{a}| \le (1 - c)r$. We can put these results together to show that if $|\mathbf{x} - \mathbf{a}| \le r$, then

$$|\mathbf{g}(\mathbf{x}, \mathbf{y}) - \mathbf{a}| \le |g(\mathbf{x}, \mathbf{y}) - \mathbf{g}(\mathbf{a}, \mathbf{y})| + |\mathbf{g}(\mathbf{a}, \mathbf{y}) - \mathbf{a}| \le c|\mathbf{x} - \mathbf{a}| + (1 - c)r \le r. \quad (1.130)$$

So the map $\mathbf{x} \to \mathbf{g}(\mathbf{x}, \mathbf{y})$ is a contraction mapping that sends a complete metric space (the closed ball of $\mathbf{x}$ with $|\mathbf{x} - \mathbf{a}| \le r$) into itself. This shows that for all $\mathbf{y}$ with $|\mathbf{y} - \mathbf{b}| < s$ there is a fixed point $\mathbf{x} = \mathbf{h}(\mathbf{y})$. This proves that $\mathbf{g}(\mathbf{h}(\mathbf{y}), \mathbf{y}) = \mathbf{h}(\mathbf{y})$ and hence that $\mathbf{f}(\mathbf{h}(\mathbf{y}), \mathbf{y}) = 0$.

It is easy to see from the contraction mapping principle that the fixed point $\mathbf{h}(\mathbf{y})$ is a continuous function of the parameter $\mathbf{y}$. Consider $\mathbf{y}'$ near $\mathbf{y}$. Then $\mathbf{g}(\mathbf{h}(\mathbf{y}), \mathbf{y}) = \mathbf{h}(\mathbf{y})$ and $\mathbf{g}(\mathbf{h}(\mathbf{y}'), \mathbf{y}') = \mathbf{h}(\mathbf{y}')$. So $\mathbf{h}(\mathbf{y}') - \mathbf{h}(\mathbf{y}) = \mathbf{g}(\mathbf{h}(\mathbf{y}'), \mathbf{y}') - \mathbf{g}(\mathbf{h}(\mathbf{y}), \mathbf{y}) = \mathbf{g}(\mathbf{h}(\mathbf{y}'), \mathbf{y}') - \mathbf{g}(\mathbf{h}(\mathbf{y}), \mathbf{y}') + \mathbf{g}(\mathbf{h}(\mathbf{y}), \mathbf{y}') - \mathbf{g}(\mathbf{h}(\mathbf{y}), \mathbf{y})$. This gives $|\mathbf{h}(\mathbf{y}') - \mathbf{h}(\mathbf{y})| \le c|\mathbf{h}(\mathbf{y}') - \mathbf{h}(\mathbf{y})| + |\mathbf{g}(\mathbf{h}(\mathbf{y}), \mathbf{y}') - \mathbf{g}(\mathbf{h}(\mathbf{y}), \mathbf{y})|$. Write this as $(1 - c)|\mathbf{h}(\mathbf{y}') - \mathbf{h}(\mathbf{y})| \le |\mathbf{g}(\mathbf{h}(\mathbf{y}), \mathbf{y}') - \mathbf{g}(\mathbf{h}(\mathbf{y}), \mathbf{y})|$. Then as $\mathbf{y}' \to \mathbf{y}$ we have $\mathbf{g}(\mathbf{h}(\mathbf{y}), \mathbf{y}') \to \mathbf{g}(\mathbf{h}(\mathbf{y}), \mathbf{y})$, and hence $\mathbf{h}(\mathbf{y}') \to \mathbf{h}(\mathbf{y})$.

It remains to show that $\mathbf{h}$ has a continuous derivative. Let $\mathbf{u}$ be a fixed vector. The directional derivative $\mathbf{h}'(\mathbf{y})\mathbf{u}$ is a vector computed as the limit of the difference quotient $(\mathbf{h}(\mathbf{y} + k\mathbf{u}) - \mathbf{h}(\mathbf{y}))/k$ as $k \to 0$. In order to get a handle on this, compute

$$0 = \mathbf{f}(\mathbf{h}(\mathbf{y} + k\mathbf{u}), \mathbf{y} + k\mathbf{u}) - \mathbf{f}(\mathbf{h}(\mathbf{y}), \mathbf{y}). \quad (1.131)$$

Each term on the right is zero. Then expand

$$0 = \mathbf{f}(\mathbf{h}(\mathbf{y}+k\mathbf{u}), \mathbf{y}+k\mathbf{u}) - \mathbf{f}(\mathbf{h}(\mathbf{y}), \mathbf{y}) = \int_0^1 \frac{d}{dt}\mathbf{f}(t\mathbf{h}(\mathbf{y}+k\mathbf{u}) + (1-t)\mathbf{h}(\mathbf{y}), t(\mathbf{y}+k\mathbf{u}) + (1-t)\mathbf{y})\, dt. \quad (1.132)$$

By the chain rule for partial derivatives the right hand side is the sum of two terms, one from the I derivatives, and one from the J derivatives. The first term is given by $A(k)(\mathbf{h}(\mathbf{y} + k\mathbf{u}) - \mathbf{h}(\mathbf{y}))$, where

$$A(k) = \int_0^1 [\mathbf{f}_I'(t\mathbf{h}(\mathbf{y} + k\mathbf{u}) + (1 - t)\mathbf{h}(\mathbf{y}), t(\mathbf{y} + k\mathbf{u}) + (1 - t)\mathbf{y})]\, dt. \quad (1.133)$$

The second term is $B(k)k\mathbf{u}$, where

$$B(k) = \int_0^1 [\mathbf{f}_J'(t\mathbf{h}(\mathbf{y} + k\mathbf{u}) + (1 - t)\mathbf{h}(\mathbf{y}), t(\mathbf{y} + k\mathbf{u}) + (1 - t)\mathbf{y})]\, dt. \quad (1.134)$$

This gives the solution

$$\frac{\mathbf{h}(\mathbf{y} + k\mathbf{u}) - \mathbf{h}(\mathbf{y})}{k} = -A(k)^{-1}B(k)\mathbf{u}. \quad (1.135)$$

The quantities $A(k)$ and $B(k)$ are continuous in k, because of the continuity of $\mathbf{h}$. Now let $k \to 0$ and apply the dominated convergence theorem to justify

taking the limit inside the integral. The integrands become independent of k, and we obtain the desired formula for this vector, namely

$$\mathbf{h}'(\mathbf{y})\mathbf{u} = -A(0)^{-1}B(0)\mathbf{u} = -\mathbf{f}'_I(\mathbf{h}(\mathbf{y}),\mathbf{y})^{-1}\mathbf{f}'_J(\mathbf{h}(\mathbf{y}),\mathbf{y})\mathbf{u}. \tag{1.136}$$

That is, the directional derivative is the matrix

$$\mathbf{h}'(\mathbf{y}) = -\mathbf{f}'_I(\mathbf{h}(\mathbf{y}),\mathbf{y})^{-1}\mathbf{f}'_J(\mathbf{h}(\mathbf{y}),\mathbf{y}). \tag{1.137}$$

The right hand side of this is continuous in $\mathbf{y}$. This shows that the left hand side is continuous in $\mathbf{y}$. As a consequence, $\mathbf{h}(\mathbf{y})$ as a function of $\mathbf{y}$ is differentiable, and the derivative $\mathbf{h}'(\mathbf{y})$ as a function of $\mathbf{y}$ is continuous. □

The last part of the above proof seems complicated but is actually a straightforward application of the technique of the mean value theorem. It follows unpublished notes of Joel Feldman.

The implicit function theorem has a geometric interpretation. Consider the case $m < n$ and a function $\mathbf{f}(\mathbf{x},\mathbf{y})$ for $\mathbf{x}$ in $\mathbf{R}^m$ and $\mathbf{y}$ in $\mathbf{R}^{n-m}$, where the function values are in $\mathbf{R}^m$. A surface of dimension $n-m$ in $\mathbf{R}^n$ is given implicitly by $\mathbf{f}(\mathbf{x},\mathbf{y}) = \mathbf{c}$, that is, $\mathbf{f}(\mathbf{x},\mathbf{y}) - \mathbf{c} = 0$. The theorem says that we can write $\mathbf{x} = \mathbf{h}(\mathbf{y})$, where $\mathbf{h}$ is a function from $\mathbf{R}^{n-m}$ to $\mathbf{R}^m$, such that $\mathbf{f}(\mathbf{h}(\mathbf{y}),\mathbf{y}) = \mathbf{c}$. Thus $\mathbf{x} = \mathbf{h}(\mathbf{t}), \mathbf{y} = \mathbf{t}$ is a parametric representation of the surface.
Example: Consider the example from Rudin with

$$\begin{aligned} p = f_1(x,y,u,v,w) &= 2e^x + yu - 4v + 3 \\ q = f_2(x,y,u,v,w) &= y\cos(x) - 6x + 2u - w. \end{aligned} \tag{1.138}$$

The problem is to solve $f(x,y,u,v,w) = 0, g(x,y,u,v,w) = 0$ near (0,1,3,2,7) for x, y in terms of u, v, w. For this to be possible in principle, one needs to have x, y involved in the equation in a non-trivial way. It is sufficient to have the linearization

$$\begin{bmatrix} \frac{\partial p}{\partial x} & \frac{\partial p}{\partial y} \\ \frac{\partial q}{\partial x} & \frac{\partial q}{\partial y} \end{bmatrix} = \begin{bmatrix} 2e^x & u \\ -y\sin(x) - 6 & \cos(x) \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ -6 & 1 \end{bmatrix} \tag{1.139}$$

have an inverse. But the inverse is

$$\begin{bmatrix} 2 & 3 \\ -6 & 1 \end{bmatrix}^{-1} = \frac{1}{20}\begin{bmatrix} 1 & -3 \\ 6 & 2 \end{bmatrix}. \tag{1.140}$$

So one can use the iteration function

$$\begin{aligned} r = g_1(x,y;u,v,w) &= x - \frac{1}{20}(f(x,y,z,u,w) - 3g(x,y,z,u,w)) \\ s = g_2(x,y;u,v,w) &= y - \frac{1}{20}(6f(x,y,z,u,w) + 2g(x,y,z,u,w)). \end{aligned} \tag{1.141}$$

For each fixed value of u, v, w near $(3, 2, 7)$ this should have a fixed point near (0,1). |

Example: Here is an R program to carry this out for the input $(u, v, w) = (3, 2, 6)$, which one hopes is near enough to $(u, v, w) = (3, 2, 7)$ where we know the solution.

```
f1 <- function (x,y,u,v,w) 2 * exp(x) + y * u - 4 * v + 3
f2 <- function (x,y,u,v,w) y * cos(x) - 6 * x + 2 * u - w
x <- 0
y <- 1
u <- 3
v <- 2
w <- 6
g1 <- function (x,y) x - ( f1(x,y,u,v,w) - 3 * f2(x,y,u,v,w) )/20
g2 <- function (x,y) y - ( 6 * f1(x,y,u,v,w) + 2 * f2(x,y,u,v,w) )/20
for (n in 1:40) {
r <- g1(x,y)
s <- g2(x,y)
x <- r
y <- s }
x
   [1] 0.1474038
y
   [1] 0.8941188
```

This result is for input (3,2,6), and it is reasonably close to the point (0,1) that one would get for input (3,2,7). Of course to get a better idea of the function this computation needs to be repeated for a variety of inputs near (3,2,7). Even then one only gets an idea of the function at inputs sufficiently near this point. At inputs further away fixed point iteration may fail, and the behavior of the function is harder to understand. |

Theorem 1.23 (Inverse function theorem) *Let $\mathbf{y} = \mathbf{f}(\mathbf{x})$ be a function from $\mathbf{R}^m$ to $\mathbf{R}^m$. Here $\mathbf{x}$ is in $\mathbf{R}^m$ and $\mathbf{y}$ is in $\mathbf{R}^m$. Suppose $\mathbf{f}$ is continuous with continuous derivative near some point $\mathbf{x} = \mathbf{a}$. Suppose that the m by m matrix $\mathbf{f}(\mathbf{a})$ has an inverse matrix. Then there is a continuous function $\mathbf{h}(\mathbf{y})$ defined for $\mathbf{y}$ near $\mathbf{b} = \mathbf{f}(\mathbf{a})$ with $\mathbf{f}(\mathbf{h}(\mathbf{y})) = \mathbf{y}$. In fact, the function $\mathbf{h}(\mathbf{y})$ has a continuous derivative, and*

$$\mathbf{h}'(\mathbf{y}) = \mathbf{f}'(\mathbf{h}(\mathbf{y}))^{-1}. \tag{1.142}$$

The inverse function theorem is the special case of the implicit function theorem when $\mathbf{f}(\mathbf{x}, \mathbf{y})$ is of the form $\mathbf{f}(\mathbf{x}) - \mathbf{y}$. It is of great importance. For example, if a system is described by variables $x_1, \ldots, x_n$, and $u_i = f_i(x_1, \ldots, x_n)$ gives new variables, we might want to describe the system by $u_1, \ldots, u_n$. This would be a passive transformation. But can one recover the original variables? If the matrix $\partial u_i / \partial x_j$ is non-singular, then the inverse functions says that it should be possible. That is, we have $x_j = h_j(u_1, \ldots, u_n)$.

1.8 Second order partial derivatives

In this section we consider scalar functions of several variables. For instance, consider $u = f(\mathbf{x}) = f(x, y, z)$. The first order partial derivatives

$$[f'_{,1}(\mathbf{x}), f'_{,2}(\mathbf{x}), f'_{,3}(\mathbf{x})] = \left[\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial x}\right] \tag{1.143}$$

naturally form a covector. How about the second order partial derivatives? These can be arranged in a matrix

$$f''(\mathbf{x}) = f''(x, y, z) = \begin{bmatrix} f''_{,11}(\mathbf{x}) & f''_{,12}(\mathbf{x}) & f''_{,13}(\mathbf{x}) \\ f''_{,21}(\mathbf{x}) & f''_{,22}(\mathbf{x}) & f''_{,23}(\mathbf{x}) \\ f''_{,31}(\mathbf{x}) & f''_{,32}(\mathbf{x}) & f''_{,33}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 u}{\partial x^2} & \frac{\partial^2 u}{\partial x \partial y} & \frac{\partial^2 u}{\partial x \partial z} \\ \frac{\partial^2 u}{\partial y \partial x} & \frac{\partial^2 u}{\partial y^2} & \frac{\partial^2 u}{\partial z \partial x} \\ \frac{\partial^2 u}{\partial z \partial x} & \frac{\partial^2 u}{\partial z \partial y} & \frac{\partial^2 u}{\partial z^2} \end{bmatrix}. \tag{1.144}$$

One ordinarily expects that this is a symmetric matrix, that is,

$$f''_{,12}(\mathbf{x}) = \frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 u}{\partial y \partial x} = f''_{,21}(\mathbf{x}) \tag{1.145}$$

and similarly for the other three cases.

The symmetry of this matrix is due to the equality of mixed partial derivative. This fact is not completely obvious. Here is a strategy for showing that it is true. There is little loss of generality in considering the case of a function $f(x, y)$ of two variables. Define $\Delta_1(h)f(x, y) = f(x + h, y) - f(x, y)$ and $\Delta_2(k)f(x, y) = f(x, y + k) - f(x, y)$. Then $\Delta_1(h)f(x, y)/h \to f_{,1}(x, y)$ as $h \to 0$ and $\Delta_2(k)f(x, y)/k \to f_{,2}(x, y)$ as $k \to 0$.

The key to the mixed partial derivatives identity is the algebraic identity

$$\Delta_1(h)\Delta_2(k)f(x, y) = \Delta_2(k)\Delta_1(h)f(x, y). \tag{1.146}$$

Then the obvious attempt at a proof is to write

$$\lim_{h \to 0} \frac{\Delta_2(k)\Delta_1(h)f(x, y)}{hk} = \frac{\Delta_2(k)f'_{,1}(x, y)}{k} \tag{1.147}$$

and

$$\lim_{k \to 0} \frac{\Delta_2(k)f'_{,1}(x, y)}{k} = f''_{,21}(x, y). \tag{1.148}$$

Then

$$f''_{,12}(x, y) = \lim_{h \to 0} \lim_{k \to 0} \frac{\Delta_1(h)\Delta_2(k)f(x, y)}{hk} = \lim_{k \to 0} \lim_{h \to 0} \frac{\Delta_2(k)\Delta_1(h)f(x, y)}{hk} = f''_{,21}(x, y). \tag{1.149}$$

The trouble with this is that it is possible to find clever examples where the interchange of limits does not work. However if we make a continuity assumption, then this problem goes away.

Theorem 1.24 (Equality of mixed partial derivatives) *Assume that $f(x,y)$ and $f'_{,1}(x,y)$ and $f'_{,2}(x,y)$ and $f''_{,21}(x,y)$ are continuous. Then it follows that $f''_{,12}(x,y) = f''_{,21}(x,y)$.*

Proof: By the mean value theorem

$$\frac{\Delta_2(k)\Delta_1(h)f(x,y)}{hk} = \frac{\Delta_2(k)f'_{,1}(x+h^*,y)}{k} = f''_{,21}(x+h^*,y+k^*). \quad (1.150)$$

Here h^* is between x and $x+h$ and k^* is between y and $y+k$. Consider $\epsilon > 0$. Choose h and k so small that $f''_{,21}(x+h^*,y+k^*)$ is a distance less than ϵ from $f''_{,21}(x,y)$. Thus $\Delta_2(k)\Delta_1(h)f(x,y)/(hk)$ is a distance less than ϵ from $f''_{,21}(x,y)$. Now let $k \to 0$. We conclude that for h sufficiently small (depending on ϵ) we have that $\Delta_1(h)f'_{,2}(x,y)/h$ is a distance less than or equal to ϵ from $f''_{,21}(x,y)$. This proves that $f''_{,12}(x,y) = f''_{,21}(x,y)$. □

The conclusion is that, apart from technicalities, we may safely assume the equality of mixed partial derivatives. This has a number of important conclusions.

Theorem 1.25 *Consider a covector of functions $p_1 = h(x_1,\ldots,x_n),\ldots,p_n = h_n(x_1,\ldots,x_n)$ that each has continuous derivatives. Suppose that this covector is integrable, in the sense that there is a function $u = f(x_1,\ldots,x_n)$ with continuous second partial derivatives satisfying*

$$du = p_1\,dx_1 + \cdots + p_n\,dx_n. \quad (1.151)$$

Then for each i,j the integrability condition

$$\frac{\partial p_i}{\partial x_j} = \frac{\partial p_j}{\partial x_i} \quad (1.152)$$

is satisfied.

Proof: The proof depends on the equality of mixed partial derivatives. For each k we have

$$p_k = \frac{\partial u}{\partial x_k}. \quad (1.153)$$

Therefore

$$\frac{\partial p_i}{\partial x_j} = \frac{\partial}{\partial x_j}\frac{\partial}{\partial x_i}u = \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}u = \frac{\partial p_j}{\partial x_i}. \quad (1.154)$$

□

Another important application is to the first and second derivative tests.

Theorem 1.26 (First derivative test) *Assume that $u = f(x_1,\ldots,x_n)$ has continuous first partial derivatives. Suppose that there is a point where there is a local minimum or a local maximum. Then at that point the differential*

$$du = \frac{\partial u}{\partial x_1}\,dx_1 + \cdots + \frac{\partial u}{\partial x_n}\,dx_n \quad (1.155)$$

is zero, that is, each partial derivative is zero at that point.

A point where the differential du vanishes is called a *critical point*. The value of u at a critical point is called a *critical value*. At a critical point, the symmetric matrix of second partial derivatives has a special significance. In this context it is is called the *Hessian matrix*.

Theorem 1.27 (Second derivative test) *Assume that $u = f(x_1, \ldots, x_n)$ has continuous second partial derivatives. Suppose that there is a point where the derivative $f'(\mathbf{x})$ covector with entries*

$$f'_{,i}(\mathbf{x}) = \frac{\partial u}{\partial x_i} \tag{1.156}$$

vanishes. Consider the second derivative Hessian matrix $f''(\mathbf{x})$ with entries

$$f''_{,ij}(\mathbf{x}) = \frac{\partial^2 u}{\partial x_i \, \partial x_j}. \tag{1.157}$$

This is a symmetric matrix with real eigenvalues. If at the given point all eigenvalues of this matrix are strictly positive, then the function has a local minimum. Similarly, if all eigenvalues of this matrix are strictly negative, then the function has a local maximum.

1.9 Problems

Problems 1: Fixed point iteration

1. Let $g(x) = \cos(x/2)$. It has a stable fixed point $r > 0$ with $g(r) = r$. Use fixed point iteration to find a numerical value for r. Also find $g'(r)$.

2. There is a theorem that says that if $|g'(x)| \leq c < 1$ in an interval of the form $[p - r, p + r]$, and if $|g(p) - p| \leq (1 - c)r$, then g has a fixed point in the interval. Use this to give a proof that $g(x) = \cos(x/2)$ has a fixed point in the interval $[\pi/6, \pi/2]$.

3. Say that $g : [a, b] \to [a, b]$ is a continuous function. Suppose that g is increasing, in the sense that $x \leq y$ implies $g(x) \leq g(y)$. Show that fixed point iteration starting with a always leads to a fixed point r.

4. Let $g(x) = (1/2)(x^2 + 2x^3 - x^4)$. This has four fixed points $r_1 < r_2 < r_3 < r_4$. Find them, and specify which ones are stable. Compute everything exactly.

5. In the preceding problem, prove that if $r_1 < x < r_3$, then fixed point iteration starting at x converges to r_2. Give a detailed discussion. Hint: It may help to carefully draw a graph and use the graphical analysis of fixed point iteration. Do not make assumptions about the graph that are not justified.

6. In physics one calculates the spontaneous magnetization m by the fixed point equation

$$m = \tanh(Jm). \tag{1.158}$$

Here $J > 0$ is a fixed parameter. This equation may be solved by fixed point iteration using the function $g(x) = \tanh(Jx)$. Notice that $g'(x) = J\mathrm{sech}^2(Jx)$ satisfies $0 < g'(x) < J$ except at $x = 0$ where $g'(0) = J$. In particular it is increasing. Similarly, one can compute that $g''(x) < 0$ for $x > 0$ and $g''(x) > 0$ for $x < 0$. Also remember that $g(x) \to \pm 1$ as $x \to \pm\infty$.

(a) Suppose $J > 1$. Describe the fixed points. Discuss stability. Sketch orbits.

(b) For each of the fixed points, describe the set of x such that fixed point iteration starting at x converges to that fixed point.

Recitation 1

1. Use fixed point iteration to numerically find the largest root r of $f(x) = x^3 - 5x^2 + 3x + 1 = 0$. Use $g(x) = x - f(x)/f'(s)$, where s is chosen close to the unknown root r. (Since $f(4) = -3$ is not very large, perhaps s could be near 4.) Start the iteration near the root.

2. Consider a smooth function $f(x)$ with a simple root r, that is, $f(r) = 0$ and $f'(r) \neq 0$. Let $g(x) = x - f(x)/f'(x)$. Find $g'(x)$. Find $g'(r)$.

3. Use the iteration function of the previous problem to numerically find the largest root for the example of the first problem.

4. Suppose that $g : [a, b] \to [a, b]$ is an increasing function: $x \leq y$ implies $g(x) \leq g(y)$. Prove or disprove the following general assertion: There exists s in $[a, b]$ such that s is not a fixed point and iteration starting at s converges to a fixed point.

5. Suppose that $g : [a, b] \to [a, b]$ is an increasing function: $x \leq y$ implies $g(x) \leq g(y)$. Prove or disprove the following general assertion: The function g has a fixed point.

Problems 2: The size of a matrix

1. Let H be the real symmetric matrix

$$H = \begin{bmatrix} 10 & 0 & 2 \\ 0 & 10 & 4 \\ 2 & 4 & 2 \end{bmatrix}. \tag{1.159}$$

This matrix has determinant zero, so one eigenvalue is zero. Find all eigenvalues. Find the corresponding eigenvectors, as column vectors. (Are they orthogonal?) Produce a matrix P with the normalized eigenvectors

as columns. Show that $P^T P = I$. Show by explicit computation that $HP = P\Lambda$, where Λ is diagonal. Find the spectral representation of H.

2. Let A be the real matrix

$$A = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix}. \tag{1.160}$$

Find the Lipschitz norm of A (the square root of the largest eigenvalue of $A^T A$). Find the 2 norm of A (the square root of sum of squares of entries, or, equivalently, the square root of the trace of $A^T A$). Compare them.

3. This problem deals with the Lipschitz norm. Say that A is a real square matrix. The claim is that it is always true that $\|A^2\| = \|A\|^2$. Prove or disprove.

4. Let R be the real symmetric matrix

$$R = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}. \tag{1.161}$$

Find the Lipschitz norm $\|R\|$.

5. Find all real square matrices A such that $\|A\| = \|A\|_2$. If you need a hint, see below.

Hint: Consider a vector $\mathbf{x}$ that is not the zero vector, and another vector $\mathbf{a}$. The Schwarz inequality says that the inner product $\mathbf{a}\cdot\mathbf{x}$ satisfies $|\mathbf{a}\cdot\mathbf{x}| \leq |\mathbf{a}||\mathbf{x}|$ with equality only when $\mathbf{a} = c\mathbf{x}$. (Since $\mathbf{a} \cdot \mathbf{x} = |\mathbf{a}||\mathbf{x}| \cos(\theta)$, this is when $\cos(\theta) = \pm 1$, so the vectors are either pointing in the same or opposite direction.)

Use the Schwarz inequality for each i to prove

$$|A\mathbf{x}|^2 = \sum_i (\sum_j a_{ij} x_j)^2 \leq \sum_i (\sum_j a_{ij}^2 \sum_k x_k^2) = \|A\|_2^2 |\mathbf{x}|^2. \tag{1.162}$$

When is this an equality? (Consider the situation for each fixed i.) Once you have the form of the matrix you can calculate $A^T A$ and evaluate the norms.

Recitation 2

1. Describe all 2 by 2 matrices with only one eigenvalue that are not diagonalizable.

2. Consider the matrix
$$A = \begin{bmatrix} -3 & 16 \\ -1 & 5 \end{bmatrix}. \tag{1.163}$$
Find the eigenvalue. Let $\delta \neq 0$. Find an invertible matrix P with $AP = PJ$, where
$$J = \begin{bmatrix} \lambda & \delta \\ 0 & \lambda \end{bmatrix}. \tag{1.164}$$
The matrix P will depend on δ.

3. Consider $0 \leq r \leq s$. Must there be a 2 by 2 matrix with spectral radius r and norm s? Prove that your answer is correct.

4. Let
$$R = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix} = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) \end{bmatrix} - \begin{bmatrix} \sin(\theta) \\ -\cos(\theta) \end{bmatrix} \begin{bmatrix} \sin(\theta) & -\cos(\theta) \end{bmatrix}. \tag{1.165}$$
Check this identity. Find the eigenvalues and eigenvectors. Find R^2.

5. Find the eigenvalues and eigenvectors of
$$A = \begin{bmatrix} 1 & \frac{1}{10} \\ \frac{1}{10} & 1 \end{bmatrix} \tag{1.166}$$
Use a computer program to show a trajectory heading toward the origin near the stable direction ($0 < \lambda < 1$) and then heading out along the unstable direction ($1 < \lambda$). This gives a pictorial way of seeing eigenvectors!

Here is a sample program. Try it with other initial conditions!

```
x <- vector()
y <- vector()
u <- 10
v <- -10 + 1/10
for (i in 1:40) {
u <- u + v/10
v <- u/10 + v
x[i] <- u
y[i] <- v }
frame()
plot(x,y)
lines(x,y)
```

Problems 3: The derivative as a matrix

1. The problem is to solve $f(u, v) = 0$ for $u = h(v)$ as a function of v. Suppose $f(u_0, v_0) = 0$. If we want to find a solution with $u_0 = h(v_0)$, we can set $m = \partial f(u, v)/\partial u \neq 0$ evaluated at the point. Assume $m \neq 0$. Then we can set $g(u, v) = u - (1/m)f(u, v)$. Then $\partial g(u, v)/\partial u$ will be close to zero near the point. (Why?) So fixed point iteration $u \mapsto g(u, v)$ with v fixed near v_0 should work to produce a solution.

   Say that instead we want to solve $f(u, v) = 0$ for $v = k(u)$ as a function of u. Specify the assumption that is needed, and describe the procedure.

2. This continues the previous problem. Consider $f(u, v) = u^3 - v^2$. Describe all points u, v for which the above results give a solution for u as a function of v. Describe all points u, v for which the above results give a solution for v as a function of u. Graph the equation $f(u, v) = 0$ with some care. Indicate the functions $u = h(v)$ and $v = k(u)$ that you get. To what extent are they uniquely specified.

3. Say that $u = x^3 - 3xy^2$ and $v = 3x^2y - y^3$. Find the 2 by 2 derivative matrix. What is the determinant of this matrix? What is its inverse?

4. This continues the previous problem. Say that $s = \sin(ue^v)$. Find $\partial s/\partial u$ and $\partial s/\partial v$. Find s as a function of x and y. Use the chain rule to evaluate the two entries of the derivative matrix (row covector) $\partial s/\partial x$ and $\partial s/\partial y$.

5. Let
$$u = f(x, y) = \frac{x^3}{x^2 + y^2} \tag{1.167}$$
   with $f(0, 0) = 0$ at the origin.

   a) Show that u is continuous at the origin by direct calculation using the definition of continuity.

   b) Evaluate $\partial u/\partial x = f'_{,1}(x, y)$ and $\partial u/\partial y = f'_{,2}(x, y)$ away from the origin. Evaluate $\partial u/\partial x$ and $\partial u/\partial y$ at the origin, using the definition of the (one-dimensional) derivative.

   c) Is $u = f(x, y)$ a C^1 function? That is, are the partial derivatives $\partial u/\partial x$ and $\partial u/\partial y$ both continuous? Prove that your answer is correct by direct calculation.

   c)The condition for $u = f(x, y)$ to be differentiable at the origin is that
$$f(h, k) = f(0, 0) + f'_{,1}(0, 0)h + f'_{,2}(0, 0)k + r(0, 0; h, k) \tag{1.168}$$
   with $|r(0, 0; h, k)|/\sqrt{h^2 + k^2} \to 0$ as $\sqrt{h^2 + k^2} \to 0$. Is the function differentiable? Using only this definition, prove that your answer is correct.

Recitation 3

1. Define

$$\begin{aligned} g_1(x,y) &= xy - 2x - 2y + 6 \\ g_2(x,y) &= xy - 2x + 1. \end{aligned} \tag{1.169}$$

Find all fixed points by explicit computation. (Hint: Start by eliminating the xy term.) Find the derivative matrix at each fixed point. Discuss eigenvalues and stability.

2. The obvious general mean value theorem would say that if $\mathbf{f}(\mathbf{x})$ is a continuously differentiable function from a convex domain in $\mathbf{R}^n$ to $\mathbf{R}^m$, then there exists t with $0 \leq t \leq 1$ such that

$$\mathbf{f}(\mathbf{q}) - \mathbf{f}(\mathbf{p}) = \mathbf{f}'((1-t)\mathbf{q} + t\mathbf{p})(\mathbf{q} - \mathbf{p}). \tag{1.170}$$

But is this true? Hint: Take $n = 1$ and $m = 2$ and $\mathbf{f}(x) = [x^2, x^3]^T$.

3. Say that a function is C^k if it has k derivatives, all continuous.

a) For each k, find a function f from the real line to itself that is C^k but not C^{k+1}. Hint: It may help to take $f(x) = 0$ for $x \leq 0$.

b) A function is C^∞ if it is C^k for every k. Find a function with $f(x) = 0$ for $x \leq 0$ that is C^∞ and is not the zero function.

4. a) Let X and Y be metric spaces. Let $f : X \to Y$ be a function. Suppose that f is Lipschitz with Lipschitz constant c. This means that the distance from $f(x)$ to $f(y)$ is bounded by c times the distance from x to y. (The Lipschitz norm of f is the least such constant.) Prove that f is uniformly continuous.

b) Let $X = Y = [0, 1]$. Find a function $f : X \to Y$ that is uniformly continuous but not Lipschitz. Hint: Take $f(x) = \sqrt{x}$. Show that if $x < y$ and $y - x < \epsilon^2$, then $\sqrt{y} - \sqrt{x} < \epsilon$.

Problems 4: Implicit function theorem

1. Consider the surface in $\mathbf{R}^3$ given by

$$s = x^4 + x^2y^2 + y^4 + y^2z^2 + z^4 + z^2x^2 = 1. \tag{1.171}$$

(a) Calculate the differential of the function defining the surface. For which points on the surface does the differential vanish? (b) For which points on the surface does the implicit function theorem define at least one of the variables as a function of the other two near the point?

2. (a) In the preceding problem it should be possible to solve for y in terms of x, z near the point (0,1,0). Find a function $g(y; x, z)$ such that fixed point iteration $y \mapsto g(y; x, z)$ with this function (for fixed x, z) gives the

corresponding value y. Express this function in terms of $s-1 = f(x,y,z)$. (b) Answer the same question if y is to be expressed in terms of x, z near $(0,-1,0)$.

3. Consider the function given by

$$u = f_1(x,y) = x^4 - 6x^2y^2 + y^4 \tag{1.172}$$
$$v = f_2(x,y) = 4x^3y - 4y^3x. \tag{1.173}$$

   (a) Find the derivative matrix.

   (b) Find the determinant of the derivative matrix. Find all points at which this determinant vanishes.

   (c) Show that the four points $(x,y) = (\pm 1, 0)$, $(x,y) = (0, \pm 1)$ all map to $(u,v) = (1,0)$. Near which of these point is there an inverse function?

4. The function of the previous problem maps the point $(0,1)$ to $(1,0)$. There is an inverse function that sends points (u,v) near $(1,0)$ to points (x,y) near $(0,1)$. Find functions $g_1(x,y;u,v)$ and $g_2(x,y;u,v)$ such that fixed point iteration with these functions (for fixed (u,v)) give the corresponding inverse values x, y. Express these functions in terms of $f_1(x,y)$, $f_2(x,y)$ and u, v. (Use the algorithm involving the derivative matrix of $f_1(x,y)$, $f_2(x,y)$ evaluated at the point $(0,1)$.

5. Consider the iteration function $g_1(x,y;u,v)$, $g_2(x,y;u,v)$ found in the previous problem. Show that if $|x| < 1/100$ and $|y-1| < 1/100$, then the linearization at such a point has norm bounded by $1/2$. (Hint: Bound the 2-norm.)

6. (a) Consider an equation $f(p,v,t) = 0$. Consider a particular point in $\mathbf{R}^3$ that gives a solution of this equation. Describe the condition that guarantees that one can solve for v as a function of p, t near this point.

   (b) In physics the quantities p, v, t are related to pressure, volume, and temperature. An important equation relating these quantities is

$$f(p,v,t) = \left(p + \frac{3}{v^2}\right)\left(v - \frac{1}{3}\right) - \frac{8}{3}t = 0. \tag{1.174}$$

   Show that $p = 1$, $v = 1$, $t = 1$ is a solution.

   (c) Can the implicit function theorem be used to specify v as a function of p, t near this point where $p = 1, v = 1, t = 1$? Justify your answer.

Recitation 4

1. (a) Is $xy^4\,dx + 2x^2y^3\,dy$ exact? Justify your answer.

   (b) Is $3x^2y^2\,dx + 4x^3y\,dy$ exact? Justify your answer.

(c) Can $3x^2y^2\,dx + 4x^3y\,dy$ be multiplied by a power y^m to produce an exact differential? Justify your answer.

2. Let

$$u = f(x, y) = x^2y^2 - 2xy^2 + 5x^2 - 10xy + 7y^2 - 10x + 10y. \quad (1.175)$$

There is a point on the line $y = 0$ where $du = 0$. Find it, and find the corresponding value of u. Compute the Hessian matrix. Apply the second derivative test to establish whether it is a local minimum, local maximum, or saddle point (or something else).

3. The second derivative. Say that we want to minimize $w = h(u, v)$ by application of the first derivative test and the second derivative test. However we want to use coordinates x, y related to the u, v by $u = f(x, y)$, $v = g(x, y)$. The corresponding 2 by 2 derivative matrix is assumed to be invertible.

Show that at a point where $\partial w/\partial u = 0$ and $\partial w/\partial v = 0$ it is also true that $\partial w/\partial x = 0$ and $\partial w/\partial y = 0$.

Consider such a point where the first partial derivatives vanish. Suppose that at this point the quadratic form

$$\frac{\partial^2 w}{\partial u^2}p^2 + 2\frac{\partial^2 w}{\partial u \partial v}pq + \frac{\partial^2 w}{\partial v^2}q^2 > 0 \quad (1.176)$$

for all p, q other than $0, 0$. Show that $\partial^2 w/\partial x^2 > 0$.

4. Let $g(x, y) = (x^2 - y^2)/(x^2 + y^2)$ and let $g(0, 0) = 0$. Then $g(x, y)$ is a bounded function that is C^2 away from (0,0). Show that $g(x, 0) = 1$ for $x \neq 0$, $g(0, y) = -1$ for $y \neq 0$. (Thus g is not continuous at (0,0).) Show also that $x\partial g(x, y)/\partial x$ and $y\partial g(x, y)/\partial y$ are bounded functions in the region away from $(0, 0)$. In the following use only these general properties of $g(x, y)$.

Let $f(x, y) = xyg(x, y)$. Calculate $\partial f(x, y)/\partial x$ and $\partial f(x, y)/\partial y$ at the origin. Show that $f(x, y)$ is C^1. Compute $\partial f(x, y)/\partial x$ along $y = 0$. Calculate $\partial f(x, y)/\partial y$ along $x = 0$. Show that your answers also work at the origin. Compute the mixed second partial derivatives at the origin.

Chapter 2

Integration

2.1 The Riemann integral

The rest of these lectures centers on integration, and here the foundational result is the change of variables theorem. In dimension greater than one this is a non-trivial result; the limiting process that defines the derivative is related to the limiting process that defines the integral in a rather subtle way. This chapter on the Riemann integral is directed toward making this relation explicit.

In the following the integral is defined in terms of upper sums and lower sums. The notions of *supremum* (least upper bound) and *infimum* (greatest lower bound) play a role. For easy reference, here are the definitions.

Let S be a set of real numbers. Then $\sup S$ is the least upper bound of S. That is, it is an upper bound:

$$\forall x \in S \;\; x \leq \sup S, \tag{2.1}$$

and it is the least upper bound: For every b we have

$$\forall x \in S \;\; x \leq b \Rightarrow \sup S \leq b. \tag{2.2}$$

Equivalently,

$$b < \sup S \Rightarrow \exists x \in S \;\; b < x. \tag{2.3}$$

Similarly, $\inf S$ is the number with the property that is the greatest lower bound of S. That is, it is a lower bound:

$$\forall x \in S \;\; \inf S \leq x, \tag{2.4}$$

and it is the greatest lower bound: For every b we have

$$\forall x \in S \;\; b \leq x \Rightarrow b \leq \inf S. \tag{2.5}$$

Equivalently,

$$\inf S < b \Rightarrow \exists x \in S \;\; x < b. \tag{2.6}$$

If f is a real function defined on some set, then the supremum and infimum of the function are the supremum and infimum of the set of values of the function.

An *interval* is a subset of $\mathbf{R}$ that is connected. It is *degenerate* if it is empty or consists of only one point. An n-dimensional *cell* is a subset of $\mathbf{R}^n$ that is a product of n intervals. An n-dimensional cell is bounded if and only if each of the n intervals is bounded. For a bounded cell we may define the n-dimensional *volume* by

$$m_n(I) = \prod_{i=1}^{n} \Delta x_i, \tag{2.7}$$

where $\Delta x_i \geq 0$ is the length of side i.

An n dimensional cell is closed if and only if each of the n intervals is closed. We say that an n dimensional cell is *non-degenerate* if each of the n intervals is non-degenerate.

In mathematics a *set partition* of a set A is a collection of non-empty non-overlapping subsets whose union is A. In the following we need a slightly different notion of partition. In this version the subsets are allowed to overlap, but only in lower dimension. Let C be a subset of $\mathbf{R}^n$. Then $\mathcal{P}$ is a *partition* (more precisely, cell partition) of C provided that $\mathcal{P}$ is a finite set of closed, bounded, non-degenerate cells, each pair of cells in $\mathcal{P}$ has an intersection that is a closed bounded degenerate cell, and C is the union of the cells in $\mathcal{P}$. From now on we consider a set C that has at least one partition. Such a set will be called a *rectangular set*.

A partition $\mathcal{P}$ is *finer* than partition $\mathcal{Q}$ if for each I in $\mathcal{P}$ there is a J in $\mathcal{Q}$ with $I \subseteq J$.

If $\mathcal{Q}$ and $\mathcal{R}$ are partitions, then there is always a partition $\mathcal{P}$ that is finer than $\mathcal{Q}$ and also finer than $\mathcal{R}$. This can be found by taking the cells of $\mathcal{P}$ to be of the form $I = J \cap K$, where J is a cell of $\mathcal{Q}$ and K is a cell of $\mathcal{R}$.

Let $f : C \to \mathbf{R}$ be a bounded function. For each partition $\mathcal{P}$ of A define the *lower sum* by

$$L(f, \mathcal{P}) = \sum_{I \in \mathcal{P}} \inf f_I \; m(I) \tag{2.8}$$

and the *upper sum* by

$$U(f, \mathcal{P}) = \sum_{I \in \mathcal{P}} \sup f_I \; m(I). \tag{2.9}$$

Here f_I denotes the restriction of f to the closed bounded non-degenerate cell I.

It is easy to see that when $\mathcal{P}$ is a refinement of $\mathcal{Q}$, then

$$L(f, \mathcal{Q}) \leq L(f, \mathcal{P}) \leq U(f, \mathcal{P}) \leq U(f, \mathcal{Q}). \tag{2.10}$$

Furthermore, we can take more and more refined partitions and get corresponding lower and upper integrals. More precisely, we have the *lower integral*

$$L(f) = \sup_{\mathcal{P}} L(f, \mathcal{P}) \tag{2.11}$$

and the *upper integral*

$$U(f) = \inf_{\mathcal{P}} U(f, \mathcal{P}). \tag{2.12}$$

Here $\mathcal{P}$ ranges over all partitions of the set A. It is not hard to show that $L(f) \leq U(f)$.

If $L(f) = U(f)$, then we say that f is *Riemann integrable* with *integral* $I(f) = L(f) = U(f)$. Warning: There are functions that are not Riemann integrable; for such functions $L(f) < U(f)$.

The reader will recall examples from the case $n = 1$. If f is defined on $[a, b]$ and f is monotone increasing (or monotone decreasing), then f is Riemann integrable. On the other hand, if $f(x) = 1$ for x rational and $f(x) = 0$ for x irrational, then $L(f) = 0$, while $U(f) = b - a$.

The lower and upper integrals are always defined, but in general they have a major defect: they need not be additive. The following proposition states what is true in general: the lower integral is *superadditive (on functions)*, and the upper integral is *subadditive*.

Proposition 2.1 *The lower integral is superadditive:*

$$L(f+g) \geq L(f) + L(g), \tag{2.13}$$

while the upper integral is subadditive:

$$U(f+g) \leq U(f) + U(g). \tag{2.14}$$

Proof: It is sufficient to prove this for the case of the lower integral. We have for each x in I the inequality

$$\inf_{x \in I} f(x) + \inf_{x \in I} g(x) \leq f(x) + g(x), \tag{2.15}$$

so the left hand side is an lower bound. Therefore the greatest lower bound satisfies

$$\inf_{x \in I} f(x) + \inf_{x \in I} g(x) \leq \inf_{x \in I} (f(x) + g(x)). \tag{2.16}$$

Consider partitions $\mathcal{Q}$ and $\mathcal{R}$. There is a partition $\mathcal{P}$ that is finer than either of them. As a consequence, we have that

$$L(f, \mathcal{Q}) + L(g, \mathcal{R}) \leq L(f, \mathcal{P}) + L(g, \mathcal{P}) \leq L(f+g, \mathcal{P}) \leq L(f+g). \tag{2.17}$$

So we have an upper bound for the $L(f, \mathcal{Q})$. The least upper bound $L(f)$ must then satisfy

$$L(f) + L(g, \mathcal{R}) \leq L(f+g). \tag{2.18}$$

Similarly, we have an upper bound for the $L(f, \mathcal{R})$. The least upper bound $L(f)$ must then satisfy

$$L(f) + L(g) \leq L(f+g). \tag{2.19}$$

□

Theorem 2.2 *The Riemann integral is additive for Riemmann integrable functions:*

$$I(f+g) = I(f) + I(g). \tag{2.20}$$

The above theorem is the main reason for defining the Riemann integral as the common value of the lower and upper integrals. Things can go wrong when the upper and lower integrals differ. The reader will find it not difficult to produce an example in one dimension where $L(f+g) > L(f) + L(g)$.

Let A be an arbitrary bounded subset of $\mathbf{R}^n$. All these notions may be extended to the case of a bounded function $f : A \to \mathbf{R}$. We merely have to choose a rectangular set C such that $A \subseteq C$. Then we can define $\bar{f}$ to be equal

to f on A and to be equal to 0 on the complement of A in C. The integral $I(f)$ of f over A may be defined to be the integral $I(\bar{f})$ of $\bar{f}$ over C.

There are various notations for the integral. For instance, if we represent the function $f : A \to \mathbf{R}$ by $\mathbf{x} \mapsto f(\mathbf{x})$ with $\mathbf{x}$ ranging over A, then we could write

$$I(f) = I(\mathbf{x} \mapsto f(\mathbf{x})). \tag{2.21}$$

A variation that will be convenient in the following is

$$I(f) = I(f(\mathbf{x}); \mathbf{x}). \tag{2.22}$$

The most common notation is something like

$$I(f) = \int f(\mathbf{x})\, d\mathbf{x}. \tag{2.23}$$

The official definition of Riemann integral used here is in terms of partitions of a set into finitely many closed bounded non-degenerate cells. As a consequence, every Riemann integrable function is not only bounded, but it also vanishes outside of a bounded set. In practice one often wants to integrate unbounded functions or functions defined on all of Euclidean space. There are two situations that are dramatically different.

- The integral is an absolutely convergent sum of absolutely convergent Riemann integrals. For instance, f has an integral over all of Euclidian space if we write the space as an infinite union of closed bounded non-degenerate cells C_k, and the restrictions f_k to C_k satisfy

$$\sum_k I(|f_k|) < +\infty. \tag{2.24}$$

  In that case we can define

$$I(f) = \sum_k I(f_k) \tag{2.25}$$

  without any ambiguity.

- The integral is represented as a conditionally convergent sum. In this case the question is the order in which the sum is taken. By changing the order one can get any result.

Sometime people speak about "improper Riemann integrals" in a way that confuses these two notions. But the distinction that matters is between absolute convergence and conditional convergence.

2.2 Jordan content

Consider a rectangular set C, and $A \subseteq C$. Let 1_A be the *indicator function* of A, equal to 1 on A and 0 on $C \setminus A$. (The indicator function is sometimes called the characteristic function, but this term has a conflicting use in probability theory.) If 1_A is Riemann integrable, then A is said to be *Jordan measurable*. The *Jordan content* (or volume) of A is then defined to be

$$m(A) = I(1_A). \tag{2.26}$$

Jordan content is additive. This means that if $A = \bigcup_{i=1}^{k} A_i$ is a union of disjoint Jordan measurable sets A_i, then A is Jordan measurable, and $m(A) = \sum_{i=1}^{n} m(A_i)$.

It is fairly easy to compute the Jordan content of a cell. However the formula for a ball is more difficult. For the record, the volume of a ball of radius r in n dimensions is $v_n r^n$, where

$$v_n = \frac{1}{n} \frac{2\pi^{n/2}}{\Gamma(n/2)}. \tag{2.27}$$

The most familiar cases are the interval of length $2r$, the disk of radius r with area πr^2, and the ball in three dimensions of radius r and volume $\frac{4}{3}\pi r^3$. The area of the $n-1$ sphere of radius r in n dimensions is given by $a_n r^{n-1} = dv_n r^n/dr$, so

$$a_{n-1} = n v_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}. \tag{2.28}$$

The most familiar cases are 2 points, the circle of radius r with length $2\pi r$, and the 2-sphere in three dimensions of radius r and area $4\pi r^2$. For reference the relevant values of the Γ function are $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, $\Gamma(1) = 1$, and $\Gamma(\frac{3}{2}) = \frac{1}{2}\sqrt{\pi}$.

2.3 Approximation of Riemann integrals

The notions of *interior* and *boundary* will occur; here is a quick review. If A is a subset of a metric space, then the interior $\mathrm{int}(A)$ is the largest open subset of A. The boundary $\mathrm{bdy}(A)$ is the set of points that do not belong to the interior of A and also do not belong to the interior of the complement of A. Thus the closure of A is the union of $\mathrm{int}(A)$ and $\mathrm{bdy}(A)$.

If f is a real function on $C \subseteq \mathbf{R}^n$, and $A \subseteq C$, define the *oscillation* of f on A by

$$\mathrm{osc}_A(f) = \sup_{\mathbf{z} \in A} f(\mathbf{z}) - \inf_{\mathbf{w} \in A} f(\mathbf{w}). \tag{2.29}$$

The oscillation is a special case of notion of *diameter* of a set. In fact, we have $\mathrm{osc}_A(f) = \sup_{\mathbf{z},\mathbf{w} \in A} |f(\mathbf{z}) - f(\mathbf{w})|$, which shows that the oscillation is the diameter of the image of the restriction of f to A.

Define the *oscillation* of f at $\mathbf{x}$ by

$$\mathrm{osc}_{\mathbf{x}}(f) = \inf\{\mathrm{osc}_A(f) \mid \mathbf{x} \in \mathrm{int}(A)\}. \tag{2.30}$$

It may be shown that the function f is continuous at $\mathbf{x}$ if and only if $\mathrm{osc}_{\mathbf{x}}(f) = 0$.

There is a very interesting connection between oscillation and the Riemann integral. In fact we have

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) = \sum_{I \in \mathcal{P}} \mathrm{osc}_I(f) m_n(I). \tag{2.31}$$

This implies that the integral exists if and only if the infimum over all these oscillation sums is zero.

Theorem 2.3 *Suppose that f is Riemann inegrable and that h is a Lipschitz function. Then the composition of $h \circ f$ is also Riemann integrable.*

Proof: To say that h is Lipschitz is to say that there is a constant M with $|h(y) - h(z)| \leq M|y - x|$. It follows that $\mathrm{osc}_I(h \circ f) \leq M \mathrm{osc}_I(f)$. So

$$U(h \circ f, \mathcal{P}) - L(h \circ f, \mathcal{P}) \leq M(U(f, \mathcal{P}) - L(f, \mathcal{P})). \tag{2.32}$$

If f is integrable, then so is $h \circ f$. □

Every C^1 function is Lipschitz on every bounded set. So the above result applies to functions such as $h(y) = y^2$.

Theorem 2.4 *Suppose that f and g are Riemann integrable. Then so is the product $f \cdot g$.*

Proof: This follows from the identity

$$f \cdot g = \frac{1}{4}\left((f+g)^2 - (f-g)^2\right). \tag{2.33}$$

Since by the previous result the square of a Riemann integrable function is Riemann integrable, the right hand size defines a Riemann integrable function. □

Theorem 2.5 *Suppose that f is Riemann inegrable and that h is a one-to-one continuous function whose inverse is Lipschitz. Then the composition $f \circ h$ is also Riemann integrable.*

This theorem will be proved in a later chapter. It follows that a change of variables that is C^1 with C^1 inverse preserves the property of being Riemann integrable. It follows that it also preserves the property of being Jordan measurable.

Given a function f on $\mathbf{R}^n$, there is a set of $\mathbf{x}$ where $f(\mathbf{x}) \neq 0$. The closure of this set is the *support* of f. Thus to say that f has compact support is equivalent to saying that f vanishes outside of a bounded set.

Another important property of lower and upper integrals is approximation by continuous functions. The lemma below applies in particular to Riemann integrals, since every Riemann integral is both a lower and an upper integral.

Theorem 2.6 *Let A be a bounded set. Let f be a bounded real function on A. Then for every $\epsilon > 0$ there exists a continuous function g with $g \leq f$ such that the lower integral satisfies $L(f) - I(g) < \epsilon$. There is a similar result for upper integrals.*

Proof: Choose a partition $\mathcal{P}$ such that

$$\sum_{I \in \mathcal{P}} \sup f_I m_n(I) > L(f) - \frac{\epsilon}{2}. \tag{2.34}$$

Let $\mathcal{P}_+$ be the set of I with $f_I \geq 0$, while $\mathcal{P}_-$ is the set of I with $f_I < 0$. For I in $\mathcal{P}_+$ define continuous g_I with support in I and with $0 \leq g_I \leq \sup f_I \leq f$ on I and with $f_I m_n(I) - I(g_I)$ very small. Thus g_I can be a continuous trapezoidal function with very steep sides. For I in $\mathcal{P}_-$ define continuous $g_I \leq 0$ with compact support and with $g_I \leq \sup f_I \leq f$ on I and with $f_I m_n(I) - I(g_I)$ very small. Again g_I can be a continuous trapezoidal function with very steep sides, however now it has constant value on all of I and consequently has a slightly larger support. Let $g = \sum_{I \in \mathcal{P}} g_I$. It is not difficult to show that $g \leq f$. Furthermore, we can arrange it so that

$$I(g) > \sum_{I \in \mathcal{P}} \sup f_I m_n(I) - \frac{\epsilon}{2}. \tag{2.35}$$

Thus this is the desired function. □

2.4 Fubini's theorem

There is a famous theorem of integration theory called *Fubini's theorem* that has a version for the Riemann integral. It says that a multi-dimensional integral may be written as an *iterated integral*. The usual formulation says that if f is a bounded function on a bounded subset $\mathbf{R}^{m+n}$, then under suitable conditions

$$\int f(\mathbf{x}, \mathbf{y})\, d\mathbf{x}\, d\mathbf{y} = \int \left[\int f(\mathbf{x}, \mathbf{y})\, d\mathbf{x} \right] d\mathbf{y} \tag{2.36}$$

and also

$$\int f(\mathbf{x}, \mathbf{y})\, d\mathbf{x}\, d\mathbf{y} = \int \left[\int f(\mathbf{x}, \mathbf{y})\, d\mathbf{y} \right] d\mathbf{x}. \tag{2.37}$$

Here $\mathbf{x}$ ranges over a subset of $\mathbf{R}^m$, and $\mathbf{y}$ ranges over a subset of $\mathbf{R}^n$. The left hand side is an ordinary Riemann integral over a subset of $\mathbf{R}^{m+n}$. The right hand side is an iterated integral. Thus in the first case for each fixed $\mathbf{y}$ there is a corresponding m dimensional Riemann integral. These integrals define a function of $\mathbf{y}$, and the n-dimensional integral of this function gives the final iterated integral.

In the following theoretical development we shall use a somewhat different notation. The reason for this is that we shall be comparing lower and upper

sums, and the variant notation makes it easy to incorporate these notions. In particular, the above formulas will be written

$$I(f(\mathbf{x},\mathbf{y});\mathbf{x},\mathbf{y}) = I(I(f(\mathbf{x},\mathbf{y});\mathbf{x});\mathbf{y}) \tag{2.38}$$

and

$$I(f(\mathbf{x},\mathbf{y});\mathbf{x},\mathbf{y}) = I(I(f(\mathbf{x},\mathbf{y});\mathbf{y});\mathbf{x}). \tag{2.39}$$

We shall see that in certain circumstances these formulas are true. However there are technical issues. For example, suppose that the Riemann integral $I(f(\mathbf{x},\mathbf{y});\mathbf{x},\mathbf{y})$ exists. Then it is not guaranteed that for fixed $\mathbf{y}$ the integral $I(f(\mathbf{x},\mathbf{y});\mathbf{x})$ exists. Nor is it guaranteed that for each fixed $\mathbf{x}$ that the integral $I(f(\mathbf{x},\mathbf{y});\mathbf{y})$ exists.
Example: Consider the following function defined on $[0,1]\times[0,1]$. Let $f(x,y) = 1$ when x is rational and $y = 1/2$, but $f(x,y) = 0$ elsewhere. This is Riemann integrable with integral zero. However for $y = 1/2$ the function $x \mapsto f(x,y)$ is not Riemann integrable. In fact, its lower integral is 0, while its upper integral is 1. |

Theorem 2.7 (Fubini's theorem for lower and upper integrals) *For lower integrals*

$$L(f(\mathbf{x},\mathbf{y});\mathbf{x},\mathbf{y}) \le L(L(f(\mathbf{x},\mathbf{y});\mathbf{x});\mathbf{y}), \tag{2.40}$$

while for upper integrals

$$U(U(f(\mathbf{x},\mathbf{y});\mathbf{x});\mathbf{y}) \le U(f). \tag{2.41}$$

There are similar results where the roles of $\mathbf{x}$ and $\mathbf{y}$ are reversed.

Proof: It is sufficient to prove the result for lower integrals. The result for upper integrals is proved in the same way, but with the inequalities reversed.

For the proof, it is useful to have the concept of a product partition. Let $C = C_1 \times C_2$ be the cell over which the integration takes place. The $\mathbf{x}$ variables range over C_1, while the $\mathbf{y}$ variables range over C_2. If $\mathcal{P}_1$ is a partition of C_1, and $\mathcal{P}_2$ is a partition of C_2, then the product partition $\mathcal{P}_1 \times \mathcal{P}_2$ is the partition of C consisting of all $I_1 \times I_2$ with I_1 from C_1 and I_2 from C_2. Given an arbitrary partition $\mathcal{P}$ of C, then there is a product partition that is finer than $\mathcal{P}$. So it is reasonable to first deal with product partitions.

First we need a simple lemma that only involves sums, not integrals. This states that

$$L(f,\mathcal{P}_1 \times \mathcal{P}_2) \le L(L(f(\mathbf{x},\mathbf{y});\mathbf{x},\mathcal{P}_1);\mathbf{y},\mathcal{P}_2). \tag{2.42}$$

The proof of the lemma uses $\inf_{(\mathbf{x},\mathbf{y})\in I_1\times I_2} f(\mathbf{x},\mathbf{y}) = \inf_{\mathbf{y}\in I_2} \inf_{\mathbf{x}\in I_1} f(\mathbf{x},\mathbf{y})$. The key ingredient is then the product property $m_{m+n}(I_1\times I_2) = m_m(I_1)m_n(I_2)$. We have

$$L(f,\mathcal{P}_1 \times \mathcal{P}_2) = \sum_{I_2\in\mathcal{P}_2} \sum_{I_1\in\mathcal{P}_1} \inf_{\mathbf{y}\in I_2} \inf_{\mathbf{x}\in I_1} f(\mathbf{x},\mathbf{y}) m_m(I_1) m_n(I_2). \tag{2.43}$$

From the general principle that $\sum_I \inf_{\mathbf{y}} h_I(\mathbf{y}) \leq \inf_{\mathbf{y}} \sum_I h_I(\mathbf{y})$ we get

$$L(f, \mathcal{P}_1 \times \mathcal{P}_2) \leq \sum_{I_2 \in \mathcal{P}_2} \inf_{\mathbf{y} \in I_2} \sum_{I_1 \in \mathcal{P}_1} \inf_{\mathbf{x} \in I_1} f(\mathbf{x}, \mathbf{y}) m_m(I_1) m_n(I_2). \tag{2.44}$$

This translates to

$$L(f, \mathcal{P}_1 \times \mathcal{P}_2) \leq \sum_{I_2 \in \mathcal{P}_2} \inf_{\mathbf{y} \in I_2} L(f(\mathbf{x}, \mathbf{y}); \mathbf{x}, \mathcal{P}_1) m_n(I_2). \tag{2.45}$$

This leads easily to the statement of the lemma. The proof of the lemma is complete.

Since lower sums are bounded by the lower integral, the lemma gives

$$L(f, \mathcal{P}_1 \times \mathcal{P}_2) \leq L(L(f(\mathbf{x}, \mathbf{y}); \mathbf{x}, \mathcal{P}_1); \mathbf{y}, \mathcal{P}_2) \leq L(L(f(\mathbf{x}, \mathbf{y}); \mathbf{x}), \mathbf{y}, \mathcal{P}_2). \tag{2.46}$$

Again for the same reason

$$L(f, \mathcal{P}_1 \times \mathcal{P}_2) \leq L(L(f(\mathbf{x}, \mathbf{y}); \mathbf{x}); \mathbf{y}). \tag{2.47}$$

Given an arbitrary partition $\mathcal{P}$, there is a finer product partition $\mathcal{P}_1 \times \mathcal{P}_2$, so we must also have

$$L(f, \mathcal{P}) \leq L(L(f(\mathbf{x}, \mathbf{y}); \mathbf{x}); \mathbf{y}). \tag{2.48}$$

That is, the iterated lower integral is an upper bound for the lower sums. Since $L(f)$ is the least of all such upper bounds, we have

$$L(f) \leq L(L(f(\mathbf{x}, \mathbf{y}); \mathbf{x}); \mathbf{y}) \tag{2.49}$$

as desired. □

Example: The theorem above gives a kind of Fubini theorem that works for the lower and upper integrals. Here is an example that shows that equality is not guaranteed. Consider the case of the upper integral of a function f defined on the cell $[0, 1] \times [0, 1]$. Suppose that there is a countable dense set D such that f is one on that set, zero on its complement. Then $U(f(x, y); x, y) = 1$. Now suppose that the set D has the property that for every horizontal line, there is at most one point on the line that is in D. Then for each y the function $x \mapsto f(x, y)$ has upper integral $U(f(x, y); x) = 0$. Thus $U(U(f(x, y), x), y) = 0$. So the iterated upper integral is smaller than the upper integral.

How can we find such a set D? First consider the set E of points in the plane with both coordinates rational. Consider all lines in the plane with fixed angle θ from the x axis, so that the slope is $m = \tan(\theta)$. Suppose that m is an irrational number. For instance, we could take lines at an angle $\theta = \pi/6$ from the x axis, with slope $m = 1/\sqrt{3}$. Every such line intersects E in at most one point. (Why?) Now rotate the picture by angle $-\theta$, so that we get a set D that consists of E rotated by this angle, and such that the lines become horizontal lines. |

Theorem 2.8 (Fubini's theorem for the Riemann integral) *Suppose that the Riemann integral $I(f(\mathbf{x},\mathbf{y});\mathbf{x},\mathbf{y})$ exists. Then for each fixed $\mathbf{y}$ the lower integral and upper integral are automatically defined and satisfy $L(f(\mathbf{x},\mathbf{y});\mathbf{x}) \leq U(f(\mathbf{x},\mathbf{y});\mathbf{x})$. Furthermore, as functions of $\mathbf{y}$ these each define Riemann integrable functions. Finally, we have both the formulas*

$$I(f(\mathbf{x},\mathbf{y});\mathbf{x},\mathbf{y}) = I(L(f(\mathbf{x},\mathbf{y});\mathbf{x});\mathbf{y}) \tag{2.50}$$

and

$$I(f(\mathbf{x},\mathbf{y});\mathbf{x},\mathbf{y}) = I(U(f(\mathbf{x},\mathbf{y});\mathbf{x});\mathbf{y}). \tag{2.51}$$

The result of course works in the other order. For the sake of completeness here is an explicit statement.

Theorem 2.9 (Fubini's theorem for the Riemann integral) *Suppose that the Riemann integral $I(f(\mathbf{x},\mathbf{y});\mathbf{x},\mathbf{y})$ exists. Then for each fixed $\mathbf{x}$ the lower integral and upper integral are automatically defined and satisfy $L(f(\mathbf{x},\mathbf{y});\mathbf{y}) \leq U(f(\mathbf{x},\mathbf{y});\mathbf{y})$. Furthermore, as functions of $\mathbf{x}$ these each define Riemann integrable functions. Finally, we have both the formulas*

$$I(f(\mathbf{x},\mathbf{y});\mathbf{x},\mathbf{y}) = I(L(f(\mathbf{x},\mathbf{y});\mathbf{y});\mathbf{x}) \tag{2.52}$$

and

$$I(f(\mathbf{x},\mathbf{y});\mathbf{x},\mathbf{y}) = I(U(f(\mathbf{x},\mathbf{y});\mathbf{y});\mathbf{x}). \tag{2.53}$$

Proof: The two preceding theorems are essentially the same; it is sufficient to prove the first one. The proof uses the results that relate lower integrals to iterated lower integrals and upper integrals to iterated upper integrals. Once we have these results, we are almost done. We have

$$L(f) \leq L(L(f(\mathbf{x},\mathbf{y});\mathbf{x});\mathbf{y}) \leq U(L(f(\mathbf{x},\mathbf{y});\mathbf{x});\mathbf{y}) \leq U(U(f(\mathbf{x},\mathbf{y});\mathbf{x});\mathbf{y}) \leq U(f). \tag{2.54}$$

If $L(f) = U(f)$, then $L(L(f(\mathbf{x},\mathbf{y});\mathbf{x});\mathbf{y}) = U(L(f(\mathbf{x},\mathbf{y});\mathbf{x});\mathbf{y})$. This proves the integrability of the function that sends $\mathbf{y}$ to $L(f(\mathbf{x},\mathbf{y});\mathbf{x})$.

Similarly, we have

$$L(f) \leq L(L(f(\mathbf{x},\mathbf{y});\mathbf{x});\mathbf{y}) \leq L(U(f(\mathbf{x},\mathbf{y});\mathbf{x});\mathbf{y}) \leq U(Uf(\mathbf{x},\mathbf{y});\mathbf{x});\mathbf{y}) \leq U(f). \tag{2.55}$$

If $L(f) = U(f)$, then $L(U(f(\mathbf{x},\mathbf{y});\mathbf{x});\mathbf{y}) = U(Uf(\mathbf{x},\mathbf{y});\mathbf{x});\mathbf{y})$. This proves the integrability of the function that sends $\mathbf{y}$ to $U(f(\mathbf{x},\mathbf{y});\mathbf{x})$. □

In the above proof it does not seem to matter whether one used the lower integral or the upper integral. This is clarified by the following remark. Define the difference function

$$h(\mathbf{y}) = U(f(\mathbf{x},\mathbf{y});\mathbf{x}) - L(f(\mathbf{x},\mathbf{y});\mathbf{x}). \tag{2.56}$$

Then $h \geq 0$. If f is Riemann integrable, then the analysis in the above proof shows that h is Riemann integrable and $I(h) = 0$. Thus h is negligible from

the Riemann integral point of view. For instance, if h is continuous, then it is identically zero. In such a case the Fubini theorem is true in the form stated at the outset, that is, without any special consideration of upper and lower integrals.

The above exposition of Fubini's theorem follows unpublished notes on the Riemann integral by Mariusz Wodzicki [20].

2.5 Uniform convergence

Say that f_n is a sequence of functions on a bounded set A, and f is another such function. We would like conditions that ensure the convergence of the integrals $I(f_n) \to I(f)$. The requirement of uniform convergence is certainly sufficient.

We review the basic definitions. Here is *pointwise convergence*. We say that f_n converges to f pointwise on A if for all $\mathbf{x}$ in A and for all $\epsilon > 0$ there is an N such that for $n \geq N$ we have that $|f_n(\mathbf{x}) - f(\mathbf{x})| < \epsilon$. In other words, the requirement is that for all $\mathbf{x}$ the limit as $n \to \infty$ of $f_n(\mathbf{x})$ is equal to $f(\mathbf{x})$. Sometimes this is written

$$\forall \mathbf{x} \, \forall \epsilon > 0 \, \exists N \, \forall n \geq N \, |f_n(\mathbf{x}) - f(\mathbf{x})| < \epsilon. \tag{2.57}$$

Notice that the $\forall \mathbf{x}$ may occur anywhere to the left of the $\exists N$. This says that the N may depend on $\mathbf{x}$.

Contrast that with the much stronger requirement of *uniform convergence*. We say that f_n converges to f uniformly on A if for all $\epsilon > 0$ there is an N such that for $n \geq N$ and for all $\mathbf{x}$ in A we have that $|f_n(\mathbf{x}) - f(\mathbf{x})| < \epsilon$. In other words, the requirement is that as n goes to infinity the function f_n approaches the function f in the sense that throughout the set A the deviation of f_n from f becomes arbitrary small. Sometimes this is written

$$\forall \epsilon > 0 \, \exists N \, \forall n \geq N \, \forall \mathbf{x} \, |f_n(\mathbf{x}) - f(\mathbf{x})| < \epsilon. \tag{2.58}$$

Notice that the $\forall \mathbf{x}$ may occur anywhere to the right of the $\exists N$ (but before the final inequality). Thus the N may not depend on $\mathbf{x}$. Clearly uniform convergence implies pointwise convergence.

The most famous result about uniform convergence is the following one. It is proved by a standard three ϵ argument.

Theorem 2.10 *Suppose that each f_n is continuous and that $f_n \to f$ uniformly. Then f is continuous.*

Sometimes it is necessary to find examples of functions that do not converge pointwise or do not converge uniformly. The condition for a function to not converge pointwise is

$$\exists \mathbf{x} \, \exists \epsilon > 0 \, \forall N \, \exists n \geq N \, |f_n(\mathbf{x}) - f(\mathbf{x})| \geq \epsilon. \tag{2.59}$$

The condition for a function to not converge uniformly is

$$\exists \epsilon > 0 \, \forall N \, \exists n \geq N \, \exists \mathbf{x} \, |f_n(\mathbf{x}) - f(\mathbf{x})| \geq \epsilon. \tag{2.60}$$

It is easier to prove that a function is not uniformly convergent, since the $\mathbf{x}$ is allowed to depend on N. In many cases it is possible to take $n = N$.

The following theorem involving uniform convergence and integration is elementary.

Theorem 2.11 *Suppose that all the f_n and f are Riemann integrable on the bounded set A. If f_n converges to f uniformly on A, then $I(f_n)$ converges to $I(f)$.*

One way to prove this is to first prove that $I(|f_n - f|)$ converges to zero.

There is a remarkable theorem of Dini that shows that under certain very special circumstances uniform convergence is automatic. The context is that of a sequence of functions that is decreasing. (There is an obvious variant with a sequence of functions that is increasing).

Theorem 2.12 (Dini's theorem) *Let A be a compact set. Suppose that f_n is a decreasing sequence of continuous functions on A. Suppose that $f_n \downarrow f$ converges pointwise. Finally, suppose that f is also continuous. Then f_n converges uniformly to f.*

Proof: Let $h_n = f_n - f$. Then $h_n \to 0$ is also decreasing, and it has pointwise limit zero. Furthermore, h_n is continuous. Let $\epsilon > 0$. By continuity, for each n the set U_n where $h_n < \epsilon$ is open. The open sets U_n are increasing. Now we use pointwise convergence. For each $\mathbf{x}$ in A, there are n sufficiently large (depending on $\mathbf{x}$) such that $h_n(x) < \epsilon$. Hence for such n we have $x \in U_n$. This shows that the union of the U_n for all n includes A. In other words, the U_n form an open cover of A. Since A is compact, there is a finite subcover. This implies that there is an N such that U_N includes A. Also the U_n for $n \geq N$ include A. In other words, for all $n \geq N$ the set where $h_n < \epsilon$ includes all of A. This is uniform convergence. □

2.6 Dominated convergence

It is rather amazing that uniform convergence is not necessary for convergence of the integrals. This section presents the dominated convergence theorem for the Riemann integral, first proved by Arzela in 1885.

Let A be a bounded set. Let f_n be a sequence of real Riemann integrable functions on A that are bounded by a fixed constant M. Thus there is an M such that $|f_n(x)| \leq M$ for all x in A and all n. Then the sequence of functions is said to be *dominated*. It is dominated in two ways: by the fixed bounded set A on which the functions are all defined, and by the fixed constant M by which they are all bounded.

Suppose that the sequence of functions f_n is dominated, and suppose that $f_n \to f$ pointwise as $n \to \infty$, where f is also Riemann integrable. This is called *dominated convergence* (for the Riemann integral). . The theorem says that in this situation the integrals converge: $I(f_n) \to I(f)$ as $n \to \infty$.

We begin with a lemma about approximation of lower integrals via continuous functions. This is a special case of a result proved earlier, but it is convenient to record it here.

Lemma 2.13 *Let A be a bounded set. Let $f \geq 0$ be a bounded real function on A. Then for every $\epsilon > 0$ there exists a continuous function g with $0 \leq g \leq f$ and $L(f) - I(g) < \epsilon$.*

Proof: From the definition of the lower integral it follows that there is a step function k with $0 \leq k \leq f$ and with $L(f) - I(k) < \epsilon/2$. However one can approximate each step by a continuous trapezoid with very steep sides, so that the resulting trapezoidal function g satisfies $0 \leq g \leq k$ and $I(k) - I(g) < \epsilon/2$. Then $I(f) - I(g) < I(f) - I(k) + I(k) - I(g) < \epsilon/2 + \epsilon/2 = \epsilon$. □

We now turn to an important result on monotone convergence; it will be used to prove the dominated convergence theorem. In the monotone convergence result there is no assumption that the functions are Riemann integrable. However they have lower integrals, so the result is stated in terms of lower integrals.

Lemma 2.14 (Monotone convergence for the lower Riemann integral) *Let A be a bounded set. Let p_n be a sequence of real functions on A such that there is a constant M with $0 \leq p_n \leq M$. Suppose that $p_n \downarrow 0$ pointwise as $n \to \infty$. Then the lower integrals converge: $L(p_n) \to 0$ as $n \to \infty$.*

Proof: The assumption is that p_n is a sequence of bounded functions with $p_n \downarrow 0$, that is, the functions are decreasing pointwise to zero. If Dini's theorem applied, then we would have $I(p_n) \downarrow 0$. However, there are obvious problems. The p_n need not be continuous, so Dini's theorem does not apply. Not only that, the p_n need not be Riemann integrable. In this case $I(p_n)$ is not even defined, and we must work with $L(p_n)$.

Ultimately the proof will reduce to Dini's theorem. Dini's theorem relies on the assumption that the functions are defined on a compact set, but that set can be taken to be a closed bounded non-degenerate cell C with $A \subseteq C$. (In fact this is how the Riemann integral is defined.) Uniform convergence is hiding somewhere in the world of pointwise convergence, but it is well hidden. We first use the fact that we can approximate by continuous functions. The essential idea is to approximate better and better as we go along the sequence.

Consider $\epsilon > 0$. Choose g_i continuous with $0 \leq g_i \leq p_i$ and

$$L(p_i) - I(g_i) \leq \epsilon \frac{1}{2^i}. \tag{2.61}$$

Unfortunately, there is no guarantee that the functions g_i are decreasing. To fix this, let

$$h_n = \min(g_1, \ldots, g_n). \tag{2.62}$$

Then the $h_n \downarrow 0$ are decreasing pointwise to zero, and each h_n is continuous. Hence by Dini's theorem $I(h_n) \downarrow 0$. This looks promising.

The problem is that in general the integral of a minimum can be much smaller than the integrals of the individual functions. So we need to use special features of the choices made above to ensure that $I(g_n) - I(h_n)$ is small. We have for each $j = 1, \ldots, n$ that

$$g_n - g_j \leq \max(g_j, g_n) - g_j \leq \sum_{i=1}^{n-1} (\max(g_i, g_n) - g_i) \tag{2.63}$$

since each $\max(g_i, g_n) - g_i \geq 0$. The sum on the right hand side does not depend on j, so it is an upper bound for all of the $g_n - g_j$. By definition of h_n we then have

$$g_n - h_n \leq \sum_{i=1}^{n-1} (\max(g_i, g_n) - g_i). \tag{2.64}$$

The above inequality only involves Riemann integrable functions. Hence we are allowed to use the additivity of the integral to write the integral of the right hand side as the sum of the integrals. This gives

$$I(g_n) - I(h_n) \leq \sum_{i=1}^{n-1} (I(\max(g_i, g_n)) - I(g_i)). \tag{2.65}$$

For $i = 1, \ldots, n$ we have $g_i \leq p_i$ and $g_n \leq p_n \leq p_i$, so $\max(g_i, g_n) \leq p_i$. Therefore we have

$$I(g_n) - I(h_n) \leq \sum_{i=1}^{n-1} (L(p_i) - I(g_i)). \tag{2.66}$$

Hence,

$$I(g_n) - I(h_n) \leq \sum_{i=1}^{n-1} \epsilon \frac{1}{2^i} = \epsilon \left(1 - \frac{1}{2^{n-1}}\right). \tag{2.67}$$

This is the result that is needed. We conclude by noting that

$$L(p_n) - I(h_n) = L(p_n) - I(g_n) + I(g_n) - I(h_n) \leq \epsilon \frac{1}{2^n} + \epsilon \left(1 - \frac{1}{2^{n-1}}\right). \tag{2.68}$$

This gives

$$L(p_n) \leq I(h_n) + \epsilon \left(1 - \frac{1}{2^n}\right) < I(h_n) + \epsilon. \tag{2.69}$$

So when n is so large that $I(h_n)$ is less than ϵ, then $L(p_n)$ is less than 2ϵ. □

Remark: In the above proof it could be tempting to use $g_n - g_j \leq p_j - g_j$ for $j \leq n$ to prove $g_n - h_n \leq \sum_{i=1}^{n} (p_i - g_i)$. The problem would be that the right hand side only has a lower integral. Furthermore, the lower integral is only known to be superadditive, so the lower integral of the sum could be much larger than the sum of the lower integrals. This was avoided in the proof by using $\max(g_i, g_n)$ in place of p_i. |

Theorem 2.15 (Dominated convergence for the Riemann integral) *Let A be a bounded set. Let f_n be a sequence of real Riemann integrable functions on A that are dominated by a fixed constant M. Thus there is an M such that $|f_n(x)| \leq M$ for all x in A and all n. Suppose that $f_n \to f$ pointwise as $n \to \infty$, where f is also Riemann integrable. Then the integrals converge: $I(f_n) \to I(f)$ as $n \to \infty$.*

Remark: It actually suffices to prove an apparently weaker version of the theorem. This theorem says that if $\bar{f}_n \geq 0$ and $\bar{f}_n \leq 2M$ and $\bar{f}_n \to 0$ pointwise, then $I(\bar{f}_n) \to 0$. If we set $\bar{f}_n = |f_n - f|$, we see that $I(|f_n - f|) \to 0$. Hence $|I(f_n) - I(f)| = |I(f_n - f)| \leq I(|f_n - f|) \to 0$. |

Proof: Suppose that $\bar{f}_n \geq 0$ with $\bar{f}_n \leq 2M$ and $\bar{f}_n \to 0$ pointwise. It is sufficient to show that $I(\bar{f}_n) \to 0$. Here is the strategy. For each n let $p_n = \sup_{k \geq n} \bar{f}_k$ be the pointwise supremum of the $\bar{f}_k$ for all $k \geq n$. This is an infinite supremum! Then $0 \leq \bar{f}_n \leq p_n$. Furthermore, $p_n \downarrow 0$ is decreasing pointwise to zero. (Why?) There is an apparent problem: p_n need not be Riemann integrable. In this case $I(p_n)$ is not defined. However the lemma on monotone convergence proves that the lower integrals $L(p_n) \to 0$. But then $I(\bar{f}_n) = L(\bar{f}_n) \leq L(p_n) \to 0$, so we have the desired result. □

This dominated convergence theorem for the Riemann integral was first proved by Arzela in 1885. It states that if one has a dominated sequence of Riemann integrable functions that converges pointwise to a Riemann integrable function, then the limit of the integrals is the integral of the limit. This was long before the introduction of the Lebesgue integral; Lebesgue's thesis was published in 1902. The Lebesgue theory gives a much more powerful result. It says that if one has a dominated sequence of Lebesgue integrable functions that converge pointwise, then the limit is automatically Lebesgue integable, and the limit of the integrals is the integral of the limit.

This proof given above is elegant but somewhat strange. The theorem is a result about a sequence of Riemann integrable functions, but the proof uses a sequence of functions that need not in general be Riemann integrable. In fact, the proof uses ideas that are perhaps more natural in the context of the Lebesgue integral. It is taken from a paper by Luxemburg [10]. In this paper Luxemburg says that his proof is essentially the same as Hausdorff's proof published in 1927. However at one point Hausdorff gives an incorrect argument. This needs to be replaced by other reasoning; this is supplied by Luxemburg.

2.7 Differentiating a parameterized integral

In rough summary, the dominated convergence theorem says that if A is bounded and the functions $f_n(\mathbf{x})$ are uniformly bounded ($|f_n(\mathbf{x})| \leq M$), then the condition that $\lim_{n\to\infty} f_n(\mathbf{x}) = f(\mathbf{x})$ pointwise in $\mathbf{x}$ implies that

$$\lim_{n\to\infty} \int_A f_n(\mathbf{x})\, d\mathbf{x} = \int_A f(\mathbf{x})\, d\mathbf{x}. \tag{2.70}$$

There is a variant form of the dominated convergence theorem in which the functions depend on a continuous parameter. This says that if A is bounded and the functions $f(\mathbf{x}, \mathbf{y})$ are uniformly bounded ($|f(\mathbf{x}, \mathbf{y})| \leq M$), then the condition that $\lim_{\mathbf{y}\to\mathbf{y}_0} f(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}, \mathbf{y}_0)$ pointwise in $\mathbf{x}$ implies that

$$\lim_{\mathbf{y}\to\mathbf{y}_0} \int_A f(\mathbf{x}, \mathbf{y})\, d\mathbf{x} = \int_A f(\mathbf{x}, \mathbf{y}_0)\, d\mathbf{x}. \tag{2.71}$$

This variant form is a consequence of the original theorem and of a general fact about continuity in the metric space setting. The general fact is that if for every sequence $\mathbf{a}_n$ with $\lim_n \mathbf{a}_n = \mathbf{y}_0$ we have $\lim_n g(\mathbf{a}_n) = g(\mathbf{y}_0)$, then $g(\mathbf{y})$ is continuous at $\mathbf{y}_0$. We can then apply this to $g(\mathbf{y}) = \int_A f(\mathbf{x}, \mathbf{y})\, d\mathbf{x}$.

Theorem 2.16 *Let A be a bounded set that is Jordan measurable, and U be an open set. Suppose that $f(\mathbf{x}, \mathbf{y})$ is C^1, and write $f_2'(\mathbf{x}, \mathbf{y})$ for the covector of partial derivatives with respect to the $\mathbf{y}$ variables. Suppose also that there is a constant M with $|f_2'(\mathbf{x}, \mathbf{y})| \leq M$. Let*

$$g(\mathbf{y}) = \int_A f(\mathbf{x}, \mathbf{y})\, d\mathbf{x}. \tag{2.72}$$

Then

$$g'(\mathbf{y})\mathbf{h} = \int_A f_2'(\mathbf{x}, \mathbf{y})\mathbf{h}\, d\mathbf{x}. \tag{2.73}$$

Proof: Write

$$g(\mathbf{y} + \mathbf{h}) - g(\mathbf{y}) - \int_A f_2'(\mathbf{x}, \mathbf{y})\mathbf{h}\, d\mathbf{x} = \int_A (f(\mathbf{x}, \mathbf{y} + \mathbf{h}) - f(\mathbf{x}, \mathbf{y}) - f_2'(\mathbf{x}, \mathbf{y})\mathbf{h})\, d\mathbf{x}. \tag{2.74}$$

This can also be written

$$g(\mathbf{y} + \mathbf{h}) - g(\mathbf{y}) - \int_A f_2'(\mathbf{x}, \mathbf{y})\mathbf{h}\, d\mathbf{x} = \int_A \int_0^1 (f_2'(\mathbf{x}, \mathbf{y} + t\mathbf{h}) - f_2'(\mathbf{x}, \mathbf{y}))\mathbf{h}\, dt\, d\mathbf{x}. \tag{2.75}$$

This has absolute value bounded by

$$\epsilon(\mathbf{y}, \mathbf{h}) = \int_A \int_0^1 |f_2'(\mathbf{x}, \mathbf{y} + t\mathbf{h}) - f_2'(\mathbf{x}, \mathbf{y})|\, dt\, d\mathbf{x}. \tag{2.76}$$

times $|\mathbf{h}|$. All that remains is to show that $\epsilon(\mathbf{y}, \mathbf{h}) \to 0$ as $\mathbf{h} \to 0$. For fixed $\mathbf{x}$ and t the integrand approaches zero as $\mathbf{h} \to 0$. The conclusion follows from the dominated convergence theorem. □

This theorem gives a practical condition for differentiating an integral depending on a parameter with respect to the parameter. For the theorem to apply it is important that the bound on $f_2'(\mathbf{x}, \mathbf{y})$ be independent of $\mathbf{x}$ and of $\mathbf{y}$.

2.8 Approximate delta functions

Consider a function $\delta_1(\mathbf{x}) \geq 0$ that is positive and has integral one. Suppose for convenience that it is continuous and has compact support. For instance, it might vanish outside the closed ball of radius $c > 0$.

Construct a family of functions $\delta_\epsilon(\mathbf{x}) \geq 0$ defined for each $\epsilon > 0$ by scaling according to the rule

$$\delta_\epsilon(\mathbf{x}) = \delta_1(\frac{\mathbf{x}}{\epsilon})\frac{1}{\epsilon^n}. \tag{2.77}$$

Then it is easy to see that for each $\epsilon > 0$ the n-dimensional integral

$$\int \delta_\epsilon(\mathbf{x})\, d\mathbf{x} = 1. \tag{2.78}$$

This family of functions (considered for $\epsilon > 0$ small) will be called a family of *approximate delta functions*. Notice that δ_ϵ vanishes outside the closed ball of radius ϵc.

Theorem 2.17 *Let f be a bounded continuous function, and let δ_ϵ be a family of approximate delta functions. Then*

$$\lim_{\epsilon \to 0} \int f(\mathbf{x})\delta_\epsilon(\mathbf{x})\, d\mathbf{x} = f(0). \tag{2.79}$$

Proof: Write

$$\int f(\mathbf{x})\delta_\epsilon(\mathbf{x})\, d\mathbf{x} = \int f(\mathbf{x})\delta_1(\frac{\mathbf{x}}{\epsilon})\frac{1}{\epsilon^n}\, d\mathbf{x}. \tag{2.80}$$

Make the change of variable $\mathbf{x} = \epsilon\mathbf{u}$. This gives

$$\int f(\mathbf{x})\delta_\epsilon(\mathbf{x})\, d\mathbf{x} = \int f(\epsilon\mathbf{u})\delta_1(\mathbf{u})\, d\mathbf{u}. \tag{2.81}$$

The integrand converges pointwise in $\mathbf{u}$ to $f(0)\delta_1(\mathbf{u})$ on the closed ball of radius c and is bounded by a constant independent of ϵ. By the dominated convergence theorem the integral converges to $f(0)$. □

Remark: The amazing thing about this theorem is that the right hand side is independent of the particular choice of approximate delta function. For this reason, it is customary to write it in the form

$$\int f(\mathbf{x})\delta(\mathbf{x})\, d\mathbf{x} = f(0). \tag{2.82}$$

Of course, there is no such *delta function* $\delta(\mathbf{x})$ with this property, but it is still convenient to describe its properties. While the left hand side does not have a literal meaning, it gives an easy way to remember the result. Furthermore, it allows one to summarize various useful facts, such as

$$\int \delta(\mathbf{x})\, d\mathbf{x} = 1 \tag{2.83}$$

and

$$f(\mathbf{x})\delta(\mathbf{x}) = f(0)\delta(\mathbf{x}). \tag{2.84}$$

Also, the delta function is even

$$\delta(-\mathbf{x}) = \delta(\mathbf{x}) \tag{2.85}$$

and transforms under a change of scale $a \neq 0$ by

$$|a^n|\delta(a\mathbf{x}) = \delta(\mathbf{x}). \tag{2.86}$$

More generally, we have

$$|a^n|\delta(a\mathbf{x} - \mathbf{y}) = \delta(\mathbf{x} - \frac{1}{a}\mathbf{y}). \tag{2.87}$$

The reader may check that each of these suggests a meaningful statement about approximate delta functions. |

The integrals involving approximate delta functions are often of the form

$$\int h(\mathbf{y})\delta_\epsilon(\mathbf{z} - \mathbf{y})\, d\mathbf{y} = \int h(\mathbf{z} - \mathbf{x})\delta_\epsilon(\mathbf{x})\, d\mathbf{x}. \tag{2.88}$$

The two integrals expressions are equivalent after the change of variable $\mathbf{y} = \mathbf{z} - \mathbf{x}$. The new feature is that we look at the result as a function of $\mathbf{z}$. If $|\mathbf{h}(\mathbf{y})| \leq M$, then each integral above as a function of $\mathbf{z}$ has magnitude bounded by M. Furthermore, it is a continuous function of $\mathbf{z}$. Also, suppose that $\mathbf{h}$ has compact support K. Then this integral as a function of $\mathbf{z}$ has compact support in the set $K_{\epsilon c}$ of points a distance at most ϵc from K. The result may be stated in this context as follows.

Theorem 2.18 *Let h be a bounded continuous function, and let δ_ϵ be a family of approximate delta functions. Then*

$$\lim_{\epsilon \to 0} \int h(\mathbf{y})\delta_\epsilon(\mathbf{z} - \mathbf{y})\, d\mathbf{y} = \lim_{\epsilon \to 0} \int h(\mathbf{z} - \mathbf{x})\delta_\epsilon(\mathbf{x})\, d\mathbf{x} = h(\mathbf{z}). \tag{2.89}$$

2.9 Linear algebra (determinants)

If A is a square matrix, then there is an associated number $\det(A)$, the *determinant* of A. The determinant of a product of matrices is the product of the determinants. The determinant of the identity matrix I is $\det(I) = 1$. A square matrix A has an inverse matrix A^{-1} if and only if $\det(A) \neq 0$.

Various linear algebra operations may be expressed in terms of *elementary matrices*. These are square matrices E. For each such matrix there is a corresponding linear transformation $\mathbf{x} \mapsto E\mathbf{x}$. There are three kinds of elementary matrix:

1. A *scaling* matrix that is obtained by multiplying row i of the identity matrix by a constant $a \neq 0$. The corresponding linear transformation sends x_i to ax_i and leaves the other coordinates unchanged. The determinant is a.

2. An *interchange* matrix is obtained by interchanging rows i and j of the identity matrix. The corresponding linear transformation sends x_i to x_j and x_j to x_i and leaves the other coordinates unchanged. The determinant is -1.

3. A *shear* matrix is obtained by modifying row i of the identity matrix by adding a constant c times row j. The corresponding linear transformation sends x_i to $x_i + cx_j$ and leaves the other coordinates unchanged. The determinant is 1.

Every invertible matrix may be written as a product of elementary matrices. This gives a way of computing the determinant. The significance of the determinant is that its absolute value gives a factor by which volumes are multiplied. This is reflected in the ways that integrals are computed.

1. For a scaling the integral transforms by

$$|a| \int f(ax_i)\, dx_i = \int f(y_i)\, dy_i. \tag{2.90}$$

   Volumes are multiplied by the absolute value $|a|$ of the scale factor, which is the absolute value of the determinant.

2. For an interchange the integral transforms by

$$\int f(x_i, x_j)\, dx_i\, dx_j = \int f(y_j, y_i)\, dy_i\, dy_j. \tag{2.91}$$

   The absolute value of the determinant is 1. Volumes are left unchanged.

3. For an shear use Fubini's theorem to integrate with respect to x_i with x_j fixed.

$$\int f(x_i + cx_j, x_j)\, dx_i\, dx_j = \int f(y_i, x_j)\, dy_i\, dx_j = \int f(y_i, y_j)\, dy_i\, dy_j. \tag{2.92}$$

   The determinant is 1. Volumes are left unchanged

Theorem 2.19 *Under an invertible linear transformation an integral transforms by*

$$\int f(\mathbf{y})\, d\mathbf{y} = |\det(A)| \int f(A\mathbf{x})\, d\mathbf{x}. \tag{2.93}$$

Proof: Write $A = E_1 \cdots E_k$ as a product of elementary matrices. Then repeated use of the integral identities above gives

$$\int f(\mathbf{y})\, d\mathbf{y} = |\det(E_1)| \cdots |\det(E_k)| \int f(E_1 \cdots E_k \mathbf{x})\, d\mathbf{x}. \tag{2.94}$$

By the multiplicative properties of absolute value and determinant we have

$$|\det(E_1)| \cdots |\det(E_k)| = |\det(E_1) \cdots \det(E_k)| = |\det(E_1 \cdots E_k)|. \tag{2.95}$$

This gives the result. □

In general a change of variable is not given by a fixed linear transformation, and so the formula is not quite this simple. However for the case of an approximate delta function the result is essentially the same.

Theorem 2.20 *Consider an open subset V of $\mathbf{R}^n$ and a change of variable function $\mathbf{g} : V \to \mathbf{R}^n$. Suppose that $\mathbf{g}$ is one-to-one and C^1. Furthermore, suppose that $\det \mathbf{g}'(\mathbf{x}) \neq 0$. Let h be a bounded continuous function. Let δ_ϵ be a family of approximate delta functions. Let K be a compact subset of $\mathbf{g}(V)$ and consider $\mathbf{y}$ in K. Then*

$$\lim_{\epsilon \to 0} \int_V h(\mathbf{x}) \delta_\epsilon(\mathbf{g}(\mathbf{x}) - \mathbf{y})\, d\mathbf{x} = \frac{1}{|\det \mathbf{g}'(\mathbf{g}^{-1}(\mathbf{y}))|} h(\mathbf{g}^{-1}(\mathbf{y})). \tag{2.96}$$

Proof: First it is helpful to take some care about the region of integration. The integral is over the set of $\mathbf{x}$ with $|g(\mathbf{x}) - \mathbf{y}| \leq \epsilon c$. Consider the set $K_{\epsilon c}$ consisting of all points with distance no greater than ϵc from K. There is an ϵ_1 such that $K_{\epsilon_1 c}$ is a subset of $\mathbf{g}(V)$. Since this is a compact set, the function $\|(\mathbf{g}^{-1})'\|$ is bounded there by some constant λ. Now suppose that $\mathbf{y}$ is in K and $|\mathbf{y}' - \mathbf{y}| \leq \epsilon c$ for some $\epsilon \leq \epsilon_1$. Then $|\mathbf{g}^{-1}(\mathbf{y}') - \mathbf{g}^{-1}(\mathbf{y})| \leq \lambda |\mathbf{y}' - \mathbf{y}| \leq \lambda c \epsilon$. In particular for $\mathbf{x}$ in the region of integration $|\mathbf{x} - \mathbf{g}^{-1}(\mathbf{y})| \leq \lambda |\mathbf{g}(\mathbf{x}) - \mathbf{y}| \leq \lambda \epsilon c$.

Make the change of variable $\mathbf{x} = \mathbf{g}^{-1}(\mathbf{y}) + \epsilon \mathbf{t}$. The integration region is now $|\mathbf{t}| \leq \lambda c$. This is a fixed bounded set, independent of ϵ. The integral on the left hand side is

$$\int h(\mathbf{x}) \delta_1 \left(\frac{\mathbf{g}(\mathbf{x}) - \mathbf{y}}{\epsilon} \right) \frac{1}{\epsilon^n}\, d\mathbf{x} = \int h(\mathbf{g}^{-1}(\mathbf{y}) + \epsilon \mathbf{t}) \delta_1 \left(\frac{\mathbf{g}(\mathbf{g}^{-1}(\mathbf{y}) + \epsilon \mathbf{t}) - \mathbf{g}(\mathbf{g}^{-1}(\mathbf{y}))}{\epsilon} \right) d\mathbf{t}. \tag{2.97}$$

The integrand is bounded, independent of ϵ. By the dominated convergence theorem the limit as $\epsilon \to 0$ of this is

$$\int h(\mathbf{g}^{-1}(\mathbf{y})) \delta_1(\mathbf{g}'(\mathbf{g}^{-1}(\mathbf{y}))\mathbf{t})\, d\mathbf{t} = \int h(\mathbf{g}^{-1}(\mathbf{y})) \frac{1}{|\det(\mathbf{g}'(\mathbf{g}^{-1}(\mathbf{y})))|} \delta_1(\mathbf{u})\, d\mathbf{u}. \tag{2.98}$$

The last step is the change of variables $\mathbf{u} = \mathbf{g}'(\mathbf{g}^{-1}(\mathbf{y}))\mathbf{t}$. This involves a matrix $\mathbf{g}'(\mathbf{g}^{-1}(\mathbf{y}))$ that only depends on the parameter $\mathbf{y}$ and so may be regarded as constant. Performing the $\mathbf{u}$ integral gives the result on the right hand side. □

Remark: For later use we note that the integral $\int_V h(\mathbf{x}) \delta_\epsilon(\mathbf{g}(\mathbf{x}) - \mathbf{y})\, d\mathbf{x}$ bas a function of $\mathbf{y}$ is uniformly bounded on K, independent of ϵ. Furthermore,

for each ϵ it is continuous in $\mathbf{y}$. (The dominated convergence also applies to integrals that depend on a real parameter such as $\mathbf{y}$.) |
Remark: Again there is a common abbreviation for this kind of result. In the present case one could write

$$\int_V h(\mathbf{x})\delta(\mathbf{g}(\mathbf{x}) - \mathbf{y})\, d\mathbf{x} = \frac{1}{|\det \mathbf{g}'(\mathbf{g}^{-1}(\mathbf{y}))|} h(\mathbf{g}^{-1}(\mathbf{y})). \tag{2.99}$$

Even more radically, one could write

$$\delta(\mathbf{g}(\mathbf{x}) - \mathbf{y}) = \frac{1}{|\det \mathbf{g}'(\mathbf{g}^{-1}(\mathbf{y}))|} \delta(\mathbf{x} - \mathbf{g}^{-1}(\mathbf{y})). \tag{2.100}$$

Even if there is no such thing as a delta function, such identities involving delta functions make perfectly good sense. |

2.10 Change of variables

Theorem 2.21 (Change of variables for the Riemann integral) *Consider an open subset V of $\mathbf{R}^n$ and a change of variable function $\mathbf{g} : V \to \mathbf{R}^n$. Suppose that $\mathbf{g}$ is one-to-one and C^1. Furthermore, suppose that $\det \mathbf{g}'(\mathbf{x}) \neq 0$. Let f be a Riemann integrable function with compact support in $\mathbf{g}(V)$. Then*

$$\int_{\mathbf{g}(V)} f(\mathbf{y})\, d\mathbf{y} = \int_V f(g(\mathbf{x}))|\det \mathbf{g}'(\mathbf{x})|\, d\mathbf{x}. \tag{2.101}$$

Proof: First we give the proof for the case when f is a continuous function. The plan is to do the proof in two parts: write the right hand side as a limit, and then write the left hand side as the same limit.

Here is the part dealing with the integral on the right hand side. Let K be the support of f. We first have a limit of single integrals

$$\int_K f(\mathbf{y})|\det \mathbf{g}'(\mathbf{g}^{-1}(\mathbf{y}))|\delta_\epsilon(\mathbf{g}(\mathbf{x}) - \mathbf{y})\, d\mathbf{y} \to f(g(\mathbf{x}))|\det \mathbf{g}'(\mathbf{x})| \tag{2.102}$$

as $\epsilon \to 0$. This result uses that continuity of f.

The integral in the above formula vanishes unless $\mathbf{g}(\mathbf{x})$ is in $K_{\epsilon c}$. Choose ϵ_1 so that $K_{\epsilon_1 c}$ is in $\mathbf{g}(V)$. So for $\epsilon \leq \epsilon_1$ we may integrate $\mathbf{x}$ over the compact set $L = \mathbf{g}^{-1}(K_{\epsilon_1 c})$. Furthermore, these integrals are uniformly bounded, independent of ϵ. By the dominated convergence theorem we get limit of double integrals

$$\int_L \int_K f(\mathbf{y})|\det \mathbf{g}'(\mathbf{g}^{-1}(\mathbf{y}))|\delta_\epsilon(\mathbf{g}(\mathbf{x}) - \mathbf{y})\, d\mathbf{y}\, d\mathbf{x} \to \int_V f(g(\mathbf{x}))|\det \mathbf{g}'(\mathbf{x})|\, d\mathbf{x} \tag{2.103}$$

as $\epsilon \to 0$.

Here is the part dealing with the integral on the left hand side. First we have a limit of single integrals. By the change of variable formula for approximate delta functions we have

$$\lim_{\epsilon \to 0} \int_L f(\mathbf{y})|\det \mathbf{g}'(\mathbf{g}^{-1}(\mathbf{y}))|\delta_\epsilon(\mathbf{g}(\mathbf{x}) - \mathbf{y})\, d\mathbf{x} \to f(\mathbf{y}) \tag{2.104}$$

as $\epsilon \to 0$.

As a function of $\mathbf{y}$ the above integral is bounded by a constant independent of ϵ and is supported on K. The dominated convergence theorem gives a limit of double integrals

$$\int_K \int_L f(\mathbf{y}) |\det \mathbf{g}'(\mathbf{g}^{-1}(\mathbf{y}))| \delta_\epsilon(\mathbf{g}(\mathbf{x}) - \mathbf{y}) \, d\mathbf{x} \, d\mathbf{y} \to \int_K f(\mathbf{y}) \, d\mathbf{y} \tag{2.105}$$

as $\epsilon \to 0$.

The proof for continuous f is concluded by noting that according to Fubini's theorem the two double integrals are the same.

The proof for general Riemann integrable functions requires additional comment. First, the fact that the right hand side is integrable follows from general properties of the Riemann integral with respect to composition and multiplication. Then one can use approximation by continuous functions. In fact, for every Riemann integrable f there are continuous functions $g \le f \le h$ with integrals arbitrarily close to the integral of f. Since the integrals on the left hand side are arbitrarily close, it follows that the integrals on the right hand side are also arbitrarily close. □

This beautiful proof is from a recent paper by Ivan Netuka [15]. (A few details of the proof have been changed.) That formulation in this paper is for the Lebesgue integral, but it also works for the Riemann integral if one recognizes that the Riemann integral also has a dominated convergence theorem.

Remark: A physicist or engineer who is familiar with delta functions might summarize the entire proof by recalling that

$$|\det \mathbf{g}'(\mathbf{g}^{-1}(\mathbf{y}))| \delta_\epsilon(\mathbf{g}(\mathbf{x}) - \mathbf{y}) = \delta(\mathbf{x} - \mathbf{g}^{-1}(\mathbf{y})). \tag{2.106}$$

So

$$\int f(g(\mathbf{x}))| \det \mathbf{g}'(\mathbf{x})| \, d\mathbf{x} = \int \int f(\mathbf{y})| \det \mathbf{g}'(\mathbf{g}^{-1}(\mathbf{y}))| \delta(\mathbf{g}(\mathbf{x}) - \mathbf{y}) \, d\mathbf{y} \, d\mathbf{x} \tag{2.107}$$

which in turn is equal to

$$\int \int f(\mathbf{y}) \delta(\mathbf{x} - \mathbf{g}^{-1}(\mathbf{y})) \, d\mathbf{x} \, d\mathbf{y} = \int f(\mathbf{y}) \, d\mathbf{y} \tag{2.108}$$

Thus the proof is immediately memorable. |

2.11 Problems

Problems 5: Dominated Convergence

1. (a) Give an example of functions f and g on $[0, 1]$ such that their lower integrals satisfy $L(f + g) > L(f) + L(g)$.

   (b) Give an example of functions f and g on $[0, 1]$ such that their upper integrals satisfy $U(f + g) < U(f) + U(g)$.

2. Consider a function $f(x, y)$ defined for $0 \leq x \leq 1, 0 \leq y \leq 1$. Suppose that $(x, y) \mapsto f(x, y)$ is Riemann integrable as a function on the square. Give an example to show that there can be y such that $x \mapsto f(x, y)$ is not Riemann integrable. In your example, calculate the upper integral $h(y)$ of $x \mapsto f(x, y)$. Does the integral of $y \mapsto h(y)$ exist? Justify your answer.

3. Consider functions $f_n(x)$ with $f_n(x) = n^2 x$ for $0 \leq x \leq 1/n$ and $f(x) = n - n^2(x - 1/n)$ for $1/n \leq x \leq 2/n$, zero elsewhere. Find the integral of the pointwise limit. Find the limit of the integrals.

4. Consider functions defined on $[a, b]$. Say that $f_n \downarrow f$ if for each $m \geq n$ we have $f_m(x) \leq f_n(x)$, and if $f(x)$ is the infimum of the $f_n(x)$. (a) Show that the limit of the $f_n(x)$ as $n \to \infty$ is $f(x)$. (b) Suppose each f_n is continuous. Prove or disprove: The convergence must be uniform.

5. Let f_n be a sequence of Riemann integrable functions on $[a, b]$ with each $f_n(x) \geq 0$, uniformly bounded. We would like to have a sequence p_n of functions such that for each n we have $f_n(x) \leq p_n(x)$ and such that $p_n(x) \downarrow p(x)$. An obvious device is to take $p_n(x) = \sup_{k \geq n} f_k(x)$. (a) Prove that the sequence $p_n(x)$ is decreasing. (b) Must each p_n be Riemann integrable? Give a careful discussion and proof.

Recitation 5

1. Evaluate
$$\int_0^2 \int_0^3 y e^{-xy} \, dy \, dx \tag{2.109}$$
without using integration by parts.

2. (a) Consider the function f defined on the unit square with $f(x, y) = 1$ if x is rational and $y = \frac{1}{2}$, zero otherwise. Is it Riemann integrable? Prove that your answer is correct.

   (b) Consider the function f defined on the unit square with $f(x, y) = 1$ if x is rational and y is rational, zero otherwise. Is it Riemann integrable? Prove that your answer is correct.

   (c) Consider the function f defined on the unit square with $f(x, y) = 1$ if x is rational and y is rational and $x = y$, zero otherwise. Is it Riemann integrable? Prove that your answer is correct.

   (d) Consider the function f defined on the unit square with $f(x, y) = 1$ if x is rational and y is rational and $x \neq y$, zero otherwise. Is it Riemann integrable? Prove that your answer is correct.

   (e) Consider the function f defined on the unit square with $f(x, y) = 1$ if x is rational or y is rational, zero otherwise. Is it Riemann integrable? Prove that your answer is correct.

3. Consider the functions $f_n(x) = x^n$ defined for $0 \leq x \leq 1$. Find the integral of the limit and the limit of the integrals. (a) Does the dominated convergence theorem apply? Explain in detail. (b) Does Dini's theorem apply? Explain in detail. (c) Is there uniform convergence? Justify your answer directly from the definition of uniform convergence.

4. For each $t > 0$ define the function

$$\delta_t(x) = \max(\frac{1}{\sqrt{t}} - \frac{1}{t}|x - \sqrt{t}|, 0) \tag{2.110}$$

Define $\delta_0(x) = 0$ and $\delta_{-t}(x) = \delta_t(x)$.

Also, for each t define

$$\phi(x, t) = t\delta_t(x). \tag{2.111}$$

(a) Compute the x integral of $\delta_t(x)$ for each $t \neq 0$. Let $t \to 0$. Compute the x integral of the pointwise limit and the limit of the x integrals.

(b) Compute the x integral of $\phi(x, t)$. Compute the t derivative of the x integral of $\phi(x, t)$.

(c) For fixed x, compute the t derivative of $\phi(x, t)$ at $t = 0$. Compute the x integral of this t partial derivative.

(d) What does this say about differentiating under the integral sign?

Problems 6: Change of Variables

1. (a) For each $k = 0, 1, 2, 3, \ldots$ find a C^k approximate delta function that vanishes outside of a bounded set? (b) Is there a C^∞ approximate δ function that vanishes outside of a bounded set? Prove or disprove.

2. Let $u = x^3 - 3xy^2$ and $v = 3x^2y - y^3$. Consider the region D of x, y in the first quadrant such that the corresponding u, v satisfy $|u| < 3$ and $|v - 2| < 1$. Evaluate

$$\int_D (x^2 + y^2)^2 \, dx \, dy. \tag{2.112}$$

3. (a) There is another interesting change of variable formula in one domension. It says that if $[a, b]$ may be partitioned into intervals I such that on each interval $g'(x) > 0$ or $g'(x) < 0$, then

$$\int_a^b h(g(x)) \, dx = \int_{-\infty}^{\infty} h(y) \sum_{\{t | g(t) = y\}} \frac{1}{|g'(t)|} \, dy. \tag{2.113}$$

The t sum is restricted to t between a and b for which $g'(t) \neq 0$. If there are no such t then the sum is zero. Hint: The sum over intervals I of the **y** integrals over $g(I)$ is the integral over y of the sum over the intervals I with y in $g(I)$.

(b) Let $a > 0$. Compute $\delta(x^2 - a^2)$ as a multiple of a sum of two delta functions of the form $\delta(x \pm a)$.

4. There is a marvelous formula for computing the $n-1$ dimensional surface area for an implicitly defined surface $g(\mathbf{x}) = c$ in some region of $\mathbf{R}^n$. The rule is to compute the n-dimensional integral of $\delta(g(\mathbf{x}) - c)|dg(\mathbf{x})|\, d\mathbf{x}$.

   (a) For $n = 2$ this is the integral of $\delta(g(x,y) - c)|dg(x,y)|\, dx\, dy$. Here

   $$dg(x,y) = \frac{\partial g(x,y)}{\partial x}\, dx + \frac{\partial g(x,y)}{\partial y}\, dy. \tag{2.114}$$

   and

   $$|dg(x,y)| = \sqrt{\left(\frac{\partial g(x,y)}{\partial x}\right)^2 + \left(\frac{\partial g(x,y)}{\partial y}\right)^2}. \tag{2.115}$$

   Evaluate this in terms of an x integral involving partial derivatives of $g(x,y)$. These partial derivatives will be evaluated at the (implicitly defined) y satisfying $g(x,y) = c$. Hint: First do the y integral.

   (b) Say that the implicit function theorem applied to $g(x,y) = c$ defines $y = h(x)$ as a function of x. Express the above result in terms of derivatives of $h(x)$. Show that this gives the usual formula for arc length.

5. Use the general formula to compute the area of the hemisphere $z \geq 0$ for the sphere $x^2 + y^2 + z^2 = a^2$. This is the three-dimensional integral of $\delta(x^2 + y^2 + z^2 - a^2) 2\sqrt{x^2 + y^2 + z^2}\, dx\, dy\, dz$. Hint: First do the z integral to express the result as an x, y integral. Then it is easy to do this x, y integral in polar coordinates.

Recitation 6

1. Evaluate

   $$\int_0^1 \int_{3y}^3 e^{-x^2}\, dx\, dy. \tag{2.116}$$

2. A previous problem gave a marvelous formula for computing the $n-1$ dimensional surface area for an implicitly defined surface $g(\mathbf{x}) = 0$ in some region of $\mathbf{R}^n$. Explain why this formula is "obvious" by drawing relevant pictures.

3. Do the integral

   $$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2+y^2}\, dx\, dy \tag{2.117}$$

   by changing to polar coordinates.

4. Do the integral

   $$\int_{-\infty}^{\infty} e^{-x^2}\, dx \tag{2.118}$$

   by applying Fubini's theorem to the previous problem.

5. The Gamma function may be written

$$\Gamma(z) = 2\int_0^\infty e^{-r^2} r^{2z-1} dr. \tag{2.119}$$

Show that $\Gamma(\frac{1}{2}) = \pi^{\frac{1}{2}}$.

6. Let a_{n-1} be the area of the unit sphere in dimension $n-1$. Prove that $\pi^{\frac{n}{2}} = a_{n-1}\frac{1}{2}\Gamma(\frac{n}{2})$. Hint: Express the n dimensional integral $\int e^{-\mathbf{x}^2}\, d\mathbf{x}$ in Cartesian coordinates and in polar coordinates.

Chapter 3

Differential Forms

3.1 Coordinates

There are two theories of integration. The first describes how to integrate a function over a set. The second explains how to integrate a differential form over an oriented surface. It is the second theory that is the natural setting for calculus. It is also a thread that runs through much of geometry and even of applied mathematics. The topic of this chapter is differential forms and their integrals, culminating with the general form of Stokes' theorem.

In the following a function is said to be *smooth* if it is C^∞. Two open sets U, V of $\mathbf{R}^n$ are said to be *diffeomorphic* if there is a one-to-one smooth function $\mathbf{f} : U \to V$ with smooth inverse function $\mathbf{f}^{-1} : V \to U$.

An n-dimensional *manifold patch* is a set M together with a collection of functions defined on M. Each such function is one-to-one from M onto some open subset of $\mathbf{R}^n$. Such a function is called a coordinate system. There are two requirements on the set of coordinate system functions.

- If $\mathbf{x} = (x_1, \ldots, x_n) : M \to U$ is a coordinate system with values in the open set U, and if $\mathbf{f} : U \to V$ is a diffeomorphism from the open set U to the open set V, then $\mathbf{f}(\mathbf{x}) : M \to U$ is also a coordinate system.

- If $\mathbf{x} = (x_1, \ldots, x_n) : M \to U$ is a coordinate system mapping M onto U, and if $\mathbf{u} = (u_1, \ldots, u_n) : M \to V$ is a coordinate system mapping M onto V, then there is a diffeomorphism $f : U \to V$ such that $\mathbf{u} = \mathbf{f}(\mathbf{x})$.

This definition of manifold patch is not a standard definition; its purpose here is to capture the idea of a set with many different coordinate systems, each of which is defined on the same set.

For $n = 1$ a typical manifold patch is something like a featureless curve. It does not have a shape, but one could think of something like the letter S, without the end points. (One should not think of a curve in the shape of an 0, because it does not match with an open interval of numbers. Also one should not have a curve that looks like the Greek α, because it has a self-intersection point.) For $n = 2$ a typical manifold patch is like a patch of cloth, but without a border. It can have holes. The case $n = 0$ is a single point.

The coordinate functions serve to attach numbers to the points of M. This can be done in many ways, and there is no one coordinate system superior to all the others, at least not without further consideration. Furthermore, in many treatments (including here) the points in M do not have names. If we want to specify a point, we say that it is the point in M such that in a particular coordinate system the coordinate values are certain specified numbers.

The concept of manifold patch is natural in geometry and in many applications of mathematics. A manifold patch M is (at least at the outset) assumed to be featureless, except for the fact that it can have various coordinate systems. These coordinate systems are supposed to be on a democratic status; one is as good as another. Since the notions of open set, closed set, compact set, continuous function, smooth function, smooth curve, and so on are independent of coordinate system, they make sense for a manifold patch. On the other hand,

notions of distance, angle, congruence, and so on are not defined (although they may be introduced later). It is amazing how much useful mathematics can be done in this general context.

The manifold patch concept is particularly suited to the *local* notion of geometry, that is, it gives a description of what happens near a point. This is because a manifold patch is modeled on an open subset of $\mathbf{R}^n$. There is a more general concept of *manifold* that consists of many manifold patches joined together. This is a *global* or big picture notion of geometry, and it is a fascinating topic in advanced mathematics. Here we focus on the local story, with only brief mention of possible global issues.

Example: Here is a typical example of a manifold patch in applied mathematics. Consider a box of gas with pressure p and volume V and temperature T. These are related by an equation of state $f(p, V, T) = 0$. In nice cases this equation may be solved for one variable in terms of the others via the implicit function theorem. Finding the equation of state is a major task of physics. But even after this is determined, it is not clear which coordinates will be most convenient to use. In this case the manifold patch M is the set of possible states of the gas. One possible coordinate system is p, V. Another is p, T. Yet another is V, T. Physicists and chemists and geologists use whichever coordinate system is appropriate to the problem under consideration. |

3.2 Scalar fields

There are three notions that are important at the outset, that of scalar field, vector field, and differential 1-form. We look at each of them in turn. In general, we shall require that these be defined on a manifold patch M.

A *scalar field* is a function $s = h(\mathbf{u}) = h(u_1, \ldots, u_n)$ from M to $\mathbf{R}$. Here $\mathbf{u} = (u_1, \ldots, u_n)$ is some coordinate system mapping M onto the open set V, and h is a smooth function on V. Usually we picture such a function by drawing contour lines of s on M. This sketch may indicate maxima and minima and saddle points of s. Notice that while we can express $s = h(\mathbf{u})$, we can also express $s = h(\mathbf{f}(\mathbf{x}))$ in terms of some other coordinate system. The notion of scalar field is independent of coordinate system.

Example: Here is a simple example that illustrates the notion of manifold patch in an elementary application. Consider the problem of making a box with a given amount of material to contain the maximum volume. The box will have five sides, a base and four vertical sides. It is open at the top. In this case the manifold patch is the set M of possible boxes made with this material.

Say that the side lengths of the base are u, v and the height is w. The amount of material available is a fixed number A. Thus $uv + 2uw + 2vw = A$. There are various possible coordinate systems to describe the boxes. One possibility is u, v. Another is u, w. Yet another is v, w. These each send M to a coordinate patch. Each coordinate system may be expressed in terms of each other coordinate system. For instance, we may express u, v as in terms of u, w by noting that $v = (A - 2uw)/(u + 2w)$ for $0 < 2uw < A$.

The volume $V = uvw$ is a function on M that is a scalar field. For every constant $c > 0$ the set of points in M (box shapes) where $V = c$ is either a curve (the set of boxes with volume c), a single point (the box where V assumes its maximum value), or empty. Later on we shall solve the problem of finding the box shape for which V assumes its maximum value. The answer turns out to be that $w/u = \frac{1}{2}$ and $w/v = \frac{1}{2}$ for this shape. That is, the open top box with maximal volume has a square base and a height half the length of each side of the base.

The solution follows a typical strategy for applications of mathematics. Consider the set M of objects of interest. Attach various numerical values to the points in M and label these by variable names. Use what you know about M to establish various relations between these variable. Finally, make some (perhaps arbitrary) choice of which variables will serve as coordinates, and eliminate the other variables. Then the mathematical calculation begins. |

3.3 Vector fields

A *vector field* is a differential operator of the form

$$X = \sum_{j=1}^{n} a_j \frac{\partial}{\partial x_j}. \tag{3.1}$$

Here each a_j is a scalar field. The differential operator acts on a scalar field s to give another scalar field

$$X\, s = \sum_{j=1}^{n} a_j \frac{\partial s}{\partial x_j}. \tag{3.2}$$

Again, the notion of vector field is independent of the coordinate system. Thus we can also write

$$X = \sum_{i=1}^{n} \bar{a}_i \frac{\partial}{\partial u_i}. \tag{3.3}$$

Here

$$\bar{a}_i = \sum_{j=1}^{n} \frac{\partial u_i}{\partial x_j} a_j = \sum_{i=1}^{n} f'_{i,j}(\mathbf{x}) a_j. \tag{3.4}$$

One can picture a vector field in the following way. At each point of M one draws an arrow. This arrow is not to be thought of as a displacement in M, but as a kind of rate of change at this point of M. For instance, if M is thought of a region where there is a fluid flow, then the vector field might be something like the fluid velocity. More precisely, the components of the vector field are velocities. The vector field itself describes how quantities change with time as they move with the fluid.

Giving a vector field is equivalent to giving a system of ordinary differential equations. More precisely, it is equivalent to giving an *autonomous system* of

first order ordinary differential equations. (The word autonomous means that the vector field does not change with time.) The equations corresponding to vector field $\sum_j g_j(\mathbf{x})\partial/\partial x_j$ are

$$\frac{dx_j}{dt} = g_j(x_1, \ldots, x_n). \tag{3.5}$$

Of course this can also be written in the abbreviated form $d\mathbf{x}/dt = \mathbf{g}(\mathbf{x})$. A solution of such an equation is given by functions $h_j(t)$ with

$$\frac{dh_j(t)}{dt} = g_j(h_1(t), \ldots, h_n(t)). \tag{3.6}$$

Again this has a brief form $d\mathbf{h}(t)/dt = \mathbf{g}(\mathbf{h}(t))$.

Locally, a vector field is a fairly boring object, with one exception. This is at a point in the manifold patch M where the vector field X vanishes, that is, where each a_j has the value zero. Away from such points the vector field is doing nothing more interesting than uniform motion.

Theorem 3.1 (Straightening out theorem) . *If*

$$X = \sum_{i=1}^{n} a_i \frac{\partial}{\partial x_i} \neq 0 \tag{3.7}$$

is a vector field that is non-zero near some point, then near that point there is another coordinate system $u_1, \ldots, u_n$ in which it has the form

$$X = \frac{\partial}{\partial u_j}. \tag{3.8}$$

Proof: Here is the idea of the proof of the straightening out theorem. Say that $a_j \neq 0$. Solve the system of differential equations

$$\frac{dx_i}{dt} = a_i \tag{3.9}$$

with initial condition 0 on the surface $x_j = 0$. This can be done locally, by the existence theorem for systems of ordinary differential equations with smooth coefficients. The result is that x_i is a function of the coordinates x_i for $i \neq j$ restricted to the surface $x_j = 0$ and of the time parameter t. Furthermore, since $dx_j/dt \neq 0$, the condition $t = 0$ corresponds to the surface $x_j = 0$. So if $x_1, \ldots, x_n$ corresponds to a point in M near the given point, we can define for $i \neq j$ the coordinates u_i to be the initial value of x_i on the $x_j = 0$, and we can define $u_j = t$. In these coordinates the differential equation becomes

$$\frac{du_i}{dt} = 0, i \neq j, \tag{3.10}$$

$$\frac{du_j}{dt} = 1. \tag{3.11}$$

□
Example: Consider the vector field

$$-y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} \tag{3.12}$$

away from the origin. The corresponding system is

$$\frac{dx}{dt} = -y \tag{3.13}$$

$$\frac{dy}{dt} = x. \tag{3.14}$$

Take the point to be $y = 0$, with $x > 0$. Take the initial condition to be $x = r$ and $y = 0$. Then $x = r\cos(t)$ and $y = r\sin(t)$. So the coordinates in which the straightening out takes place are polar coordinates r, t. Thus if we write $x = r\cos(\phi)$ and $y = r\sin(\phi)$, we have

$$-y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} = \frac{\partial}{\partial \phi}, \tag{3.15}$$

where the partial derivative with respect to ϕ is taken with r held fixed. |
Example: Consider the Euler vector field

$$x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} = r\frac{\partial}{\partial r}, \tag{3.16}$$

where the partial derivative with respect to r is taken with fixed ϕ. We need to stay away from the zero at the origin. If we let $t = \ln(r)$, then this is

$$x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} = r\frac{\partial}{\partial r} = \frac{\partial}{\partial t}, \tag{3.17}$$

where the t derivative is taken with ϕ fixed. |

Say that a vector field defining a system of ordinary differential equations has an isolated zero. Thus the equation $d\mathbf{x}/dt = \mathbf{g}(\mathbf{x})$ has a point $\mathbf{x}^*$ where $\mathbf{g}(\mathbf{x}^*) = 0$. Then it becomes interesting to approximate this equation near the point by

$$\frac{d\bar{\mathbf{x}}}{dt} = \mathbf{g}'(\mathbf{x}^*)\bar{\mathbf{x}}, \tag{3.18}$$

where $\bar{\mathbf{x}} = \mathbf{x} - \mathbf{x}^*$. This is called the *linearization* of the vector field at the point $\mathbf{x}^*$. The behavior of the linearization is studied by finding the eigenvalues of the matrix $\mathbf{g}'(\mathbf{x}^*)$.

This story is already of interest in the case $n = 2$. Here are some common cases.

Stable node Real eigenvalues with $\lambda_1 < 0, \lambda_2 < 0$.

Unstable node Real eigenvalues with $\lambda_1 > 0, \lambda_2 > 0$.

Hyperbolic fixed point (saddle) Real eigenvalues with $\lambda_1 < 0 < \lambda_2$.

Stable spiral Nonreal eigenvalues with $\lambda = \mu \pm i\omega$, $\mu < 0$.

Unstable spiral Nonreal eigenvalues with $\lambda = \mu \pm i\omega$, $\mu > 0$.

Elliptic fixed point (center) Nonreal eigenvalues $\lambda = \pm i\omega$.

There are yet other cases when one of the eigenvalues is zero.
Example: Here is a fairly typical analysis of a vector field via fixed points and linearizations. Consider the system

$$\begin{aligned} \frac{du}{dt} &= u(v-1) \\ \frac{dv}{dt} &= 4 - u^2 - v^2. \end{aligned} \tag{3.19}$$

There are fixed points in the u, v plane at $(0, 2)$ and $(0, -2)$ and $(\sqrt{3}, 1)$ and $(-\sqrt{3}, 1)$. We can compute the linearizations at each of these fixed points. The first two are hyperbolic fixed points with vertical and horizontal eigenspaces. The eigenvalues at $(0, -2)$ are -3 in the horizontal direction and 6 in the vertical direction. The eigenvalues at $(0, 2)$ are 1 in the horizontal direction and -4 in the vertical direction. By setting $u = 0$ in the original equation one can see that v moves along the vertical axis from the fixed point at $(0, -2)$ to the fixed point at $(0, 2)$.

The eigenvalues at $(\sqrt{3}, 1)$ and $(-\sqrt{3}, 1)$ are both stable spirals with eigenvalues $-1 \pm \sqrt{5}i$. The first one spirals in clockwise, while the other one spirals in counter-clockwise. There are orbits that approach $(0, -2)$ from either side. Points below these orbits never reach the spirals. Everything above on the left gets attracted to the left spiral, while everything above on the right gets attracted to the right spiral.

Suppose that the same system were viewed in another coordinate system. Many details would differ. However there would still be four fixed points, and the eigenvalues of the linearizations would be the same. So the qualitative picture would be much the same, but perhaps less symmetric. |
Example: A classic example from physics is the pendulum

$$\frac{dq}{dt} = \frac{1}{m}p \tag{3.20}$$

$$\frac{dp}{dt} = -mg\sin(\frac{1}{a}q). \tag{3.21}$$

Here $q = a\theta$ represents displacement, and p represents momentum. The zeros are at $\theta = n\pi$. When n is even this is the pendulum at rest in a stable position; when n is odd this is the pendulum at rest upside down, in a very unstable position. The linearization at a zero is

$$\frac{dq}{dt} = \frac{1}{m}p \tag{3.22}$$

$$\frac{dp}{dt} = -\frac{mg}{a}(-1)^n q. \tag{3.23}$$

In matrix form this is

$$\frac{d}{dt}\begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{m} \\ -(-1)^n \frac{mg}{a} & 0 \end{pmatrix} \begin{pmatrix} \tilde{q} \\ \tilde{p} \end{pmatrix}. \tag{3.24}$$

The eigenvalues λ are given by $\lambda^2 = -(-1)^n \frac{g}{a}$. When n is even we get an elliptic fixed point, while when n is odd we get a hyperbolic fixed point.

The big picture is seen by examining the scalar

$$H = \frac{1}{2m}p^2 - mga\cos(\frac{1}{a}q). \tag{3.25}$$

This is the energy, and it is constant for each solution. While the energy does not describe the time dependence of the solutions, it does show the shape of the solutions in the phase plane. |

The following question is natural. Suppose that a vector field has an isolated zero. At that zero it has a linearization. Is it possible to choose coordinates nearby so that the vector field is given in those new coordinates by its linearization? It turns out that this can often be done. The answer to the question is negative in general. See Nelson [14] for a discussion of this delicate matter.

3.4 Fluid velocity and the advective derivative

It may be puzzling to think of a vector field as a differential operator. This section is a digression to point out that this kind of construction is quite natural in the context of fluid dynamics.

The velocity field of a fluid is an important example of a vector field. Consider a fluid with a velocity vector field that is independent of time. For simplicity, consider a two-dimensional case; think of the velocity of the water on the surface of a river. Just because the velocity does not depend on time, this does not mean that there is no motion. If one follows a particular particle, it is transported along a path. If the position of the particle is described by coordinates u, v, then the motion of the particle is given by

$$\begin{aligned} \frac{du}{dt} &= a = f(u,v) \\ \frac{dv}{dt} &= b = g(u,v). \end{aligned} \tag{3.26}$$

Here a, b are the velocity vector field components with respect to the coordinates u, v.

Now let $s = h(u,v)$ be some time-independent quantity. For instance, it could be the temperature of the fluid at each point in space. If we follow this quantity along a particle, then it does change in time, according to

$$\frac{ds}{dt} = \frac{\partial s}{\partial u}\frac{du}{dt} + \frac{\partial s}{\partial v}\frac{dv}{dt} = \left(a\frac{\partial}{\partial u} + b\frac{\partial}{\partial v}\right)s. \tag{3.27}$$

In fluid dynamics the differential operator on the right represents the effect of the fluid flow given by the velocity vector field. It is the derivative following the motion of the particle. It is so important that it has many names: *advective derivative*, *particle derivative*, *material derivative*, *substantial derivative*, *Lagrangian derivative*, and so on. The components a, b of the velocity vector field depend on the coordinate system. If now one changes coordinates, say to w, z, then the equation becomes

$$\begin{aligned} \frac{dw}{dt} &= p = m(z, w) \\ \frac{dz}{dt} &= q = n(z, w), \end{aligned} \tag{3.28}$$

where

$$\begin{aligned} p &= a\frac{\partial w}{\partial u} + b\frac{\partial w}{\partial v} \\ q &= a\frac{\partial z}{\partial u} + b\frac{\partial z}{\partial v}. \end{aligned} \tag{3.29}$$

Then a straightforward calculation gives

$$p\frac{\partial}{\partial w} + q\frac{\partial}{\partial z} = a\left(\frac{\partial w}{\partial u}\frac{\partial}{\partial w} + \frac{\partial z}{\partial u}\frac{\partial}{\partial z}\right) + b\left(\frac{\partial w}{\partial v}\frac{\partial}{\partial w} + \frac{\partial z}{\partial v}\frac{\partial}{\partial z}\right). \tag{3.30}$$

By the chain rule this is

$$p\frac{\partial}{\partial w} + q\frac{\partial}{\partial z} = a\frac{\partial}{\partial u} + b\frac{\partial}{\partial v}. \tag{3.31}$$

Since the advective derivative represents the rate of change along the motion of the particle, it is independent of the coordinate system. Specifying the advective derivative is thus a particularly attractive way of specifying the vector field.

3.5 Differential 1-forms

A *differential 1-form* ω assigns to each vector field X a corresponding scalar field $\langle \omega \mid X \rangle$. There is a very important way of constructing a differential 1-form ds from a scalar field s. This is called the *differential* of the scalar field. The definition is

$$\langle ds \mid X \rangle = X\, s. \tag{3.32}$$

This is perhaps the fundamental definition in the entire subject. Take

$$X = \sum_{j=1}^{n} a_j \frac{\partial}{\partial x_j}. \tag{3.33}$$

If we write out the definition explicitly, we get

$$\langle ds \mid X \rangle = \sum_{j=1}^{n} a_j \frac{\partial s}{\partial x_j}. \tag{3.34}$$

If we apply this to the scalar $s = x_i$ that is one of the coordinates, then we get

$$\langle dx_i \mid X \rangle = a_i. \tag{3.35}$$

It follows that

$$\langle ds \mid X \rangle = \sum_{j=1}^{n} \frac{\partial s}{\partial x_j} \langle dx_j, X \rangle = \langle \sum_{j=1}^{n} \frac{\partial s}{\partial x_j} \, dx_j, X \rangle. \tag{3.36}$$

The final result is that the differential ds is given by

$$ds = \sum_{j=1}^{n} \frac{\partial s}{\partial x_j} \, dx_j. \tag{3.37}$$

This is the most basic computational tool of the theory. The coordinate basis forms dx_i are sometimes called the *dual basis* of the coordinate basis vector fields $\partial/\partial x_j$.

The general 1-form may be written in the form

$$\omega = \sum_{j=1}^{n} p_j \, dx_j. \tag{3.38}$$

Here the p_j are scalar fields. Its value on the vector field X is the scalar field

$$\langle \omega \mid X \rangle = \sum_{j=1}^{n} p_j a_j. \tag{3.39}$$

Remark: There is a dramatic difference between vector field bases and 1-form bases. The notation $\partial/\partial z$ does not make sense unless z is a variable that belongs to a given coordinate system. For instance, if the coordinate system is q, z, s, then $\partial/\partial z$ means to differentiate with respect to z holding q, s both constant. On the other hand, a differential dy makes sense for an arbitrary scalar field y, whether or not it belongs to a coordinate system. |

Example: Here is an illustration of some of these ideas. Consider the problem of making a box with a given amount of material to contain the maximum volume. The box will have five sides, a base and four vertical sides. It is open at the top. In this case the manifold patch is the set M of possible boxes made with this material.

Say that the side lengths of the base are u, v and the height is w. The amount of material available is a fixed number A. Thus $uv + 2uw + 2vw = A$. Since A is a constant, we have

$$(v + 2w) \, du + (u + 2w) \, dv + 2(u + v) \, dw = 0. \tag{3.40}$$

This relation is valid on all of M. We are interested in the point of M (that is, in the particular shape of box) with the property that the volume $V = uvw$ is maximized. At this point we have

$$dV = vw \, du + uw \, dv + uv \, dw = 0. \tag{3.41}$$

Now it is time to choose a coordinate system to work with, and it is convenient to choose u, v. Thus we eliminate dw from the system. Multiplication by $2(u+v)/(uv)$ gives

$$(2\frac{vw}{u} + 2w)\,du + (2\frac{uw}{v} + 2w)\,dv + 2(u+v)\,dw = 0 \tag{3.42}$$

at the point. Subtracting the equations gives

$$v(1 - 2\frac{w}{u})\,du + u(1 - 2\frac{w}{v})\,dv = 0 \tag{3.43}$$

at the point. Since u, v is a coordinate system, the coefficients must be zero at the point. This gives $w = u/2$ and $w = v/2$ as the dimensions of the box. The box with maximal volume has a square base and a height half the length of each side of the base. |

The 1-form ω is said to be *exact* in an open set if $\omega = ds$ for some scalar field s defined on that open set. It is also sometimes called an *exact differential*. The 1-form ω is said to be *closed* if for each $j \neq k$ we have

$$\frac{\partial p_j}{\partial x_k} = \frac{\partial p_k}{\partial x_j}. \tag{3.44}$$

These are $\binom{n}{2}$ conditions. The following theorem is both trivial and supremely important.

Theorem 3.2 *If ω is an exact 1-form in some open set, then ω is also a closed form.*

The notion of differential 1-form is coordinate invariant. If

$$\alpha = \sum_{k=1}^{n} q_k\,du_k \tag{3.45}$$

is a differential 1-form, and if $\mathbf{u} = \mathbf{f}(\mathbf{x})$, then

$$\alpha = \sum_{j=1}^{n} \bar{q}_j\,dx_j \tag{3.46}$$

with

$$\bar{q}_j = \sum_{k=1}^{n} q_k \frac{\partial u_k}{\partial x_j} = \sum_{k=1}^{n} q_k \mathbf{f}'_{k,j}(\mathbf{x}) \tag{3.47}$$

expresses the same form in the other coordinate system. It may be shown by calculation that the criterion for being exact or closed is the same in either coordinate system.

The coordinate invariance is a reflection of the fact that the pairing of differential 1-form and vector field gives rise to a well-defined scalar. Explicitly, we have

$$\langle \alpha \mid X \rangle = \sum_{j=1}^{n} \bar{q}_j a_j = \sum_{k=1}^{n} q_k \bar{a}_k. \tag{3.48}$$

The **x** coordinates and the **u** coordinates give the same result.

An exact differential 1-form $\omega = ds$ may be pictured by the contour surfaces of s. At points where $\omega \neq 0$ these are hypersurfaces of dimension $n-1$. It is sometimes helpful to indicate which contour lines have larger values of s and which have smaller values of s. It is harder to picture a differential 1-form that is not exact. The idea is to draw fragments of contour surfaces. These $n-1$ dimensional fragments end in $n-2$ dimensional surfaces.
Example: Take an example when $n = 2$. A typical example of a differential 1-form that is not exact is $y\,dx$. The fragmented contour lines are all vertical. The form indicates increase to the right in the upper half plane and increase to the left in the lower half plane. As the x axis is approached the density of these fragmented contour lines must diminish at constant rate. Some of the lines have end points at their lower ends (in the upper half plane) or at their upper ends (in the lower half plane). |

3.6 Polar coordinates

Polar coordinates provide a useful example of these ideas. Consider the manifold patch M that is the Euclidean plane. The points in the planes are not numbers; they are geometrical objects. Let $M^\bullet$ be M with a point removed. This is the *punctured plane*. Let $M^\dagger$ be $M^\bullet$ after removal of a half-line that starts at the missing point. This is the *cut plane*.

We may choose a Cartesian coordinate system x, y on M. Suppose that the point where $x = 0$ and $y = 0$ is the point that is removed to make $M^\bullet$. Suppose that the line where $x \leq 0$ and $y = 0$ is the half-line in the definition of $M^\dagger$. Then on $M^\dagger$ the polar coordinates r, θ are related to the Cartesian coordinates by

$$\begin{aligned} x &= r\cos(\theta) \\ y &= r\sin(\theta). \end{aligned} \tag{3.49}$$

Here θ ranges from $-\pi$ to π, with a jump in value across the half-line. These equations are identities saying the the scalars on the left are equal to the scalars on the right as real functions on $M^\dagger$. Similarly, we have the identity $r^2 = x^2 + y^2$.

Another way of thinking of the relation between the two coordinate systems is to define open subsets U and V of $\mathbf{R}^2$ by taking $U = \{(a, b) \mid a > 0, -\pi < b < \pi\}$ and $V = \mathbf{R}^2 \setminus \{(p, q) \mid p \leq 0, q = 0\}$. Then the coordinate system (x, y) maps $M^\dagger$ to V, and the coordinate system (r, θ) maps $M^\dagger$ to U. The change of coordinates is a smooth one-to-one function $\mathbf{f}$ from U to V. The result is that $(x, y) = \mathbf{f}(r, \theta)$ as functions from $M^\dagger$ to V. The two coordinate systems provide two numerical descriptions of the same object $M^\dagger$.

Taking the differential and then eliminating the trig functions gives

$$dx = \frac{x}{r}\,dr - y\,d\theta \tag{3.50}$$

$$dy = \frac{y}{r}\,dr + x\,d\theta. \tag{3.51}$$

It follows that

$$\begin{aligned} x\,dx + y\,dy &= r\,dr \\ x\,dy - y\,dx &= r^2\,d\theta. \end{aligned} \tag{3.52}$$

In particular, the differential form

$$\omega^\dagger = \frac{x\,dy - y\,dx}{x^2 + y^2} = d\theta \tag{3.53}$$

as an identity on $M^\dagger$. This shows that $\omega^\dagger$ on the cut plane $M^\dagger$ is an exact form.

The form on the cut plane is not a particularly natural object. Instead, the *angle form*

$$\omega^\bullet = \frac{x\,dy - y\,dx}{x^2 + y^2} \tag{3.54}$$

makes sense on all of $M^\bullet$. Furthermore, it is a closed form on $M^\bullet$. The price to pay is that it is not an exact form on the punctured plane $M^\bullet$. The interpretation of this form is that near every point in $M^\bullet$ it represents angular change. In fact, locally we can always define an angle variable such that $\omega^\bullet = d\chi$. But there is no such angle variable defined on all of $M^\bullet$.

There is a similar computation for vector fields. The chain rule followed by elimination of trig functions gives

$$\begin{aligned} \frac{\partial}{\partial r} &= \frac{x}{r}\frac{\partial}{\partial x} + \frac{y}{r}\frac{\partial}{\partial y} \\ \frac{\partial}{\partial \theta} &= -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} \end{aligned} \tag{3.55}$$

as an identity on the cut plane $M^\dagger$.

Consider the vector fields

$$\begin{aligned} E &= x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} \\ R &= x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}. \end{aligned} \tag{3.56}$$

These are both defined on the plane M. The vector field E is sometimes called the *Euler vector field*; it plays a role in the study of homogeneous functions. The vector field R is the *rotation vector field*. It describes rotations of constant angular frequency. The origin is a zero of the vector field R. Rotations of the origin just leave it fixed.

If we restrict the vector fields E and R to the punctured plane $E^\bullet$, then they can be locally straightened out. In fact this is true for arbitrary non-zero vector fields. For differential forms the condition for local exactness is that the form be a closed form. There is no such restriction for vector fields.

For the Euler operator we can take $r = e^t$ and get $E = r\partial/\partial r = \partial/\partial t$ in all of $E^\bullet$. For the rotation operator the straightening out is only local. For every

point in $E^\bullet$ we can choose an angular variable χ that is defined near the point, and then $R = \partial/\partial\chi$ near the point.

This discussion shows that it is useful to think of Cartesian coordinates and polar coordinates as scalars defined on a manifold patch. Then the equations above are identities, either for scalars or for differential forms or for vector fields. The closed differential form ω defined on the punctured plane is a fundamental mathematical object; for instance it underlies many of the calculations in complex variable theory. The rotation operator R has an associated system of differential equations. These are the equations for a linear oscillator, a basic system that occurs throughout applied mathematics. These examples also illustrate that differential forms and vector fields are quite different objects.

The example in this section is not typical in one respect: there are natural notions of length and angle. Thus in Cartesian coordinates dx and dy are orthogonal unit forms. In polar coordinates dr and $r\,d\theta$ are orthogonal unit forms. (Note: The form $r\,d\theta$ is not a closed form.) There is a similar story for vector fields. In Cartesian coordinates $\partial/\partial x$ and $\partial/\partial y$ are orthogonal unit vectors. In polar coordinates $\partial/\partial r$ and $1/r\,\partial/\partial\theta$ are orthogonal unit vectors. The reason for this is that the underlying manifold M is the Euclidean plane, which has natural notions of length and angle. Such special structure need not be present in other manifolds.

3.7 Integrating factors and canonical forms

A classic application of these ideas is ordinary differential equations in the plane. Such an equation is often written in the form

$$p\,dx + q\,dy = 0. \tag{3.57}$$

Here $p = f(x, y)$ and $q = g(x, y)$ are functions of x, y. This means that a solution of the equation is a curve where the differential form $p\,dx + q\,dy$ is zero. There can be many such curves.

The equation is determined by the differential form $\alpha = p\,dx + q\,dy$, but two different forms may determine equivalent equations. For example, if $\mu = h(x, y)$ is a non-zero scalar, then the form $\mu\alpha = \mu p\,dx + \mu q\,dy$ is a quite different form, but it determines an equivalent differential equation.

If $\alpha = p\,dx + q\,dy$ is exact, then $p\,dx + q\,dy = dz$, for some scalar z depending on x and y. Each solution of the differential equation is then given implicitly by $z = c$, where c is the constant of integration.

If $\alpha = p\,dx + q\,dy$ is not exact, then one looks for an *integrating factor* μ such that

$$\mu\alpha = \mu(p\,dx + q\,dy) = dz \tag{3.58}$$

is exact. Once this is done, again the general solution of the differential equation is then given implicitly by $z = c$, where c is constant of integration.

Theorem 3.3 *Suppose that $\alpha = p\,dx + q\,dy$ is a differential form in two dimensions that is non-zero near some point. Then α has a non-zero integrating factor μ near the point, so $\mu\alpha = dv$ for some scalar field v.*

Proof: Consider the non-zero vector field $X = q\frac{\partial}{\partial x} - p\frac{\partial}{\partial y}$. By the straightening out theorem, there is a new coordinate system u, v such that $X = \frac{\partial}{\partial u}$. This means that $\frac{\partial x}{\partial u} = q$ and $\frac{\partial y}{\partial u} = -p$. It is easy to check that

$$\alpha = p\,dx + q\,dy = (p\frac{\partial x}{\partial v} + q\frac{\partial y}{\partial v})\,dv = w\,dv, \tag{3.59}$$

where w is a non-zero scalar. We can then take $\mu = 1/w$. □

Finding an explicit integrating factor may be no easy matter. However, there is a strategy that may be helpful.

Recall that if a differential form is exact, then it is closed. So if μ is an integrating factor, then

$$\frac{\partial \mu p}{\partial y} - \frac{\partial \mu q}{\partial x} = 0. \tag{3.60}$$

This condition may be written in the form

$$p\frac{\partial \mu}{\partial y} - q\frac{\partial \mu}{\partial x} + \left(\frac{\partial p}{\partial y} - \frac{\partial q}{\partial x}\right)\mu = 0. \tag{3.61}$$

Say that by good fortune there is an integrating factor μ that depends only on x. Then this gives a linear ordinary differential equation for μ that may be solved by integration.

Example: Consider the standard problem of solving the linear differential equation

$$\frac{dy}{dx} = -ay + b, \tag{3.62}$$

where a, b are functions of x. Consider the differential form $(ay - b)\,dx + dy$. Look for an integrating factor μ that depends only on x. The differential equation for μ is $-d\mu/dx = a\mu$. This has solution $\mu = e^A$, where A is a function of x with $dA/dx = a$. Thus

$$e^A(ay - b)\,dx + e^A\,dy = d(e^A y - S), \tag{3.63}$$

where S is a function of x with $dS/dx = e^A b$. So the solution of the equation is $y = e^{-A}(S + c)$. |

The theory of differential forms is extraordinarily different from the theory of vector fields. A nonzero vector field may always be straightened out locally. For differential forms this is only possible if the form is closd (and hence locally exact).

Theorem 3.4 *Consider a differential form $\alpha = p\,dx + q\,dy$ in two dimensions. Suppose that near some point α is not zero. Then*

- *If α is closed near this point, then there is a scalar field z with $\alpha = dz$.*

- *If α is not closed, then there there is a new coordinate system w, v with $\alpha = w\,dv$.*

Proof: Since $\alpha = p\,dx + q\,dy$ is not zero, there is a new coordinate system u, v in which it has the form $\alpha = w\,dv$. In this coordinate system the condition that α is a closed form is that $\frac{\partial w}{\partial u} = 0$.

If α is closed, then w is a function of v. Thus $\alpha = w\,dv$ has an integral z that is a function of v.

If α is not closed, then the matrix that expresses the partial derivatives of w, v in terms of u, v is non-singular. By the inverse function theorem w, u is also a coordinate system. □

A theorem of Darboux gives a list of standard representations of 1-forms in higher dimensions. The differential equations book by Ince [7] treats the three dimensional situation. The treatise by Şuhubi [19] gives a full discussion for n dimensions.

3.8 The second differential

This section deals criteria with the critical points of a scalar function. These concern the coordinate invariant versions of the first derivative test and the second derivative test. The first derivative test involves first partial derivatives, that is, the differential. The second derivative test involves the Hessian matrix of second partial derivatives, which at a critical point gives a second differential.

Theorem 3.5 (Coordinate invariance of first derivative test) *Suppose that M is a manifold patch and z is a scalar field on M. If z has a local maximum or local minimum at a certain point, then at that point*

$$dz = \sum_{i=1}^{n} \frac{\partial z}{\partial x_i}\, dx_i = 0. \tag{3.64}$$

This condition may be expressed in any coordinate system.

Theorem 3.6 (Coordinate invariance of second derivative test) *Suppose that M is a manifold patch and z is a scalar field on M. Consider a point where $dz = 0$. Then at that point*

$$d^2 z = \sum_{i=1}^{n} \sum_{\ell=1}^{n} \frac{\partial^2 z}{\partial x_i \partial x_\ell}\, dx_i\, dx_\ell. \tag{3.65}$$

If the Hessian matrix on the right is positive definite (negative definite), then the function z has a local minimum (local maximum). This condition may be expressed in any coordinate system.

The computation that underlies these results begins with

$$\frac{\partial z}{\partial y_i} = \sum_{j=1}^{n} \frac{\partial z}{\partial x_j} \frac{\partial x_j}{\partial y_i}. \tag{3.66}$$

If we differentiate again, we get

$$\frac{\partial^2 z}{\partial y_i \partial y_k} = \sum_{j=1}^{n} \sum_{\ell=1}^{n} \frac{\partial^2 z}{\partial x_j \partial x_\ell} \frac{\partial x_j}{\partial y_i} \frac{\partial x_\ell}{\partial y_k} + \sum_{j=1}^{n} \frac{\partial z}{\partial x_j} \frac{\partial^2 x_j}{\partial y_i \partial y_k}. \tag{3.67}$$

The second derivative in the second term on the right is a rather complicated factor. But if the first derivatives $\partial z/\partial x_j = 0$ for $j = 1, \ldots, n$ at a certain point, then we are left with the Hessian matrix at this point transformed by the coordinate transformations on left and right. This is a matrix congruence, so it preserves the positive definite or negative definite property.

In the case of a function of two variables, there is a simple criterion for application of the second derivative test. Suppose that $z = h(x, y)$ is a smooth function. Consider a point where the first derivative test applies, that is, the differential $dz = d\,h(x, y)$ is zero. Consider the case when the Hessian is non-degenerate, that is, has determinant not equal to zero. Suppose first that the determinant of the Hessian matrix is strictly positive. Then the function has either a local minimum or a local maximum, depending on whether the trace is positive or negative. Alternatively, suppose that the determinant of the Hessian matrix is strictly negative. Then the function has a saddle point.

The case of n dimensions is more complicated. The Hessian matrix may be transformed by matrix congruence transformations to a diagonal matrix with entries ϵ_j that are $+1$, -1, or 0. In the non-degenerate case the entries are ± 1. If they are all $+1$ then we have a local minimum, while if they are all -1 we have a local maximum. Otherwise we have a saddle.

There is a more powerful insight into these results that comes from changing to a new coordinate system. The first result states that away from a critical point nothing interesting happens.

Theorem 3.7 *Let $z = f(x_1, \ldots, x_n)$ be a smooth function on an n-dimensional manifold patch such that at a certain point $dz \neq 0$. Then there is a new coordinate system $u_1, \ldots, u_n$ near the point such that $z = u_1$.*

Proof: We may assume without loss of generality that $\partial z/\partial x_1 \neq 0$. Let $u_1 = z$ and let $u_j = x_j$ for $j = 2, \ldots, n$. Then the matrix of partial derivatives $\partial u_i/\partial x_j$ is non-singular. So by the inverse function theorem the x_j may be expressed in terms of the u_j. □

The next result says that even when the first derivative vanishes, there are common circumstances when there is nothing interesting going on with the second derivative. See Milnor [11] for a proof.

Theorem 3.8 (Morse lemma) *Let z be a smooth function on an n-dimensional manifold such that dz vanishes at a certain point. Let z_0 be the value of the*

function at that point. Suppose that the Hessian is non-degenerate at this point. Then there is a coordinate system $u_1, \ldots, u_n$ near the point with

$$z = z_0 + \sum_{i=1}^{n} \epsilon_i u_i^2, \tag{3.68}$$

where ϵ_i are constants that each have the value ± 1.

3.9 Regular surfaces

Consider a manifold patch N with coordinates $x_1, \ldots, x_n$. It is often of interest to consider a surface of dimension $k < n$. The nicest kind of surface is called a *regular surface*. A regular k-surface is a subset $S \subseteq N$ with the following property. Near every point of S there is a coordinate system $u_1, \ldots, u_n$ for N such that the nearby part of S is defined by $u_{k+1} = 0, \ldots, u_n = 0$. (In advanced treatments a regular surface is sometimes called an *embedded manifold*.)

The most classical case is a 2-dimensional surface in 3 dimensions. A 1-dimensional surface is called a *curve*; in many cases it can be treated with the same techniques. A regular curve may often be visualized as something like the letter S, but there is at least one other possibility: it may be like the letter O. In this latter situation near every point there is a coordinate system mapping the letter O to an open interval, but at least two such coordinate systems are required to describe the curve. A 0-dimensional regular surface consists of isolated points.

An *implicit* representation of a surface is to give it as the solution of equations $g_p(x_1, \ldots, x_n) = c_p$, for $p = k+1, \ldots, n$. We can also write this as $\mathbf{g}(\mathbf{x}) = \mathbf{c}$. The derivative $\mathbf{g}'(\mathbf{x})$ is a $n - k$ by n matrix. The largest rank it can have is $n - k$, and it is natural to require that it has this rank. The $n - k$ differential forms

$$dg_p(\mathbf{x}) = \sum_{i=1}^{n} g'_{p,i}(\mathbf{x})\, dx_i \tag{3.69}$$

are then linearly independent. In this case the surface defined by $\mathbf{g}(\mathbf{x}) = \mathbf{c}$ will be called a *regular implicit surface*. It is clear that every regular surface is a regular implicit surface.

Theorem 3.9 *Every regular implicit surface is a regular surface.*

Proof: Suppose the surface is given by $g_{k+1}(\mathbf{x}) = 0, \ldots, g_n(\mathbf{x}) = 0$. Without loss of generality we may assume that the $n - k$ by $n - k$ matrix of partial derivatives with respect to $x_{k+1}, \ldots, x_n$ is non-singular. Set $u_i = \mathbf{h}_i(\mathbf{x}) = x_i$ for $i = 1, \ldots, k$ and $u_i = \mathbf{h}_i(\mathbf{x}) = \mathbf{g}_i(\mathbf{x})$ for $i = k+1, \ldots, n$. Then the n by n matrix of partial derivatives for $\mathbf{h}'(\mathbf{x})$ is non-singular. So $\mathbf{u} = \mathbf{h}(\mathbf{x})$ locally has an inverse function $\mathbf{x} = \mathbf{h}^{-1}(\mathbf{u})$. Thus $u_1, \ldots, u_n$ is a coordinate system, and S is given by setting the last $n - k$ variables equal to zero. □

Consider a manifold patch N of dimension n with coordinates $x_1, \ldots, x_n$. Introduce another manifold patch P of dimension k with coordinates $u_1, \ldots, u_k$. A *parameterized k-surface* is a mapping ϕ given by $x_i \leftarrow f_i(u_1, \ldots, u_k)$, for $i = 1, \ldots, n$, from P to N. Such a surface can be singular in various ways; in particular the mapping need not be one-to-one. For this reason one wants to think of P and and its image in N as quite distinct objects. In dealing with parameterized surfaces in this generality one mainly works with the mapping ϕ. In the case of curves, we can think of a curve where the image looks like the letter S, but we can also think of it looking like the Greek letter α, with self-intersection.

There is also a useful notion of *regular locally parameterized surface*. For $S \subseteq N$ to be such a surface we require that for every point in S there is an open subset $U \subseteq N$, a k-dimensional manifold patch P, and a smooth mapping $\phi : P \to N$ given by $\mathbf{x} \leftarrow \mathbf{f}(\mathbf{x})$. Furthermore,

1. The mapping ϕ is one-to-one and has image $\phi(P) = S \cap U$.

2. The derivative mapping $\mathbf{f}'(\mathbf{u})$ has rank k at every point.

3. The mapping ϕ sends open sets in P to relatively open sets in $\phi(P)$.

The third requirement is a self-avoiding requirement. It says that for every open subset $W \subseteq P$ there is an open subset $V \subseteq N$ such that $\phi(W) = S \cap V$. In effect every small piece of the surface is isolated from remote parts of the surface. Spivak [18] gives an example of a curve in the plane that fails to be self-avoiding. The parameter space is an open interval, and it maps into a figure that looks like the number 6. There is one point on the curve that is not isolated from remote parts of the curve. Small enough parameter intervals around that point give only the left part of the 6. But the right part of the 6 gets arbitrarily close to the point.

Every regular surface is a regular locally parameterized surfaces. In fact, if the surface is given near the point by setting $u_{k+1} = 0, \ldots, u_n = 0$, then the surface near the point can be parameterized by $u_1, \ldots u_k$. In such a case one may choose to think of the parameter space as part of the surface, so the $u_1, \ldots, u_k$ are coordinates on the surface $S \cap U$. In fact, each $S \cap U$ becomes a manifold patch in its own right.

Theorem 3.10 *Every locally parameterized regular surface is a regular surface.*

Proof: Consider a point on the surface. Suppose the parametric representation near this point is $x_i = f_i(u_1, \ldots, u_k)$. Without loss of generality suppose that the partial derivatives with respect to the variables $x_1, \ldots, x_k$ form a k by k invertible matrix. Define a new function $x_i = g_i(u_1, \ldots, u_n)$ as follows. Take $x_i = g_i(\mathbf{u}) = f_i(u_1, \ldots, u_k)$ for $i = 1, \ldots, k$, and take $x_i = g_i(\mathbf{u}) = f_i(u_1, \ldots, u_k) + u_i$ for $i = k+1, \ldots, u_n$. Then the derivative of $\mathbf{g}(\mathbf{u})$ is an invertible n by n matrix. So we may locally express $\mathbf{u} = \mathbf{g}^{-1}(\mathbf{x})$ by the inverse function theorem. We would like to show that near the given point on the surface it is obtained by setting $u_{k+1} = 0, \ldots, u_n = 0$. Clearly when this is satisfied

we have $x_i = f_i(u_1, \ldots, u_k)$, and so the corresponding point is on the surface. On the other hand, we have the self-avoiding condition. Consider a parameter region W around the point so small that it is in the region where the inverse function theorem applies. Then there is an open subset V in N near the point such that that every $\mathbf{x}$ in V that is also in $S \cap U$ is of the form $\mathbf{x} = \mathbf{f}(u_1, \ldots, u_k)$ for $u_1, \ldots, u_k$ in W. In other words, $\mathbf{x} = \mathbf{g}(u_1, \ldots, u_k, 0, \ldots, 0)$. Since $\mathbf{g}$ is one-to-one, this means that nearby points $\mathbf{x}$ on the surface $S \cap U$ have u_i coordinates satisfying $u_{k+1} = 0, \ldots, u_n = 0$. $\square$

We could also define a *regular parameterized surface* to be one where for each point in S the open set U is all of N. Then $S = \phi(P)$ needs only one coordinate patch. This is a regular surface that happens also to be a manifold patch. Often this is the natural setting for stating local results.

If $s = h(x_1, \ldots, x_n)$ is a scalar, then there is a natural pullback scalar expressed in terms of $u_1, \ldots, u_k$. This is

$$h(\mathbf{x})(\mathbf{x} \leftarrow \mathbf{f}(\mathbf{u})) = h(\mathbf{f}(\mathbf{u})). \tag{3.70}$$

If we differentiate this equation with respect to u_α, we get a quantity

$$X_\alpha h = \frac{\partial}{\partial u_\alpha} h(\mathbf{f}(\mathbf{u})) = \sum_{i=1}^{n} h'_{,i}(\mathbf{f}(\mathbf{u})) f'_{i,\alpha}(\mathbf{u}). \tag{3.71}$$

This suggests that we define vectors that differentiate scalars:

$$X_\alpha = \sum_{i=1}^{n} f'_{i,\alpha}(\mathbf{u}) \frac{\partial}{\partial x_i}. \tag{3.72}$$

The notation requires explanation. To get the proper value for $X_\alpha h$ we have to perform the partial derivatives to get $h'_{,i}(\mathbf{x})$, but after that we have to substitute $\mathbf{x} \leftarrow \mathbf{f}(\mathbf{u})$ in the result.

The vectors X_α for $\alpha = 1, \ldots, k$ are not the usual kind of vector field; instead each X_α is a *vector field along the parameterized surface*. That is, the input to X_α is given by the $\mathbf{u}$, while the output corresponds to a vector at the point $\mathbf{f}(\mathbf{u})$ on the surface. Each such vector is a *tangent vector* to the surface.

Consider now a surface with both a parametric and an implicit representation. In that case we have $\mathbf{g}(\mathbf{x})(\mathbf{x} \leftarrow \mathbf{f}(\mathbf{u})) = \mathbf{g}(\mathbf{f}(\mathbf{u})) = \mathbf{c}$. Explicitly,

$$g_p(\mathbf{f}(\mathbf{u})) = c_p \tag{3.73}$$

for $p = k+1, \ldots, n$. Differentiation with respect to $\mathbf{u}_\alpha$ gives

$$\sum_{i=1}^{n} g'_{p,i}(\mathbf{f}(\mathbf{u})) f'_{i,\alpha}(\mathbf{u}) = 0. \tag{3.74}$$

This result may also be written in terms of differential forms and vector fields in the form

$$\langle dg_p(\mathbf{x}) \mid X_\alpha \rangle = 0. \tag{3.75}$$

The notion does not make this explicit, but there is an assumption that there is a replacement $\mathbf{x} = \mathbf{f}(\mathbf{u})$ in the coefficients of the differential forms. The significance of this equation is that the differentials of the functions defining the surface vanish on the tangent vectors to the surface. Since there are k independent tangent vectors and $n - k$ independent differential forms, the differential forms $dg_p(\mathbf{x})$, $p = k + 1, \ldots, n$ form a basis for the space of differential forms that vanish on the tangent vectors.

3.10 Lagrange multipliers

The topic in this section is *constrained optimization*. The problem is to maximize or minimize a function restricted to a surface. This *constraint surface* is given implicitly.

Theorem 3.11 (Lagrange multiplier theorem) *Consider a regular k-surface given implicitly by $g_p(\mathbf{x}) = c_p$, $p = k + 1, \ldots, n$, Suppose that $h(\mathbf{x})$ is a smooth function whose restriction to the surface has a local minimum or a local maximum at a certain point given by $\mathbf{x}$. Then there are unique coefficients $\lambda_1, \ldots, \lambda_p$ such that*

$$dh(\mathbf{x}) = \sum_{p=k+1}^{n} \lambda_p \, dg_p(\mathbf{x}) \tag{3.76}$$

at that point.

Proof: Take a parametric representation $\mathbf{x} \leftarrow \mathbf{f}(\mathbf{u})$ near the point. The function $h(\mathbf{x})$ pulls back to $h(\mathbf{f}(\mathbf{u}))$. The first derivative test gives

$$dh(\mathbf{f}(\mathbf{u})) = 0. \tag{3.77}$$

More explicitly,

$$\sum_{i=1}^{n} h'_{,i}(\mathbf{f}(\mathbf{u}))\mathbf{f}'_{i,\alpha}(\mathbf{u}) = 0 \tag{3.78}$$

for $\alpha = 1, \ldots, k$. We can also write this as

$$\langle dh(\mathbf{x}) \mid X_\alpha \rangle = 0 \tag{3.79}$$

for $\alpha = 1, \ldots, k$. This says that $dh(\mathbf{x})$ belongs to the space of forms that vanish on the tangent vectors. It follows that it is a linear combination of the forms $dg_p(\mathbf{x})$ that form a basis for this space. □

The coefficients λ_p are called *Lagrange multipliers*. This result is intuitive. It says that if h has a local maximum on the surface, then the only way it can be made larger is by moving off the surface by relaxing the constraint that the surface is defined by constants. The Lagrange multiplier λ_p itself is the partial derivative of the critical value with respect to a change in the parameter c_p.

Example: Say that we want to maximize or minimize $u = x + y + 2z$ subject to $v = x^2 + y^2 + z^2 = 1$. The manifold in this case is the unit sphere. The Lagrange multiplier condition says that

$$du = dx + dy + 2\,dz = \lambda dv = \lambda(2x\,dx + 2y\,dy + 2z\,dz). \tag{3.80}$$

Thus $1 = 2\lambda x$, $1 = 2\lambda y$, and $2 = 2\lambda z$. Insert these in the constraint equation $x^2 + y^2 + z^2 = 1$. This gives $(1/4) + (1/4) + 1 = \lambda^2$, or $\lambda = \pm\sqrt{3/2}$. So $x = \pm\sqrt{2/3}/2$, $y = \pm\sqrt{2/3}/2$, $z = \pm\sqrt{2/3}$. |

Example: Say that we want to maximize or minimize $u = x - 4y + 3z + z^2$ subject to $v = x - y = 0$ and $w = y - z = 0$. The manifold in this case is just a line through the origin. The Lagrange multiplier condition says that

$$dx - 4\,dy + (3 - 2z)\,dz = \lambda(dx - dy) + \mu(dy - dz). \tag{3.81}$$

Thus $1 = \lambda$, $-4 = -\lambda + \mu$, and $(3 - 2z) = -\mu$. When we solve we get $\mu = -3$ and so $z = 0$.

Of course we could also solve this example without Lagrange multipliers. Since the manifold is $x = y = z$, the function to be maximized or minimized is $u = z^2$, and this has its minimum at $z = 0$. The utility of the Lagrange multiplier technique in more complicated problems is that it is not necessary to do such a preliminary elimination before solving the problem. |

Example: Here is a simple example to emphasize the point that the Lagrange multiplier technique is coordinate independent. Say that one wants to maximize or minimize z subject to $x^2 + y^2 + z^2 = 1$. The Lagrange multiplier method says to write $dz = \lambda(2x\,dx + 2y\,dy + 2z\,dz)$. This says that $x = y = 0$, and so $z = \pm 1$. In spherical polar coordinates this would be the problem of maximizing $r\cos(\theta)$ subject to $r^2 = 1$. This would give $dr\cos(\theta) - r\sin(\theta)\,d\theta = \lambda 2r\,dr$. Thus $\sin(\theta) = 0$, and the solution is $\theta = 0$ or $\theta = \pi$. |

3.11 Differential k-forms

The algebra and differential calculus of differential k-forms may be unfamiliar, but fortunately it is an easy subject. The main properties stated in the following sections may be proved by checking that the definitions take the same form after a change of coordinate system. These proofs tend to be dull, and many of them will be omitted. The book of Rudin [17] takes such a computational approach, but he does not stress the invariance under coordinate changes, which is the most wonderful aspect of the subject. There are alternate proofs that are more interesting and conceptual, but they are also more abstract. More advanced books [2, 12] give an idea of such an approach.

Example: We begin with a quick summary of the algebraic aspect of differential forms. Consider the case of a three-dimensional space with arbitrary coordinates u, v, w. The 0-forms are the scalars. The 1-forms are $p\,du + q\,dv + r\,dw$. These will eventually be integrated over curves. The 2-forms are $a\,dvdw + b\,dw\,du + c\,du\,dv$. These are integrated over surfaces. The 3-forms are $s\,du\,dv\,dw$. These

are good for integrals over 3-dimensional regions. However for the moment we are only concerned with the differential forms, not with their integrals.

These forms have an algebra. The fundamental law is the anticommutative law for 1-forms. Thus for instance $dw\,du = -\,du\,dw$. Since 1-forms anticommute with themselves, we have $du\,du = 0$, and so on.

The algebra here is called the exterior product. In theoretical discussions it is denoted by a wedge symbol, so that we would write $dv \wedge dw$ instead of the shorter form $dv\,dw$. Sometimes it is a good idea to use such a notation, since it reminds us that there is a rather special kind of algebra, different from ordinary multiplication. Practical computations tend to leave it out. |

A more theoretical approach to the definitions relates differential forms to vector fields. A differential k-form ω on a coordinate patch is a quantity that depends on k vector fields. We write it as $\langle \omega \mid X_1, \ldots, X_k \rangle$. One way to get such a k-form is to multiply together k 1-forms and then anti-symmetrize. The multiplication operation that accomplishes this is often written $\wedge$ and is called the *exterior product*. In the simplest case of a differential 2-form $\omega = \alpha \wedge \beta$ this is given by the determinant

$$\langle \alpha \wedge \beta \mid X, Y \rangle = \det \begin{bmatrix} \langle \alpha \mid X \rangle & \langle \beta \mid X \rangle \\ \langle \alpha \mid Y \rangle & \langle \beta \mid Y \rangle \end{bmatrix} \tag{3.82}$$

When we write this out we get

$$\langle \alpha \wedge \beta \mid X, Y \rangle = \langle \alpha \mid X \rangle \langle \beta \mid Y \rangle - \langle \alpha \mid Y \rangle \langle \beta \mid X \rangle. \tag{3.83}$$

This can be thought of as a kind of signed area attached to the vectors X, Y. This product anti-commutes: For 1-forms α, β we always have

$$\alpha \wedge \beta = -\beta \wedge \alpha. \tag{3.84}$$

In particular

$$\alpha \wedge \alpha = 0. \tag{3.85}$$

The general formula for a product of k 1-forms is also given by a determinant

$$\langle \alpha_1 \wedge \cdots \wedge \alpha_k \mid X_1, \ldots, X_k \rangle = \det \left(\langle \alpha_i \mid X_j \rangle \right). \tag{3.86}$$

This can be thought of as a kind of signed volume attached to the vectors $X_1, \ldots, X_k$.

The general definition of a k-form ω is that it associates to each $X_1, \ldots, X_k$ a number $\langle \omega \mid X_1, \ldots, X_k \rangle$. This expression is suppose to be multi-linear, that is, it is linear in each X_i with the other X_j for $j \neq i$ held fixed. It is also supposed to be alternating, in that interchanging two vectors gives a minus sign.
Remark: There is also a general definition of the exterior product of forms. If θ is a p-form and λ is a q form, then $\theta \wedge \lambda$ is a $p + q = k$ form given by

$$\langle \theta \wedge \lambda \mid X_1, \ldots, X_k \rangle = \sum_{\sigma} \operatorname{sign}(\sigma) \langle \theta \mid X_{\sigma(1)}, \ldots, X_{\sigma(p)} \rangle \langle \lambda \mid X_{\sigma(p+1)}, \ldots, X_{\sigma(k)} \rangle, \tag{3.87}$$

where the sum is over all permutations such that $\sigma(1), \ldots, \sigma(p)$ are in increasing order and $\sigma(p+1), \ldots, \sigma(k)$ are in increasing order.

A simple example is the product of a 2-form θ with a 1-form β. Then

$$\langle \theta \wedge \beta \mid X, Y, Z \rangle = \langle \theta \mid X, Y \rangle \langle \beta, Z \rangle - \langle \theta \mid X, Z \rangle \langle \beta, Y \rangle + \langle \theta \mid Y, Z \rangle \langle \beta, X \rangle. \quad (3.88)$$

|

The multiplicative properties are summarized as follows.

Associative law

$$(\omega \wedge \sigma) \wedge \tau = \omega \wedge (\sigma \wedge \tau). \quad (3.89)$$

are equal as $n + m + p$ forms.

Distributive law

$$\omega \wedge (\sigma + \tau) = \omega \wedge \sigma + \omega \wedge \tau. \quad (3.90)$$

Commutative law for even degree forms If either ω or σ is an even degree forms, then

$$\omega \wedge \sigma = \sigma \wedge \omega. \quad (3.91)$$

Anticommutative law for odd degree forms If both ω or σ are odd degree forms, then

$$\omega \wedge \sigma = -\sigma \wedge \omega. \quad (3.92)$$

The way to remember this is that even degree forms commute with everything. On the other hand, odd degree forms anticommute with each other. In particular, if ω has odd degree, then $\omega \wedge \omega = 0$.

3.12 The exterior derivative

Example: Say that we have a three-dimensional space with coordinates x, y, z. If we have a scalar like $x^2 z$, then we know how to take its differential. In this case we get

$$dx^2 z = 2xz\,dx + x^2\,dz. \quad (3.93)$$

Say that we have a differential form like $x^2 z\,dy\,dz$. The rule is to compute the differential by putting the scalar on the left, as we have done. Then the derivative is obtained by taking the differential of the scalar part. Thus

$$d(x^2 z\,dy\,dz) = d(x^2 z)\,dy\,dz. \quad (3.94)$$

When we compute this, we get

$$d(x^2 z\,dy\,dz) = (2xz\,dx + x^2\,dz)\,dy\,dz = 2xz\,dx\,dy\,dz + x^2\,dz\,dy\,dz. \quad (3.95)$$

But $dz\,dy\,dz = -dy\;dz\,dz = 0$ since $dz\,dz = 0$. So the final result is

$$d(x^2 z\,dy\,dz) = 2xz\,dx\,dy\,dz. \tag{3.96}$$

|

Now to a more theoretical treatment. We already know that the differential of a scalar is

$$du = \frac{\partial u}{\partial x_1}\,dx_1 + \cdots + \frac{\partial u}{\partial x_n}\,dx_n. \tag{3.97}$$

Every k form may be written as a sum of forms of the form

$$\omega = u\,dx_{i_1} \wedge \cdots \wedge dx_{i_k}. \tag{3.98}$$

We can define the *exterior derivative* or *differential* by

$$d\omega = du \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}. \tag{3.99}$$

It is important that the du goes on the left.

Here are the main properties of the exterior derivative.

Additivity

$$d(\omega + \sigma) = d\omega + d\sigma. \tag{3.100}$$

Product property If ω is a k-form and σ is an ℓ-form, then

$$d(\omega \wedge \sigma) = d\omega \wedge \sigma + (-1)^k \omega \wedge d\sigma. \tag{3.101}$$

Differential of a differential

$$dd\omega = 0. \tag{3.102}$$

If we think of d as a degree one quantity, then the sign in the product property makes sense. Also, in this context the differential of a differential property also makes sense.

A k-form σ is called *closed* if $d\sigma = 0$. A k-form σ is called *exact* in an open set if $\sigma = d\alpha$ for some $k-1$ form α in the open set. It follows from the differential of a differential property that every exact form is closed.

It is useful to look at these quantities in low dimensions. For instance, in three dimensions one might have a differential 2-form such as

$$\sigma = a\,dy \wedge dz + b\,dz \wedge dx + c\,dx \wedge dy. \tag{3.103}$$

Here x, y, z are arbitrary coordinates, and a, b, c are smooth functions of x, y, z. Similarly, in three dimensions a typical 3-form might have the form

$$\tau = s\,dx \wedge dy \wedge dz. \tag{3.104}$$

Notice that these forms are created as linear combinations of exterior products of 1-forms.

Since these expressions are so common, it is customary in many contexts to omit the explicit symbol for the exterior product. Thus the forms might be written

$$\sigma = a\,dy\,dz + b\,dz\,dx + c\,dx\,dy \tag{3.105}$$

and

$$\tau = s\,dx\,dy\,dz. \tag{3.106}$$

The exterior derivative of an r-form α is an $r+1$-form $d\alpha$. It is defined by taking the differentials of the coefficients of the r-form. For instance, for the 1-form

$$\alpha = p\,dx + q\,dy + r\,dz \tag{3.107}$$

the differential is

$$d\alpha = dp\,dx + dq\,dy + dr\,dz. \tag{3.108}$$

This can be simplified as follows. First, note that

$$dp = \frac{\partial p}{\partial x}\,dx + \frac{\partial p}{\partial y}\,dy + \frac{\partial p}{\partial z}\,dz. \tag{3.109}$$

Therefore

$$dp\,dx = \frac{\partial p}{\partial y}\,dy\,dx + \frac{\partial p}{\partial z}\,dz\,dx = -\frac{\partial p}{\partial y}\,dx\,dy + \frac{\partial p}{\partial z}\,dz\,dx. \tag{3.110}$$

Therefore, the final answer is

$$d\alpha = d(p\,dx + q\,dy + r\,dz) = \left(\frac{\partial r}{\partial y} - \frac{\partial q}{\partial z}\right)\,dy\,dz + \left(\frac{\partial p}{\partial z} - \frac{\partial r}{\partial x}\right)\,dz\,dx + \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right)\,dx\,dy. \tag{3.111}$$

Similarly, suppose that we have a 2-form

$$\sigma = a\,dy\,dz + b\,dz\,dx + c\,dx\,dy. \tag{3.112}$$

Then

$$d\sigma = da\,dy\,dz + db\,dz\,dx + dc\,dx\,dy = \frac{\partial a}{\partial x}\,dx\,dy\,dz + \frac{\partial b}{\partial y}\,dy\,dz\,dx + \frac{\partial c}{\partial z}\,dz\,dx\,dy. \tag{3.113}$$

This simplifies to

$$d\sigma = d(a\,dy\,dz + b\,dz\,dx + c\,dx\,dy) = \left(\frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z}\right)\,dx\,dy\,dz. \tag{3.114}$$

There are ways of picturing differential forms. If s is a 0-form, that is, a scalar field, then it is a function from some set to the real numbers. The set could be two-dimensional (or three-dimensional). So it certainly makes sense to talk about a curve (or a surface) where s has a particular constant value. These are the contour curves (surfaces) of the scalar field s. The scalar field does not change along these curves (surfaces). Closely spaced contour curves (surfaces) indicate a rapid increase or decrease in the values of s.

For an exact 1-form ds one uses the same picture of contour curves (surfaces), but the philosophy is a bit different. One magnifies the region near a given point, and one notes that the magnified curves (surfaces) are nearly lines (planes). So at a small scale they are approximately the contour lines (planes) of linear functions. The linear functions (forms) do not change much along such lines (planes).

For a 1-form that is not exact the picture looks almost the same, except that the the small scale contour lines (planes) can have end points (end curves). They don't come from a big scale contour curve (surface) picture. The end points (end curves) have an orientation that indicates the direction of increase for the small scale contour lines (planes). The differential of the 1-form corresponds to this cloud of end points (end curves). A line integral of a 1-form along a curve represents the cumulative change as the curve crosses the contour lines (planes).

More precisely, a differential 1-form assigns to each point and to each tangent vector at that point a real number. The form is pictured by indicating those spaces of tangent vectors at particular points on which the form gives the value zero. Such a tangent space at a particular point is a line (plane).

If there is a metric, then you can picture a 1-form by vectors that are perpendicular to the lines (planes) defining the form. Many people like to do this. But it complicates calculations. And for many applications there is not a natural choice of metric.

Example: Consider the 1-form $y\,dx$ in two dimensions. This is represented by vertical contour lines that terminate at points in the plane. The density of these lines is greater as one gets farther from the x axis. The increase is to the right above the x axis, and it is to the left below the y axis. The differential of $y\,dx$ is $dy\,dx = -dx\,dy$. This 2-form represents the cloud of terminating points, which has a uniform density. The usual convention is that the positive orientation is counterclockwise. So the orientations of these source points are clockwise. This is consistent with the direction of increase along the contours lines. |

To understand the picture for 2-forms, we can look at 3-dimensional space. If we look at a 1-form, it assigns numbers to little vectors. We picture the 1-form by the vectors (forming little planes) on which it is zero. If we look at a 2-form, it assigns numbers to little oriented parallelograms. We picture it by looking at the intersection of all the little parallelograms for which it has the value zero. These determine a line. So one can think of the 2-form locally as given by little lines. Globally they form curves, with a kind of spiral orientation. They may have oriented end points. The differential of a 2-form corresponds to the cloud of oriented end points. The integral of a 2-form on an oriented surface depends on what the 2-form assigns to little oriented parallelograms formed by tangent vectors to the surface. The non-zero contributions correspond to when the curves representing the 2-form are transverse to the surface.

More precisely, a differential 2-form in 3-dimensional space assigns to each point and to each ordered pair of tangent vectors at that point a real number. The form is pictured by those tangent vectors that belong to a pair giving the value zero. Such a tangent space at a particular point is a line.

When there is a metric, it is possible to picture the 2-form as a vector field

along the direction of the lines. In that case the surface integral represents the amount by which the vectors penetrate the surface.
Example: Say the 2-form is $dx\,dy$. It is zero on the vector pair $\frac{\partial}{\partial z}, \frac{\partial}{\partial x}$ and on the vector pair $\frac{\partial}{\partial y}, \frac{\partial}{\partial z}$. In other words, it is zero on a pair including the vector $\frac{\partial}{\partial z}$. So we picture it by where it is zero, that is, by lines in the z direction. If we integrate it along a surface where z is constant, then the little parallelograms in the surface are spanned by vectors like $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$, and so we get a non-zero result. |
Remark: Consider the case of three dimensions. Anyone familiar with vector analysis will notice that if s is a scalar, then the formula for ds resembles the formula for the gradient in cartesian coordinates. Similarly, if α is a 1-form, then the formula for $d\alpha$ resembles the formula for the curl in cartesian coordinates. The formula $d\,ds = 0$ then corresponds to the formula curl grad $s = 0$.

In a similar way, if σ is a 2-form, then the formula for $d\sigma$ resembles the formula for the divergence of a vector field $\mathbf{v}$ in cartesian coordinates. The formula $d\,d\alpha = 0$ then corresponds to the formula div curl $\mathbf{v} = 0$.

There are, however, important distinctions. First, the differential form formulas take the same form in arbitrary coordinate systems. This is not true for the formulas for the gradient, curl, and divergence. The reason is that the usual definitions of gradient, curl, and divergence are as operations on vector fields, not on differential forms. This leads to a much more complicated theory, except for the very special case of cartesian coordinates on Euclidean space. Later on we shall examine this issue in detail.

Second, the differential form formulas have natural formulations for manifolds of arbitrary dimension. While the gradient and divergence may also be formulated in arbitrary dimensions, the curl only works in three dimensions.

This does not mean that notions such as gradient of a scalar (a vector field) or divergence of a vector field (a scalar) are not useful and important. Indeed, in some situations they play an essential role. However one should recognize that these are relatively complicated objects.

The same considerations apply to the purely algebraic operations, at least in three dimensions. The exterior product of two 1-forms resembles in some way the cross product of vectors, while the exterior product of a 1-form and a 2-form resembles a scalar product of vectors. Thus the exterior product of three 1-forms resembles the triple scalar product of vector analysis. Again these are not quite the same thing |

3.13 The Poincaré lemma

In the following it will be convenient to have a notion of a set in which nothing interesting can happen. A subset U of $\mathbf{R}^n$ will here be called a *nice region* if it is diffeomorphic to an open ball. This is not standard terminology, but it will be convenient here.

Proposition 3.12 *The following are nice regions, that is, diffeomorphic to the*

open ball B centered at zero with radius one.

1. *The space $\mathbf{R}^n$.*
2. *The space $\mathbf{R}^n_+$ consisting of all points in $\mathbf{R}^n$ with strictly positive coordinates.*
3. *The open cube C_n with all coordinates between 0 and 1.*
4. *The interior of the simplex Δ_n consisting of all points $\mathbf{x}$ in C_n with $\sum_{i=1}^n x_i < 1$.*

Proof:

1. There is a map from $\mathbf{x}$ in B to $\mathbf{y}$ in $\mathbf{R}^n$ given by $\mathbf{y} = \mathbf{x}/\sqrt{1 - |\mathbf{x}|^2}$. The inverse map is given by $\mathbf{x} = \mathbf{y}/\sqrt{1 + |\mathbf{y}|^2}$.
2. There is a map from $\mathbf{R}^n$ to $\mathbf{R}^n_+$ given by $z_i = e^{y_i}$. The inverse map is $y_i = \log(z_i)$.
3. There is a map from C_n to $\mathbf{R}^n_+$ given by $z_i = u_i/\sqrt{1 - u_i^2}$. The inverse map is $u_i = z_i/\sqrt{1 + z_i^2}$.
4. There is a map from Δ_n to $\mathbf{R}^n_+$ given by $\mathbf{z} = \mathbf{x}/(1 - \sum_i x_i)$. The inverse map is $\mathbf{x} = \mathbf{z}/(1 + \sum_i z_i)$.

□

An n-dimensional *local manifold patch* is a manifold patch with a coordinate system $\mathbf{x} : M \to U$, where $U \subseteq \mathbf{R}^n$ is a nice region. In the case $n = 0$ a local manifold patch is a single point. The terminology used here is not standard, but the idea is that a local manifold patch has no interesting global features. In fact, for each n it is an essentially unique object.
Example: An example of a 2-dimensional manifold patch with global features is one modeled on a plane with a point removed. Another example is one modeled on a plane with two points removed. These last two examples are not only not diffeomorphic to the plane, but they are also not diffeomorphic to each other. In fact, they are very different as global object. |
Example: If ω is defined on a manifold patch, that is, if $\omega = d\alpha$ in the region, then ω is closed: $d\omega = 0$. The converse is not true in general. Here is a two dimensional example. Let

$$\omega = \frac{1}{x^2 + y^2}(x\,dy - y\,dx). \tag{3.115}$$

in the plane with the origin removed. Then ω is closed, but not exact. If we remove a line running from the origin to infinity, then the resulting region is a local manifold patch. In this smaller region ω is exact, in fact, $\omega = d\phi$, where $x = r\cos(\phi)$ and $y = r\sin(\phi)$. |

What is true is that if ω is closed, then ω is locally exact. In fact, it is exact on every local manifold patch. This will be proved in the following famous *Poincaré lemma*.

Theorem 3.13 (Poincaré lemma) *Consider a local manifold patch of dimension n. Let $1 \leq k \leq n$ and suppose that the form ω is a closed k-form, that is, $d\omega = 0$. Then ω is exact on this local manifold patch, that is, there is a form α defined on it with $\omega = d\alpha$.*

Proof: We may as well assume that the coordinate system sends the local manifold patch to an open ball centered at the origin. This implies that if $x_1, \ldots, x_n$ are coordinates of a point in the region, and if $0 \leq t \leq 1$, then $tx_1, \ldots, tx_n$ are coordinates of a point in the region.

If ω is a k-form, then we may obtain a form $\bar{\omega}$ by substituting tx_i for x_i everywhere. In particular, expressions dx_i become $d(tx_i) = x_i\, dt + t\, dx_i$. Every differential form σ involving dt and other differentials may be written $\sigma = \sigma_1 + \sigma_2$, where σ_1 is the static part, depending on t but with no factors of dt, and $\sigma_2 = dt\, \beta$ is the remaining dynamic part, with β depending on t but with no factors of dt. Define $K(\sigma) = K(\sigma_2) = \int_0^1 dt\, \beta$, where $\sigma_2 = dt\, \beta$. The claim is that

$$K(d\bar{\omega}) + dK(\bar{\omega}) = \omega. \tag{3.116}$$

This is proved in two parts. The first part is to prove a result for ω_1. By definition $K(\bar{\omega}_1) = 0$. We show that $K(d\bar{\omega}_1) = \omega$. But $(d\bar{\omega}_1)_2$ only involves t derivatives of the coefficients, so by the fundamental theorem of calculus $K(d\bar{\omega}_1) = K((d\bar{\omega}_1)_2) = \omega$.

The second part is that $K(d\bar{\omega}_2) = -dK(\bar{\omega}_2)$. But $d\bar{\omega}_2 = -dt\, d\beta = -dt\, (d\beta)_1$, so $K(d\bar{\omega}_2) = -K(dt\, (d\beta)_1) = -dK(\bar{\omega}_2)$. These two parts establish the claim.

The result follows from the claim. If $d\omega = 0$, then $d\bar{\omega} = 0$, and so $\omega = dK(\bar{\omega})$. □

Remark: The algorithm is simple, provided that one can do the integrals. Start with a closed differential form ω defined in terms of $x_1, \ldots, x_n$. Replace x_i by tx_i everywhere, including in differentials. Collect all terms that begin with dt. Put the dt in front. Integrate from 0 to 1 (with respect to t, keeping everything else fixed). The result is a form α with $d\alpha = \omega$. |

Example: Consider the closed 1-form $\omega = x\, dy + y\, dx$. Then $\bar{\omega} = t^2 x\, dy + t^2 y\, dx + 2xyt\, dt$. The integral of $2xyt\, dt$ is $\alpha = xy$. |

Example: Consider the closed form $\omega = dx\, dy$. Then $\bar{\omega} = t^2 dx\, dy + tx\, dt\, dy - ty\, dt\, dx$. The integral of $t\, dt\, (x\, dy - y\, dx)$ is $\alpha = (1/2)(x\, dy - y\, dx)$. |

3.14 Substitution and pullback

There is a useful distinction in analysis between two kinds of objects. The first kind is a function that sends numbers to numbers. In the one dimensional case one has examples such as sin and $\sqrt{}$. We often use variables to define such functions. The sine function may be written $u \mapsto \sin(u)$ or $w \mapsto \sin(w)$. The function that squares, adds one, and then takes the square root may be written $y \mapsto \sqrt{y^2 + 1}$ or $t \mapsto \sqrt{t^2 + 1}$. In such expressions the variables are only place markers. In logic such a variable is called a *bound variable*. The term *dummy variable* is sometimes used. Functions may be composed. For instance, the

composition $(y \mapsto \sqrt{y^2+1}) \circ (u \mapsto \sin(u)) = (w \mapsto \sqrt{\sin^2(w)+1})$. On the other hand, the composition $(u \mapsto \sin(u)) \circ (y \mapsto \sqrt{y^2+1}) = (z \mapsto \sin(\sqrt{z^2+1}))$. In general $(y \mapsto f(y)) \circ (u \mapsto g(u))$ is just another name for $t \mapsto f(g(t))$ which is itself just the composition $f \circ g$. In many instances we leave out the composition symbol and just write fg.

The other kind of object is an expression that explicitly involves variables. In logic this corresponds to the notion of *free variable*. For instance, $\sin(z)$ and $\sin(t)$ are different expressions. There is an important operation called *substitution* of an expression for a variable. An example is $u \leftarrow \sin(t)$. This means to substitute $\sin(t)$ for u. As an example, $u^2 \circ (u \leftarrow \sin(t)) = \sin^2(t)$. Substitutions may be composed. Thus, for instance, $(u \leftarrow \sin(t)) \circ (t \leftarrow \sqrt{z^2+1}) = (u \leftarrow \sin(\sqrt{z^2+1}))$. And $u^2 \circ (u \leftarrow \sin(\sqrt{z^2+1})) = \sin^2(\sqrt{z^2+1})$. Again we often leave out the composition symbol. There are general identities such as $h(x)(x \leftarrow g(t)) = h(g(t))$ and $(x \leftarrow g(t))(t \leftarrow f(u)) = (x \leftarrow g(f(u)))$. Composition of substitutions and composition of functions are clearly closely related.

The substitution operation occurs in various forms and has various names, including replacement and change of variables . In computer science there is a related notion called *assignment*. An assignment $x \leftarrow g(t)$ makes a change in the machine state. It takes the number stored under the label t, computes $g(t)$, and then stores this result under the label x.

There is another notation that is very useful: In place of $h(x)(x \leftarrow g(t)) = h(g(t))$ one instead writes $(x \leftarrow g(t))^* h(x) = h(g(t))$. This notation at first seems strange, but it is very useful. The idea is that the substitution $x \leftarrow g(t)$ is thought of as a operation that acts on expressions $h(x)$, converting them to other expressions $h(g(t))$. This kind of operation is called *pullback*.

If we write $(u \leftarrow \sin(t))^* u^2 = \sin^2(t)$, then it seems natural to define the expression $(u \leftarrow \sin(t))^* du^2 = d\sin^2(t) = 2\sin(t)\cos(t)\,dt$. It also seems natural to define $(u \leftarrow \sin(t))^* 2u\,du = 2\sin(t) d\sin(t) = 2\sin(t)\cos(t)\,dt$. That fact that these give the same answer should be a source of satisfaction. Substitution is somewhat more general than first appears; it has a natural application to differential forms. In this context it is particularly natural to call it a pullback.

The same ideas extend to several variables. Take, for instance, the situation when we have two variables x, y. A function such as xy^2 is a function on the plane. Say that we want to perform the substitution ψ given by $x \leftarrow t^2, y \leftarrow t^3$. Then we use ψ to pull back xy^2 to a function t^8 on the line. We can write $\psi^* xy^2 = t^8$. If we think of ψ as a parameterized curve, then the pullback is the function on the curve expressed in terms of the parameter.

We can also pull back a differential form such as $d(xy^2) = y^2\,dx + 2xy\,dy$ via the same ψ. The result using the right hand side is $t^6\,dt^2 + 2t^5\,dt^3 = 2t^7\,dt + 6t^7\,dt = 8t^7\,dt$. Of course using the left hand side we also get $dt^8 = 8t^7\,dt$.

For a form like $\omega = y\,dx + 2xy\,dy$ that is not exact, pulling it back by ψ gives the result $2t^4\,dt + 6t^7\,dt = (2t^4 + 6t^7)\,dt = d(\frac{2}{5}t^5 + \frac{3}{4}t^8)$. Thus $\psi^*\omega$ is exact, though ω is not exact. The pullback is a non-trivial operation on differential forms.

In the following we will give a precise definition of an operation called manifold mapping that corresponds to substitution or assignment. With this we can give a rigorous definition of the pullback of a differential form or the pushforward of a vector field.

3.15 Pullback of a differential form

Now suppose that N is a manifold patch of dimension n and M is a manifold patch of dimension m. Suppose $\phi : N \to M$ is a smooth function. We shall call such a ϕ a *manifold mapping*. Sometimes we may just say *mapping*.

Here is a framework relating the notion of manifold mapping to other concepts:

- $\mathbf{x} : N \to \mathbf{R}^n$ is a coordinate system on N.
- $\mathbf{y} : M \to \mathbf{R}^m$ is a coordinate system on M,
- $\mathbf{g}$ is a function from a nice region in $\mathbf{R}^n$ to $\mathbf{R}^m$.
- $\mathbf{g}(\mathbf{x}) : N \to \mathbf{R}^m$ is a function from N to $\mathbf{R}^m$.
- $\phi : N \to M$ is the manifold mapping equal to $\mathbf{y} \leftarrow \mathbf{g}(\mathbf{x})$.

The notation $\mathbf{y} \leftarrow \mathbf{g}(\mathbf{x})$ is a way of defining the manifold mapping ϕ in terms of coordinate systems. Thus ϕ takes a point in N, reads the numbers $\mathbf{x}$, computes the numbers $\mathbf{g}(\mathbf{x})$, and then finds the point in M where $\mathbf{y}$ has this value.

Let u be a scalar field on M. Define the *pullback* $\phi^* u = u \circ \phi$ as a scalar field on N. Thus if $u = h(\mathbf{y})$, then the pullback is

$$(\mathbf{y} \leftarrow \mathbf{g}(\mathbf{x}))^* h(\mathbf{y}) = h(\mathbf{g}(\mathbf{x})). \tag{3.117}$$

Similarly, define the pullback of an exact 1-form du by $\phi^* \, du = d\phi^* u$. Thus

$$(\mathbf{y} \leftarrow \mathbf{g}(\mathbf{z}))^* dh(\mathbf{y}) = dh(\mathbf{g}(\mathbf{x})) = \sum_{j=1}^{n} \sum_{i=1}^{m} h'_{,i}(\mathbf{g}(\mathbf{x})) \mathbf{g}'_{i,j}(\mathbf{x}) \, dx_j. \tag{3.118}$$

In particular,

$$(\mathbf{y} \leftarrow \mathbf{g}(\mathbf{x}))^* dy_i = d(g_i(\mathbf{x})) = \sum_{j=1}^{n} g'_{i,j}(\mathbf{x}) \, dx_j. \tag{3.119}$$

Every k form may be written as a sum of forms of the form

$$\omega = h(\mathbf{y}) \, dy_{i_1} \wedge \cdots \wedge dy_{i_k}. \tag{3.120}$$

We can define the *pullback* by

$$(\mathbf{y} \leftarrow \mathbf{g}(\mathbf{x}))^* \omega = h(\mathbf{g}(\mathbf{x})) \wedge dg_{i_1}(\mathbf{x}) \wedge \cdots \wedge dg_{i_k}(\mathbf{x}). \tag{3.121}$$

This can then be written in terms of products of the dx_j. If the products are arranged in order, then the resulting coefficient is a determinant.

The result of the above computation may be written more succinctly if we use the definition $d\mathbf{y}_I = dy_{i_1} \wedge \cdots dy_{i_k}$, where $I = \{i_1, i_2, \ldots, i_k\}$ and $i_1 < i_2 < \cdots < i_k$. Then we write

$$\omega = \sum_I h_I(\mathbf{y})\, d\mathbf{y}_I, \tag{3.122}$$

where the sum is over sets of I with k elements. Its pullback is

$$\phi^*\omega = \sum_J \sum_I h_I(\mathbf{g}(\mathbf{x})) \det \mathbf{g}'_{I,J}(\mathbf{x})\, d\mathbf{x}_J \tag{3.123}$$

The sums are over sets J and I with k elements.
Example: Consider the mapping $\phi = (z \leftarrow x^2 y, w \leftarrow xy^3)$. Then

$$\phi^*(dz\, dw) = (2xy\, dx + x^2\, dy)(y^3\, dx + 3xy^2\, dy) = 5x^2y^3\, dx\, dy. \tag{3.124}$$

The $5x^2y^3$ factor is the determinant of the derivative of the transformation that relates z, w to x, y. We have already encountered similar formulas when we change coordinates. In this special case we think of the mapping ϕ as an equality, and we write the result as an equality. However in general we want to keep the possibility that z, w describe one set of objects and x, y describe another set of objects. In this case we must write ϕ explicitly to describe how one goes from the objects described by x, y to the objects described by z, w. |

The general properties of the pullback are extraordinarily nice.

Additivity

$$\phi^*(\omega + \sigma) = \phi^*\omega + \phi^*\sigma. \tag{3.125}$$

Product property

$$\phi^*(\omega \wedge \sigma) = \phi^*\omega \wedge \phi^*\sigma. \tag{3.126}$$

Derivative

$$\phi^*(d\omega) = d(\phi^*\omega). \tag{3.127}$$

Composition

$$(\phi \circ \chi)^*\omega = \chi^*\phi^*\omega. \tag{3.128}$$

Notice the reversal in the composition property. We have $(\mathbf{y} \leftarrow \mathbf{g}(\mathbf{x}))(\mathbf{x} \leftarrow \mathbf{f}(\mathbf{u})) = (\mathbf{y} \leftarrow \mathbf{g}(\mathbf{f}(\mathbf{u})))$. So the identity says that $(\mathbf{y} \leftarrow \mathbf{g}(\mathbf{f}(\mathbf{u})))^* = (\mathbf{y} \leftarrow \mathbf{g}(\mathbf{x}))^*(\mathbf{x} \leftarrow \mathbf{f}(\mathbf{u}))^*$.

The pullback has two interpretations. In the case of a *passive tranformation* there are the same number of y_i coordinates as x_j coordinates, and the substitution $\mathbf{y} = \mathbf{g}(\mathbf{x})$ is just giving a different description of the same points. In the special case of a passive transformation $M = N$, and the transformation ϕ given by $\mathbf{y} \leftarrow \mathbf{g}(\mathbf{x})$ is the identity. All that is happening is that $\mathbf{y}$ coordinates are being expressed in terms of $\mathbf{x}$ coordinates. In this case it is appropriate to

write $\mathbf{y} = \mathbf{g}(\mathbf{x})$ as an actual equality of functions on M. One writes everything as an equality, for instance,

$$dy_i = \sum_j \frac{\partial y_i}{\partial x_j} \, dx_j. \tag{3.129}$$

The corresponding equation for vector fields is

$$\frac{\partial}{\partial y_p} = \sum_q \frac{\partial x_q}{\partial y_p} \frac{\partial}{\partial x_q}. \tag{3.130}$$

The two matrices in the last two equations are inverses of each other. This is the reason for the coordinate invariance of the pairing of 1-forms with vector fields.

In the case of an *active transformation* the coordinates y_i are describing one situation, and the coordinates x_j are describing some other situation. For instance, the x_j could be parameters describing a singular surface in the space described by the y_i. There is no reason to require that N and M have the same dimension. The transformation $\mathbf{y} \leftarrow \mathbf{g}(\mathbf{x})$ takes points of N to points of a different space M. In this case it is best to write the pullback explicitly.

In the special case of an active transformation with $M = N$, the transformation ϕ given by $\mathbf{x} \leftarrow \mathbf{g}(\mathbf{x})$ makes sense. In that case the transformation can be iterated. This is equivalent to iterating the function $\mathbf{g}$, since $\mathbf{x}(\mathbf{x} \leftarrow \mathbf{g}(\mathbf{x}))^n = \mathbf{g}^n(\mathbf{x})$. It is important to emphasize that the mapping $\mathbf{x} \leftarrow \mathbf{g}(\mathbf{x})$ is not the same as the function $\mathbf{g}$. In fact, a common notation for the function $\mathbf{g}$ is $\mathbf{x} \mapsto \mathbf{g}(\mathbf{x})$ with the arrow going the other direction.

There is an intermediate situation that can be confusing. This is when N is a subset of M, and ϕ sends each point in N into the same point, but now regarded as a point in M. In this situation many people write equations $\mathbf{y} = \mathbf{g}(\mathbf{x})$. There has to be enough context to indicate that this means that the restriction of $\mathbf{y}$ to N is $\mathbf{g}(\mathbf{x})$.

3.16 Pushforward of a vector field

There is a *pushforward* notion for vector fields, but it is not as nice. The idea is that if $\phi : N \to M$ and Z is a vector field on N, then $\phi_* Z$ is a mapping from N to vectors tangent to the image of ϕ in M. This is not a vector field in the ordinary sense, but it is a somewhat more general object, a vector field along the mapping ϕ. This takes a rather concrete form in a coordinate representation. Say that ϕ is the map $\mathbf{x} \leftarrow \mathbf{f}(\mathbf{t})$. Let

$$Z = \sum_{\alpha=1}^{k} a_\alpha \frac{\partial}{\partial t_\alpha} \tag{3.131}$$

be a vector field on N. Then $\phi_* Z$ is obtained as follows. Consider a scalar $h(\mathbf{x})$ on M. Use the chain rule to compute

$$\phi_* Z\, h(\mathbf{x}) = \sum_{\alpha=1}^{k} a_\alpha \frac{\partial}{\partial t_\alpha} h(\mathbf{f}(\mathbf{t})) = \sum_{i=1}^{n} \left(\sum_{\alpha=1}^{k} f'_{i,\alpha}(\mathbf{t}) a_\alpha \right) h'_{,i}(\mathbf{f}(\mathbf{t})). \tag{3.132}$$

In other words,

$$\phi_* \frac{\partial}{\partial t_\alpha} = \sum_{i=1}^{n} f'_{i,\alpha}(\mathbf{t}) \frac{\partial}{\partial x_i} (\ \ ) \,|_{\mathbf{x} \leftarrow \mathbf{f}(\mathbf{t})} \,. \tag{3.133}$$

The vectors on the right hand side are tangent to the image, since after the differentiation is performed one substitutes a point in the image. In other words, if one regards $\mathbf{x} \leftarrow \mathbf{f}(\mathbf{t})$ as a parameterized surface in M, then these are tangent vectors at various points on the surface.

Since this is a relatively awkward notion, it is not so common to make it explicit. Typically one merely talks of the vector with components

$$\frac{\partial x_i}{\partial t_\alpha} = f'_{i,\alpha}(\mathbf{t}). \tag{3.134}$$

It is understood that this is a way of talking about certain tangent vectors to the surface.

Remark: For those interested in theoretical considerations, it is worth noting that the notions of pushforward and pullback are related. The relationship is somewhat complicated, so it may be wise to omit the following discussion on a first reading.

To understand it, we need the notion of vector field W along a mapping ϕ from N to M. This sends each point in N to a differential operator that differentiates scalars on M, such that at each point in N the the derivative is evaluated at the image point in M under ϕ. In coordinates this would have the form $W = \sum_j k_j(\mathbf{t}) \partial/\partial x_j$, where $k_j(\mathbf{t})$ is a scalar on N and the x_j are coordinates on M. If Z is a vector field on N, then the pushforward $\phi_* Z$ is a vector field along ϕ. We also need the notion of differential k-form γ along a mapping ϕ from N to M. This acts as an antisymmetric multilinear function on vector fields along ϕ, so $\langle \gamma \mid W_1, \ldots, W_k \rangle$ is a scalar on N. In the case of a 1-form γ could take the form $\gamma = \sum_i f_i(\mathbf{t})\, dx_i$. We would have $\langle \gamma, W \rangle = \sum_j f_j(\mathbf{t}) k_j(\mathbf{t})$; there are corresponding formulas for k-forms. If ω is a differential form on M, then there is a corresponding object $\omega \circ \phi$ that is a differential form along ϕ. It may be defined by pulling back the scalar coefficients of ω. Thus in the 1-form case, if $\omega = \sum_j h_j(\mathbf{x})\, dx_j$, then $\omega \circ \phi = \sum_j h_j(\mathbf{f}(\mathbf{t}))\, dx_j$.

The object of interest is $\phi^* \omega$, the pullback of the differential form ω. Suppose ω is a k-form on M and $Z_1, \ldots, Z_k$ are vector fields on N and ϕ is a map from N to M. Then the pullback $\phi^* \omega$ is a k-form on N given by

$$\langle \phi^* \omega \mid Z_1, \ldots, Z_k \rangle = \langle \omega \circ \phi \mid \phi_* Z_1, \ldots, \phi_* Z_k \rangle. \tag{3.135}$$

This is an equality of scalar fields on N. The $\phi_* Z_i$ are vector fields along ϕ; they map points in N to tangent vectors at the image points in M. The $\omega \circ \phi$

is a differential k-form along ϕ. In the case $k = 1$ the formula says

$$\langle \phi^* \omega \mid Z \rangle = \sum_j h_j(\mathbf{f}(\mathbf{t})) \sum_\alpha f'_{j,\alpha}(\mathbf{t}) a_\alpha. \tag{3.136}$$

This is consistent with the usual formula

$$\phi^* \omega = \sum_j h_j(\mathbf{f}(\mathbf{t})) \, df_j(\mathbf{t}) = \sum_j h_j(\mathbf{f}(\mathbf{t})) \sum_\alpha f'_{j,\alpha}(\mathbf{t}) \, dt_\alpha \tag{3.137}$$

for the pullback. In summary, the pushforward of a vector field is not an ordinary vector field; it is a vector field along the manifold mapping. However, the pullback of a differential form is another differential form. It is the pullback that is most simple and natural. |

3.17 Orientation

An *orientation* of an n dimensional vector space is determined by a list of n basis vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Two such lists determine the same orientation if they are related by a matrix with determinant > 0. They determine the opposite orientation if they are related by a matrix with determinant < 0. There are always two orientations.

Sometimes it is useful to have new lists of basis vectors related to the old sets of vectors by a determinant that has the value ± 1. This can be done in various ways, but here is a simple special case. Consider a variant $\bar{\mathbf{u}}_i = s_i \mathbf{u}_{\tau(i)}$, where τ is a permutation of $\{1, \ldots, n\}$ and each $s_i = \pm 1$. Then the determinant is the product of the s_i times the sign of the permutation τ. We shall use the term *variant* as a technical term for such a new basis.

In one dimension an orientation is determined by specifying one of two directions. So if $\mathbf{u}$ is a vector, every strictly positive multiple of $\mathbf{u}$ determines the same orientation, while every strictly negative multiple of $\mathbf{u}$ determines the opposite orientation. So the two variants $\mathbf{u}$ and $-\mathbf{u}$ determine the two orientations.

In two dimensions an orientation is specified by taking two linearly independent vectors $\mathbf{u}, \mathbf{v}$ in order. In fact, the same orientation is specified by the variants $\mathbf{u}, \mathbf{v}$ and $\mathbf{v}, -\mathbf{u}$ and $-\mathbf{u}, -\mathbf{v}$ and $-\mathbf{v}, \mathbf{u}$. The opposite orientation would be given by any of the variants $\mathbf{u}, -\mathbf{v}$ and $-\mathbf{v}, -\mathbf{u}$ and $-\mathbf{u}, \mathbf{v}$ and $\mathbf{v}, \mathbf{u}$. These two orientations are often called counter-clockwise and clockwise. These are not absolute notions; a counter-clockwise orientation on a piece of paper becomes clockwise when the paper is viewed from the other side. Often the orientation is pictured by a rectangle with vectors as sides. For instance this could be $\mathbf{u}, \mathbf{v}, -\mathbf{u}, -\mathbf{v}$, which takes you back to the starting point.

In three dimensions an orientation is determined by a list of three vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$. Of course many other triples of vectors determine the same orientation. If we permute the order and change the signs, we get $3! \cdot 2^3 = 48$ variant lists, of which 24 have one orientation and 24 the opposite orientation. The two

orientations are usually called right-handed and left-handed. Again this is not absolute; a mirror will reverse the orientations. Here one draws a cell with six oriented sides. In the list of three vectors, the first vector goes from an departure side to a destination side, the second and third vectors give the orientation of the destination side. For a given orientation there are six faces the vector can point to, each has its orientation determined. Since there are four ways to find vectors determining the orientation of a given face, this gives 24 variant triples of vectors.

In higher dimensions the idea is the same. One can think of the orientation of a cell as giving a first vector from one side of the cell to the other, then giving an orientation of the destination side. For dimension zero it is helpful to think of an orientation as just a choice of a sign + or −.

Consider a parameterized cell C consisting of the vectors $\sum_{i=1}^{n} t_i \mathbf{u}_i$, where $0 \le t_i \le 1$. This cell has $2n$ boundary cells. For each k there are 2 corresponding boundary cells, related to each other by the vector $\mathbf{u}_k$. The cell C_k^- is obtained by setting $t_k = 0$ and consists of the combinations $\sum_{i \neq k} t_i \mathbf{u}_i$. The cell C_k^+ is obtained by setting $t_k = 1$ and consists of the combinations $\sum_{i \neq k} t_i \mathbf{u}_i + \mathbf{u}_k$. Notice that $\mathbf{u}_k$ takes C_k^- to C_k^+, while $-\mathbf{u}_k$ takes C_k^+ to C_k^-.

Given the orientation determined by $\mathbf{u}_1, \ldots, \mathbf{u}_n$, there is a natural orientation on each of the $2n$ boundary cells. The cell C_k^+ has orientation obtained by moving $\mathbf{u}_k$ to the front of the list, with a sign change for each interchange, and then removing it. This implies that the orientation of the cell C_k^+ is given by the orientation $\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$ when k is odd and by the opposite orientation when k is even. The cell C_k^- has the opposite orientation. This implies that the orientation of the cell C_k^- is given by the orientation $\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$ when k is even and by the opposite orientation when k is odd.

As an example, consider the case $n = 2$ with the list of vectors $\mathbf{u}, \mathbf{v}$. The orientations of the boundary cells C_1^+, C_2^+ are $\mathbf{v}$ and $-\mathbf{u}$ respectively, while the orientations of C_1^-, C_2^- are $-\mathbf{v}$ and $\mathbf{u}$ respectively.

A more challenging example is $n = 3$ with the list of vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$. The orientations of the boundary cells C_1^+, C_2^+, C_3^+ are given by $\mathbf{v}, \mathbf{w}$ and $-\mathbf{u}, \mathbf{w}$ and $\mathbf{u}, \mathbf{v}$. The orientations of C_1^-, C_2^-, C_3^- are given by $-\mathbf{v}, \mathbf{w}$ and $\mathbf{u}, \mathbf{w}$ and $-\mathbf{u}, \mathbf{v}$. Of course one can always use variants of these lists to define the same orientations.

All these ideas make sense for $n = 1$ with a single vector $\mathbf{u}$. There are two boundary cells each consisting of a single point. The convention is that C_1^+ corresponds to the point at $\mathbf{u}$ with positive orientation, while C_1^- corresponds to the point at the origin with negative orientation.

There is a corresponding notion of orientation for a connected manifold patch. Given a coordinate system $x_1, \ldots, x_n$, there is a corresponding orientation given by the list of basis vector fields $\partial/\partial x_1, \ldots, \partial/\partial x_n$. Given two coordinate systems on a connected manifold patch, they either have the same orientation or the opposite orientation.

There are other ways of giving orientations on a connected manifold patch.

Instead of using a list of basis of vector fields, one can use a list of basis 1-forms. Or one can give a single non-zero n-form. Often we suppose that an orientation is given. Then for every coordinate system, either the coordinate system is compatible with that orientation, or it is compatible with the opposite orientation.

Example: A one-dimensional connected manifold patch is something like a curve; an orientation gives a direction along the curve. A coordinate system has only one coordinate u. It is compatible with the orientation if it increases in the direction given by the orientation.

A two-dimensional connected manifold patch has a clock orientation. For dimension two there are two coordinates u, v. Consider a change first in u (keeping v fixed) and then in v (keeping u fixed) that is in the sense of this orientation. If both these changes are positive, or if both these changes are negative), then the coordinate system is compatible with the orientation.

The case of a zero-dimensional connected manifold patch is special; it is just a featureless point. Here an orientation is a + or − sign. A coordinate is a non-zero number attached to the point. It is compatible with the orientation if this number has the same sign. |

3.18 Integration of top-dimension differential forms

Suppose that P is a k-dimensional connected manifold patch with a given orientation. Consider a differential k-form ω on P. This is a top-dimensional form, since the degree of ω is equal to the dimension P. Also consider a Jordan measurable compact set $K \subseteq P$. We wish to define the integral of ω over K. To do this, choose a coordinate system $u_1, \ldots, u_k$ that determines the orientation of M. Then we may write

$$\omega = f(u_1, \ldots, u_k)\, du_1 \wedge \cdots \wedge du_k. \tag{3.138}$$

Furthermore, there is a set $B \subseteq \mathbf{R}^n$ such that $\mathbf{u}(K) = B$. The definition is

$$\int_K \omega = \int_B f(v_1, \ldots, v_n)\, dv_1 \cdots dv_k = I_B(f). \tag{3.139}$$

where $I_B(f)$ is the usual (unoriented) Riemann integral. (The variables $v_1, \ldots, v_n$ are just symbols that are available to define the function f.)

To show that this is well-defined, consider another coordinate system $y_1, \ldots, y_k$ that determines the same orientation. Then $\mathbf{u} = \mathbf{g}(\mathbf{y})$, where $\det \mathbf{g}'(\mathbf{y}) > 0$. Furthermore, $B = \mathbf{g}(A)$ for some Jordan measurable set A. The differential form

$$\omega = f(\mathbf{u})\, du_1 \wedge \cdots \wedge du_k = f(\mathbf{g}(\mathbf{y})) \det \mathbf{g}'(\mathbf{y})\, dy_1 \wedge \cdots \wedge dy_k \tag{3.140}$$

may be expressed in either coordinate system. The definition in the other coordinate system gives

$$\int_K f(\mathbf{g}(\mathbf{y})) \det \mathbf{g}'(\mathbf{y})\, dy_1 \wedge \cdots \wedge dy_k = I_A((f \circ \mathbf{g}) \det \mathbf{g}'). \tag{3.141}$$

But $I_B(f) = I_A((f \circ \mathbf{g}) \det \mathbf{g}')$ by the change of variables theorem. So the definitions using different coordinate systems are consistent. This is a classic example of a passive transformation. The same differential form ω is integrated over the same set K. But the numerical expression of the integral is different in the two coordinate systems.

3.19 Integration of forms over singular surfaces

Let P be a k-dimensional oriented manifold patch, and let $K \subseteq P$ be a Jordan measurable compact subset. Let N be an n dimensional manifold patch. We wish to define a *singular parameterized k-surface* χ in N. This is defined to be a smooth function

$$\chi : K \to N. \tag{3.142}$$

In this context to say that χ is smooth is to say that it extends to a manifold mapping from P to N.

We want to think of the compact set K as a set over which we can perform k-dimensional integrals. This is called a singular parameterized k-surface because there is no restriction on the rank, and there is also no requirement that the mapping be one-to-one. When $k = 1$ this χ is a parameterized curve, when $k = 2$ it is a parameterized surface, and so on. When $k = 0$ the set consists of only one point, so the surface merely specifies a point in N.

This is a classic example of an active transformation. In terms of coordinates, $\mathbf{u}$ could be a coordinate system on P, while $\mathbf{x}$ is a coordinate system of N. Suppose that ω is a differential k-form on N expressed in terms of the coordinates $\mathbf{x}$. If χ is given by $\mathbf{x} \leftarrow \mathbf{g}(\mathbf{u})$, then ω pulls back to a differential k-form $\chi^*\omega$ expressed in terms of $\mathbf{u}$. Since P is k-dimensional, we know how to integrate this form over K.

Suppose ω is a k-form on the n-dimensional space N. Suppose $\chi : K \to N$ is a parameterized k-surface in N. We define the integral of ω over χ by

$$\int_\chi \alpha = \int_K \chi^* \alpha. \tag{3.143}$$

The integral on the right is the integral of a top-dimensional form.
Example: Consider the parameterized curve ψ given by $x \leftarrow t^2, y \leftarrow t^3$, defined for $1 \le t \le 2$. We want to integrate the differential form $\omega = y\,dx + 2xy\,dy$ along this curve. If we pull it back by ψ we get $2t^4\,dt + 6t^7\,dt = (2t^4 + 6t^7)\,dt = d(\frac{2}{5}t^5 + \frac{3}{4}t^8)$. Thus the integral has the value $(\frac{2}{5}2^5 + \frac{3}{4}2^8) - (\frac{2}{5} + \frac{3}{4})$. |
Example: Consider the cylinder $x^2 + y^2 = 25$ with $0 \le x, 0 \le y, 0 \le z \le 3$. Say that one wants the integral of

$$\alpha = z\,dy\,dz + 2x\,dz\,dx - 4x^2 z\,dx\,dy \tag{3.144}$$

on this surface. One way to calculate this is to parameterize the surface by the corresponding rectangle in the x, z plane, with $0 \le x \le 5, 0 \le z \le 3$. The

calculation amounts to pulling back by $y \leftarrow \sqrt{25 - x^2}$. The pullback is the form

$$\alpha^* = \left(\frac{zx}{\sqrt{25 - x^2}} + 2x \right) \, dz \, dx. \tag{3.145}$$

This can be integrated over the rectangle to give 195/2. |

In order to avoid complicated considerations of change of orientation, it is customary to write $-\chi$ and define it in such a way that

$$\int_{-\chi} \omega = -\int_{\chi} \omega. \tag{3.146}$$

Intuitively this corresponds to changing to the opposite orientation. In general, one defines a *k-chain* to be a finite linear combination of singular parameterized k-surfaces with integer coefficients. Thus, for instance, $2\chi_1 - \chi_2$ consists of two copies of the chain χ_1 and one copy of the chain oppositely oriented from χ_2. Thus by definition the integral of ω over $2\chi_1 - \chi_2$ is twice the integral of ω over χ_1 minus the integral of ω over χ_2.

Theorem 3.14 (Change of variables) *Let ϕ be a smooth manifold mapping of an open set in a n-dimensional manifold patch N into an open set of an m-dimensional manifold patch M. Let χ be a singular k surface in the open subset of N. Let ω be a k-form defined on the open set in the manifold patch M. Then*

$$\int_{\phi\chi} \omega = \int_{\chi} \phi^* \omega. \tag{3.147}$$

Proof: The proof is easy.

$$\int_{\phi\chi} \omega = \int_K (\phi\chi)^* \omega = \int_K \chi^* \phi^* \omega = \int_{\chi} \phi^* \omega. \tag{3.148}$$

□

What happens if both $k = m = n$. Then the form is $\omega = f(\mathbf{y}) d\mathbf{y}$. So the theorem says

$$\int_{\phi\chi} f(\mathbf{y}) \, d\mathbf{y} = \int_{\chi} f(\mathbf{g}(\mathbf{y})) \det \mathbf{g}'(\mathbf{x}) \, d\mathbf{x}. \tag{3.149}$$

This looks like the classic change of variables theorem, but without the absolute value sign. The reason the absolute value is not needed is that the integrals are defined with respect to parameterizations χ and $\phi\chi$, and these are thus oriented integrals.

3.20 Stokes' theorem for chains of singular surfaces

For Stokes' theorem we use a more restricted notion of singular parameterized k-surface. Again consider a k-dimensional parameter patch. Consider a compact

subset Q realized as follows. Let $u_1, \ldots, u_k$ be coordinates on N that map into a non-degenerate open cell and that are compatible with the orientation. Then $Q \subseteq P$ consists of the points whose coordinates satisfy $a_j < u_j < b_j$ for $j = 1, \ldots, k$. In other words, Q maps into the bounded non-degenerate closed cell $[a_1, b_1] \times \cdots [a_k, b_k]$. We call this a *parameter cell*. A singular parameterized surface χ is a smooth mapping $\chi : Q \to N$.

The advantage of restriction to parameter cells is that one can define boundaries. For each i there is a manifold patch P_i^+ defined by $u_i = b_i$ and a manifold patch P_i^- defined by $u_i = a_i$. The manifold patch P_i^+ has coordinates u_j for $j \neq i$, in the same order as before.

In P_i^+ there is a bounded non-degenerate closed cell Q_i^+ whose coordinates satisfy $a_j < u_j < b_j$ for $j \neq i$. Similarly, the manifold patch P_i^- has coordinates u_j for $j \neq i$, and in P_i^- there is a bounded non-degenerate closed cell Q_i^- whose coordinates satisfy $a_j < u_j < b_j$ for $j \neq i$.

The definition of boundaries is via *face mappings*. For a parameter cell $Q \subseteq P$ there is a singular parameterized k-surface $I : Q \mapsto P$ that sends each point in Q to itself. Define singular parameterized $k-1$ face mapping surfaces $I_i^\pm : Q_i^\pm \to P$ sending each point in $Q_i^\pm$ to itself. The oriented boundary of a parameter cell is then defined by

$$\partial I = \sum_{i=1}^{k} (-1)^{i-1} (I_i^+ - I_i^-) \tag{3.150}$$

This is a chain of face mapping surfaces. Given a singular parameterized k-surface $\chi : Q \to N$, the *oriented boundary chain* is

$$\partial \chi = \chi \, \partial I. \tag{3.151}$$

This just means that $\partial\chi$ is the chain $\sum_{i=1}^{k} (-1)^{i-1} (\chi I_i^+ - \chi I_i^-)$. This is a chain of parameterized $k-1$ surfaces that each map to N.

If χ is a k-chain (an integer linear combination of parameterized k-surfaces), then its boundary $\partial\chi$ is a $k-1$ chain (the corresponding integer linear combination of the boundaries of the surfaces).

All of these objects have coordinate representations. The face mapping surfaces are

$$\begin{aligned} I_i^+ &= ((u_1, \ldots, u_k) \leftarrow (u_1, \ldots, u_{i-1}, b_i, u_{i+1}, \ldots, u_k)) \\ I_i^- &= ((u_1, \ldots, u_k) \leftarrow (u_1, \ldots, u_{i-1}, a_i, u_{i+1}, \ldots, u_k)). \end{aligned} \tag{3.152}$$

For a singular parameterized k-surface $\chi : Q \to N$ given by $\mathbf{x} \leftarrow \mathbf{g}(\mathbf{u})$ the faces in the i direction, $i = 1, \ldots, k$, are given as the pullbacks

$$\begin{aligned} \chi I_i^+ &= (\mathbf{x} \leftarrow \mathbf{g}(u_1, \ldots, u_{i-1}, b_i, u_{i+1}, \ldots, u_k)) \\ \chi I_i^- &= (\mathbf{x} \leftarrow \mathbf{g}(u_1, \ldots, u_{i-1}, a_i, u_{i+1}, \ldots, u_k)). \end{aligned} \tag{3.153}$$

Theorem 3.15 (Stokes' theorem) *If ω is a $k-1$ form and χ is a k-chain, then*

$$\int_\chi d\omega = \int_{\partial\chi} \omega. \tag{3.154}$$

Proof: The proof has two steps. The first step is to prove the theorem in the special case when the cell parameterizes itself, that is, when $\chi = I : Q \to Q$ sends every point to itself. The general $k-1$ form on a k-dimensional space is

$$\omega = \sum_{j=1}^{k} \omega_j = \sum_{j=1}^{k} f_j(u_1, \ldots, u_k)\, du_1 \wedge \ldots \wedge du_{j-1} \wedge du_{j+1} \wedge \ldots \wedge du_k. \tag{3.155}$$

In that case

$$d\omega = \sum_{j=1}^{k} d\omega_j = \sum_{j=1}^{k} (-1)^{j-1} \frac{\partial f_j(u_1, \ldots, u_k)}{\partial u_j}\, du_1 \wedge \ldots \wedge du_k. \tag{3.156}$$

Hence

$$\int_I d\omega = \int_Q I^* d\omega = \sum_{j=1}^{k} (-1)^{j-1} \int_A \frac{\partial f_j(u_1, \ldots, u_k)}{\partial u_j}\, du_1 \ldots du_k. \tag{3.157}$$

We can now use Fubini's theorem to write the integral on the right hand side as an iterated integral, with the du_j integral as the inside integral. The fundamental theorem of calculus say that the integral is equal to

$$\int_{A_j} [f_j(u_1, \ldots, b_k, \ldots, u_k) - f_j(u_1, \ldots, a_j, \ldots, u_k)]\, du_1 \ldots du_{j-1}\, du_{j+1} \ldots du_k. \tag{3.158}$$

Here A_j is the relevant $k-1$ cell. In other words,

$$\int_I d\omega = \sum_{j=1}^{k} (-1)^{j-1} [\int_{Q_j^+} I_j^{+*} \omega_j - \int_{Q_j^-} I_j^{-*} \omega_j]. \tag{3.159}$$

But for $i \neq j$ both $I_j^{+*}\omega_i = 0$ and $I_j^{-*}\omega_i = 0$. So this is

$$\int_I d\omega = \sum_{j=1}^{k} (-1)^{j-1} [\int_{Q_j^+} I_j^{+*} \omega - \int_{Q_j^-} I_j^{-*} \omega]. \tag{3.160}$$

Finally, this is

$$\int_I d\omega = \sum_{j=1}^{k} (-1)^{j-1} [\int_{I_j^+} \omega - \int_{I_j^-} \omega] = \int \partial I \omega. \tag{3.161}$$

The second step uses the fact that the integral of a differential form α over a chain χ may be expressed by pulling back to parameter cells. It also depends

on the result that the pullback of a differential is the differential of the pullback, that is, $\chi^*(d\omega) = d\chi^*\omega$. This gives

$$\int_\chi d\omega = \int_I \chi^*(d\omega) = \int_I d\chi^*\omega = \int_{\partial I} \chi^*\omega = \int_{\chi\partial I} \omega = \int_{\partial\chi} \omega. \tag{3.162}$$

So the properly formulated result is rather simple; it follows from the trivial case of the cell and from the remarkable transformation properties of differential forms. □

3.21 Classical versions of Stokes' theorem

The integral of a 1-form along an oriented curve is called a *line integral* . In general such an integral must be computed by pulling back to a parameter interval. However for an exact form there is a shortcut. The *fundamental theorem of calculus* is the case relating scalars to 1-forms. It says that for every scalar field s and every parameterized curve C we have

$$\int_C ds = \Delta s. \tag{3.163}$$

Here C is an oriented path from one point to another point, and Δs is the value of s at the final point minus the value of s at the initial point. Notice that the result does not depend on the choice of path. This is because ds is an exact form.

Example: Consider the form $y^2\,dx + 2xy\,dy$. Since it is exact, we have

$$\int_C y^2\,dx + 2xy\,dy = \int_C d(xy^2) = \Delta(xy^2) \tag{3.164}$$

independent of the path. |

The most common version of Stokes' theorem relates 1-forms to 2-forms. The 2-dimensional version of the theorem is *Green's theorem* . It says that

$$\int_R \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) dx\,dy = \int_{\partial R} p\,dx + q\,dy. \tag{3.165}$$

Here R is an oriented region in two dimensional space, and ∂R is the curve that is its oriented boundary.

Example: A classical application of Green's theorem is the computation of area via

$$\int_R dx\,dy = \frac{1}{2}\int_{\partial R} x\,dy - y\,dx. \tag{3.166}$$

In polar coordinates this takes the form

$$\int_R r\,dr\,d\theta = \frac{1}{2}\int_{\partial R} r^2\,d\theta. \tag{3.167}$$

In this case a typical parameter region for the integral on the left hand side may be considered in the r, θ plane as the four sided region where θ ranges from 0 to 2π and r ranges from 0 to some value depending on θ. The integral on the right is over a chain consisting of four oriented curves. However three of these curves contribute a total of zero: the contributions from $\theta = 0$ and $\theta = 2\pi$ take opposite orientations and cancel each other, while at $r = 0$ the integrand vanishes. So only one oriented curve on the right hand side contributes to the calculation of the area. |

The integral of a 2-form over a surface is called a *surface integral*. The *classical Stokes's theorem* says that for an oriented two dimensional surface S in a three dimensional space with oriented boundary curve ∂S we have

$$\int_S \left(\frac{\partial r}{\partial y} - \frac{\partial q}{\partial z}\right) dy\,dz + \left(\frac{\partial p}{\partial z} - \frac{\partial r}{\partial x}\right) dz\,dx + \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) dx\,dy = \int_{\partial S} (p\,dx + q\,dy + r\,dz). \tag{3.168}$$

This result for 2-forms has an obvious analog in n dimensions. This case of Stokes' theorem has important consequences for line integrals of closed forms.

Theorem 3.16 (Integral over a boundary) *Let ω be a closed differential 1-form. Let R be a surface on which ω is smooth, and let ∂R be the curve that is its oriented boundary. Then*

$$\int_{\partial R} \omega = 0. \tag{3.169}$$

Theorem 3.17 (Independence of path for closed forms) *Let ω be a closed differential 1-form. Let R be a surface on which ω is smooth, and let ∂R be the curve that is its oriented boundary. Suppose that C_1 and C_2 are oriented curves such that $C_1 - C_2 = \partial R$. Then*

$$\int_{C_1} \omega = \int_{C_2} \omega. \tag{3.170}$$

Theorem 3.18 (Exactness of closed forms in special regions) *Suppose ω is a closed differential 1-form that is smooth in an open set U. Suppose U has the property that whenever C is a closed curve in U, then $C = \partial R$ for some region in U. Then ω is exact in U.*

Proof: Fix an initial point and a final point, and suppose that the final point has coordinates $\mathbf{x}'$. Consider the scalar

$$s = h(\mathbf{x}') = \int_{\text{initial}}^{\text{final}(\mathbf{x}')} \omega. \tag{3.171}$$

By the property of the region U and the independence of path for closed forms, this is a well-defined scalar depending only on the final point. It is not too hard to show that $ds = \omega$ in U. □

The result in two dimensions only requires Green's theorem. Even this case is significant. Much of what is interesting in complex variables depends on the fact that

$$\alpha = \frac{x\,dy - y\,dx}{x^2 + y^2} \tag{3.172}$$

is a form (defined in the plane with one point removed) that is closed but not exact. If one considers the plane with an entire half-line from the origin removed, then this form is exact in that smaller region, in fact, $\alpha = d\phi$, where ϕ is a suitable angle. But the interest is in what happens with curves that go entirely around the origin. Since such a curve is not a boundary, it is not surprising that the result can be a non-zero multiple of 2π.

Gauss's theorem is the case relating $n-1$ forms to n forms. The classical case is when $n = 3$, so that it relates 2-forms to 3-forms. When $n = 2$ it is Green's theorem. Let W be an oriented three dimensional region, and let ∂W be the oriented surface that forms its boundary. Then the three dimensional version of Gauss's theorem states that

$$\int_W \left(\frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z}\right) dx\,dy\,dz = \int_{\partial W} a\,dy\,dz + b\,dz\,dx + c\,dx\,dy. \tag{3.173}$$

3.22 Picturing Stokes' theorem

This section treats the pictures associated with Stokes' theorem. It will appear that in n dimensional space it is easy to draw pictures of differential forms of degree 0, 1, $n-1$, and n. Since Stokes' theorem relates integrals of k forms and $k+1$ forms, we see that we will have nice pictures for $k = 0$, for $k = 1$ and $n = 2, 3$, and for $k = n-1$.

A 0-form s is a scalar field, and a scalar field is pictured by its contour hypersurfaces (surfaces of dimension $n-1$). Such surfaces have no boundaries. The differential of a scalar field s is an exact 1-form ds. It is pictured by looking close to each point; near a given point the contour hypersurfaces look like a hyperplanes.

In the case of a 1-form α the idea is to look at each point and draw contour hyperplanes in a region very near the point. These hyperplanes include the subspaces on which the form vanishes. One can also imagine nearby hyperplanes on which the form assumes a constant value. The problem is that these hyperplanes do not necessarily fit together to form a hypersurface. Nevertheless, the integral $\int_C \alpha$ over an oriented curve C is well-defined; it is obtained by integrating the changes in value as the curve crosses the little hyperplanes. If in a discrete version of the integation the spacing between the hyperplanes corresponds to constant changes, then the calculation reduces to counting the number of hyperplanes crossed, keeping track of whether the crossing is upward or downward. The Stokes' theorem for a 1-form α says that for an oriented surface S with oriented boundary ∂S we have

$$\int_S d\alpha = \int_{\partial S} \alpha. \tag{3.174}$$

For a closed 1-form $d\alpha = 0$ and the integral over every boundary is zero.

A k-form in n dimensions is a much more complicated object. A strategy to visualize them is to look at a certain subspace where the form vanishes. This does not completely characterize the form, but it gives at least some intuition about what it looks like.

Consider the case of dimension n. If X is a vector field, and ω is a k-form, the we may define the *interior product* of X with ω to be the $k-1$ form $X \rfloor \omega$ defined by

$$\langle X \rfloor \omega \mid Y_1, \ldots, Y_{k-1} \rangle = \langle \omega \mid X, Y_1, \ldots, Y_{k-1} \rangle. \tag{3.175}$$

It follows from the definition that the $k-2$ form $X \rfloor (X \rfloor \omega) = 0$.

The way to compute with the interior product is to use the fact that

$$\frac{\partial}{\partial x_j} \rfloor \, dx_j \wedge d\mathbf{x}_I = d\mathbf{x}_I. \tag{3.176}$$

Thus, for instance, the interior product of $\frac{\partial}{\partial y}$ with $dx\,dy$ is equal to the interior product of $\frac{\partial}{\partial y}$ with $-dy\,dx$ which is $-dx$.

The *characteristic subspace* of ω is the subspace of all X with $X \rfloor \omega = 0$. The condition for X to be a characteristic vector is that for every k-tuple $X, Y_1, \ldots, Y_{k-1}$ to which X belongs we have $\langle \omega \mid X, Y_1, \ldots, Y_{k-1} \rangle = 0$. If the k-form is non-zero, then the dimension of the characteristic subspace is $\leq n-k$.

If the characteristic subspace of ω has dimension $n-r$ at each point, then the form is said to have *rank* r. If the k-form is non-zero, then the rank is $\geq k$. It is not true that every non-zero k-form is of rank k. The simplest counterexample is in dimension 4 and is $\omega = dx_1\,dx_2 + dx_3\,dx_4$, which has rank 4. In this case, the characteristic subspace consists only of the zero vector. It may be shown that a non-zero k-form is of rank k if and only if it is decomposable, that is, it may be represented as a product of non-zero 1-forms. The form in the example is not decomposable. For more on decomposable forms see the books by Crampin and Pirani [5] and by Şuhubi [19].

If a k-form is of rank k, then it would seem natural to picture it by its corresponding characteristic subspace of dimension $n-k$. This may not give complete information about the form, but it will indicate its general character.

Proposition 3.19 *Consider a non-zero differential k-form in n dimensions. If $k = 1$ or if $k = n-1$, then the form has rank k.*

Proof: For the case $k = 1$ of a 1-form α the equation for a characteristic vector is

$$X \rfloor \alpha = \sum_{i=1}^{n} a_i \frac{\partial}{\partial x_i} \rfloor \sum_{j=1}^{n} p_j \, dx_j = \sum_{i=1}^{n} a_i p_i = 0. \tag{3.177}$$

If the form α is non-zero at some point, then some $p_j \neq 0$, and so the corresponding space of vectors X at the given point is $n-1$ dimensional.

The case of an $n-1$ form ω is more interesting. Let $\sigma = dx_1 \wedge \cdots dx_n$ be a volume form. The general $n-1$ form may always be written as an interior product with a vector field

$$\omega = Y \rfloor \sigma = \sum_{j=1}^{n} p_j \frac{\partial}{\partial x_j} \rfloor \sigma = \sum_{j=1}^{n} p_j (-1)^{j-1} \, dx_1 \cdots dx_{j-1} \, dx_{j+1} \cdots dx_n. \quad (3.178)$$

The characteristic equation is then $X \rfloor \omega = X \rfloor (Y \rfloor \sigma) = 0$. Suppose that Y and hence ω are non-zero. Then for every non-zero scalar $\lambda \neq 0$ we have that $X = \lambda Y$ is a non-zero solution of the characteristic equation. On the other hand, if X and Y are linearly independent, then it is easy to see that $X \rfloor (Y \rfloor \sigma) \neq 0$. □

It is tempting to represent the $n-1$ form ω by a vector field Y satisfying the characteristic equation, but this choice of vector field is not unique. (In the proof above the choice of vector field depends on the choice of volume form σ, which in turn depends on the coordinate system.) Instead, one should think of an $n-1$ form ω as a field of long very thin cylindrical tubes, such that the vector field X is tangent to the tubes. The reason for thinking of tubes instead of lines is to capture the intuition about orientation; the tubes may be thought of as having a kind of spiral orientation that gives an orientation in the space transverse to the tube. In a discrete approximation the tubes may be pictured as closely spaced, and the integral $\int_S \omega$ of the form ω over an oriented hypersurface S is then proportional to the number of tubes that penetrate the surface, taking into account a sign at each intersection point from the orientation. In general tubes will have oriented end points, and the density of these points gives a geometric representation of the derivative

$$d\omega = \sum_{j=1}^{n} \frac{\partial p_j}{\partial x_j} \, dx_1 \cdots dx_n. \quad (3.179)$$

Stokes' theorem for an $n-1$ form ω is then the Gauss theorem

$$\int_W d\omega = \int_{\partial W} \omega. \quad (3.180)$$

The left hand side represents the integral of the oriented end points over the n-dimensional oriented region W. The right hand side represents the flux of ω though the boundary surface ∂W, which in a discrete approximation is imagined as proportional to the number of tubes penetrating the surface, taking into account orientation. In other words, the output through the boundary is explained by an integral of the production inside the region. In the discrete approximation, the Gauss theorem reduces to counting, since the number of lines passing through the bounding surface ∂W described by the $n-1$ form ω corresponds to the number of points in the interior W described by the n-form $d\omega$, as usual taking signs into account. In the case when ω is closed, there is no production, and the total flow across the boundary ∂W is zero.

We shall see in the next chapter that when there is a given volume form, then the $n-1$ form in the Gauss theorem may be represented by a vector field. In this case the Gauss theorem becomes the divergence theorem.

In the case $n = 3$ a 1-form α has a derivative $d\alpha$ that is a 2-form. The 1-form may be pictured as surfaces that end in curves, and the the 2-form is represented by thin tubes. The tubes act as hinges on which the surfaces hang. Since $d\alpha$ is an exact form, the tubes have no end points. Stoke's theorem relates the 1-form α represented by the surfaces to the 2-form $d\alpha$ represented by the thin tubes. The formula represents the integral of α around the closed curve ∂S in terms of the integrated flux of the tubes given by $d\alpha$ through a surface S that this curve bounds. The result is independent of the surface. In a discrete approximation the Stoke's theorem is again the result of counting, since the number of surfaces with equal increments (taking into account increase or decrease) is equal to the number of tubes acting as hinges (taking into account orientation).
Example: Take the form $\alpha = x\,dy$ in three dimensions and an oriented surface S in a constant z plane bounded by two values of y and by a positive and a negative value of x. The boundary ∂S is taken oriented so that y increases for the positive value of x and y decreases for the negative value of x. Then the integral along the boundary is positive. This is explained by the tubes representing $d\alpha = dx\,dy$, which are vertical, of constant density, and have an orientation compatible with the orientation of S. |

The case $n = 2$ gives rise to two equivalent pictures. The first picture is that of Green's theorem. Represent the 1-form $\alpha = p\,dx + q\,dy$ near each point by the line where it vanishes. Then α represents the increase or decrease as one crosses such lines along an oriented curve. It may help to think of the lines as double lines representing a step up or a step down in a given direction. The differential $d\alpha = (\partial q/\partial x - \partial p/\partial y)\,dx\,dy$ represents the density of hinge points where the lines begin. So Green's theorem says that the total increase or decrease is completely explained by this cloud of hinge points. When the form α is closed, there are no hinge points, and the integral around every closed curve is zero.
Example: Take the form $\alpha = x\,dy$ in two dimensions. This is represented by lines of constant y whose spacing decreases and reverses sign as x passes through zero. Consider a region S bounded by two values of y and by a positive and a negative value of x. The boundary ∂S is taken oriented so that y increases for the positive value of x and y decreases for the negative value of x. One thinks of α as indicating some sort of subjective change in vertical distance. The integral along ∂S is positive, since one is going uphill along the entire closed curve. In essence this is the famous picture of the Penrose stairs. (The most famous illustration of these stairs is the Ascending and Descending lithograph print by Escher.) |

The other picture for $n = 2$ is closer to the Gauss theorem. It is suggested by writing $\alpha = p\,dy - q\,dx$ and consider it as a $2-1$ form. In dimension 2 the tubes representing the form have a transverse orientation. These can be thought of as double lines, where the orientation goes from one line to its neighbor. So α represents the amount that these lines cross an oriented curve, taking into

account the orientation. The differential $d\alpha = (\partial p/\partial x + \partial q/\partial y)\, dx\, dy$ represents the density of the oriented points at which these tubes begin. This picture suggests a conservation law, where the amount of flux across the boundary is explained by the rate of generation within the region. If the form is closed, then there is no generation within, and so the flow is compensated by the flow out.
Example: It is important that the boundary ∂S is the entire boundary of the region S where the form is defined. The classical example is the form $\alpha = (x\, dy - y\, dx)/(x^2 + y^2)$ defined in the plane with the origin removed. This 1-form is described by radial (constant angle) half-lines. Say that S is the annulus between the circles $x^2 + y^2 = 1$ and $x^2 + y^2 = 4$. Then ∂S consists of the two circles, with opposite orientations. Each line that enters the annulus on one circle leaves it on the other circle. So there is complete cancellation, and the integral of α over ∂S is zero. (Actually, the computation of the boundary ∂S is slightly more complicated. If S is regarded as a chain parameterized by a rectangular region, then the boundary ∂S is a chain consisting of the two circles with opposite orientations, plus two segments along $y = 0$ with $1 < x < 2$. These last two segments have opposite orientations, so their contributions cancel, and they may be ignored.) |

3.23 Electric and magnetic fields

The power of the differential forms concept is that it works in spaces without much structure, other than a notion of differentiability. Euclidean space is quite special in this regard, in that there is considerable structure, in particular notions such as length and orthogonality. For applications of forms to problems involving Euclidean space, it is desirable to use basic forms that are of unit length and are mutually orthogonal. The most obvious example is to take Cartesian coordinates, where the basic forms are dx, dy, and dz, taken in that order. Once we have such basic forms, we can convert from 1-forms to vector fields, just by taking the same components. We can also convert from 1-forms to 2-forms, by replacing dx, dy and dz by $dy\, dz$, $dz\, dx$, and $dx\, dy$.

This also works in other coordinate systems. The most important example is spherical polar coordinates r, θ, ϕ, where $x = r\cos(\phi)\sin(\theta)$, $y = r\sin(\phi)\sin(\theta)$, and $z = r\cos(\theta)$. Here $r^2 = x^2 + y^2 + z^2$. This is the distance from the origin, so surfaces of constant r are spheres. The quantities θ and ϕ are co-latitude and longitude. The basic forms are dr, $r\, d\theta$, and $r\sin(\theta)\, d\phi$.

Sometimes it is useful to use cylindrical coordinates ρ, ϕ, z. Here $x = \rho\cos(\phi)$, $y = \rho\sin(\phi)$, and $z = z$. This time $\rho^2 = x^2 + y^2$. This is the distance from the z axis, so surfaces of constant ρ are cylinders. The basic forms are $d\rho$, $\rho\, d\phi$, and dz.

An *electric field* is most often modeled as a vector field. However in this section we will think of it as a 1-form. Thus, for example, the electric field of a point charge at the origin is

$$E = \frac{1}{4\pi r^2} \frac{x\, dx + y\, dy + z\, dz}{r}. \tag{3.181}$$

This can also be written in spherical polar coordinates as

$$E = \frac{1}{4\pi r^2}\, dr. \tag{3.182}$$

This 1-form is

$$E = -d\phi, \tag{3.183}$$

where ϕ is a 0-form called the *electric potential* . For the electric field of a point charge at the origin the electric potential is

$$\phi = \frac{1}{4\pi r}. \tag{3.184}$$

The surfaces of constant potential are spheres. In most physics texts the electric field E is represented by the gradient vector field, which gives vectors in the radial direction, orthogonal to these spheres.

There is another kind of electric field often considered in physics. It is called the *electric displacement* (also called *electric flux density*) and is written D. For our purposes we can think of it as the same field, but considered as a 2-form. Thus for the electric field of a point charge this is

$$D = \frac{1}{4\pi r^2}\frac{x\, dy\, dz + y\, dz\, dx + z\, dx\, dy}{r}. \tag{3.185}$$

This has a nice expression in spherical polar coordinates as

$$D = \frac{1}{4\pi r^2} r\, d\theta\; r\sin(\theta)\, d\phi = \frac{1}{4\pi}\sin(\theta)\, d\theta\, d\phi. \tag{3.186}$$

This is represented geometrically by lines coming radially from the origin.

The fundamental equations of electrostatics are

$$dE = 0 \tag{3.187}$$

and

$$dD = R. \tag{3.188}$$

Here R is a 3-form called the charge-density.

These equation have integral forms. The first equation says that for every surface W the integral around the closed curve ∂W that is its boundary is zero:

$$\int_{\partial W} E = 0. \tag{3.189}$$

The second says that for every region Ω the integral over the surface $\partial\Omega$ of D is equal to the total charge in the region:

$$\int_{\partial\Omega} D = \int_{\Omega} R. \tag{3.190}$$

As an illustration of these ideas, here is the computation of the electric displacement field of a charge that is uniform on a ball of radius ϵ about the

origin. For D we have the same formula as before for $r \geq \epsilon$. However for $r \leq \epsilon$ we have

$$D_{\text{in}} = \frac{r^3}{\epsilon^3} D. \tag{3.191}$$

One way to see that this works is to compute the exterior derivative. This is for $r \leq \epsilon$

$$d\, D_{\text{in}} = \frac{3r^2}{\epsilon^3}\, dr\, \frac{1}{4\pi} \sin(\theta)\, d\theta\, d\phi = \left(\frac{4\pi\epsilon^3}{3}\right)^{-1} r^2 \sin(\theta)\, dr\, d\theta\, d\phi = \left(\frac{4\pi\epsilon^3}{3}\right)^{-1} dx\, dy\, dz. \tag{3.192}$$

Indeed this a constant charge density within the ball with total charge equal to one. Thus the radial lines representing the electric displacement field begin inside this ball. If we restrict the form to the region outside the charged ball, then D is a closed 2-form that is not exact.

The corresponding electric field for $r \leq \epsilon$ is

$$E_{\text{in}} = \frac{r^3}{\epsilon^3} E = \frac{1}{4\pi} \frac{r}{\epsilon^3}\, dr. \tag{3.193}$$

The potential is

$$\phi_{\text{in}} = \frac{1}{4\pi\epsilon^3} \frac{1}{2} (3\epsilon^2 - r^2). \tag{3.194}$$

The constant term makes the potential continuous at $r = \epsilon$.

A *magnetic field* (also called *magnetic flux density*) is most often modeled as a vector field, but in many ways it is more natural to model it as a differential 2-form. For example, consider the case of a wire bearing current along the z axis. The magnetic flux density in this case is most easily expressed in cylindrical coordinates via

$$B = \frac{1}{2\pi}\, dz\, \frac{x\, dx + y\, dy}{x^2 + y^2} = \frac{1}{2\pi\rho}\, dz\, d\rho. \tag{3.195}$$

The lines of magnetic flux density circle around the wire; their density drops off with distance from the wire.

There is another kind of magnetic field, sometimes called the *magnetic field intensity*. It is written H. For our purposes we can think of it as the same field, but considered as a 1-form. Thus for the magnetic field intensity of the wire we have

$$H = \frac{1}{2\pi} \frac{x\, dy - y\, dx}{x^2 + y^2} = \frac{1}{2\pi\rho}\, \rho\, d\phi = \frac{1}{2\pi}\, d\phi \tag{3.196}$$

in cylindrical coordinates. The surfaces of constant ϕ are planes with one side on the wire. Notice that this is not an exact 1-form, since the integral around a closed curve surrounding the wire is 1. In physics the magnetic field intensity is often represented by a vector field orthogonal to the planes, again circling the wire.

The fundamental equations of magnetostatics are

$$dB = 0 \tag{3.197}$$

and

$$dH = J. \tag{3.198}$$

Here J is a 2-form called the current-density. The fact that J is exact represents current conservation: the lines representing the 2-form J have no end points.

These equations have integral forms. The first says that for every region Ω the integral over the surface $\partial\Omega$ of B is zero.

$$\int_{\partial\Omega} B = 0. \tag{3.199}$$

Magnetic flux lines never end. The second equation says that for every surface W the integral of H around the closed curve ∂W that is its boundary is the current passing through the surface.

$$\int_{\partial W} H = \int_W J. \tag{3.200}$$

As an illustration of these ideas, here is the computation of the magnetic intensity field of a current that is uniform on a cylinder of radius ϵ about the z axis. For H we have the same formula as before for $\rho \geq \epsilon$. However for $\rho \leq \epsilon$ we have

$$H_{\text{in}} = \frac{\rho^2}{\epsilon^2} H. \tag{3.201}$$

One way to see that this works is to compute the exterior derivative. This is for $r \leq \epsilon$

$$d\, H_{\text{in}} = \frac{2\rho}{\epsilon^2}\, d\rho\, \frac{1}{2\pi}\, d\phi = \left(\pi\epsilon^2\right)^{-1} \rho\, d\rho\, d\phi = \left(\pi\epsilon^2\right)^{-1}\, dx\, dy. \tag{3.202}$$

Indeed this is a constant current density within the cylinder with total current equal to one. Thus the planes representing the H field end in lines inside the cylinder. If we restrict H to the region outside the wire, then it is an example of a closed 1-form that is not exact.

The corresponding magnetic flux density for $\rho \leq \epsilon$ is

$$B_{\text{in}} = \frac{\rho^2}{\epsilon^2} B = \frac{1}{2\pi}\frac{\rho}{\epsilon^2}\, dz\, d\rho. \tag{3.203}$$

Since $dB = 0$, it seems reasonable that it should have a 1-form *magnetic potential* A with $dA = B$. Such a potential is of course not unique, since one may add to it a 1-form of the form ds and get the same magnetic flux density. As an example, for the case of the wire the vector potential outside the wire may be taken to be

$$A = -\frac{1}{2\pi} \log(\frac{\rho}{\epsilon})\, dz. \tag{3.204}$$

The reason for writing it with the $\epsilon > 0$ is that it is convenient to have the magnetic potential be zero at the surface of the wire. The corresponding expression inside is then

$$A_{\text{in}} = -\frac{1}{2\pi}\frac{1}{2}\left(\frac{\rho^2}{\epsilon^2} - 1\right) dz. \tag{3.205}$$

This also is zero at the surface of the wire.

Problems 7: Vector fields

1. Straightening out. A vector field that is non-zero at a point can be transformed into a constant vector field near that point by a change of coordinate system. Pick a point away from the origin, and find coordinates u, v so that
$$-\frac{y}{x^2+y^2}\frac{\partial}{\partial x}+\frac{x}{x^2+y^2}\frac{\partial}{\partial y}=\frac{\partial}{\partial u}. \tag{3.206}$$

2. Linearization. Consider the vector field
$$\mathbf{u}=x(4-x-y)\frac{\partial}{\partial x}+(x-2)y\frac{\partial}{\partial y}. \tag{3.207}$$
Find its zeros. At each zero, find its linearization. For each linearization, find its eigenvalues. Use this information to sketch the vector field.

3. Nonlinearity. Consider the vector field
$$\mathbf{v}=(1+x^2+y^2)y\frac{\partial}{\partial x}-(1+x^2+y^2)x\frac{\partial}{\partial y}. \tag{3.208}$$
Find its linearization at zero. Show that there is no coordinate system near 0 in which the vector field is expressed by its linearization. Hint: Solve the associated system of ordinary differential equations, both for $\mathbf{v}$ and for its linearization. Find the period of a solution in both cases.

4. Nonlinear instability. Here is an example of a fixed point where the linear stability analysis gives an elliptic fixed point, but changing to polar coordinates shows the unstable nature of the fixed point:
$$\frac{dx}{dt} = -y+x(x^2+y^2) \tag{3.209}$$
$$\frac{dy}{dt} = x+y(x^2+y^2). \tag{3.210}$$
Change the vector field to the polar coordinate representation, and solve the corresponding system of ordinary differential equations.

5. A predator-prey system. Fix $\alpha>0$. In the region with $0<u$ and $0<v$ consider the system
$$\frac{du}{dt} = u(1-v) \tag{3.211}$$
$$\frac{dv}{dt} = \alpha v(u-1).$$
The u variable represents prey; the v variable represents predators. (a) Sketch this vector field. Find the zero. What kind of linearization is there at this zero?

(b) Show that each solution satisfies

$$\alpha v(u-1)\,du + u(v-1)\,dv = 0. \tag{3.212}$$

Show that $1/(uv)$ is an integrating factor for this differential form.

(c) Integrate this form to find an equation for the solution curves in the u, v plane. Show that these are compatible with your sketch. What value of the constant of integration corresponds to the fixed point?

Recitation 7

1. Exact differentials. Is $(x^2 + y^2)\,dx + 2xy\,dy$ an exact differential form? If so, write it as the differential of a scalar.

2. Exact differentials. Is $(1 + e^x)\,dy + e^x(y - x)\,dx$ an exact differential? If so, write it as the differential of a scalar.

3. Exact differentials. Is $e^y\,dx + x(e^y + 1)\,dy$ an exact differential? If so, write it as the differential of a scalar.

4. Constant differential forms. A differential form usually cannot be transformed into a constant differential form, but there are special circumstances when that can occur. Is it possible to find coordinates u and v near a given point (not the origin) such that

$$-y\,dx + x\,dy = du? \tag{3.213}$$

5. Constant differential forms. A differential form usually cannot be transformed into a constant differential form, but there are special circumstances when that can occur. Is it possible to find coordinates u and v near a given point (not the origin) such that

$$-\frac{y}{x^2 + y^2}\,dx + \frac{x}{x^2 + y^2}\,dy = du? \tag{3.214}$$

6. Ordinary differential equations. Solve the differential equation $(xy^2 + y)\,dx - x\,dy = 0$ by finding an integrating factor that depends only on y.

Problems 8: Differential forms

1. Say that the differential 1-form $\alpha = p\,dx + q\,dy + r\,dz$ has an integrating factor $\mu \neq 0$ such that $\mu\alpha = ds$. Prove that $\alpha \wedge d\alpha = 0$. Also, express this condition as a condition on p, q, r and their partial derivatives.

2. Show that $\alpha = dz - y\,dx - dy$ has no integrating factor.

3. Show that the differential 1-form $\alpha = yz\,dx + xz\,dy + dz$ passes the test for an integrating factor.

4. In the previous problem it might be difficult to guess the integrating factor. Show that $\mu = e^{xy}$ is an integrating factor, and find s with $\mu\alpha = ds$.

5. The differential 2-form $\omega = (2xy - x^2)\,dx\,dy$ is of the form $\omega = d\alpha$, where α is a 1-form. Find such an α. Hint: This is too easy; there are many solutions.

Recitation 8

1. The differential 3-form $\sigma = (yz + x^2z^2 + 3xy^2z)\,dx\,dy\,dz$ is of the form $\sigma = d\omega$, where ω is a 2-form. Find such an ω. Hint: Many solutions.

2. Let $\sigma = xy^2z\,dy\,dz - y^3z\,dz\,dx + (x^2y + y^2z^2)\,dx\,dy$. Show that this 2-form σ satisfies $d\sigma = 0$.

3. The previous problem gives hope that $\sigma = d\alpha$ for some 1-form α. Find such an α. Hint: This may require some experimentation. Try α of the form $\alpha = p\,dx + q\,dy$, where p, q are functions of x, y, z. With luck, this may work. Remember that when integrating with respect to z the constant of integration is allowed to depend on x, y.

Problems 9: Stokes' theorem

1. Let C be the curve $x^2 + y^2 = 1$ in the first quadrant from $(1, 0)$ to $(0, 1)$. Evaluate
$$\int_C xy\,dx + (x^2 + y^2)\,dy. \tag{3.215}$$

2. Let C be a curve from $(2, 0)$ to $(0, 3)$. Evaluate
$$\int_C 2xy\,dx + (x^2 + y^2)\,dy. \tag{3.216}$$

3. Consider the problem of integrating the differential form
$$\alpha = -\frac{y}{x^2 + y^2}\,dx + \frac{x}{x^2 + y^2}\,dy \tag{3.217}$$
from $(1, 0)$ to $(-1, 0)$ along some curve avoiding the origin. There is an infinite set of possible answers, depending on the curve. Describe all such answers.

4. Let R be the region $x^2 + y^2 \le 1$ with $x \ge 0$ and $y \ge 0$. Let ∂R be its boundary (oriented counterclockwise). Evaluate directly
$$\int_{\partial R} xy\,dx + (x^2 + y^2)\,dy. \tag{3.218}$$

5. This continues the previous problem. Verify Green's theorem in this special case, by explicitly calculating the appropriate integral over the region R.

Recitation 9

1. Let
$$\alpha = -y\,dx + x\,dy + xy\,dz. \tag{3.219}$$
Fix $a > 0$. Consider the surface S that is the hemisphere $x^2 + y^2 + z^2 = a^2$ with $z \geq 0$. Integrate α over the boundary ∂S of this surface (a counterclockwise circle in the x, y plane).

2. This continues the previous problem. Verify Stokes's theorem in this special case, by explicitly calculating the appropriate integral over the surface S.

3. Let $\sigma = xy^2z\,dy\,dz - y^3z\,dz\,dx + (x^2y + y^2z^2)\,dx\,dy$. Integrate σ over the sphere $x^2 + y^2 + z^2 = a^2$. Hint: This should be effortless.

Chapter 4

The Metric Tensor

4.1 The interior product

The subject of differential forms is general and beautiful, but it fails to capture the ideas of length, area, and volume. These all depend on a more complicated kind of calculation, where, in the spirit of Pythagoras, one takes a square root of a sum of squares. We begin this part of the notes with an easy case, when one simply assumes that one has an understanding of volume. This leads to an important reformulation of Stokes' theorem called the divergence theorem. Then we look more seriously at the extra ingredient that is needed to deal with these topics, that is, with the metric tensor.

There is an operation called *interior product* (or *contraction*). In the case of interest to us, it is a way of defining the product of a vector field $\mathbf{u}$ with a k-form ω to get a $k-1$ form $\mathbf{u}\rfloor\omega$. (The exterior product is also called the wedge product; the interior product is sometimes called the hook product. In the following we call it the interior product.) The definition of interior product that is used in practice is very simple and may be illustrated by an example. Consider a basis vector field $\frac{\partial}{\partial u_j}$ and a differential form $du_i\, du_j\, du_k$. The rule is to move the du_j to the left and then remove it. So the result in this case is

$$\frac{\partial}{\partial u_j}\rfloor\, du_i\, du_j\, du_k = -\frac{\partial}{\partial u_j}\rfloor\, du_j\, du_i\, du_k = -du_i\, du_k. \tag{4.1}$$

There is a general theoretical definition that is also illuminating. Thus for $k \geq 1$ the interior product of the vector field $\mathbf{u}$ with the k-form σ is defined by

$$\langle \mathbf{u}\rfloor\sigma \mid \mathbf{v}_1, \ldots, \mathbf{v}_{k-1}\rangle = \langle \sigma \mid \mathbf{u}, \mathbf{v}_1, \ldots, \mathbf{v}_{k-1}\rangle. \tag{4.2}$$

When $k = 1$ this is already familiar. For a 1-form α the interior product $\mathbf{u}\rfloor\alpha$ is the scalar $\langle \alpha \mid \mathbf{v}\rangle$. For a scalar field s we take $\mathbf{u}\rfloor s = 0$.

One interesting property of the interior product is that if α is an r-form and β is an s-form, then

$$\mathbf{u}\rfloor(\alpha \wedge \beta) = (\mathbf{u}\rfloor\alpha) \wedge \beta + (-1)^r \alpha \wedge (\mathbf{u}\rfloor\beta). \tag{4.3}$$

This is a kind of triple product identity.

In particular, we may apply this when $r = 1$ and $s = n$. Since β is an n-form, it follows that $\alpha \wedge \beta = 0$. Hence we have in this special case

$$\langle \alpha \mid \mathbf{u}\rangle\beta = \alpha \wedge (\mathbf{u}\rfloor\beta). \tag{4.4}$$

Another application is with two 1-forms β and γ. In this case it gives

$$\mathbf{a}\rfloor(\beta \wedge \gamma) = \langle \beta \mid \mathbf{a}\rangle\gamma - \langle \gamma \mid \mathbf{a}\rangle\beta. \tag{4.5}$$

So the interior product of a vector with $\beta \wedge \gamma$ is a linear combination of β and γ.

Later we shall see the connection with classical vector algebra in three dimensions. The exterior product $\beta \wedge \gamma$ is an analog of the cross product, while

$\alpha \wedge \beta \wedge \gamma$ is an analog of the triple scalar product. The combination $-\mathbf{a}\rfloor(\beta \wedge \gamma)$ will turn out to be an analog of the triple vector product.

The general formula that is used for computations is that for a vector field $\mathbf{v}$ and 1-forms $\alpha_1, \ldots, \alpha_k$ we have

$$\mathbf{v}\rfloor(\alpha_1 \wedge \cdots \wedge \alpha_k) = \sum_{i=1}^{k} (-1)^{i-1} \langle \alpha \mid \mathbf{v} \rangle \alpha_1 \wedge \cdots \wedge \alpha_{i-1} \wedge \alpha_{i+1} \wedge \cdots \wedge \alpha_k. \quad (4.6)$$

In a coordinate representation this implies the following identity for the interior product of a vector field with an n-form:

$$\left(\sum_{j=1}^{n} a_j \frac{\partial}{\partial u_j} \right) \rfloor \, du_1 \cdots du_n = \sum_{i=1}^{n} a_i (-1)^{i-1} \, du_1 \cdots \, du_{i-1} \, du_{i+1} \cdots \, du_n. \quad (4.7)$$

4.2 Volume

Consider an n-dimensional manifold. The new feature is a given n-form, taken to be never zero. We denote this *volume form* by vol. In coordinates it is of the form

$$\mathrm{vol} = \sqrt{g} \, du_1 \cdots du_n. \quad (4.8)$$

This coefficient $\sqrt{g}$ depends on the coordinate system. The choice of the notation $\sqrt{g}$ for the coefficient will be explained in the following chapter. (Then $\sqrt{g}$ will be the square root of the determinant of the matrix associated with the Riemannian metric for this coordinate system.) It is typical to make the coordinate system compatible with the orientation, so that volumes work out to be positive.

The most common examples of volume forms are the volume in $\mathrm{vol} = dx \, dy \, dz$ in Cartesian coordinates and the same volume $\mathrm{vol} = r^2 \sin(\theta) \, dr \, d\theta \, d\phi$ in spherical polar coordinates. The convention we are using for spherical polar coordinates is that θ is the co-latitude measured from the north pole, while ϕ is the longitude. We see from these coordinates that the $\sqrt{g}$ factor for Cartesian coordinates is 1, while the $\sqrt{g}$ factor for spherical polar coordinates is $r^2 \sin(\theta)$.

In two dimensions it is perhaps more natural to call this area. So in Cartesian coordinates $\mathrm{area} = dx \, dy$, while in polar coordinates $\mathrm{area} = r \, dr \, d\phi$.

For each scalar field s there is an associated n-form $s \, \mathrm{vol}$. The scalar field and the n-form determine each other in an obvious way. They are said to be dual to each other, in a certain special sense.

For each vector field $\mathbf{v}$ there is an associated $n-1$ form given by $\mathbf{v}\rfloor\mathrm{vol}$. The vector field and the $n-1$ form are again considered to be dual to each other, in the same sense. If $\mathbf{v}$ is a vector field, then $\mathbf{v}\rfloor\mathrm{vol}$ might be called the corresponding *flux density*. It is an $n-1$ form that describes how much $\mathbf{v}$ is penetrating a given $n-1$ dimensional surface. In coordinates we have

$$\left(\sum_{j=1}^{n} a_j \frac{\partial}{\partial u_j} \right) \rfloor \, \mathrm{vol} = \sum_{i=1}^{n} a_i (-1)^{i-1} \sqrt{g} \, du_1 \cdots \, du_{i-1} \, du_{i+1} \cdots \, du_n. \quad (4.9)$$

In two dimensions a vector field is of the form

$$\mathbf{u} = a\frac{\partial}{\partial u} + b\frac{\partial}{\partial v}. \tag{4.10}$$

The area form is

$$\text{area} = \sqrt{g}\, du\, dv. \tag{4.11}$$

The corresponding flux is

$$\mathbf{u}\rfloor\text{area} = \sqrt{g}(a\, dv - b\, du). \tag{4.12}$$

In three dimensions a vector field is of the form

$$\mathbf{u} = a\frac{\partial}{\partial u} + b\frac{\partial}{\partial v} + c\frac{\partial}{\partial w}. \tag{4.13}$$

The volume form is

$$\text{vol} = \sqrt{g}\, du\, dv\, dw. \tag{4.14}$$

The corresponding flux is

$$\sqrt{g}(a\, dv\, dw + b\, dw\, du + c\, du\, dv). \tag{4.15}$$

4.3 The divergence theorem

Consider a vector field $\mathbf{v}$ in a space with a volume element vol. Define the *flux density* of the vector field to be the $n-1$ form $\mathbf{v}\rfloor\text{vol}$. Define the *divergence* of a vector field $\mathbf{v}$ to be the scalar div $\mathbf{v}$ such that

$$d(\mathbf{v}\rfloor\text{vol}) = \text{div } \mathbf{v}\, \text{vol}. \tag{4.16}$$

In other words, it is the dual of the differential of the dual. The general divergence theorem then takes the following form.

Theorem 4.1 *Consider an n-dimensional region for which there is a volume form. Consider a vector field $\mathbf{v}$ with its associated $n-1$ form flux density $\mathbf{v}\rfloor\text{vol}$, and consider the scalar field* div $\mathbf{v}$ *associated with the exterior derivative of this form. Then Stokes' theorem gives*

$$\int_W \text{div } \mathbf{v}\, \text{vol} = \int_{\partial W} \mathbf{v}\rfloor\text{vol}. \tag{4.17}$$

The right hand side in the divergence theorem is called the *flux* of the vector field through the surface. The left hand suggests that the flux is produced by the divergence of the vector field in the interior.

Suppose that

$$\mathbf{v} = \sum_{j=1}^{n} a_j \frac{\partial}{\partial u_j}. \tag{4.18}$$

To compute the integral over the boundary we have to pull back the differential form $\mathbf{v}\rfloor\text{vol}$ to the parameter space. Say that $t_1, \ldots, t_{n-1}$ are the parameters. Then the differential form pulls back to

$$\phi^*(\mathbf{v}\rfloor\text{vol}) = \sum_{j=1}^{n} a_j \nu_j \sqrt{g}\, dt_1 \cdots dt_{n-1}, \tag{4.19}$$

where ν_i is $(-1)^{i-1}$ times the determinant of the matrix obtained by deleting the ith row from the matrix $\partial u_j/\partial t_\alpha$. The quantity $\sqrt{g}$ must also be expressed in terms of the t_i. The explicit form of the theorem is then

$$\int_W \frac{1}{\sqrt{g}} \sum_{j=1}^{n} \frac{\partial}{\partial u_j}(\sqrt{g} a_j)\sqrt{g}\, du_1 \cdots du_n = \int \sum_{j=1}^{n} a_j \nu_j \sqrt{g}\, dt_1 \cdots dt_{n-1}. \tag{4.20}$$

If one compares this with other common formulations of the divergence theorem, one sees that there is no need to normalize the components ν_j, and there is also no need to compute the surface area of ∂W. Both of these operations can be a major nuisance; it is satisfying that they are not necessary.

There is another closely related way of looking at the surface integral in the divergence theorem. The terminology is not standardized, but here is one choice. Define the *transverse surface element* to be the *interior pullback* of the volume as

$$\text{element} = \phi_1^*\text{vol} = \phi_1^*\sqrt{g}\, du_1, \cdots, du_n = \sum_{i=1}^{n} \sqrt{g}\nu_i\, dx_i\, dt_1, \cdots, dt_{n-1}. \tag{4.21}$$

Notice that this is just the pullback where one freezes the du_i and only pulls back the other du_j. Then the flux density pulled back to the surface may be written

$$\phi^*(\mathbf{v}\rfloor\text{vol}) = \mathbf{v}\rfloor\text{element} = \sqrt{g}\sum_{i=1}^{n} a_i \nu_i\, dt_1, \cdots, dt_{n-1}. \tag{4.22}$$

The interior product on the left is a vector field with an n-form, while the interior product on the right is a vector field with an 1-form (freezing the dt_i). So the vector field paired with this surface element is the object that must be integrated.

In two dimensions the divergence theorem says that

$$\int_R \frac{1}{\sqrt{g}}\left(\frac{\partial\sqrt{g}a}{\partial u} + \frac{\partial\sqrt{g}b}{\partial v}\right)\text{area} = \int_{\partial R} \sqrt{g}(a\, dv - b\, du). \tag{4.23}$$

Here the area form is $\sqrt{g}\, du\, dv$, where the particular form of $\sqrt{g}$ is that associated with the u, v coordinate system. Notice that the coefficients in the vector field are expressed with respect to a coordinate basis. We shall see in the next part of this book that this is not the only possible choice. This theorem involves a kind of line integral, but the starting point is a vector field instead of a 1-form.

A marvellous application of the divergence theorem in two dimensions is the formula

$$\int_R dx\,dy = \frac{1}{2}\int_{\partial R} x\,dy - y\,dx. \tag{4.24}$$

This says that one can determine the area by walking around the boundary. It is perhaps less mysterious when one realizes that $x\,dy - y\,dx = r^2\,d\phi$.

In three dimensions the divergence theorem says that

$$\int_W \frac{1}{\sqrt{g}}\left(\frac{\partial\sqrt{g}a}{\partial u} + \frac{\partial\sqrt{g}b}{\partial v} + \frac{\partial\sqrt{g}c}{\partial w}\right)\mathrm{vol} = \int_{\partial W}\sqrt{g}(a\,dv\,dw + b\,dw\,du + c\,du\,dv). \tag{4.25}$$

Here the volume form is $\sqrt{g}\,du\,dv\,dw$, where the particular form of $\sqrt{g}$ is that associated with the u, v, w coordinate system. Again the coefficients a, b, c of the vector field are expressed in terms of the coordinate basis vectors $\partial/\partial u, \partial/\partial v, \partial/\partial w$. This is not the the only possible kind of basis for a vector field, so in some treatments the formulas will appear differently. They will be ultimately equivalent in terms of their geometrical meaning. This result involves a kind of surface integral, but the starting point is a vector field instead of a 2-form.

The divergence theorem says that the integral of the divergence of a vector field over W with respect to the volume is the integral of the flux of the vector field across the bounding surface ∂W. A famous application in physics is when the vector field represents the electric field, and the divergence represents the density of charge. So the amount of charge in the region determines the flux of the electric field through the boundary.

An important identity for a differential form ω and a scalar field d is

$$d(s\omega) = ds \wedge \omega + s\,d\omega. \tag{4.26}$$

This gives an *integration by parts* formula

$$\int_W ds \wedge \omega + \int_W s\,d\omega = \int_{\partial W} s\omega. \tag{4.27}$$

Apply this to $\omega = \mathbf{u}\rfloor\mathrm{vol}$ and use $ds \wedge \mathbf{u}\rfloor\mathrm{vol} = \langle ds \mid \mathbf{u}\rangle\,\mathrm{vol}$. This gives the divergence identity

$$\mathrm{div}\ (s\mathbf{u}) = ds \cdot \mathbf{u} + s\,\mathrm{div}\ \mathbf{u}. \tag{4.28}$$

From this we get another important integration by parts identity

$$\int_W \langle ds \mid \mathbf{u}\rangle\,\mathrm{vol} + \int_W s\,\mathrm{div}\ \mathbf{u}\,\mathrm{vol} = \int_{\partial W} s\,\mathbf{u}\rfloor\mathrm{vol}. \tag{4.29}$$

4.4 Conservation laws

A conservation law *conservation law* is an assertion that the total amount of some quantity (such as mass) does not change. A *continuity equation* is a local form of a conservation law. It says that any change in the amount of the quantity

in some region is compensated by flow across the boundary of the region. In mathematical form it is expressed by an equation of the form

$$\frac{\partial R}{\partial t} + dJ = 0. \tag{4.30}$$

Here R is an n-form, the mass in kilograms, and J is an $n-1$ form, the mass flux in kilograms per second). The coefficients of these two forms have units kilograms per cubic meter and kilograms per second per square meter.) The integral form of such a conservation law is

$$\frac{d}{dt}\int_W R + \int_{\partial W} J = 0. \tag{4.31}$$

It says that the rate of change of the amount of substance inside the region W plus the net (outward minus inward) flow through the boundary ∂W is zero. Thus, for instance, the amount in the region can only decrease if there is a compensating outward flow. In fluid dynamics the flux J of mass is $J = \mathbf{v}\rfloor R$, where $\mathbf{v}$ is the fluid velocity vector field. Since the coefficients of $\mathbf{v}$ are in meters per second, and the basis vectors are in inverse meters, the units of $\mathbf{v}$ itself is in inverse seconds.

Often one writes

$$R = \rho\,\mathrm{vol} \tag{4.32}$$

Here the coefficient ρ is a scalar density (in kilograms per cubic meter). In this case the conservation law reads

$$\frac{\partial \rho}{\partial t} + \mathrm{div}\ (\rho\mathbf{v}) = 0. \tag{4.33}$$

The corresponding integral form is

$$\frac{d}{dt}\int_W \rho\,\mathrm{vol} + \int_{\partial W} \rho\mathbf{v}\rfloor\mathrm{vol} = 0. \tag{4.34}$$

The units for this equation are kilograms per second. For a fluid it is the law of conservation of mass. The theory also applies when $\mathbf{v}$ is a time-dependent vector field.

The conservation law also appears in the equivalent form

$$\frac{\partial \rho}{\partial t} + \mathbf{v}\rfloor d\rho + \mathrm{div}\ \mathbf{v}\,\rho = 0 \tag{4.35}$$

or, more concretely,

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{n} v_i \frac{\partial \rho}{\partial x_i} = -\mathrm{div}\ \mathbf{v}\,\rho. \tag{4.36}$$

Here v_i is the component of velocity in the direction corresponding to x_i. The left hand side is a material derivative, that is, the change in density following a typical particle. This suggests a method of solving the differential equation by

integrating the vector field to find particle trajectories. The change in density along a particle trajectory is driven by the right hand side, which is a measure of how fast particles are compressing together.

A solution for this conservation law describes how ρ at $t = t_0$ determines ρ for later $t > t_0$. The solution method uses the space-time curves $c_{t'}$ for $t_0|leqt' \leq t$ that define the particle trajectories. Fix a space-time point given by $\mathbf{x}$ and t. The curve is required to pass through this point, that is, the particle will reach $\mathbf{x}$ at time t. Then $c_{t'}$ describes where it came from, that is, its location in space-time at time $t' \leq t$. The space coordinates of $c_{t'}$ are $\mathbf{x} \circ c_{t'}$. The time component of $c_{t'}$ is just t'. The curve solves the ordinary differential equations

$$\frac{dx_i \circ c_{t'}}{dt'} = v_i \circ c_{t'}. \tag{4.37}$$

This says that the particle moves according to the fluid velocity.

From the chain rule and the ordinary differential equation for the particle trajectory we get

$$\frac{d\rho \circ c_{t'}}{dt'} = \frac{\partial \rho}{\partial t} \circ c_{t'} + \sum_{i=1}^{n} \frac{dx_i \circ c_{t'}}{dt'} \frac{\partial \rho}{\partial x_i} \circ c_{t'} = \frac{\partial \rho}{\partial t} \circ c_{t'} + \left(\sum_{i=1}^{n} v_i \frac{\partial \rho}{\partial x_i} \right) \circ c_{t'}. \tag{4.38}$$

Using the partial differential equation, we obtain an ordinary differential equation for the density along the particle trajectory in the form

$$\frac{d\rho \circ c_{t'}}{dt'} = -\text{div } \mathbf{v} \circ c_{t'}\, \rho \circ c_{t'}. \tag{4.39}$$

Suppose that we know the restriction ρ_0 of ρ to $t = t_0$. Solving the ordinary differential equation leads to

$$\rho = \exp(-\int_{t_0}^{t} \text{div } \mathbf{v} \circ c_{t'}\, dt')\, \rho_0 \circ c_{t_0} \tag{4.40}$$

This is the solution for ρ at any point given by $\mathbf{x}, t$. The particles at time t have been transported from where they were at time t_0. However they may be more (or less) spread out because they are diverging (or converging) geometrically, and this will decrease (or increase) the density.

The object of interest is really the differential form $R = \rho\,\text{vol}$ that gives the mass in a given region. For each t' there is a map $c_{t'}$ from the time t space to the time t' space. The solution says that R is the pullback under c_{t_0} of $R_0 = \rho_0\,\text{vol}$. The exponential factor in ρ comes from pulling back the volume form. All that happens is that mass is transported to new locations.

Example: Consider an exploding star, at a scale at which the original star is a point. If the point is the origin, and the explosion was at time zero, then the particles found at $\mathbf{x}, t$ will have travelled a distance $\mathbf{x}$ in time t and will then have (constant) velocity $\mathbf{v} = \mathbf{x}/t$. If we look at particles at the same time but at greater or lesser distance, then they will not be the same particles, so they will have correspondingly greater or lesser velocities.

The explosion need not be symmetric; the profile at time $t > 0$ is $\rho = f(\mathbf{x}, t)$. The volume form is $d^n\mathbf{x} = dx_1 \cdots dx_n$, and so the divergence of the vector field is div $\mathbf{v} = n/t$. The exponential factor is thus $(t_0/t)^n$. (For the exploding star example the dimension of space is $n = 3$, but it is just as easy to do the calculation for arbitrary dimension.) For fixed $\mathbf{x}$ and t the solution of the ordinary differential equation $d\mathbf{x}'/dt' = \mathbf{x}'/t'$ is $\mathbf{x}' = (\mathbf{x}/t)t'$. The space component of the solution curve $c_{t'}$ is therefore $\mathbf{x} \circ c_{t'} = (\mathbf{x}/t)t'$. So if $\rho_0 = f(\mathbf{x}, t_0)$, then $\rho_0 \circ c_{t_0} = f((\mathbf{x}/t)t_0, t_0)$. The solution to the partial differential equation is

$$\rho = f(\mathbf{x}, t) = \left(\frac{t_0}{t}\right)^n f(t_0 \frac{\mathbf{x}}{t}, t_0). \tag{4.41}$$

It consists of a profile that is the same for all particles with the same constant velocity $\mathbf{x}/t$, modified by a factor that says that they are flying apart.

For such an explosion the mass form $R = \rho\, d^n\mathbf{x}$ is the pullback under the map $\mathbf{x} \leftarrow (t_0/t)\mathbf{x}$ of $R_0 = f(\mathbf{x}, t_0)\, d^n\mathbf{x}$. This leads to an equivalent solution formula $R = f((t_0/t)\mathbf{x}, t_0)\, d^n(t_0/t)\mathbf{x} = f((t_0/t)\mathbf{x}, t_0)(t_0/t)^n\, d^n\mathbf{x}$. |

4.5 The metric

In geometry one measures distance in a straight line using the Euclidean distance. In coordinates this is computed by using the theorem of Pythagoras. However with some calculus one can also measure distance along a curve. It is even possible to do this in arbitrary coordinate systems. This motivates a very general notion of squared distance, given by the *metric tensor*

$$\mathsf{g} = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}\, dx_i\, dx_j. \tag{4.42}$$

Here the product $dx_i\, dx_j$ is not the anti-symmetric exterior project, but instead is a symmetric tensor product. The metric tensor is determined in a particular coordinate system by functions g_{ij} forming the matrix of coefficients. This matrix is required to be symmetric and positive definite. The distance along a regular parameterized curve C is then given by

$$\text{distance} = \int_C \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}\, dx_i\, dx_j}. \tag{4.43}$$

This can be computed via a parametrization of C via

$$\text{distance} = \int_{t_0}^{t_1} \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} \frac{dx_i}{dt}\frac{dx_j}{dt}}\, dt. \tag{4.44}$$

Because of the square root the integrals can be nasty. In this expression it is helpful to think of $\frac{dx_i}{dt}$ as the components of the *velocity*. The square root

expression is then interpreted as the *speed* $\frac{ds}{dt}$. So the computation of *arc length* of the curve comes down to integrating ds.

If we have the metric tensor in one set of coordinates, then we have it in any other set of coordinates. Thus

$$\mathsf{g} = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{\alpha=1}^{n}\frac{\partial x_i}{\partial u_\alpha}\,du_\alpha\right) g_{ij}\left(\sum_{\beta=1}^{n}\frac{\partial x_j}{\partial u_\beta}\,du_\beta\right) = \sum_{\alpha=1}^{n}\sum_{\beta=1}^{n}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial x_i}{\partial u_\alpha}g_{ij}\frac{\partial x_j}{\partial u_\beta}\right) du_\alpha\,du_\beta. \tag{4.45}$$

So the matrix in the new coordinates is

$$g_{\alpha\beta} = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial x_i}{\partial u_\alpha}g_{ij}\frac{\partial x_j}{\partial u_\beta}. \tag{4.46}$$

The case $g_{ij} = \delta_{ij}$ is the special case when the x_i are Cartesian coordinates. In this case

$$\mathsf{g} = \sum_{j=1}^{n} dx_i^2. \tag{4.47}$$

However even if we start with Cartesian coordinates, the coefficients of the metric tensor takes a more complicated form in other coordinate systems. If the coefficients come from Cartesian coordinates via a change of coordinates, then the metric is said to be a *flat metric*.

A familiar example of a flat metric is the metric in three dimensional Euclidean space. Then we have

$$\mathsf{g} = dx^2 + dy^2 + dz^2 = dr^2 + r^2\,d\theta^2 + r^2\sin^2(\theta)\,d\phi^2. \tag{4.48}$$

The latter is in spherical polar coordinates.

The notion of metric tensor goes at least part way to erasing the distinction between differential 1-forms and vector fields. It provides a correspondence, one that is sometimes awkward, but that always exists. Let g_{ij} be the coefficients of the metric tensor for the coordinate system $u_1, \ldots, u_n$. If

$$X = \sum_{k=1}^{n} a_k\frac{\partial}{\partial u_k} \tag{4.49}$$

is a vector field, then the associated differential 1-form is

$$\mathsf{g}X = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_i g_{ij}\right) du_j. \tag{4.50}$$

One can also go the other direction. If

$$\omega = \sum_{k=1}^{n} p_k\,du_k \tag{4.51}$$

is a differential 1-form, then the associated vector field is

$$\mathsf{g}^{-1}\omega = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} g^{ij} p_j \right) \frac{\partial}{\partial u_i}. \tag{4.52}$$

Here we are using the perhaps unfamiliar notation that g^{ij} is the inverse matrix to g_{ij}. (This notation is standard in this context.)

Another quantity associated with the metric tensor g is the volume form

$$\text{vol} = \sqrt{g}\, du_1 \cdots du_n. \tag{4.53}$$

Here g denotes the determinant of the matrix g_{ij}. (This notation is also standard.) The interpretation of this as volume is left to a later section.

There is a very important construction that produces new metrics. Suppose that the n dimensional space has coordinates $x_1, \ldots, x_n$, and there is a k-dimensional regular parametrized surface with coordinate $u_1, \ldots, u_k$. Start with the metric

$$\mathsf{g} = \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij}\, dx_i\, dx_j. \tag{4.54}$$

The pullback of the metric to the surface is

$$\mathsf{g}^* = \sum_{\alpha=1}^{k} \sum_{\beta=1}^{k} g^*_{\alpha\beta}\, du_\alpha\, du_\beta. \tag{4.55}$$

Here

$$g^*_{\alpha\beta} = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial x_i}{\partial u_\alpha} g_{ij} \frac{\partial x_j}{\partial u_\beta}. \tag{4.56}$$

A simple example is the pullback of the Euclidean metric given above to the sphere $x^2 + y^2 + z^2 = a^2$. The metric pulls back to

$$\mathsf{g}^* = a^2\, d\theta^2 + a^2 \sin^2(\theta)\, d\phi^2. \tag{4.57}$$

This is not a flat metric. Even if one only considers a small open subset of the sphere, it is still impossible to find coordinates u, v such that $\mathsf{g}^* = du^2 + dv^2$.

Remark: The words *tensor* and *tensor field* can refer to a number of kinds of objects. Strictly speaking, a tensor is defined at a particular point, and a tensor field is a function that assigns to each point a tensor at that point. More precisely, a tensor of type (p, q) at a point is a real multi-linear function whose inputs consist of q vectors in the tangent space at the point and p vectors in the dual space to the tangent space at the point (covectors). When $p = 0$ and all the inputs are vectors it is called a *covariant tensor*. . When $q = 0$ and all the inputs are covectors it is called a *contravariant tensor*. When both kinds of vectors are allowed as inputs, it is called a *mixed tensor*. A tensor field assigns (in a smooth way) to each point in a manifold patch a corresponding tensor at

that point. People are often careless and use the word tensor to mean tensor field.

The most basic tensor fields are scalar fields of type (0,0), vector fields of type (1,0), and differential 1-forms of type (0,1) There are also more complicated tensor fields. A differential k-form assigns to each point a real multi-linear function on k-tuples of tangent vectors at the point, so it is of type $(0, k)$. A metric tensor field assigns to each pair an inner product on tangent vectors at the point, so it is of type $(0, 2)$. For the more complicated tensors one can also impose symmetry conditions. Thus one distinguishes between the *anti-symmetric tensor* case (differential k-forms) and the *symmetric tensor* case (the metric tensor). The metric tensor is required not only to be symmetric, but also positive definite. The inverse of the metric tensor is a tensor of type $(2, 0)$; it is also symmetric and positive definite.

The only example in these lecture of a mixed tensor is a (1,1) tensor, that is, a linear transformation. An example is the linear transformation associated with a vector field at a zero. This should be distinguished from the quadratic form associated with a scalar field at a critical point, which is a symmetric covariant tensor of type (0,2).

The study of tensors at a point is called *tensor algebra*, while the study of tensor fields is *tensor calculus*. The metric tensor field provides a particularly rich ground to explore. A choice of metric tensor field is the beginning of a subject called *Riemannian geometry*. The metric tensor field and related objects are fundamental to Einstein's general relativity. |

4.6 Twisted forms

There is a variation on the idea of differential form that comes up in various contexts, including discussions of volume. The new kind of object is known as a *twisted form* or a *pseudoform*. In the following we consider the notion of twisted form in top dimension. The main ingredient is that when an integral involving the twisted form is expressed in terms of a new coordinate system, the expression for the new integral involves the absolute value of the Jacobian determinant.

Suppose that P is a k-dimensional manifold patch. Consider a twisted differential k-form Ω on M. Also consider a Jordan measurable compact set $K \subseteq P$. We wish to define the integral of Ω over K. To do this, choose a coordinate system $u_1, \ldots, u_k$. Write

$$\Omega = f(u_1, \ldots, u_k)\, |\, du_1 \wedge \cdots \wedge\, du_k|. \tag{4.58}$$

There is a set $B \subseteq \mathbf{R}^n$ such that $\mathbf{u}(K) = B$. The definition is

$$\int_K \Omega = \int_B f(v_1, \ldots, v_n)\, dv_1 \cdots\, dv_k = I_B(f). \tag{4.59}$$

where $I_B(f)$ is the usual (unoriented) Riemann integral. (The variables $v_1, \ldots, v_n$ are just symbols that are available to define the function f.)

To show that this is well-defined, consider another coordinate system $y_1, \ldots, y_k$. Then $\mathbf{u} = \mathbf{g}(\mathbf{y})$. Also, $B = \mathbf{g}(A)$ for some Jordan measurable set A. The twisted differential form

$$\Omega = f(\mathbf{u})\,|\,du_1 \wedge \cdots \wedge\, du_k| = f(\mathbf{g}(\mathbf{y}))|\det \mathbf{g}'(\mathbf{y})|\,|\,dy_1 \wedge \cdots \wedge dy_k| \tag{4.60}$$

may be expressed in either coordinate system. The definition in the other coordinate system gives

$$\int_K \Omega = \int_K f(\mathbf{g}(\mathbf{y}))|\det \mathbf{g}'(\mathbf{y})|\,|\,dy_1 \wedge \cdots \wedge dy_k| = I_A((f \circ \mathbf{g})|\det \mathbf{g}'|). \tag{4.61}$$

But $I_B(f) = I_A((f \circ \mathbf{g})|\det \mathbf{g}'|)$ by the change of variables theorem. So the definitions using different coordinate systems are consistent.

The twisted form concept does not depend on orientation. Furthermore, it makes sense to say that a twisted form is positive. For this reason, it is natural to interpret the volume form $\text{Vol} = \sqrt{g}\,|\,du_1 \wedge \cdots \wedge du_n|$ as a twisted form instead of as an ordinary form. The same applies in other applications, such as when the form represents a mass density or a probability density.

The twisted property for the volume may be seen as follows. Suppose that the matrix for the metric tensor in the $\mathbf{u}$ coordinate system is G. Then the matrix in the $\mathbf{y}$ coordinate system is $\mathbf{g}'(\mathbf{y})^T G \mathbf{g}'(\mathbf{y})$. If the determinant in the $\mathbf{u}$ system is $g = \det G$, then the determinant in the $\mathbf{y}$ coordinate system is $g(\det \mathbf{g}'(\mathbf{y}))^2$. The volume may thus be represented in either coordinate system, with

$$\text{Vol} = \sqrt{g}\,|\,du_1 \wedge \cdots \wedge\, du_n| = \sqrt{g}\,|\det \mathbf{g}'(\mathbf{y})|\,|\,dy_1 \wedge \cdots \wedge\, dy_n|. \tag{4.62}$$

This is exactly what one would expect from a twisted form.

It may seem to be a nuisance to have two kinds of differential forms, the usual ones and the twisted ones. When we are integrating over an oriented region the distinction is subtle, and it may be easier to deal with usual differential forms. For example, we can consider the volume form as a usual differential form, but remember to choose an orientation so that the integral of the volume form $\text{vol} = \sqrt{g}\,dx_1 \wedge \cdots\, dx_n$ over the region is positive. For more information about twisted forms, see the book of Burke [4].

4.7 The gradient and divergence and the Laplace operator

The *gradient* of a scalar is the vector field

$$\nabla s = \text{grad}\ s = \mathbf{g}^{-1}\, ds. \tag{4.63}$$

In coordinates the gradient has the form

$$\nabla s = \text{grad}\ s = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} g^{jk} \frac{\partial s}{\partial u_k} \right) \frac{\partial}{\partial u_j} \tag{4.64}$$

The *Laplacian* of s is the divergence of the gradient. Thus

$$\nabla^2 s \,\mathrm{vol} = \mathrm{div}\ \ \mathrm{grad}\ \ s \,\mathrm{vol} = d(\nabla\, s \rfloor \mathrm{vol}). \tag{4.65}$$

In coordinates this is

$$\nabla^2\, s = \frac{1}{\sqrt{g}} \sum_{i=1}^{n} \frac{\partial}{\partial u_i} \left(\sqrt{g} \sum_{k=1}^{n} g^{ik} \frac{\partial s}{\partial u_k} \right). \tag{4.66}$$

Theorem 4.2 (Green's first identity) *If s and u are scalars defined in the bounded region Ω, then*

$$\int_\Omega s\, \nabla^2 u \,\mathrm{vol} + \int_\Omega \nabla s\, \mathsf{g}\, \nabla u \,\mathrm{vol} = \int_{\partial\Omega} s\, \nabla u \rfloor \mathrm{vol}. \tag{4.67}$$

Proof: This is just integration by parts. By the product rule we have

$$d(s\, \nabla u \rfloor \,\mathrm{vol}) = ds\, \nabla u \rfloor \,\mathrm{vol} + s\, d(\nabla u \rfloor \,\mathrm{vol}). \tag{4.68}$$

However $ds \nabla u \rfloor \,\mathrm{vol} = \langle ds \mid \nabla u \rangle \mathrm{vol} = \nabla s\, \mathsf{g}\, \nabla u \,\mathrm{vol}$. So this is

$$d(s\, \nabla u \rfloor \,\mathrm{vol}) = \nabla s\, \mathsf{g}\, \nabla u \,\mathrm{vol} + s\, \nabla^2 u \,\mathrm{vol}. \tag{4.69}$$

Integrate and use Stokes' theorem. □

This identity is often used in cases when either s vanishes on $\partial\Omega$ or ∇u vanishes on $\partial\Omega$, In that case it says that

$$-\int_\Omega s\, \nabla^2 u \,\mathrm{vol} = \int_\Omega \nabla s\, \mathsf{g}\, \nabla u \,\mathrm{vol}. \tag{4.70}$$

In particular,

$$-\int_\Omega u\, \nabla^2 u \,\mathrm{vol} = \int_\Omega \nabla u\, \mathsf{g}\, \nabla u \,\mathrm{vol} \geq 0. \tag{4.71}$$

This suggests that in some sense $-\nabla^2$ is a positive operator.

The remaining objects are in three dimensions. The *cross product* of two vectors $\mathbf{v}$ and $\mathbf{w}$ is defined as the unique vector $\mathbf{v} \times \mathbf{w}$ such that

$$(\mathbf{v} \times \mathbf{w}) \rfloor \mathrm{vol} = \mathsf{g}\mathbf{v} \wedge \mathsf{g}\mathbf{w}. \tag{4.72}$$

In other words, it is the operation on vectors that corresponds to the exterior product on forms. The *curl* of a vector field $\mathbf{v}$ is defined by

$$(\mathrm{curl}\ \mathbf{v}) \rfloor \mathrm{vol} = d(\mathsf{g}\mathbf{v}). \tag{4.73}$$

It is easy to see that curl grad $f = 0$ and that div curl $\mathbf{v} = 0$. In this language Stokes's theorem says that

$$\int_S \mathrm{curl}\ \mathbf{v} \rfloor \mathrm{vol} = \int_{\partial S} \mathsf{g}\mathbf{v}. \tag{4.74}$$

4.8 Orthogonal coordinates and normalized bases

In many applications it is possible to use orthogonal coordinates to simplify the calculations. This section presents some of the common formulas for this case. It is mainly for reference and for comparison with other treatments. All of the formulas that follow are a consequence of the general theory. It is very convenient to choose coordinates so that the Riemannian metric is diagonal with respect to this coordinate system. Such a coordinate system is called a system of *orthogonal coordinates*. In this case it has the form

$$\mathsf{g} = h_1^2\, du_1^2 + h_2^2\, du_2^2 + \cdots + h_n^2\, du_n^2. \tag{4.75}$$

Here each coefficient h_i is a function of the coordinates $u_1, \ldots, u_n$.

Consider a manifold with a given Riemannian metric. For instance, it could be a k dimensional surface in some Euclidean space of larger dimension n. If the manifold has dimension at most three, then near every point there is always a new coordinate system that is a system of orthogonal coordinates. In the case of three dimensions this is not a particularly obvious fact, but it is a consequence of the Cartan-Kähler theorem. There is a treatment in the book of Bryant and coauthors [3].

When we have orthogonal coordinates, it is tempting to make the basis vectors have length one. Thus instead of using the usual coordinate basis vectors $\frac{\partial}{\partial u_i}$ one uses the normalized basis vectors $\frac{1}{h_i}\frac{\partial}{\partial u_i}$. Similarly, instead of using the usual coordinate differential forms du_i one uses the normalized differentials $h_i\, du_i$. Then

$$\mathsf{g}\left(a_1 \frac{1}{h_1}\frac{\partial}{\partial u_1} + \cdots a_n \frac{1}{h_n}\frac{\partial}{\partial u_n}\right) = a_1 h_1\, du_1 + \cdots a_n h_n\, du_n \tag{4.76}$$

With normalized basis vectors the coefficients of vector fields and the corresponding differential forms are the same. This makes it very easy to confuse vector fields with differential forms.

In orthogonal coordinates the volume is given in terms of normalized differentials by

$$\mathrm{vol} = h_1\, du_1 \wedge \cdots \wedge h_n\, du_n. \tag{4.77}$$

A simple example of orthogonal coordinates is that of *polar coordinates* r, ϕ in the plane. These are related to Cartesian coordinates x, y by

$$x = r\cos(\phi) \tag{4.78}$$

$$y = r\sin(\phi) \tag{4.79}$$

The Riemannian metric is expressed as

$$\mathsf{g} = dr^2 + r^2\, d\phi^2. \tag{4.80}$$

The normalized basis vectors are $\frac{\partial}{\partial r}$ and $\frac{1}{r}\frac{\partial}{\partial \phi}$. The normalized basis forms are dr and $r\, d\phi$. The area form is $r\, dr \wedge d\phi$. Warning: Even though coordinate

forms like $d\phi$ are closed forms, a normalized form like $r\,d\phi$ need not be a closed form. In fact, in this particular case $d(r\,d\phi) = dr \wedge d\phi \neq 0$.

Another example of orthogonal coordinates is that of *spherical polar coordinates* r, θ, ϕ. These are related to Cartesian coordinates x, y, z by

$$x = r\cos(\phi)\sin(\theta) \tag{4.81}$$
$$y = r\sin(\phi)\sin(\theta) \tag{4.82}$$
$$z = r\cos(\theta) \tag{4.83}$$

The Riemannian metric is expressed as

$$\mathsf{g} = dr^2 + r^2\,d\theta^2 + r^2\sin^2(\theta)\,d\phi^2. \tag{4.84}$$

The normalized basis vectors are $\frac{\partial}{\partial r}$ and $\frac{1}{r}\frac{\partial}{\partial \theta}$ and $\frac{1}{r\sin(\theta)}\frac{\partial}{\partial \phi}$. The normalized basis forms are dr and $r\,d\theta$ and $r\sin(\theta)\,d\phi$. The volume form is $r^2\sin(\theta)\,dr \wedge d\theta \wedge d\phi$.

If f is a scalar field, then its gradient is

$$\nabla f = \sum_{i=1}^{n} \frac{1}{h_i}\frac{\partial f}{\partial u_i}\frac{1}{h_i}\frac{\partial}{\partial u_i}. \tag{4.85}$$

If $\mathbf{u}$ is a vector field, then its divergence $\nabla \cdot \mathbf{u}$ is a scalar field. Say that $\mathbf{u}$ has an expression in terms of normalized basis vectors of the form

$$\mathbf{u} = \sum_{i=1}^{n} a_i \frac{1}{h_i}\frac{\partial}{\partial u_i}. \tag{4.86}$$

Then

$$\text{div } \mathbf{u} = \nabla \cdot \mathbf{u} = \sum_{i=1}^{n} \frac{1}{h_1 \cdots h_n}\frac{\partial}{\partial u_i}\left(\frac{h_1 \cdots h_n}{h_i} a_i\right). \tag{4.87}$$

In coordinates the Laplacian has the form

$$\nabla^2 f = \frac{1}{h_1 \cdots h_n}\sum_{i=1}^{n} \frac{\partial}{\partial u_i}\left(\frac{h_1 \cdots h_n}{h_i^2}\frac{\partial f}{\partial u_i}\right) \tag{4.88}$$

For example, in three dimensions with Cartesian coordinates it is

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}. \tag{4.89}$$

In spherical polar coordinates it is usually written

$$\nabla^2 f = \frac{1}{r^2}\frac{\partial}{\partial r} r^2 \frac{\partial f}{\partial r} + \frac{1}{r^2}\left[\frac{1}{\sin(\theta)}\frac{\partial}{\partial \theta}\sin(\theta)\frac{\partial f}{\partial \theta} + \frac{1}{\sin^2(\theta)}\frac{\partial^2 f}{\partial \phi^2}\right]. \tag{4.90}$$

We conclude by recording the explicit form of the divergence theorem and Stokes' theorem in the context of orthogonal coordinates. The first topic is

the divergence theorem in two dimensions. Say that the vector field $\mathbf{v}$ has an expression in terms of normalized basis vectors of the form

$$\mathbf{v} = a\frac{1}{h_u}\frac{\partial}{\partial u} + b\frac{1}{h_v}\frac{\partial}{\partial v}. \tag{4.91}$$

Recall that the area form is

$$\text{area} = h_u h_v \, du \, dv. \tag{4.92}$$

Then the corresponding differential 2-form is

$$\mathbf{v} \rfloor \text{area} = a h_v \, dv - b h_u \, du. \tag{4.93}$$

The divergence theorem in two dimensions is obtained by applying Greens's theorem for 1-forms to this particular 1-form. The result is

$$\int_R \left[\frac{1}{h_u h_v} \left(\frac{\partial}{\partial u}(h_v a) + \frac{\partial}{\partial v}(h_u b) \right) \right] h_u h_v \, du \, dv = \int_{\partial R} a h_v \, dv - b h_u \, du. \tag{4.94}$$

The expression in brackets on the left is the divergence of the vector field. On the right the integrand measures the amount of the vector field crossing normal to the curve.

The next topic is the divergence theorem in three dimensions. Say that the vector field $\mathbf{v}$ has an expression in terms of normalized basis vectors of the form

$$\mathbf{v} = a\frac{1}{h_u}\frac{\partial}{\partial u} + b\frac{1}{h_v}\frac{\partial}{\partial v} + c\frac{1}{h_w}\frac{\partial}{\partial w}. \tag{4.95}$$

Recall that the volume form is

$$\text{vol} = h_u h_v h_w \, du \, dv \, dw. \tag{4.96}$$

Then the corresponding differential 2-form is

$$\mathbf{v} \rfloor \text{vol} = a h_v h_w \, dv \, dw + b h_w h_u \, dw \, du + c h_u h_v \, du \, dv. \tag{4.97}$$

The divergence theorem in three dimensions is obtained by applying Gauss's theorem for 2-forms to this particular 2-form. The result is

$$\int_V \left[\frac{1}{h_u h_v h_w} \left(\frac{\partial}{\partial u}(h_v h_w a) + \frac{\partial}{\partial v}(h_w h_u b) + \frac{\partial}{\partial w}(h_u h_v c) \right) \right] h_u h_v h_w \, du \, dv \, dw \quad = \int_{\partial V} a h_v h_w \, dv \, dw + b h_w h_u \, dw \, du + c h_u h_v \, du \, dv. \tag{4.98}$$

The expression in brackets is the divergence of the vector field.

The final topic is the classical Stokes's theorem in three dimensions. Say that the vector field $\mathbf{v}$ has an expression in terms of normalized basis vectors as above. There is a corresponding differential 1-form

$$\mathbf{g}\mathbf{v} = a h_u \, du + b h_v \, dv + c h_w \, dw. \tag{4.99}$$

The classical Stokes's theorem in three dimensions is obtained by applying Stokes's theorem for 1-forms to this particular 1-form. This gives on the left hand side

$$\int_S \left[\frac{1}{h_v h_w}\left(\frac{\partial h_w c}{\partial v} - \frac{\partial h_v b}{\partial w}\right)\right] h_v h_w \, dv\, dw + \left[\frac{1}{h_u h_w}\left(\frac{\partial h_u a}{\partial w} - \frac{\partial h_w c}{\partial u}\right)\right] h_w h_u \, dw\, du \quad + \left[\frac{1}{h_u h_v}\left(\frac{\partial h_v b}{\partial u} - \frac{\partial h_u a}{\partial v}\right)\right] h_u h_v \, du\, dv \tag{4.100}$$

and on the right hand side

$$\int_{\partial S} a h_u \, du + b h_v \, dv + c h_w \, dw. \tag{4.101}$$

The terms in square brackets are the components of the curl of the vector field expressed in terms of normalized basis vectors.

4.9 Linear algebra (the Levi-Civita permutation symbol)

There is a powerful algebraic method to get results in linear and multi-linear algebra. This is the use of tensor algebra and the Levi-Civita symbol. This method is usually terrible for numerical calculation, and it is not very useful for giving geometric insight. The advantage is that it often produces an answer by a straightforward calculation.

It is convenient to use certain algebraic conventions for manipulating tensor coefficients. The common practice is to use lower indices for coefficients of covariant tensors and upper indices for coefficients of contravariant tensors. Repeated upper and lower indices indicate summation. (This is often called the Einstein summation convention.) Such a summation is also called a *contraction*. Thus if A is a matrix a^i_j and B is a matrix b^h_k, then the product $C = AB$ is the matrix $c^i_k = a^i_j b^j_k$ obtained by contraction. The trace $\mathrm{tr}(C) = \mathrm{tr}(AB) = \mathrm{tr}(BA)$ of the matrix is the result of a second contraction, that is, $c^i_i = a^i_j b^j_i$.

The *Levi-Civita permutation symbol* may be written in two forms: $\epsilon^{j_1 \cdots j_n} = \epsilon_{j_1 \cdots j_n}$. By definition this is equal to 0 if there are repeated indices, equal to 1 if the indices form an even permutation of $1, \ldots, n$, and equal to -1 if the indices form an odd permutation of $1, \ldots, n$. The advantage of the Levi-Civita symbol is that the sums are over all values of the indices; the symbol itself enforces the permutation condition. This is particularly useful in dealing with determinants. In this section we state some properties of determinants using the Levi-Civita symbol. Proofs are indicated in the problems. The following section gives applications of the Levi-Civita symbol to the formulation of various expressions for volume and area.

A general definition of determinant is

$$\det(A) = \epsilon^{j_1 \cdots j_n} a^1_{j_1} \cdots a^n_{j_n}. \tag{4.102}$$

This says that the determinant is a sum of products, each product having coefficient 0, 1, or -1. There are n^n such products, most of them equal to zero. Each product with a non-zero coefficient corresponds to picking a distinct element from each row and multiplying them together. The number of products with non-zero coefficient is $n!$, which is still a very large number for computational purposes.

The determinant formula above depends on the fact that the rows are taken in order $1, \ldots, n$. If we instead take them in the order $i_1, \ldots, i_n$ we get

$$\epsilon^{j_1 \cdots j_n} a^{i_1}_{j_1} \cdots a^{i_n}_{j_n} = \epsilon^{i_1 \cdots i_n} \det(A). \tag{4.103}$$

There is an even more complicated but considerably more symmetric formula for the determinant:

$$\det(A) = \frac{1}{n!} \epsilon_{i_1 \cdots i_n} \epsilon^{j_1 \cdots j_n} a^{i_1}_{j_1} \cdots a^{i_n}_{j_n}. \tag{4.104}$$

This formula leads to particularly straightforward derivations of identities such as $\det(A) = \det(A^T)$ and $\det(AB) = \det(A)\det(B)$.

Cramer's rule is a formula for the inverse of a matrix given in terms of determinants. If a^i_j is a matrix, define its *cofactor* matrix to be

$$C^j_k = \frac{1}{(n-1)!} \epsilon_{k\, i_2 \cdots i_n} \epsilon^{j\, j_2 \cdots j_n} a^{i_2}_{j_2} \cdots a^{i_n}_{j_n}. \tag{4.105}$$

(This is actually the transpose of the usual matrix of cofactors.) In tensor algebra language Cramer's rule may be stated as

$$a^i_j C^j_k = \delta^i_k \det(A). \tag{4.106}$$

In the matrix version this says that $AC = \det(A) I$. Cramer's rule thus has the succinct statement $A^{-1} = (1/\det(A))C$. While Cramer's rule is quite impractical for numerical calculation, it does give considerable insight into the structure of the inverse matrix.

4.10 Linear algebra (volume and area)

Consider vectors X_α in $\mathbf{R}^n$ for $\alpha = 1, \ldots, n$. The convex combinations of these vectors together with the zero vector form a solid, a kind of n-dimensional parallelepiped. There are two formulas for the volume of this solid. The Euclidean volume is

$$n\text{−volume}(X) = |\det X|. \tag{4.107}$$

This is the absolute value of the determinant of the entries in the vectors. There is an alternative formula that looks quite different, but is equivalent. This is

$$n\text{−volume}(X) = \sqrt{\det X^T X}. \tag{4.108}$$

The matrix $X^T X$ is the matrix of pairwise Euclidean inner products of the vectors. Sometimes this is called a *Gram matrix*.

One can generalize this to the case when there is an inner product on $\mathbf{R}^n$ given by an n by n positive definite matrix G. In that case we have

$$n-\text{volume}_G(X) = \sqrt{\det X^T G X} = \sqrt{g}\,|\det X|. \tag{4.109}$$

Here $X^T G X$ is the matrix of the pairwise inner products of vectors, a generalization of the Gram matrix. Also $g = \det G$ is the determinant of G.

Now consider vectors X_α in $\mathbf{R}^n$ for $\alpha = 1, \ldots, k$. The convex combinations of these vectors together with the zero vector form k-dimensional parallelepiped. The matrix is no longer square. It is natural to define the k dimensional area as

$$k-\text{area}_G(X) = \sqrt{\det X^T G X} \tag{4.110}$$

The Gram matrix $X^T G X$ is again the matrix of pairwise inner products of the vectors.

The following result is a generalization of the Cauchy-Binet theorem.

Theorem 4.3 *Let X be an n by k matrix. Let G be an n by n symmetric matrix. For each sequence $i_1, \ldots, i_k$ of rows of X,*

$$J^{i_1 \cdots i_k} = \epsilon^{\alpha_1 \cdots \alpha_k} X^{i_1}_{\alpha_1} \cdots X^{i_k}_{\alpha_k}. \tag{4.111}$$

represent the determinant of the corresponding k by k minor obtained by retaining only those rows. Then

$$\det(X^T G X) = \frac{1}{k!} J^{i_1 \cdots i_k} g_{i_1\, j_1} \cdots g_{i_k\, j_k} J^{j_1 \cdots j_k}. \tag{4.112}$$

In this formula it is understood that there is summation over repeated indices.

Proof: In the following there are always sums over repeated indices. We have

$$\det(X^T G X) = \frac{1}{k!} \epsilon^{\alpha_1 \cdots \alpha_k} \epsilon^{\beta_1 \cdots \beta_k} (X^T G X)_{\alpha_1\, \beta_1} \cdots (X^T G X)_{\alpha_k\, \beta_k}. \tag{4.113}$$

However

$$(X^T G X)_{\alpha\, \beta} = X^i_\alpha g_{ij} X^j_\beta. \tag{4.114}$$

So

$$\det(X^T G X) = \frac{1}{k!} \epsilon^{\alpha_1 \cdots \alpha_k} \epsilon^{\beta_1 \cdots \beta_k} X^{i_1}_{\alpha_1} \cdots X^{i_k}_{\alpha_k} g_{i_1\, j_1} \cdots g_{i_k\, j_k} X^{j_1}_{\beta_1} \cdots X^{j_k}_{\beta_k}. \tag{4.115}$$

From the definition of J we get the result as stated. □

Corollary 4.4 *Let X be an n by k matrix. Let G be an n by n diagonal matrix. For each sequence $i_1, \ldots, i_k$ of rows of X, let*

$$J^{i_1 \cdots i_k} = \epsilon^{\alpha_1 \cdots \alpha_k} X^{i_1}_{\alpha_1} \cdots X^{i_k}_{\alpha_k}. \tag{4.116}$$

represent the determinant of the corresponding k by k minor obtained by retaining only those rows. Then

$$\det(X^T X) = \sum_K g_K (J^K)^2. \tag{4.117}$$

Here $K = \{i_1, \ldots, i_k\}$ is a subset of $\{1, \ldots, n\}$. We write $(J^K)^2 = (J^{i_1 \cdots i_k})^2$ and $g_K = \prod_{i \in K} g_{ii}$.

Now consider vectors X_α in $\mathbf{R}^n$ for $\alpha = 1, \ldots, n-1$. This is the special case of codimension one. In this case, there is a nice simplification of these results.

Theorem 4.5 *Let X be an n by $n-1$ matrix. Let G be an n by n symmetric invertible matrix. For each sequence $j_2, \ldots, j_n$ of rows of X, let*

$$J^{j_2 \cdots j_n} = \epsilon^{\alpha_1 \cdots \alpha_{n-1}} X^{j_2}_{\alpha_1} \cdots X^{j_n}_{\alpha_{n-1}}. \tag{4.118}$$

represent the determinant of the corresponding $n-1$ by $n-1$ minor obtained by retaining only those rows. Let

$$\nu_j = \frac{1}{(n-1)!} \epsilon_{j\, j_2 \cdots j_n} J^{j_2 \cdots j_n} \tag{4.119}$$

be the components of a row vector that represents the determinant of the minor that does not include row j. Then

$$\det(X^T G X) = g\, \nu G^{-1} \nu^T. \tag{4.120}$$

Here g is the determinant of the matrix G.

Proof: First we need the identity

$$J^{j_2 \cdots j_n} = \nu_j \epsilon^{j\, j_2 \cdots j_n}. \tag{4.121}$$

This follows from

$$\nu_j \epsilon^{j\, j_2 \cdots j_n} = \frac{1}{(n-1)!} \epsilon^{j\, j_2 \cdots j_n} \epsilon_{j\, h_2 \cdots h_n} J^{h_2 \cdots h_n} = J^{j_2 \cdots j_n}. \tag{4.122}$$

The last identity requires some thought. Fix $j_2, \ldots, j_n$ distinct. In the sum over j the factor $\epsilon^{j\, j_2 \cdots j_n}$ only contributes when j is the one index distinct from $j_2, \ldots, j_n$. The corresponding factor $\epsilon_{j\, h_2 \cdots h_n}$ then only matters when the indices $h_2, \ldots, h_n$ are a permutation of the $j_2, \ldots, j_n$. However, both $\epsilon_{j\, h_2 \cdots h_n}$ and $J^{h_2 \cdots h_n}$ are antisymmetric in the indices $h_1, \ldots, h_n$. It follows that the product $\epsilon_{j\, h_2 \cdots h_n} J^{h_2 \cdots h_n}$ is the same for each permutation $h_1, \ldots, h_n$. When we sum over these permutations we get $(n-1)!$ terms all equal to $J^{j_2 \cdots j_n}$.

We can use the previous theorem to write

$$\det(X^T G X) = \frac{1}{(n-1)!} J^{i_2 \cdots i_n} g_{i_2\, j_2} \cdots g_{i_n\, j_n} J^{j_2 \cdots j_n}. \tag{4.123}$$

The identity shows that this is equal to

$$\det(X^T G X) = \frac{1}{(n-1)!} \nu_i \nu_j \epsilon^{i\, i_2 \cdots i_n} \epsilon^{j\, j_2 \cdots j_n} g_{i_2\, j_2} \cdots g_{i_n\, j_n}. \tag{4.124}$$

However the cofactor C^{ij} of g_{ij} in G is given by a determinant

$$C^{ij} = \frac{1}{(n-1)!} \epsilon^{i\, i_2 \cdots i_n} \epsilon^{j\, j_2 \cdots j_n} g_{i_2\, j_2} \cdots g_{i_n\, j_n}. \tag{4.125}$$

Since by Cramer's rule the inverse matrix is $g^{ij} = (1/g) C^{ij}$, we get

$$\det(X^T G X) = \nu_i C^{ij} \nu_j = g\, \nu_i g^{ij} \nu_j. \tag{4.126}$$

□

Corollary 4.6 *Let X be an n by $n-1$ matrix. Let G be a diagonal invertible matrix. For each sequence $j_2, \ldots, j_n$ of rows of X, let*

$$J^{j_2 \cdots j_n} = \epsilon^{\alpha_1 \cdots \alpha_{n-1}} X^{j_2}_{\alpha_1} \cdots X^{j_n}_{\alpha_{n-1}}. \tag{4.127}$$

represent the determinant of the corresponding $n-1$ by $n-1$ minor obtained by retaining only those rows. Let

$$\nu_j = \frac{1}{(n-1)!} \epsilon_{j\, j_2 \cdots j_n} J^{j_2 \cdots j_n} \tag{4.128}$$

be the components of a row vector that represents the determinant of the minor that does not include row j. Then

$$\det(X^T G X) = \sum_{i=1}^{n} \frac{g}{g_{ii}} \nu_i^2. \tag{4.129}$$

This general result is dramatic even in the case $n = 3$ and G the identity matrix. In that case it says that the square of the area of a parallelogram is the sum of the squares of the areas of the three parallelograms obtained by projecting on the three coordinate planes. This is a remarkable generalization of the theorem of Pythagoras. [The most common version of this observation is in the context of the cross product. There are vectors X^1 and X^2 in $\mathbf{R}^3$. They span a parallelogram with a certain area. The cross product is a vector ν perpendicular to X^1 and X^2 whose Euclidean length is this area.]

4.11 Surface area

Say that we are in n dimensions with metric

$$\mathsf{g} = \sum_{i=1}^{n} \sum_{i=1}^{n} g_{ij}\, dx_i\, dx_j. \tag{4.130}$$

The n-dimensional volume is given by integrating

$$\mathrm{vol}_n = \sqrt{g}\, dx_1 \cdots dx_n, \tag{4.131}$$

where $g = \det(g_{ij})$.

Consider a k-dimensional regular parametrized surface S with parameters $u_1, \ldots, u_k$. This parametrization is one-to-one, so that $u_1, \ldots, u_k$ may be thought of as coordinates on the surface. It seems reasonable to compute the k-dimensional *surface area* using the pull-back of g to the surface. This is

$$\mathsf{g}^* = \sum_{\alpha=1}^{k} \sum_{\beta=1}^{k} g^*_{\alpha,\beta}\, du_\alpha\, du_\beta = \sum_{\alpha=1}^{k} \sum_{\beta=1}^{k} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij} \frac{\partial x_i}{\partial u_\alpha} \frac{\partial x_j}{\partial u_\beta} \right) du_\alpha\, du_\beta. \tag{4.132}$$

Let $X^i_\alpha = \partial x_i / \partial u_\alpha$. Then we have a collection of tangent vectors X_α. It follows that

$$g^*_{\alpha\beta} = X^T_\alpha G X_\beta. \tag{4.133}$$

Notice that $g^*_{\alpha\beta}$ has the form of a Gram matrix. Let $g^* = \det(g^*_{\alpha\beta})$ be the corresponding Gram determinant. Then the area is given by integrating

$$\mathrm{area}^*_k = \sqrt{g^*}\, du_1 \cdots du_k. \tag{4.134}$$

In some sense this is the end of the story. One computes a Gram determinant, takes the square root, and integrates. Because of the square root the integrals tend to be quite nasty. In principle, though, we have a nice notion of area and of integration with respect to area. That is, we have

$$\int_S h(\mathbf{x}) \mathrm{area}_k(\mathbf{x}) = \int h(\mathbf{f}(\mathbf{u})) \sqrt{g^*}\, du_1 \cdots du_k. \tag{4.135}$$

This may be computed by any convenient parametrization $\mathbf{x} \leftarrow \mathbf{f}(\mathbf{u})$ of the surface.

The main complication in discussions of surface area is that there are alternative formulas for the Gram determinant. The general formula is hard to work with, though it is not so bad when the coordinates $x_1, \ldots, x_n$ are orthogonal coordinates. The classic approach is to restrict attention to a hypersurface, when $k = n - 1$. In this case it is useful to consider the forms

$$\pi_i = (-1)^{i-1}\, dx_1 \cdots dx_{i-1}\, dx_{i+1} \ldots\, dx_n. \tag{4.136}$$

These occur in expressions involving the interior product of the vector field $Y = \sum_i a^i \partial/\partial x_i$ with the volume form. This interior product is

$$Y \rfloor \mathrm{vol} = \sqrt{g} \sum_{i=1}^{n} a^i \pi_i. \tag{4.137}$$

If we pull back π_j to the surface, we get a form

$$\pi^*_j = \nu_j\, du_1 \cdots du_k. \tag{4.138}$$

Here ν_j is $(-1)^{j-1}$ times the determinant of the $n-1$ by $n-1$ matrix obtained by removing the jth row from the n by $n-1$ matrix $\partial x_i/\partial u_\alpha$. So the interior product pulls back to

$$(Y \rfloor \mathrm{vol})^* = \sqrt{g} \sum_{j=1}^{n} a^j \pi_j^* = \sum_{j=1}^{n} a^j \nu_j \sqrt{g}\, du_1 \cdots du_k. \tag{4.139}$$

The coefficients ν_j satisfy for each β the relation

$$\sum_{j=1}^{n} \nu_j \frac{\partial x_j}{\partial u_\beta} = 0. \tag{4.140}$$

This says that the ν_j are the coefficients of a 1-form that vanishes on the tangent vectors to the surface. In other words, the $\sum_j a^j \nu_j$ is giving a numerical indication of the extent to which the vector field is failing to be tangent, that is, of the extent to which it is penetrating the surface.

The alternate formula for the area form follows from the linear algebra treated in the previous section. It says that

$$\mathrm{area}^*_{n-1} = \sqrt{g^*}\, du_1 \cdots du_n = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} \nu_i g^{ij} \nu_j}\sqrt{g}\, du_1 \cdots du_{n-1}. \tag{4.141}$$

In this equation g^{ij} is the inverse matrix of the metric matrix g_{ik}. The ν_j is $(-1)^{j-1}$ times the determinant of the $n-1$ by $n-1$ matrix obtained by removing the jth row from the n by $n-1$ matrix $\partial x_i/\partial u_\alpha$.

The form coefficients ν_j define the coefficients of a vector with coefficients $N^i = \sum_{j=1}^{n} g^{ij}\nu_j$. The identity says that this vector is orthogonal to the surface. This vector of course depends on the parametrization. Sometimes people express the area in terms of this corresponding vector as

$$\mathrm{area}^*_{n-1} = \sqrt{g^*}\, du_1 \cdots du_n = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} N^i g_{ij} N^j}\sqrt{g}\, du_1 \cdots du_{n-1}. \tag{4.142}$$

Somewhat astonishingly, it is common to write the flux in the form

$$(Y \rfloor \mathrm{vol})^* = \sum_{i=1}^{n}\sum_{j=1}^{n} a^j g_{ij} \hat{N}^j\, \mathrm{area}_{n-1}. \tag{4.143}$$

Here $\hat{N}_k$ indicates the normalization of the vector to have length one. What is amazing is that if one writes it out, this says

$$(Y \rfloor \mathrm{vol})^* = \sum_{i=1}^{n}\sum_{j=1}^{n} a^i g_{ij} \frac{N^j}{\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} N^i g_{ij} N^j}} \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} N^i g_{ij} N^j}\sqrt{g}\, du_1 \cdots du_{n-1}. \tag{4.144}$$

There are two complicated factors involving square roots, one to normalize the orthogonal vector, the other to calculate the area. These factors cancel. They never need to be computed in a flux integral.

It may be helpful to summarize the result for hypersurface area in the form of a theorem. This is stated in a way that makes clear the connection with the divergence theorem. Recall that the transverse surface element

$$\text{element} = \phi_1 * \text{vol} = \phi_1^* \sqrt{g}\, dx_1 \cdots dx_n = \sqrt{g} \sum_{i=1}^{n} \nu_i\, dx_i\, du_1 \cdots du_{n-1} \quad (4.145)$$

is the interior pullback of the volume. The 1-form part involving the dx_i is the part that was not pulled back. It measures the extent to which a vector field is transverse to the surface, as in the setting of the divergence theorem. Since this is a form and not a vector, its norm is computed via the inverse of the metric tensor.

Theorem 4.7 *Consider an n-dimensional manifold patch and a regular parameterized $n-1$ surface with parameters $u_1, \ldots, u_n$. Consider also the transverse surface element*

$$\text{element} = \sqrt{g} \sum_{i=1}^{n} \nu_i\, dx_i\, du_1 \cdots du_{n-1} \quad (4.146)$$

that measures the extent to which a vector is transverse to the surface. Then the area element is the length of the transverse surface element:

$$\text{area} = |\text{element}| = \sqrt{g} \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \nu_i g^{ij} \nu_j}\, du_1 \cdots du_{n-1}. \quad (4.147)$$

The textbook case is that where $n = 3$ and $k = 2$, that is, a surface in 3-space. The most common coordinates are Cartesian coordinates x, y, z for which the metric is given by $dx^2 + dy^2 + dz^2$. However we might want some other set of coordinates, so temporarily we think of x, y, z as some choice of orthogonal coordinates with metric

$$\mathsf{g} = h_x^2\, dx^2 + h_y^2\, dy^2 + h_z^2\, dz^2. \quad (4.148)$$

With Cartesian coordinates we simply have $h_x = h_y = h_z = 1$.

Say that we have a surface parameterized by u, v. Then the metric on this surface is

$$\mathsf{g}^* = E\, du^2 + 2F\, du\, dv + G\, dv^2. \quad (4.149)$$

Here E, F, G are functions of u, v. They of course depend on the choice of coordinates. What is required is that $E > 0, G > 0$ and the determinant $EF - G^2 > 0$. Explicitly, the coefficients are

$$E = h_x^2 \left(\frac{\partial x}{\partial u}\right)^2 + h_y^2 \left(\frac{\partial y}{\partial u}\right)^2 + h_z^2 \left(\frac{\partial z}{\partial u}\right)^2, \quad (4.150)$$

and

$$F = h_x^2 \frac{\partial x}{\partial u}\frac{\partial x}{\partial v} + h_y^2 \frac{\partial y}{\partial u}\frac{\partial y}{\partial v} + h_z^2 \frac{\partial z}{\partial u}\frac{\partial z}{\partial v}, \tag{4.151}$$

and

$$G = h_x^2 \left(\frac{\partial x}{\partial v}\right)^2 + h_y^2 \left(\frac{\partial y}{\partial v}\right)^2 + h_z^2 \left(\frac{\partial z}{\partial v}\right)^2. \tag{4.152}$$

The formula for the area of a surface is

$$A = \int_S \text{area} = \int_S \sqrt{g}\, du\, dv = \int_S \sqrt{EG - F^2}\, du\, dv. \tag{4.153}$$

Here $g = EG - F^2$ is the determinant of the metric tensor.

The alternative expression for surface area is sometimes convenient. This is

$$A = \int_S \text{area} = \int_S \sqrt{h_y^2 h_z^2 \nu_x^2 + h_z^2 h_x^2 \nu_y^2 + h_x^2 h_y^2 \nu_z^2}\, du\, dv. \tag{4.154}$$

Here $\nu_x = J^{yz}$, $\nu_y = J^{zx}$, and $\nu_z = J^{xy}$. An expression such as J^{yz} indicates the determinant of the matrix of partial derivatives of y, z with respect to u, v.

As an example, take the surface given in Cartesian coordinates by $z = x^2 + y^2$ with $z \le 1$. Use x, y as parameters. Then $E = 1 + 4x^2$, $F = 4xy$, and $G = 1 + 4y^2$. So with these parameters the area form is $\sqrt{EG - F^2}\, dx\, dy = \sqrt{1 + 4x^2 + 4y^2}\, dx\, dy$. This is integrated over the region $x^2 + y^2 = 1$. With the alternative calculation $\nu_x = -2x$ and $\nu_y = -2y$ and $\nu_z = 1$. So the area form is $\sqrt{\nu_x^2 + \nu_y^2 + \nu_z^2}\, dx\ dy = \sqrt{4x^2 + 4y^2 + 1}\, dx\, dy$, exactly the same.

Other parametrizations are possible. Take, for instance, $x = r\cos(\phi)$, $y = r\sin(\phi)$, $z = r^2$. Then $E = 1 + 4r^2$, $F = 0$, and $G = r^2$. So with these parameters the area form is $\sqrt{EG - F^2}\, dr\, d\phi = r\sqrt{1 + 4r^2}\, dr\, d\phi$. We may as well go ahead and compute the area. It is 2π times the integral from 0 to 1 of $r\sqrt{1 + 4r^2}\, dr$. The area is thus $(\pi/6)(5^{\frac{3}{2}} - 1)$.

Problems 10: The divergence theorem

1. Let $r^2 = x^2 + y^2 + z^2$, and let

$$\mathbf{v} = \frac{1}{r^3}\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}\right). \tag{4.155}$$

Let $\text{vol} = dx\, dy\, dz$. Show that the *solid angle form*

$$\sigma = \mathbf{v} \rfloor \text{vol} = \frac{1}{r^3}(x\, dy\, dz + y\, dz\, dx + z\, dx\, dy). \tag{4.156}$$

2. In the preceding problem, show directly that $d\sigma = 0$ away from $r = 0$.

3. Find σ in spherical polar coordinates. Hint: This can be done by blind computation, but there is a better way. Express $\mathbf{v}$ in spherical polar coordinates, using Euler's theorem

$$r\frac{\partial}{\partial r} = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}. \tag{4.157}$$

Then use $\text{vol} = r^2 \sin(\theta)\, dr\, d\theta\, d\phi$ to calculate $\sigma = \mathbf{v} \rfloor \text{vol}$.

4. In the preceding problem, show that $d\sigma = 0$ away from $r = 0$ by a spherical polar coordinate calculation.

5. Let S be the sphere of radius $a > 0$ centered at the origin. Calculate the integral of σ over S.

6. Let Q be the six-sided cube with side lengths $2L$ centered at the origin. Calculate the integral of σ over Q. Prove that your answer is correct. Hint: Given the result of the previous problem, this should be effortless.

Recitation 10

The setting for these problems is Euclidean space $\mathbf{R}^n$ with $n \geq 3$ with Cartesian coordinates $x_1, \ldots, x_n$. We write $r^2 = x_1^2 + \cdots + x_n^2$ and $\text{vol} = dx_1 \cdots dx_n$. The gradient of u is the vector

$$\nabla u = \text{grad } u = \sum_{i=1}^{n} \frac{\partial u}{\partial x_i} \frac{\partial}{\partial x_i}. \tag{4.158}$$

Then

$$\nabla u \rfloor \text{vol} = \sum_{i=1}^{n} (-1)^{i-1} \frac{\partial u}{\partial x_i}\, dx_1 \cdots dx_{i-1}\, dx_{i+1} \cdots dx_n. \tag{4.159}$$

In Cartesian coordinates the Laplacian $\nabla^2 u$ is defined by

$$\nabla^2 u \,\text{vol} = \text{div grad } u \,\text{vol} = d(\nabla u \rfloor \text{vol}) = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2} \,\text{vol} \tag{4.160}$$

Often $\nabla^2 = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$ is called the Laplace operator.

1. Define the Euler operator $E = \nabla \frac{1}{2} r^2$. Show that

$$E = \sum_{i=1}^{n} x_i \frac{\partial}{\partial x_i}. \tag{4.161}$$

2. Define the form $\omega = E \rfloor \text{vol}$. Show that

$$\omega = \sum_{i=1}^{n} (-1)^{i-1} x_i\, dx_1 \cdots dx_{i-1}\, dx_{i+1} \cdots dx_n \tag{4.162}$$

3. Define the *solid angle form* $\sigma = \frac{1}{r^n}\omega$. Show that

$$r^{n-1}\, dr\, \sigma = \frac{1}{r} dr\, \omega = \text{vol}. \tag{4.163}$$

4. Show that $d\omega = n\,\text{vol}$.

5. Let B_a be the ball of radius a with volume $\alpha(n)a^n$. Let S_n be the sphere of radius a with area $n\alpha(n)a^{n-1}$. (For instance $\alpha(3) = (4/3)\pi$ occurs in the volume formula for $n = 3$, and hence $3\alpha(3) = 4\pi$ is in the area formula.) Show that

$$\int_{S_a} \omega = n\alpha(n)a^n. \tag{4.164}$$

Hint: Apply the divergence theorem.

6. For $r \neq 0$ define the fundamental solution by

$$\phi(\mathbf{x}) = \frac{1}{n\alpha(n)} \frac{1}{n-2} \frac{1}{r^{n-2}}. \tag{4.165}$$

(For $n = 3$ this is $\phi(\mathbf{x}) = 1/(4\pi)1/r$.) Show that

$$-\nabla\phi(\mathbf{x}) = \frac{1}{n\alpha(n)} \frac{1}{r^n} E. \tag{4.166}$$

7. Show that

$$-\nabla\phi(\mathbf{x}) \rfloor \text{vol} = \frac{1}{n\alpha(n)} \frac{1}{r^n} \omega. \tag{4.167}$$

8. Show that this is a closed form, and hence $\nabla^2\phi(\mathbf{x}) = 0$ away from $r = 0$.

9. Show that for every $a > 0$ the flux of the fundamental solution is

$$\int_{S_a} -\nabla\phi(\mathbf{x}) \rfloor \text{vol} = 1. \tag{4.168}$$

Problems 11: The Laplacian

The setting for these problems is Euclidean space $\mathbf{R}^n$ with $n \geq 3$ with Cartesian coordinates $x_1, \ldots, x_n$. We write $r^2 = x_1^2 + \cdots + x_n^2$ and $\text{vol} = dx_1 \cdots dx_n$. The gradient of u is the vector

$$\nabla u = \text{grad } u = \sum_{i=1}^{n} \frac{\partial u}{\partial x_i} \frac{\partial}{\partial x_i}. \tag{4.169}$$

Then

$$\nabla u \rfloor \text{vol} = \sum_{i=1}^{n} (-1)^{i-1} \frac{\partial u}{\partial x_i}\, dx_1 \cdots dx_{i-1}\, dx_{i+1} \cdots dx_n. \tag{4.170}$$

In Cartesian coordinates the Laplacian $\nabla^2 u$ is defined by

$$\nabla^2 u \,\mathrm{vol} = \mathrm{div}\ \mathrm{grad}\ u \,\mathrm{vol} = d(\nabla u \rfloor \mathrm{vol}) = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2} \,\mathrm{vol} \tag{4.171}$$

Often $\nabla^2 = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$ is called the Cartesian coordinate Laplace operator.

The fundamental solution of the Laplace equation $\nabla^2 u = 0$ is

$$\phi(\mathbf{x}) = \frac{1}{n\alpha(n)} \frac{1}{n-2} \frac{1}{r^{n-2}}. \tag{4.172}$$

The goal is to establish the following amazing identity: The fundamental solution satisfies

$$-\nabla^2 \phi(\mathbf{x}) = \delta(\mathbf{x}). \tag{4.173}$$

This shows that the the behavior of the fundamental solution at $r = 0$ can also be understood. It turns out that this is the key to solving various problems involving the Laplace operator. The following problems are intended to give at least some intuitive feel for how such an identity can come about.

The method is to get an approximation $\phi_\epsilon(\mathbf{x})$ whose Laplacian is an approximate delta function $\delta_\epsilon(\mathbf{x})$. To this end, define $r_\epsilon = \sqrt{r^2 + \epsilon^2}$ and let

$$\phi_\epsilon(\mathbf{x}) = \frac{1}{n\alpha(n)} \frac{1}{n-2} \frac{1}{r_\epsilon^{n-2}}. \tag{4.174}$$

1. Prove that $dr_\epsilon^2 = dr^2$ and hence $dr_\epsilon/dr = r/r_\epsilon$.

2. (a) Prove that
$$-\nabla \phi_\epsilon(\mathbf{x}) = \frac{1}{n\alpha(n)} \frac{1}{r_\epsilon^n} E. \tag{4.175}$$
(b) Prove that
$$-\nabla \phi_\epsilon(\mathbf{x}) \rfloor \,\mathrm{vol} = \frac{1}{n\alpha(n)} \frac{1}{r_\epsilon^n} \omega. \tag{4.176}$$

3. Show that
$$-\nabla^2 \phi_\epsilon(\mathbf{x}) = \delta_\epsilon(\mathbf{x}), \tag{4.177}$$
where $\delta_\epsilon(\mathbf{x})$ is a constant times a power of ϵ times an inverse power of r_ϵ. Find the explicit form of $\delta_\epsilon(\mathbf{x})$.

4. Show that
$$\delta_\epsilon(\mathbf{x}) = \delta_1\left(\frac{\mathbf{x}}{\epsilon}\right) \frac{1}{\epsilon^n}. \tag{4.178}$$
What is the explicit form for the function $\delta_1(\mathbf{x})$?

5. To show that this is an approximate delta function, we need to show that $\int \delta_\epsilon(\mathbf{x}) \,\mathrm{vol} = 1$. For each a compute $\int_{B_a} \delta_\epsilon(\mathbf{x}) \,\mathrm{vol}$ as an explicit function of a and ϵ and n. Hint: Use the divergence theorem.

6. Show that for fixed $\epsilon > 0$ we have

$$\int_{B_a} \delta_\epsilon(\mathbf{x}) \,\mathrm{vol} \to 1 \tag{4.179}$$

as $a \to \infty$.

Recitation 11

1. In the following we use tensor algebra notation where repeated upper and lower indices indicate summation. Prove the following identity for the Levi-Civita permutation symbols:

$$\epsilon^{j_1 \cdots j_n} \epsilon_{j_1 \cdots j_n} = n!. \tag{4.180}$$

2. Recall that

$$\det(A) = \epsilon^{j_1 \cdots j_n} a^1_{j_1} \cdots a^n_{j_n}. \tag{4.181}$$

Show that

$$\epsilon^{i_1 \cdots i_n} \det(A) = \epsilon^{j_1 \cdots j_n} a^{i_1}_{j_1} \cdots a^{i_n}_{j_n}. \tag{4.182}$$

3. Show that

$$\det(A) = \frac{1}{n!} \epsilon_{i_1 \cdots i_n} \epsilon^{j_1 \cdots j_n} a^{i_1}_{j_1} \cdots a^{i_n}_{j_n}. \tag{4.183}$$

4. Show that $\det(AB) = \det(A)\det(B)$. Hint: Use the preceding identity for AB.

5. The next three problems are relevant to Cramer's rule. Show that

$$\epsilon^{j\, j_2 \cdots j_n} a^i_j a^{i_2}_{j_2} \cdots a^{i_n}_{j_n} = \epsilon^{i\, i_2 \cdots i_n} \det(A). \tag{4.184}$$

6. Show that

$$\frac{1}{(n-1)!} \epsilon_{k\, i_2 \cdots i_n} \epsilon^{i\, i_2 \cdots i_n} = \delta^i_k. \tag{4.185}$$

7. If a^i_j is a matrix, define its *cofactor* matrix to be

$$C^j_k = \frac{1}{(n-1)!} \epsilon_{k\, i_2 \cdots i_n} \epsilon^{j\, j_2 \cdots j_n} a^{i_2}_{j_2} \cdots a^{i_n}_{j_n}. \tag{4.186}$$

(This is actually the transpose of the usual matrix of cofactors.) Prove *Cramer's rule*

$$a^i_j C^j_k = \delta^i_k \det(A). \tag{4.187}$$

8. Let $X^1, \ldots, X^n$ be n vectors in $\mathbf{R}^n$ forming a matrix X. The Euclidean volume spanned by these vectors is $|\det X|$. Define the Gram matrix to be the matrix of inner products $A = X^T X$. Show that the volume is given also by

$$\sqrt{\det(X^T X)} = |\det X|. \tag{4.188}$$

9. Let $n = 2$ in the above result. Let θ be the angle between the two vectors. Compute the area in terms of this angle and the lengths of the vectors.

10. Let $X^1, \ldots, X^n$ be n vectors in $\mathbf{R}^n$ forming a matrix X. Let G be a symmetric matrix with positive eigenvalues. Define the volume relative to G to be $\sqrt{\det(X^T G X)}$. Find a formula for this volume in terms of G and $|\det(X)|$.

11. L X be an n by k matrix, representing k vectors X^i_α, where $i = 1, \ldots, n$ and $\alpha = 1, \ldots, k$. Let G be an n by n symmetric matrix. We can define the Gram matrix as the k by k matrix $X^T X$, or more generally as the k by k matrix $X^T G X$. For each sequence $i_1, \ldots, i_k$ of rows of X,

$$J^{i_1 \cdots i_k} = \epsilon^{\alpha_1 \cdots \alpha_k} X^{i_1}_{\alpha_1} \cdots X^{i_k}_{\alpha_k} \tag{4.189}$$

represents the determinant of the corresponding k by k minor obtained by retaining only those rows. Show

$$\det(X^T G X) = \frac{1}{k!} J^{i_1 \cdots i_k} g_{i_1\, j_1} \cdots g_{i_k\, j_k} J^{j_1 \cdots j_k}. \tag{4.190}$$

In this formula it is understood that there is summation over repeated indices.

12. Take $G = I$ (with matrix δ_{ij}) in the above formula, and simplify the result. This gives a formula for the area $\sqrt{\det(X^T X)}$ in terms of areas of projections. This is a remarkable generalization of the theorem of Pythagoras.

13. Describe what this simplified formula says in the case $n = 3$ and $k = 2$.

Problems 12: Surface area

1. Consider an $n-1$ surface given implicitly by $w = g(x_1, \ldots, x_n) = C$ and implicitly by $x_i = f_i(u_1, \ldots, u_{n-1})$. For $\alpha = 1, \ldots, n-1$ there are $n-1$ tangent vectors to the surface with components $\partial x_i / \partial u_\alpha$. Prove that dw on the surface is zero on the tangent space, that is, prove that for each β we have

$$\sum_{j=1}^{n} \frac{\partial w}{\partial x_j} \frac{\partial x_j}{\partial u_\beta} = 0. \tag{4.191}$$

2. In the same situation, show that the form on the surface with components ν_j equal to the $(-1)^{j-1}$ times the determinant of $\partial x_i / \partial u_\alpha$ with row j deleted satisfies

$$\sum_{j=1}^{n} \nu_j \frac{\partial x_j}{\partial u_\beta} = 0. \tag{4.192}$$

Since it is also zero on the tangent space, it must be a multiple of dw on the surface. Hint: Consider the matrix with first column $\partial x_i / \partial u_\beta$ and remaining columns $\partial x_i / \partial u_\alpha$ for $\alpha = 1, \ldots, n-1$. Here β is an arbitrary choice of one of the α indices.

3. Consider the surface given by $x = uv, y = u+v, z = u-v$. Find an implicit equation for this surface. Verify that the two forms of the previous two problems are multiples of each other.

4. For the same surface, the metric $dx^2 + dy^2 + dz^2$ has a pullback given by $E\,du^2 + 2F\,du\,dv + G\,dv^2$. Calculate it.

5. For the same surface, calculate the area for parameter region $u^2 + v^2 \leq 1$ by integrating $\sqrt{EG - F^2}\,du\,dv$ over the region.

6. For the same surface, the forms $dy\,dz$ and $dz\,dx = -\,dx\,dz$ and $dx\,dy$ have pullbacks $\nu_x\,du\,dv$ and $\nu_y\,du\,dv$ and $\nu_z\,du\,dv$. Calculate them.

7. For the same surface, calculate the area for parameter region $u^2 + v^2 \leq 1$ by integrating $\sqrt{(\nu_x^2 + \nu_y^2 + \nu_z^2)}\,du\,dv$ over the region.

Chapter 5

Measure Zero

5.1 Outer content and outer measure

The Riemann integral is good for many purposes, but the Lebesgue integral is both more general and easier to work with. The problem is that it takes time and effort to develop the Lebesgue integral and to appreciate its qualities. In this part we take only a small step beyond the Riemann integral. We contrast sets of content zero (a Riemann integral topic) with sets of measure zero (a Lebesgue integral topic). Every set of content zero is of measure zero. However there are sets of measure zero that are not of content zero. Sets of measure zero are a key concept in analysis.

A high point for this course is the theorem due to Lebesgue that characterizes those bounded functions (on a bounded cell) that have a Riemann integral. The astonishing result is that they are exactly the functions that are continuous except on a set of measure zero. If we combine this with the fact that smooth change of variable functions take sets of measure zero to measure zero, we get a deeper insight into change of variables for the Riemann integral.

Another striking result is Sard's theorem, which concerns a smooth change of variables function having critical points at which the Jacobian determinant is zero and the inverse function theorem fails. The image of the set of critical points is the set of critical values. Even though the set of critical points may be quite large, the theorem says that the corresponding set of critical values has measure zero.

Recall that a cell $I \subseteq \mathbf{R}^n$ is a product of intervals. The n-dimensional volume of a bounded cell I is denoted $m(I)$. A cell is non-degenerate if each interval is non-degenerate. Suppose $A \subseteq \mathbf{R}^n$. The fundamental notion in the following is that of cell cover. A *cell cover* of A is a family $\mathcal{I}$ of bounded non-degenerate cells such that $A \subseteq \bigcup_{I \in \mathcal{I}} I$.

The *outer content* of A is defined by

$$\bar{m}(A) = \inf\{\sum_{I \in \mathcal{I}} m(I) \mid \mathcal{I} \text{ finite}, A \subseteq \bigcup_{I \in \mathcal{I}} I\}. \tag{5.1}$$

It follows that $A \subseteq B$ implies $\bar{m}(A) \leq \bar{m}(B)$. A set A with $\bar{m}(A) = 0$ is said to have *content zero*.

The *outer measure* of A is defined by

$$\bar{\mu}(A) = \inf\{\sum_{I \in \mathcal{I}} m(I) \mid \mathcal{I} \text{ countable}, A \subseteq \bigcup_{I \in \mathcal{I}} I\}. \tag{5.2}$$

It follows that $A \subseteq B$ implies $\bar{\mu}(A) \leq \bar{\mu}(B)$. A set A with $\bar{\mu}(A) = 0$ is said to have *measure zero*.

Proposition 5.1 *For every $A \subseteq \mathbf{R}^n$ we have*

$$\bar{\mu}(A) \leq \bar{m}(A). \tag{5.3}$$

Theorem 5.2 (Finite sub-additivity of outer content) *If $\mathcal{A}$ is a finite family of subsets of $\mathbf{R}^n$, then*

$$\bar{m}\left(\bigcup_{A\in\mathcal{A}} A\right) \leq \sum_{A\in\mathcal{A}} \bar{m}(A). \tag{5.4}$$

Theorem 5.3 (Countable sub-additivity of outer measure) *If $\mathcal{A}$ is a countable family of subsets of $\mathbf{R}^n$, then*

$$\bar{\mu}\left(\bigcup_{A\in\mathcal{A}} A\right) \leq \sum_{A\in\mathcal{A}} \bar{\mu}(A). \tag{5.5}$$

Since this theorem is the key to the entire subject, it is worth recording the proof. Let $\epsilon > 0$. Since $\mathcal{A}$ is countable, we may enumerate its elements A_n. By the fact that $\bar{\mu}(A_n)$ is defined as a greatest lower bound of sums, there is a cover I_{nk} of A_n with $\sum_k m(I_{nk}) < \bar{\mu}(A_n) + \frac{\epsilon}{2^n}$. Then $\bigcup_n A_n$ is covered by all the I_{nk}. Furthermore, $\sum_{nk} m(I_{nk}) < \sum_n \bar{\mu}(A_n) + \epsilon$. Since $\bar{\mu}(\bigcup_n A_n)$ is defined as a lower bound of sums, we have $\bar{\mu}(\bigcup_n A_n) < \sum_n \bar{\mu}(A_n) + \epsilon$. Since $\epsilon > 0$ is arbitrary, we must have $\bar{\mu}(\bigcup_n A_n) \leq \sum_n \bar{\mu}(A_n)$.

Proposition 5.4 *Suppose that A, A' are subsets with $\bar{m}(A \setminus A') = 0$ and also $\bar{m}(A' \setminus A) = 0$. Then*

$$\bar{m}(A \cap A') = \bar{m}(A) = \bar{m}(A') = \bar{m}(A \cup A'). \tag{5.6}$$

Proposition 5.5 *Suppose that A, A' are subsets with $\bar{\mu}(A \setminus A') = 0$ and also $\bar{\mu}(A' \setminus A) = 0$. Then*

$$\bar{\mu}(A \cap A') = \bar{\mu}(A) = \bar{\mu}(A') = \bar{\mu}(A \cup A'). \tag{5.7}$$

Proposition 5.6 *The outer content of A may be defined by finite open cell covers:*

$$\bar{m}(A) = \inf\{\sum_{I\in\mathcal{I}} m(I) \mid \mathcal{I} \text{ finite open}, A \subseteq \bigcup_{I\in\mathcal{I}} I\}. \tag{5.8}$$

Proposition 5.7 *The outer content of A may be defined by finite closed cell covers:*

$$\bar{m}(A) = \inf\{\sum_{I\in\mathcal{I}} m(I) \mid \mathcal{I} \text{ finite closed}, A \subseteq \bigcup_{I\in\mathcal{I}} I\}. \tag{5.9}$$

Proposition 5.8 *The outer content of A may be defined by finite closed non-overlapping cell covers:*

$$\bar{m}(A) = \inf\{\sum_{I\in\mathcal{I}} m(I) \mid \mathcal{I} \text{ finite closed non} - \text{overlapping}, A \subseteq \bigcup_{I\in\mathcal{I}} I\}. \tag{5.10}$$

Proposition 5.9 *The outer measure of A may be defined by countable open cell covers:*

$$\bar{\mu}(A) = \inf\{\sum_{I\in\mathcal{I}} m(I) \mid \mathcal{I} \text{ countable open}, A \subseteq \bigcup_{I\in\mathcal{I}} I\}. \tag{5.11}$$

Proposition 5.10 *The outer measure of A may be defined by countable closed cell covers:*

$$\bar{\mu}(A) = \inf\{\sum_{I\in\mathcal{I}} m(I) \mid \mathcal{I} \text{ countable closed}, A \subseteq \bigcup_{I\in\mathcal{I}} I\}. \tag{5.12}$$

Theorem 5.11 *Suppose $A \subseteq \mathbf{R}^n$ is compact. Then*

$$\bar{\mu}(A) = \bar{m}(A). \tag{5.13}$$

This statement is not quite as elementary as it might look, since a compact set need not be Jordan measurable. It therefore deserves a proof. Say that A is compact. We need to show that $\bar{m}(A) \leq \bar{\mu}(A)$. Let $\epsilon > 0$. Since $\bar{\mu}(A)$ is defined as a greatest lower bound of sums, there is a countable cover I_k of A such that $\sum_{k=1}^{\infty} m(I_k) < \bar{\mu}(A) + \epsilon$. Since A is compact, there is a finite subcover, say by the first p cells. Then $\bar{m}(A) \leq \sum_{k=1}^{p} m(I_k) \leq \sum_{k=1}^{\infty} m(I_k) < \bar{\mu}(A) + \epsilon$. Since for every $\epsilon > 0$ we have $\bar{m}(A) < \bar{\mu}(A) + \epsilon$, we must have $\bar{m}(A) \leq \bar{\mu}(A)$.

5.2 The set of discontinuity of a function

The purpose of this section is to establish that the set of discontinuity of a function is the countable union of a certain family of closed subsets.

If f is a real function on $C \subseteq \mathbf{R}^n$, and $A \subseteq C$, define the oscillation of f on A by

$$\mathrm{osc}_A(f) = \sup_{\mathbf{z},\mathbf{w}\in A} |f(\mathbf{z}) - f(\mathbf{w})| = \sup_{\mathbf{z}\in A} f(\mathbf{z}) - \inf_{\mathbf{w}\in A} f(\mathbf{w}). \tag{5.14}$$

Define the oscillation of f at $\mathbf{x}$ by

$$\mathrm{osc}_{\mathbf{x}}(f) = \inf\{\mathrm{osc}_U(f) \mid \mathbf{x} \in \mathrm{int}(U)\}. \tag{5.15}$$

Proposition 5.12 *The function f is continuous at $\mathbf{x}$ if and only if $\mathrm{osc}_{\mathbf{x}}(f) = 0$.*

In the following we write $\mathrm{Disc}(f)$ for the set of points at which f is not continuous.

Proposition 5.13

$$\mathrm{Disc}(f) = \{\mathbf{x} \mid \mathrm{osc}_{\mathbf{x}}(f) > 0\}. \tag{5.16}$$

Next we want to make the notion of discontinuity quantitative. For each f and each $\epsilon > 0$ define the set $D_\epsilon(f)$ by the condition

$$D_\epsilon(f) = \{\mathbf{x} \mid \mathrm{osc}_{\mathbf{x}}(f) \geq \epsilon\}. \tag{5.17}$$

Proposition 5.14 *The point $\mathbf{x}$ is in $D_\epsilon(f)$ if and only if for every U with $\mathbf{x} \in \mathrm{int}(U)$ we have $\mathrm{osc}_U(f) \geq \epsilon$.*

Proposition 5.15 *For every $\epsilon > 0$ the set $D_\epsilon(f) \subseteq \mathrm{Disc}(f)$. Furthermore, as ϵ gets smaller the sets $D_\epsilon(f)$ can only get larger.*

Proposition 5.16 *For each $\epsilon > 0$ the set $D_\epsilon(f)$ is a closed subset of $\mathbf{R}^n$.*

Proposition 5.17 *A point is a point of discontinuity if and only if it belongs to some $D_\epsilon(f)$ for some $\epsilon > 0$. Thus*

$$\mathrm{Disc}(f) = \bigcup_{\epsilon>0} D_\epsilon(f). \tag{5.18}$$

Proposition 5.18 *The set of points of discontinuity is a countable union of closed subsets:*

$$\mathrm{Disc}(f) = \bigcup_{n=1}^{\infty} D_{\frac{1}{n}}(f). \tag{5.19}$$

Corollary 5.19 *Let f be defined on a bounded set. The set $\mathrm{Disc}(f)$ has measure zero if and only if for each $\epsilon > 0$ the set $D_\epsilon(f)$ has content zero.*

5.3 Lebesgue's theorem on Riemann integrability

Theorem 5.20 (Lebesgue (1907)) *Let $C \subset \mathbf{R}^n$ be a bounded closed non-degenerate cell. Let f be a real function on C. Then f is Riemann integrable if and only if f is bounded and the set of discontinuities has measure zero, that is,*

$$\bar{\mu}(\mathrm{Disc}(f)) = 0. \tag{5.20}$$

The theorem relies on two lemmas. The first one shows that Riemann integrability implies the set of discontinuities has measure zero. This part relies heavily on the countable sub-additivity of outer measure. The other one shows that a bounded function with a set of discontinuities of measure zero is Riemann integrable. Here the remarkable thing is that measure zero is a weaker requirement than having content zero.

Lemma 5.21 *For every $\epsilon > 0$ and every partition $\mathcal{P}$ of C we have*

$$\epsilon \bar{m}(D_\epsilon) \leq U(f, \mathcal{P}) - L(f, \mathcal{P}). \tag{5.21}$$

The way the lemma is used is to note that $U(f) - L(f)$ is the greatest lower bound for $U(f, \mathcal{P}) - L(f, \mathcal{P})$. The lemma says that $\epsilon \bar{m}(D_\epsilon)$ is a lower bound, so $\epsilon \bar{m}(D_\epsilon) \leq U(f) - L(f)$. If $U(f) - L(f) = 0$, then $\bar{m}(D_\epsilon) = 0$, and hence $\bar{\mu}(D_\epsilon) = 0$. By countable subadditivity $\bar{\mu}(\mathrm{Disc}) = 0$.

Proof: Let $\mathcal{P}$ be a partition of C into closed bounded non-degenerate cells. Let $\mathcal{P}'$ be the subset of $\mathcal{P}$ consisting of sets I with $\mathrm{int}(I) \cap D_\epsilon \neq \emptyset$. Then for I in $\mathcal{P}'$ we have $\mathrm{osc}_I(f) \geq \epsilon$. Now let B be the union of the boundary points of the cells in $\mathcal{P}$, and let $D'_\epsilon = D_\epsilon \setminus B$. If $\mathbf{x}$ is in D'_ϵ, then $\mathbf{x}$ is in $D_\epsilon(f)$, and hence

$\mathbf{x}$ is in some I in $\mathcal{P}$. Since $\mathbf{x}$ is not in B, it must be in $\text{int}(I)$. Hence I is in $\mathcal{P}'$. This shows that $\mathcal{P}'$ covers D'_ϵ. So we have

$$\epsilon \bar{m}(D'_\epsilon) \leq \epsilon \sum_{I \in \mathcal{P}'} m(I) \leq \sum_{I \in \mathcal{P}'} \text{osc}_I(f)\, m(I) = U(f, \mathcal{P}') - L(f, \mathcal{P}') \leq U(f, \mathcal{P}) - L(f, \mathcal{P}). \tag{5.22}$$

Now $D_\epsilon \setminus D'_\epsilon \subseteq B$, so $\bar{m}(D_\epsilon \setminus D'_\epsilon) \leq \bar{m}(B) = 0$. This implies that $\bar{m}(D_\epsilon) = \bar{m}(D'_\epsilon)$. $\square$

Lemma 5.22 *For every $\epsilon > 0$ and every $\kappa > 0$ there exists a partition $\mathcal{P}$ such that*

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) \leq \text{osc}_C(f)(\bar{m}(D_\epsilon) + \kappa)\bar{m}(D_\epsilon) + \epsilon \bar{m}(C). \tag{5.23}$$

The way the lemma is used is to note that for every $\epsilon > 0$ and every $\kappa > 0$ we have $U(f) - L(f) \leq \text{osc}_C(f)(\bar{m}(D_\epsilon) + \kappa) + \epsilon \bar{m}(C)$. Since $\kappa > 0$ is arbitrary we get $U(f) - L(f) \leq \text{osc}_C(f)\bar{m}(D_\epsilon) + \epsilon \bar{m}(C)$. Now suppose that $\bar{\mu}(\text{Disc}) = 0$. Then for each $\epsilon > 0$ we have $\bar{\mu}(D_\epsilon) = 0$. But since D_ϵ is closed, this says that $\bar{m}(D_\epsilon) = 0$. So the right hand side is $\epsilon \bar{m}(C)$. Since $\epsilon > 0$ is arbitrary, this implies that $U(f) - L(f) = 0$.

Proof: The first term comes from the points with large oscillation, and the second term comes from the points with small oscillation. To deal with the first term, consider $\kappa > 0$. From the definition of $\bar{m}(D_\epsilon)$ there is a finite closed cover $\mathcal{I}$ of D_ϵ such that

$$\sum_{I \in \mathcal{I}} m(I) < \bar{m}(D_\epsilon) + \kappa. \tag{5.24}$$

One can thicken the cells in $\mathcal{I}$ so that each of them is open and the same estimate is satisfied. The union of the cells in the new $\mathcal{I}$ is open, and D_ϵ is a subset of this open set. One can then take the closures of the cells and get a finite closed cell cover satisfying this same estimate. By removing overlaps we can get a nonoverlapping finite closed cell family $\mathcal{P}'$ such that every cell is a subset of C. Let A' be the union of the cells in $\mathcal{P}'$. Then D_ϵ is in the interior of the closed set $A' \subseteq C$. Furthermore,

$$\bar{m}(A') = \sum_{I \in \mathcal{P}'} m(I) < \bar{m}(D_\epsilon) + \kappa. \tag{5.25}$$

Let A'' be the closure of $C \setminus A'$. It is a compact set whose intersection with D_ϵ is empty. Thus for every point $\mathbf{x}$ in A'' there is a closed bounded cell I with $\mathbf{x}$ in its interior and such that $\text{osc}_I(f) < \epsilon$. The collection of these cells forms an open cover of A''. Hence there is a finite subcover $\mathcal{I}''$. The cells $I \cap A''$ may be subdivided into subcells in such a way as to form a partition $\mathcal{P}''$ of A''. For each I in $\mathcal{P}''$ we have $\text{diam} f(I) < \epsilon$.

Now take the partition $\mathcal{P} = \mathcal{P}' \cup \mathcal{P}''$ of C. Then

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) = \sum_{I \in \mathcal{P}'} \text{osc}_I(f) m(I) + \sum_{I \in \mathcal{P}''} \text{osc}_I(f) m(I). \tag{5.26}$$

Now

$$\sum_{I \in \mathcal{P}'} \mathrm{osc}_I(f) m(I) \le \sum_{I \in \mathcal{P}'} \mathrm{osc}_C(f) m(I) \le \mathrm{osc}_I(f)(\bar{m}(D_\epsilon) + \kappa). \tag{5.27}$$

Also

$$\sum_{I \in \mathcal{P}''} \mathrm{osc}_I(f) m(I) \le \sum_{I \in \mathcal{P}''} \epsilon m(I) = \epsilon \bar{m}(A'') \le \epsilon \bar{m}(C). \tag{5.28}$$

This gives the result. □

This treatment relies heavily on unpublished notes on the Riemann integral by Mariusz Wodzicki [20].

The Lebesgue theorem has implications for functions that are restricted to complicated subsets. Let 1_A be the indicator function of A, equal to 1 on A and 0 on its complement. Then the discontinuity set of the indicator function 1_A is the boundary of A, that is, $\mathrm{Disc}(1_A) = \mathrm{bdy}(A)$.

If f is a bounded function defined on a bounded closed non-degenerate cell C, and $A \subseteq C$, then the Riemann integral of f over A is defined to be the integral of $f\,1_A$, when that Riemann integral exists. (This may also be taken to be the definition of a function that is defined only on A.) In particular, if f is a bounded function on C, and if $A \subseteq C$, then the Riemann integral of f over A exists if and only if $\mathrm{Disc}(f\,1_A)$ has measure zero. However $\mathrm{Disc}(f1_A) \subseteq \mathrm{Disc}(f) \cup \mathrm{bdy}(A)$. So if both $\mathrm{Disc}(f)$ and $\mathrm{bdy}(A)$ have measure zero, then f is integrable over A.

All this applies to the case when $f = 1$ on the bounded cell C. Then 1_A is discontinuous precisely on $\mathrm{bdy}(A)$. So 1_A is integrable if and only if $\mathrm{bdy}(A)$ has measure zero. This is precisely the situation when A is Jordan measurable. The integral of 1_A is then the content $m(A)$ of A.

While the outer content of A is defined for arbitrary subsets, the content of A is only defined when A is Jordan measurable. When a set A is Jordan measurable, its content $m(A)$ is the same as the outer content $\bar{m}(A)$. (And in this case this is the same as the outer measure $\bar{\mu}(A)$.) The point of restricting to Jordan measurable sets is that the content on Jordan measurable sets is additive, while the outer content on arbitrary sets is only subadditive. (The outer measure on arbitrary sets is also only subadditive, but in this case it takes some effort to find examples where additivity fails.)

5.4 Almost everywhere

Often we say that a property depending on $\mathbf{x}$ holds for almost every $\mathbf{x}$ if the set of $\mathbf{x}$ for which it fails has measure zero. Thus a function f on a bounded cell C is Riemann integrable if and only if it is bounded and $f(\mathbf{x})$ is continuous for almost every $\mathbf{x}$.

Sometimes we just say that the property holds almost everywhere. A function f on a bounded cell C is Riemann integrable if and only if it is bounded and also continuous almost everywhere.

Theorem 5.23 *Let $f \ge 0$ be a function with $I(f) = 0$. Then $f(\mathbf{x}) = 0$ for almost every $\mathbf{x}$.*

Proof: Suppose $f \geq 0$ and $I(f) = 0$. Let N be the set of $\mathbf{x}$ such that $f(\mathbf{x}) > 0$. If f is continuous at $\mathbf{x}$ in N, then $I(f) > 0$, which is ruled out. So $N \subseteq \text{Disc}$. Since $\bar{\mu}(\text{Disc}) = 0$, it follows that $\bar{\mu}(N) = 0$. $\square$

Corollary 5.24 *Let f, g be two Riemann-integrable functions with $I(|f - g|) = 0$. Then $f(\mathbf{x}) = g(\mathbf{x})$ for almost every $\mathbf{x}$.*

5.5 Mapping sets of measure zero

In computing with sets of measure zero it is sometimes useful to use balls instead of cells. In the following we consider closed balls $B(\mathbf{a}, r) = \{\mathbf{x} \in \mathbf{R}^n \mid |\mathbf{x}-\mathbf{a}| \leq r$, with $r > 0$.

Lemma 5.25 *Consider a subset $A \subset \mathbf{R}^n$. Then A has measure zero if and only if for every $\epsilon > 0$ there is a sequence of closed balls B_k with $A \subseteq \bigcup_{k=1}^{\infty} B_k$ and $\sum_{k=1}^{\infty} \text{vol}_n(B_k) < \epsilon$. For every $\kappa > 0$, it is possible to impose the additional requirement that the balls all have radius less than κ.*

Proof: Up to now measure zero has been defined using covering by non-degenerate bounded cells. First we show that we could use instead coverings by closed cubes. Indeed, if we have a non-degenerate cell, then it has a side with least length $L > 0$. This cell is a subset of a bigger closed cell all of whose side lengths are multiples of L. The bigger cell may be taken so that each side length is no more than 2 times the corresponding side length of the original cell. So the volume of the bigger cell is bounded by 2^n times the volume of the original cell. Furthermore, the bigger cell may be written as the union of closed cubes of side length L, and its volume is just the sum of the volumes of the individual cubes. So if we can cover by cells with small total volume, then we can also cover by closed cubes of small total volume. By further subdividing the cubes, one can make them each of side length smaller than (some small multiple) of κ.

Once we have the result for closed cubes, then we have it for closed balls. This is because a closed ball of radius r is a subset of a closed cube of side length $L = 2r$, and a closed cube of side length L is a subset of a closed ball of radius $r' = \sqrt{n}L/2$. $\square$

A function $\mathbf{h}$ is called Lipschitz continuous if there is a constant C such that it satisfies $|\mathbf{h}(\mathbf{x}') - \mathbf{h}(\mathbf{x})| \leq C|\mathbf{x}' - \mathbf{x}|$. A function $\mathbf{h}$ from an open subset U of $\mathbf{R}^n$ to $\mathbf{R}^n$ is called locally Lipschitz continuous if for every $\mathbf{x}$ there exists a neighborhood $V \subseteq U$ of $\mathbf{x}$ such that $\mathbf{h}$ restricted to V is Lipschitz continuous. It is not too hard to see that $\mathbf{h}$ is locally Lipschitz if and only if it is Lipschitz continuous on every compact subset K of U. A continuous function may map a set of measure zero to a set of non-zero measure zero. However this is impossible for a C^1 function. In fact, a C^1 function is always locally Lipschitz. So the following result is relevant.

Theorem 5.26 *Let $U \subseteq \mathbf{R}^n$ and let $\mathbf{h} : U \to \mathbf{R}^n$ be a locally Lipschitz function. Then $\mathbf{h}$ maps sets of measure zero to sets of measure zero.*

Proof: First consider a compact subset $K \subseteq U$. Consider $E \subseteq K$ with measure zero. There is a κ neighborhood of K that is also a subset of U, whose closure is also a compact subset of U. Let C be the Lipschitz constant for $\mathbf{h}$ on this set. Then $E \subset \bigcup_k B_k$, where the B_k are closed balls inside U of radius r_k and total volume $\sum_k \mathrm{vol}_n(B_k)$ arbitrarily small. If $\mathbf{x}$ is in E, then $\mathbf{x}$ is in some ball B_k with center $\mathbf{a}_k$ and radius r_k, and thus $\mathbf{h}(\mathbf{x})$ is in some ball B'_k with center $\mathbf{h}(\mathbf{a}_k)$ and radius $r'_k = Kr_k$. This shows that $h(E) \subseteq \bigcup_k B'_k$, where the B'_k are closed balls of radius $r'_k = Kr_k$. The total volume of the balls is bounded by K^n times the total volume of the original balls. This can be made arbitrarily small.

Next consider an arbitrary subset $F \subseteq U$ with measure zero. Consider a sequence of compact subsets K_n with union U. Let F_n be the intersection of F with K_n. Then each F_n has measure zero. Also, $h(F)$ is the union of the $h(F_n)$. Since each of the $h(F_n)$ has measure zero, then so does $h(F)$. $\square$

One consequence of the theorem is the result that if f is Riemann integrable, and g is one-to-one continuous with Lipschitz inverse, then $f \circ g$ is Riemann integrable. This is seen as follows. Let $E = g^{-1}(\mathrm{Disc}(f))$. Then E has measure zero. Suppose that $\mathbf{x}$ is not in E. Then $g(\mathbf{x})$ is a continuity point for f, so $\mathbf{x}$ is a continuity point for $f \circ g$. This establishes that $\mathrm{Disc}(f \circ g) \subseteq E$. Since $\mathrm{Disc}(f \circ g)$ has measure zero, the Lebesgue theorem says that $f \circ g$ must be Riemann integrable.

5.6 Sard's theorem

Sard's theorem gives a connection between ideas of differentiability and measure. It says that if $\mathbf{g}$ is a sufficiently smooth function from an open subset of $\mathbf{R}^n$ to $\mathbf{R}^m$, then the image of the set of critical points of $\mathbf{g}$ has measure zero. In other words, the set of critical values of $\mathbf{g}$ has measure zero. There are three cases, depending on whether $n = m$, $n < m$, or $n > m$. The first two are relatively easy. The third one is more difficult and will not be proved here.

Theorem 5.27 (Sard) *Let $A \subseteq \mathbf{R}^n$ be open, and let $\mathbf{g} : A \to \mathbf{R}^n$ be C^1. Let $B \subseteq A$ be the set of $\mathbf{x}$ in A for which $\mathbf{g}'(\mathbf{x})$ has rank less than n, that is, for which $\det \mathbf{g}'(\mathbf{x}) = 0$. Then $\mathbf{g}(B)$ has measure zero.*

Proof: Let C be a closed bounded non-degenerate cube with $C \subseteq U$. In fact, we may take C to be a cube with side length L. We first show that $f(B \cap C)$ has content zero. Partition C into N^n small cubes I, each of side length L/N. Suppose that one of these small cubes I intersects B. Then we can choose an $\mathbf{x}$ in I and in B. We want to show that when $\mathbf{x} + \mathbf{h}$ is in the same small cube I, then $\mathbf{g}(\mathbf{x} + \mathbf{h})$ is close to $\mathbf{g}(\mathbf{x})$, in fact, very close in at least one direction.

First we have the mean value theorem estimate

$$|\mathbf{g}(\mathbf{x} + \mathbf{h}) - \mathbf{g}(\mathbf{x})| \leq M|\mathbf{h}| \leq M\sqrt{n}(L/N)^n. \tag{5.29}$$

Here M is a bound on the norm of $\mathbf{g}'(\mathbf{x})$ with $\mathbf{x}$ in C. By the C^1 hypothesis and compactness M is finite.

Next, let $\epsilon > 0$. Since $\mathbf{g}(\mathbf{x})$ is assumed to be C^1, we can take N so large that $\mathbf{g}'(\mathbf{x})$ does not change by more than ϵ across any of the small cubes I. In particular, we have $\|\mathbf{g}'(\mathbf{x}+t\mathbf{h}) - \mathbf{g}'(\mathbf{x})\| \leq \epsilon$ for every t with $0 \leq t \leq 1$. By the integral form of the mean value theorem

$$\mathbf{g}(\mathbf{x}+\mathbf{h}) - \mathbf{g}(\mathbf{x}) - \mathbf{g}'(\mathbf{x})\mathbf{h} = \int_0^1 (\mathbf{g}'(\mathbf{x}+t\mathbf{h}) - \mathbf{g}'(\mathbf{x}))\mathbf{h}\, dt \tag{5.30}$$

Hence with $\mathbf{x}$ and $\mathbf{x}+\mathbf{h}$ in the small cube I, we have

$$|\mathbf{g}(\mathbf{x}+\mathbf{h}) - \mathbf{g}(\mathbf{x}) - \mathbf{g}'(\mathbf{x})\mathbf{h}| \leq \epsilon|\mathbf{h}| \leq \epsilon\sqrt{n}(L/N)^n. \tag{5.31}$$

Since $\mathbf{x}$ is in B, the values $\mathbf{g}'(\mathbf{x})\mathbf{h}$ range over a subspace V of dimension at most $n-1$. This means that when $\mathbf{x}+\mathbf{h}$ is also in the small cube I, then $\mathbf{g}(\mathbf{x}+\mathbf{h})$ is within $\epsilon\sqrt{n}(L/N)^n$ of the plane $\mathbf{g}(\mathbf{x}) + V$.

The conclusion of the two estimates is that when $\mathbf{x}$ is in B and in the small cube I, and $\mathbf{x}+\mathbf{h}$ is in the same small cube I, then $g(\mathbf{x}+\mathbf{h})$ is contained in a very flat cylinder. The base of this cylinder is a $n-1$ dimensional ball of radius $M\sqrt{n}(L/N)^n$. The height of the cylinder is $2\epsilon\sqrt{n}(L/N)^n$. So there is a constant c (depending on M and on n) such that the volume of the cylinder is bounded by $c(L/N)^n\epsilon$.

The result is that the volume of the image of each small cube I intersecting $B \cap C$ is bounded by $c(L/N)^n\epsilon$. There are at most N^n small cubes that intersect C, so the image of $B \cap C$ is covered by a set of volume not exceeding $N^n c(L/N)^n\epsilon = cL^n\epsilon$.

Since this works for arbitrary $\epsilon > 0$, we may conclude that the content of $g(B \cap C)$ is zero. In particular the measure of each $g(B \cap C)$ is zero. However we may write the open set A as the countable union of cubes C. It follows that $g(B)$ is the union of a countable number of sets of measure zero. Hence $g(B)$ has measure zero. □

Theorem 5.28 (Sard) *Let $A \subseteq \mathbf{R}^n$ be open. Let $n < m$, and let $\mathbf{g} : A \to \mathbf{R}^m$ with $n < m$ be C^1. Let $B \subseteq A$ be the set of points such that $\mathbf{g}'(\mathbf{x})$ has rank less than m, that is, $\det \mathbf{g}'^T(\mathbf{x})\mathbf{g}'(\mathbf{x}) = 0$. Then $B = A$, and in fact the image set $\mathbf{g}(A)$ has measure zero.*

Proof: Since $\mathbf{g}'(\mathbf{x})$ is an m by n matrix, it has rank at most n. So rank is surely less than m. Let $\mathbf{p}$ be the projection of $\mathbf{R}^m$ onto $\mathbf{R}^n$ defined setting the last $m-n+1$ components equal to zero. Let B be the set of $\mathbf{y}$ in $\mathbf{R}^m$ with $\mathbf{p}(\mathbf{y})$ in A. Let $\mathbf{h}(\mathbf{y}) = \mathbf{g}(\mathbf{p}(\mathbf{y}))$. Then $\mathbf{h} : B \to \mathbf{R}^n$ has the same image as $\mathbf{g}$. Furthermore, the derivative $\mathbf{h}'(\mathbf{y}) = \mathbf{g}'(\mathbf{p}(\mathbf{y}))\mathbf{p}'(\mathbf{y})$ has rank at most $n < m$. So by the previous theorem $\mathbf{h}(B) = \mathbf{g}(A)$ has measure zero. □

The theorem just proved deserves respect. Suppose that $A \subseteq \mathbf{R}^n$ and $\mathbf{g} : A \to \mathbf{R}^m$ with $n < m$ is continuous. One might think that the image set $\mathbf{g}(A)$ would have measure zero. However this is not true in general, as is shown by the existence of space-filling curves.

Theorem 5.29 (Sard) *Let $A \subseteq \mathbf{R}^n$ be open. Let $m < n$ and suppose $\mathbf{g} : A \to \mathbf{R}^m$ be C^{n-m+1}. Let $B \subseteq A$ be the set of $\mathbf{x}$ in A for which $\mathbf{g}'(\mathbf{x})$ has rank less than m, that is, for which $\det \mathbf{g}'\mathbf{g}'^T = 0$. Then $\mathbf{g}(B)$ has measure zero.*

This last is the most difficult case of Sard's theorem. It says that the set of $\mathbf{w}$ for which there exists $\mathbf{x}$ with $\mathbf{g}(\mathbf{x}) = \mathbf{w}$ and $\mathbf{g}'(\mathbf{x})$ having rank less than m has measure zero. In other words, for almost every $\mathbf{w}$ the surface $\mathbf{g}(\mathbf{x}) = \mathbf{w}$ has only points with $\mathbf{g}'(\mathbf{x})$ having rank m.

5.7 Change of variables

The Sard theorem is relevant to the general change of variables theorem for unoriented integrals. Here we review several such theorems, without proofs.

One version of the theorem says that if $\mathbf{g}$ is a C^1 map of an open subset of $\mathbf{R}^n$ to $\mathbf{R}^n$, then

$$\int f(\mathbf{g}(\mathbf{x}))h(\mathbf{x})|\det \mathbf{g}'(\mathbf{x})|\, d^n\mathbf{x} = \int f(\mathbf{y}) \left(\sum_{\mathbf{g}(\mathbf{x})=\mathbf{y}} h(\mathbf{x}) \right) d^n\mathbf{y}. \tag{5.32}$$

In particular we can take h as the indicator function of the set A. This gives

$$\int_A f(\mathbf{g}(\mathbf{x}))|\det \mathbf{g}'(\mathbf{x})|\, d\mathbf{x} = \int f(\mathbf{y})\#\{\mathbf{x} \in A \mid \mathbf{g}(\mathbf{x}) = \mathbf{y}\}\, d\mathbf{y}. \tag{5.33}$$

We see that if $\mathbf{g}$ is not one-to-one, then the only modification is that we need to keep track of how many times $\mathbf{g}$ assumes a certain value. According to Sard's theorem, the critical values do not matter.

Another version of change of variable is the *pushforward formula*

$$\int f(\mathbf{g}(\mathbf{x}))h(\mathbf{x})\, d^n\mathbf{x} = \int f(\mathbf{y}) \left(\sum_{\mathbf{g}(\mathbf{x})=\mathbf{y}} h(\mathbf{x}) \frac{1}{|\det \mathbf{g}'(\mathbf{x})|} \right) d^n\mathbf{y}. \tag{5.34}$$

In this situation Sard's theorem is of no use; we need to require that $\mathbf{g}'(\mathbf{x})$ is non-singular. The quantity $f(\mathbf{g}(\mathbf{x}))$ on the left is the pullback of $f(\mathbf{y})$. The function of $\mathbf{y}$ given by the sum on the right hand side is the pushforward of $h(\mathbf{x})$.

5.8 Fiber integration

There is a very useful variant of the pushforward formula in the case when $\mathbf{y} = \mathbf{g}(\mathbf{x})$ defines a function from an open subset of $\mathbf{R}^n$ to $\mathbf{R}^m$ with $0 < m < n$. Let us assume that $\mathbf{g}(\mathbf{x})$ is C^1 and that the derivative $\mathbf{g}'(\mathbf{x})$ has rank m. Then for each $\mathbf{y}$ the equation

$$\mathbf{g}(\mathbf{x}) = \mathbf{y} \tag{5.35}$$

defines a regular surface (an embedded manifold) of dimension $k = n - m$. The function $\mathbf{y} = \mathbf{g}(\mathbf{x})$ thus defines a family of disjoint surfaces. The surface corresponding to a particular value of $\mathbf{y}$ is called the *fiber* over $\mathbf{y}$. It would seem reasonable to do an iterated integral over $\mathbf{y}$ and over the surfaces (the fibers).

It will be convenient to think of $\mathbf{g}(\mathbf{x}) = \mathbf{y}$ as a system $g_i(\mathbf{x}) = y_i$ for $i = k+1, \ldots, n$. Suppose we consider a point in the space. Then near that point there is a change of variables from $x_1, \ldots, x_n$ to $u_1, \ldots, u_k, y_{k+1}, \ldots, y_n$. The $u_1, \ldots, u_k$ parameterize each surface. We write this as

$$\mathbf{x} = \mathbf{F}(\mathbf{u}, \mathbf{y}). \tag{5.36}$$

The inverse relation is

$$\begin{aligned} \mathbf{u} &= \mathbf{G}_1(\mathbf{x}) \\ \mathbf{y} &= \mathbf{G}_2(\mathbf{x}) = \mathbf{g}(\mathbf{x}). \end{aligned} \tag{5.37}$$

Then

$$\int f(g(\mathbf{x}))h(\mathbf{x})\, d\mathbf{x} = \int f(\mathbf{y}) \int h(\mathbf{F}(\mathbf{u}, \mathbf{y}))|\det \mathbf{F}'(\mathbf{u}, \mathbf{y})|\, d\mathbf{u}\, d\mathbf{y}. \tag{5.38}$$

Thus the pushforward is

$$h_*(\mathbf{y}) = \int h(\mathbf{F}(\mathbf{u}, \mathbf{y}))|\det \mathbf{F}'(\mathbf{u}, \mathbf{y})|\, d\mathbf{u}. \tag{5.39}$$

Define the *fiber form*

$$\beta^{(\mathbf{y})} = |\det \mathbf{F}'(\mathbf{u}, \mathbf{y})|\, d\mathbf{u} = \frac{1}{|\det \mathbf{G}'(\mathbf{F}(\mathbf{u}, \mathbf{y}))|}\, d\mathbf{u}. \tag{5.40}$$

The answer is now

$$h_*(\mathbf{y}) = \int_{\mathbf{g}^{-1}(\mathbf{y})} h^{(\mathbf{y})}\beta^{(\mathbf{y})}. \tag{5.41}$$

Here $h^{(\mathbf{y})}$ is the restriction of $h(\mathbf{x})$ to the surface $\mathbf{g}^{-1}(\mathbf{y})$.

It is not hard to check that the fiber form is independent of the choice of coordinates. Suppose that $\bar{\mathbf{F}}(\mathbf{w}, \mathbf{y}) = \mathbf{F}(\phi(\mathbf{w}), \mathbf{y})$. We can think of the right hand side as $\mathbf{F}(\mathbf{u}, \mathbf{y})$ with $\mathbf{u} = \phi(\mathbf{w}), \mathbf{y} = \mathbf{y}$. We can compute

$$|\det \bar{\mathbf{F}}'(\mathbf{w}, \mathbf{y})|\, d\mathbf{w} = |\det \mathbf{F}'(\phi(\mathbf{w}, \mathbf{y}))||\det \phi'(\mathbf{w})|\, d\mathbf{w} = |\det \mathbf{F}'(\mathbf{u}, \mathbf{y})|\, d\mathbf{u}. \tag{5.42}$$

The differential forms are identical.

Another way of thinking of the fiber integral is in terms of delta functions. The relation is

$$\int \delta(\mathbf{g}(\mathbf{x}) - \mathbf{y})h(\mathbf{x})\, d\mathbf{x} = \int_{\mathbf{g}^{-1}(\mathbf{y})} h^{(\mathbf{y})}\beta^{(\mathbf{y})}. \tag{5.43}$$

The delta function on the left enforces m equations, leaving an $k = n - m$ dimensional integral on the right.

The discussion in this section has avoided various technical issues. The reader may consult the fascinating article by Ponomarev [16] for further information on this subject.

5.9 Probability

The central notion of probability is that of *expectation* of a function of a vector *random variable* $\mathbf{x}$. A common case is when the expectation is given by a *probability density* $\rho(\mathbf{x})$. This is a positive function with integral one. Say that $\mathbf{y} = \mathbf{g}(\mathbf{x})$ is a random variable that is a function of $\mathbf{x}$. Then it may or may not be the case that the expectation of $f(\mathbf{y}) = f(\mathbf{g}(\mathbf{x}))$ is given by a pushed forward probability density $\rho_*(\mathbf{y})$. When this is the case, we should have

$$\int f(\mathbf{y})\rho_*(\mathbf{y})\, d\mathbf{y} = \int f(\mathbf{g}(\mathbf{x}))\rho(\mathbf{x})\, d\mathbf{x}. \tag{5.44}$$

First consider the case when n random variables are mapped to n random variables. There $\rho(\mathbf{x})$ is a joint probability density for random variables $\mathbf{x}$, $\mathbf{g}(\mathbf{x})$ is a vector of n random variables, and $f(\mathbf{g}(\mathbf{x}))$ is a function of these n random variables. The right hand side is the expectation. If one wants to write this expectation in terms of the random variables $\mathbf{y} = \mathbf{g}(\mathbf{x})$, then one has to push forward the density. The change of variables formula suggests that the new density is

$$\rho_*(\mathbf{y}) = \sum_{\mathbf{g}(\mathbf{x})=\mathbf{y}} \frac{1}{|\det \mathbf{g}'(\mathbf{x})|}\rho(\mathbf{x}). \tag{5.45}$$

This only works when the regions where $\det \mathbf{g}'(\mathbf{x}) = 0$ can be neglected, and this is not always the case. If, for instance, there is a region C of non-zero volume with $\mathbf{g}(\mathbf{x}) = \mathbf{y}^*$ for $\mathbf{x}$ in C, then the extra contribution $f(\mathbf{y}^*)\int_C \rho(\mathbf{x})\, d\mathbf{x}$ must be added to the left hand side for the identity to be valid. Sard's theorem does nothing to help, since there is no longer a factor that vanishes on the set of critical points. Even though the set of critical values has measure zero, there can be a lot of probability on a set of measure zero.
Example: An example is when $n = 1$ and the density is $\rho(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x^2)$. This is the density for a standard normal (Gaussian) distribution. Let $y = x^2$. Then $\rho_*(y) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{y}} \exp(-\frac{1}{2}y)$ for $y > 0$. This density is that of a chi-squared distribution (with one degree of freedom). |

Next consider the case when n random variables are mapped to m random variables with $m < n$ by $\mathbf{y} = \mathbf{g}(\mathbf{x})$. The pushforward formula suggests that

$$\rho_*(\mathbf{y}) = \int_{\mathbf{g}^{-1}(\mathbf{y})} \rho^{(\mathbf{y})}\beta^{(\mathbf{y})}. \tag{5.46}$$

Example: Here is an example when $n = 2$ and $m = 1$. Consider the joint normal (Gaussian) distribution with density $\rho(x_1, x_2) = \frac{1}{2\pi} \exp(-\frac{1}{2}(x_1^2 + x_2^2))$. Let $y = x_2/x_1$ be the quotient. The fiber over y is the line $x_2 = yx_1$. One choice of parameter is $u = x_1$. Then we have $x_1 = u, x_2 = yu$. The Jacobian determinant is u, so the fiber form is $\beta^{(y)} = |u|\, du$. Then $\rho^{(y)}\beta^{(y)} = \frac{1}{2\pi} \exp(-\frac{1}{2}(1+y^2)u^2)|u|\, du$. The integral of this form over the fiber is $\rho(y) = \frac{1}{\pi}\frac{1}{1+y^2}$. This is the density for the Cauchy distribution. |

Perhaps the moral of the story is that one should calculate with the original density $\rho(\mathbf{x})$. In probability theory expectations (or measures) push forward in a routine way. When you try to express them in terms of densities, then the expressions are less pleasant. Densities are functions. Functions pull back with ease, but push forward with considerable difficulty.

5.10 The co-area formula

The co-area formula relates fiber integration to area. Suppose that $\mathbf{g}$ is a smooth function from an open subset of $\mathbf{R}^n$ to $\mathbf{R}^{n-k}$ with $k > 0$. We are interested in the k-dimensional surfaces $\mathbf{g}(\mathbf{x}) = \mathbf{y}$. The derivative $\mathbf{g}'(\mathbf{x})$ is a $n - k$ by n matrix. Its rows are $n - k$ independent co-vectors (forms) that vanish on the k-dimensional tangent space to the surface. Suppose $\mathbf{F}(\mathbf{u}, \mathbf{y})$ parameterizes the surfaces $\mathbf{y} = \mathbf{g}(\mathbf{x})$. Let $\mathbf{F}_1'(\mathbf{u}, \mathbf{y})$ be the n by k matrix of partial derivatives with respect to the $\mathbf{u}$ variables. The columns of this matrix represent a basis of tangent vectors to the surface. From

$$\mathbf{g}(\mathbf{F}(\mathbf{u}, \mathbf{y})) = \mathbf{y} \tag{5.47}$$

we see that

$$\mathbf{g}'(\mathbf{x})\mathbf{F}_1'(\mathbf{u}, \mathbf{y}) = 0. \tag{5.48}$$

This expresses in matrix form the fact that the $n-k$ independent row covectors in $\mathbf{g}'(\mathbf{x})$ are zero on the k independent column vectors in $\mathbf{F}_1'(\mathbf{u}, \mathbf{y})$.

Define the *co-area factor* $C(\mathbf{x})$ by

$$C(\mathbf{x}) = \sqrt{\det \mathbf{g}'(\mathbf{x})\mathbf{g}'(\mathbf{x})^T}. \tag{5.49}$$

This is the determinant of an $n-k$ square Gram matrix. Define the *area factor* by

$$A(\mathbf{u}, \mathbf{y}) = \sqrt{\det \mathbf{F}_1'^T(\mathbf{u}, \mathbf{y})\mathbf{F}_1'(\mathbf{u}, \mathbf{y})}. \tag{5.50}$$

The determinant is that of a k square Gram matrix. The simplest form of the *co-area formula* says that

$$C(\mathbf{x})|\det \mathbf{F}'(\mathbf{u}, \mathbf{y})| = A(\mathbf{u}, \mathbf{y}), \tag{5.51}$$

where $\mathbf{x} = \mathbf{F}(\mathbf{u}, \mathbf{y})$. In the language of differential forms, this formula says that

$$C^{(\mathbf{y})}\beta^{(\mathbf{y})} = \text{area}^{(\mathbf{y})}. \tag{5.52}$$

Here $C^{(\mathbf{y})} = C(\mathbf{F}(\mathbf{u}, \mathbf{y}))$, while $\beta^{(\mathbf{y})} = |\det \mathbf{F}'(\mathbf{u}, \mathbf{y})|\, d\mathbf{u}$ and $\text{area}^{(\mathbf{y})} = A(\mathbf{u}, \mathbf{y})\, d\mathbf{u}$.

Fiber integration gives an integral version of the co-area formula. This says that

$$\int f(\mathbf{g}(\mathbf{x}))h(\mathbf{x})C(\mathbf{x})\, d^n\mathbf{x} = \int f(\mathbf{y}) \left(\int_{\mathbf{g}(\mathbf{x})=\mathbf{y}} h(\mathbf{x})\text{area}_{n-m}^{(y)}(\mathbf{x}) \right) d^m\mathbf{y}. \tag{5.53}$$

In particular we can take h as the indicator function of the set A. This gives

$$\int_A f(\mathbf{g}(\mathbf{x}))C(\mathbf{x})\,d^n\mathbf{x} = \int f(\mathbf{y})\,\mathrm{area}^{(y)}_{n-m}(\{\mathbf{x}\in A \mid \mathbf{g}(\mathbf{x}) = \mathbf{y}\})\,d^m\mathbf{y}. \tag{5.54}$$

The co-area formula may also be thought of as a formula for area integrals in terms of delta functions. Thus

$$\int h(\mathbf{x})\delta(\mathbf{g}(\mathbf{x}) - \mathbf{y})C(\mathbf{x})\,d^n\mathbf{x} = \int_{\mathbf{g}(\mathbf{x})=\mathbf{y}} h(\mathbf{x})\mathrm{area}^{(y)}_k(\mathbf{x}). \tag{5.55}$$

In particular

$$\int_A \delta(\mathbf{g}(\mathbf{x}) - \mathbf{y})C(\mathbf{x})\,d^n\mathbf{x} = \mathrm{area}^{(y)}_{n-m}(\{\mathbf{x}\in A \mid \mathbf{g}(\mathbf{x}) = \mathbf{y}\}). \tag{5.56}$$

Example: The co-area formula may seem unfamiliar, but there is a case when it becomes quite transparent. Consider a single equation $g(\mathbf{x}) = y$ that defines a $k = n - 1$ dimensional hypersurface. In the integral form of the formula take $f(y) = H(s-y)$, where H is the indicator function of the positive real numbers. Also take $h(x) = 1$. The result is

$$\int_{g(\mathbf{x})\le s} |g'(\mathbf{x})|\,d\mathbf{x} = \int_{-\infty}^{s} \mathrm{area}(\{\mathbf{x} \mid g(\mathbf{x}) = y\})\,dy. \tag{5.57}$$

It follows that

$$\frac{d}{ds}\int_{g(\mathbf{x})\le s} |g'(\mathbf{x})|\,d\mathbf{x} = \mathrm{area}(\{\mathbf{x} \mid g(\mathbf{x}) = s\}). \tag{5.58}$$

This formula is an elementary relation between volume and area for an implicitly defined surface. |

We proceed to a more systematic development of the co-area formula. The assumption is that the family of surfaces $\mathbf{y} = \mathbf{g}(\mathbf{x})$ has a smooth parametric representation $\mathbf{x} = \mathbf{F}(\mathbf{u}, \mathbf{y})$. This means that

$$\mathbf{g}(\mathbf{F}(\mathbf{u}, \mathbf{y})) = \mathbf{y} \tag{5.59}$$

and the function $\mathbf{F}(\mathbf{u}, \mathbf{y})$ has an smooth inverse function $G(\mathbf{x})$ with components $\mathbf{G}_1(\mathbf{x}) = \mathbf{u}$ and $\mathbf{G}_2(\mathbf{x}) = \mathbf{g}(\mathbf{x}) = \mathbf{y}$.

Theorem 5.30 (Co-area formula) *Consider an open subset of* $\mathbf{x}$ *in* $\mathbf{R}^n$ *and* k*-dimensional regular parameterized surfaces* $\mathbf{g}(\mathbf{x}) = \mathbf{y}$ *in this region, where the* $\mathbf{y}$ *are in* $\mathbf{R}^{n-k}$*. Suppose that this family of surfaces has a smooth parametric representation* $\mathbf{x} = \mathbf{F}(\mathbf{u}, \mathbf{y})$*. Then*

$$\int f(\mathbf{g}(\mathbf{x}))h(\mathbf{x})C(\mathbf{x})\,d^n\mathbf{x} = \int f(\mathbf{y}) \int h(\mathbf{F}(\mathbf{u}, \mathbf{y}))A(\mathbf{u}, \mathbf{y})\,d^k\mathbf{u}\,d^{n-k}\mathbf{y}. \tag{5.60}$$

Proof: By a change of variable $\mathbf{x} = \mathbf{F}(\mathbf{u}, \mathbf{y})$ we have

$$\int f(\mathbf{g}(\mathbf{x}))h(\mathbf{x})C(\mathbf{x})\, d^n\mathbf{x} = \int f(\mathbf{y})h(\mathbf{F}(\mathbf{u},\mathbf{y}))C(\mathbf{F}(\mathbf{u},\mathbf{y}))|\det \mathbf{F}'(\mathbf{u},\mathbf{y})|\, d^k\mathbf{u}\, d^{n-k}\mathbf{y}. \tag{5.61}$$

So all that remains is to show that $C(\mathbf{F}(\mathbf{u},\mathbf{y}))|\det \mathbf{F}'(\mathbf{u},\mathbf{y})| = A(\mathbf{u},\mathbf{y})$, and this is equivalent to

$$\det \mathbf{g}'(\mathbf{x})\mathbf{g}'(\mathbf{x})^T\, (\det \mathbf{F}'(\mathbf{u},\mathbf{y}))^2 = \det \mathbf{F}_1'^T(\mathbf{u},\mathbf{y})\mathbf{F}_1'(\mathbf{u},\mathbf{y}), \tag{5.62}$$

where $\mathbf{x} = \mathbf{F}(\mathbf{u}, \mathbf{y})$.

The functions $\mathbf{G}_1(\mathbf{x})$ and $\mathbf{G}_2(\mathbf{x})$ have derivatives $\mathbf{G}_1'(\mathbf{x})$ and $\mathbf{G}_2'(\mathbf{x})$ that are k by n and $n-k$ by n matrices. The function $\mathbf{F}(\mathbf{u}, \mathbf{y})$ has partial derivatives with respect to the $\mathbf{u}$ and $\mathbf{y}$ variables given by matrices $\mathbf{F}_1'(\mathbf{u}, \mathbf{y})$ and $\mathbf{F}_2'(\mathbf{u}, \mathbf{y})$ that are n by k and n by $n-k$ matrices. By the chain rule we have $\mathbf{G}'(\mathbf{F}(\mathbf{u}, \mathbf{y}))\mathbf{F}'(\mathbf{u}, \mathbf{y}) = I$. In the following we write this in abbreviated form as

$$\mathbf{G}'\mathbf{F}' = I. \tag{5.63}$$

It follows that

$$\mathbf{G}'\mathbf{G}'^T\mathbf{F}'^T\mathbf{F} = I. \tag{5.64}$$

More explicitly, we could write the last two equations as

$$\begin{bmatrix} \mathbf{G}_1' \\ \mathbf{G}_2' \end{bmatrix} \begin{bmatrix} \mathbf{F}_1' & \mathbf{F}_2' \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{5.65}$$

and

$$\begin{bmatrix} \mathbf{G}_1'\mathbf{G}_1'^T & \mathbf{G}_1'\mathbf{G}_2'^T \\ \mathbf{G}_2'\mathbf{G}_1'^T & \mathbf{G}_2'\mathbf{G}_2'^T \end{bmatrix} \begin{bmatrix} \mathbf{F}_1'^T\mathbf{F}_1' & \mathbf{F}_1'^T\mathbf{F}_2' \\ \mathbf{F}_2'^T\mathbf{F}_1' & \mathbf{F}_2'^T\mathbf{F}_2' \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}. \tag{5.66}$$

There is a theorem in linear algebra that gives a relation between the determinants of submatrices of matrices that are inverse to each other. The statement and proof are given below. In this case it says that

$$\det \mathbf{G}_2'\mathbf{G}_2'^T \det \mathbf{F}'^T\mathbf{F}' = \det \mathbf{F}_1'^T\mathbf{F}_1'. \tag{5.67}$$

This can be written

$$\det \mathbf{g}'\mathbf{g}'^T(\det \mathbf{F}')^2 = \det \mathbf{F}_1'^T\mathbf{F}_1'. \tag{5.68}$$

This gives the required identity. □

In the co-area formula the function $\mathbf{y} = \mathbf{g}(\mathbf{x})$ sends an open subset of $\mathbf{R}^n$ into $\mathbf{R}^m = \mathbf{R}^{n-k}$, where $0 < m < n$. In the case when $\mathbf{g}$ is $C^{n-m+1} = C^{k+1}$ Sard's theorem gives useful information. It says that the image of the critical set where $C(\mathbf{x}) = 0$ is sent to a set of $\mathbf{y}$ of measure zero. Thus including these sets in the integration should make no difference. Actually, it turns out that there are versions of the co-area formula that apply in much greater generality. An account of these matters may be found in the book by Lin and Yang [8].

5.11 Linear algebra (block matrices)

This section presents a theorem on block matrices that is useful for the co-area theorem. The block matrices considered have four blocks, so they resemble in some respects 2 by 2 matrices with four entries. Some of the formulas for 2 by 2 matrices carry over to this situation, at least after appropriate modifications. We begin with a lemma about the determinant of a matrix for which one of the blocks is the zero matrix. The main theorem relates determinants of blocks of a matrix and of its inverse matrix.

Lemma 5.31 *Consider a block triangular block matrix*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}. \tag{5.69}$$

Then $\det A = \det A_{11} \det A_{22}$.

Proof: It is sufficient to consider the case when A_{11} is non-singular. Decompose

$$A = \begin{bmatrix} A_{11} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1} A_{12} \\ 0 & I \end{bmatrix}. \tag{5.70}$$

It is easy to work out each of the determinants on the right. So $\det A = \det A_{11} \cdot \det A_{22} \cdot 1$. □

Theorem 5.32 *Consider a block matrix A with inverse B, so that the product $AB = I$ has the form*

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}. \tag{5.71}$$

Then $\det A \det B = 1$. *Furthermore,*

$$\det A_{22} \det B = \det B_{11} \tag{5.72}$$

and

$$\det A_{11} \det B = \det B_{22}. \tag{5.73}$$

Proof: The inverse of A is given by the block Cramer's rule

$$\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} (A_{11} - A_{12} A_{22}^{-1} A_{21})^{-1} & -(A_{22} A_{12}^{-1} A_{11} - A_{21})^{-1} \\ -(A_{11} A_{21}^{-1} A_{22} - A_{12})^{-1} & (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} \end{bmatrix}. \tag{5.74}$$

There is a triangular factorization

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} A_{11} - A_{12} A_{22}^{-1} A_{21} & 0 \\ A_{22}^{-1} A_{21} & I \end{bmatrix}. \tag{5.75}$$

By the lemma above this gives $\det A = \det A_{22} \det B_{11}^{-1}$, which leads to the first result. We also have the triangular factorization

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ A_{21} & I \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1} A_{12} \\ 0 & A_{22} - A_{21} A_{11}^{-1} A_{12} \end{bmatrix}. \tag{5.76}$$

This gives $\det A = \det A_{11} \det B_{22}^{-1}$, which is equivalent to the second result. □

Bibliography

[1] Ilka Agricola and Thomas Friedrich, *Global analysis: Differential forms in analysis, geometry and physics*, Graduate Studies in Mathematics, no. 52, American Mathematical Society, Providence, Rhode Island, 2002.

[2] Dennis Barden and Charles Thomas, *An introduction to differentiable manifolds*, Imperial College Press, London, 2003.

[3] Robert L. Bryant, S. S. Chern, Robert B. Gardner, Hubert L. Goldschmidt, and P. A. Griffiths, *Exterior differential systems*, Mathematical Sciences Research Institute Publications, no. 18, Springer, New York, 1991, (online).

[4] William L. Burke, *Applied differential geometry*, Cambridge University Press, Cambridge, 1985.

[5] M. Crampin and F. A. E. Pirani, *Applicable differential geometry*, London Mathematical Society lecture notes 57, Cambridge University Press, Cambridge, 1986.

[6] Harley Flanders, *Differential forms with applications to the physical sciences*, Academic Press, New York, 1963.

[7] E. L. Ince, *Ordinary differential equations*, Dover Publications, New York, 1956.

[8] Fanghua Lin and Xiaoping Yang, *Geometric measure theory—an introduction*, International Press, Boston, MA, 2002.

[9] David Lovelock and Hanno Rund, *Tensors, differential forms, and variational principles*, Dover Publications, New York, 1989.

[10] W. A. J. Luxenburg, *Arzela's dominated convergence theorem for the riemann integral*, American Mathematical Monthly **78** (1971), 970–979.

[11] John Milnor, *Morse theory*, Annals of Mathematics Studies, no. 51, Princeton University Press, Princeton, New Jersey, 1969.

[12] Shigeyuki Morita, *Geometry of differential forms*, Translations of Mathematical Monographs, no. 201, American Mathematical Society, Providence, Rhode Island, 2001.

[13] Edward Nelson, *Tensor analysis*, Princeton University Press, Princeton, NJ, 1967.

[14] ———, *Topics in dynamics I: Flows*, Princeton University Press, Princeton, NJ, 1969.

[15] Ivan Netuka, *The change-of-variables theorem for the lebesgue integral*, Acta Universitatis Matthiae Belii **19** (2011), 37–42.

[16] S. P. Ponomarev, *Submersions and preimages of sets of measure zero*, Siberian Mathematical Journal **28** (1987), 153–163.

[17] Walter Rudin, *Principles of mathematical analysis*, third ed., McGraw-Hill International Editions, Princeton, NJ, 1969.

[18] Michael Spivak, *Calculus on manifolds: A modern approach to classical theorems of advanced calculus*, Addison-Wesley, Reading, Massachusetts, 1965.

[19] Erdoğan Şuhubi, *Exterior analysis: Using applications of differential forms*, Academic Press, Waltham, Massachusetts, 2013.

[20] Mariusz Wodzicki, *Notes on riemann integral, an annex to h104 etc.*, Lecture notes posted at https://math.berkeley.edu/ wodzicki/H104.F10/Integral.pdf, December 2 2010.

Mathematical Notation

Linear Algebra

$\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$	n-component column vectors
ω, μ	m-component row vectors
A, B, C	m by n matrices
A^T	transpose
A^{-1}	inverse
$\operatorname{tr}(A)$	trace
$\det(A)$	determinant
$\lvert\mathbf{x}\rvert = \sqrt{\mathbf{x}^T\mathbf{x}}$	Euclidean norm
$\lVert A\rVert$	Lipschitz norm
$\lVert A\rVert_2 = \sqrt{\operatorname{tr}(A^T A)}$	Euclidean norm

Multivariable functions

$\mathbf{f}$, $\mathbf{g}$, $\mathbf{h}$	functions from $E \subseteq \mathbf{R}^n$ to $\mathbf{R}^m$
$\mathbf{x} \mapsto \mathbf{f}(\mathbf{x})$	same as $\mathbf{f}$
$\mathbf{f}'$	derivative matrix function from open $E \subseteq \mathbf{R}^n$ to m by n matrices
$\mathbf{x} \mapsto \mathbf{f}'(\mathbf{x})$	same as $\mathbf{f}'$
$\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$	variables in $\mathbf{R}^n$
$\mathbf{y} = \mathbf{f}(\mathbf{x})$	$\mathbf{y}$ as a function $\mathbf{f}(\mathbf{x})$ of $\mathbf{x}$
$y_i = f_i(\mathbf{x})$	y_i as a function $f_i(\mathbf{x})$ of $\mathbf{x}$
$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \mathbf{f}'(\mathbf{x})$	derivative matrix (Jacobian matrix)
$\frac{\partial y_i}{\partial x_j} = \frac{\partial f_i(\mathbf{x})}{\partial x_j} = f'_{i,j}(\mathbf{x})$.	entry of derivative matrix
$d\mathbf{x} = dx_1 \wedge \cdots \wedge dx_n$	exterior product of differentials
$\frac{d\mathbf{y}}{d\mathbf{x}} = \det \frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \det \mathbf{f}'(\mathbf{x})$	determinant of derivative matrix (Jacobian determinant)
$\mathbf{g} \circ \mathbf{f}$	composite function
$(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x}))$	composite function of $\mathbf{x}$
$(\mathbf{g} \circ \mathbf{f})' = (\mathbf{g} \circ \mathbf{f})\mathbf{f}'$	chain rule
$(\mathbf{g} \circ \mathbf{f})'(\mathbf{x}) = \mathbf{g}(\mathbf{f}'(\mathbf{x}))\mathbf{f}'(\mathbf{x})$	chain rule as a function of $\mathbf{x}$
$\mathbf{p} = \mathbf{g}(\mathbf{u}), \mathbf{u} = \mathbf{f}(\mathbf{x})$	composite function expressed with variables
$\frac{\partial \mathbf{p}}{\partial \mathbf{x}} = \frac{\partial \mathbf{p}}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial \mathbf{x}}$	chain rule expressed with variables
f, g, h	functions from $E \subseteq \mathbf{R}^n$ to $\mathbf{R}$
$f_{,i\,j}(\mathbf{x})$	entries of Hessian matrix of second derivatives

Integration

I, $m(I)$	cell, volume of cell
$\mathcal{P}$	partition into cells
f_I	restriction
$L(f,\mathcal{P})$, $U(f,\mathcal{P})$	lower sum, upper sum
$L(f)$, $U(f)$, $I(f)$	lower integral, upper integral, integral
1_A	indicator function of subset A
$m(A) = I(1_A)$	content (volume) of A
$\mathrm{int}(A)$	interior of subset A
$\mathrm{bdy}(A)$	boundary of subset A
$\mathrm{osc}_A(f)$	oscillation on set A
$\mathrm{osc}_{\mathbf{x}}(f)$	oscillation at point $\mathbf{x}$
$\mathrm{Disc}(f)$	$\{\mathbf{x} \mid \mathrm{osc}_{\mathbf{x}}(f) > 0\}$
$\delta_\epsilon(\mathbf{x})$	family of approximate delta functions

Differential Forms

$\mathbf{x}, \mathbf{y}, \mathbf{u}$	coordinate systems
$s = h(\mathbf{x})$	scalar field
$X = \sum_{j=1}^n a_j \frac{\partial}{\partial x_j}$	vector field
$ds = \sum_{i=1}^n \frac{\partial s}{\partial x_i}\, dx_i$	differential of a scalar (an exact 1-form)
$\omega = \sum_{i=1}^n p_i\, dx_i$	differential 1-form
$\langle \omega \mid X \rangle = \sum_{i=1}^n p_i a_i$	scalar field from form and vector
θ	differential k-form
$\langle \theta \mid X_1, \ldots, X_k \rangle$	scalar field from form and vectors
$X \rfloor \theta$	interior product $k-1$ form
$\theta \wedge \beta$	exterior product of k-form with ℓ-form
$d\theta$	exterior derivative of θ (a $k+1$ form)
$\phi = (\mathbf{x} \leftarrow \mathbf{g}(\mathbf{u}))$	manifold mapping (parameterized surface)
$\phi^* h(\mathbf{x}) = h(\mathbf{g}(\mathbf{u}))$	pullback of a scalar field
$\phi^*\, dx_i = \sum_{\alpha=1}^k g'_{i,\alpha}(\mathbf{u})\, du_\alpha$	pullback of a basis differential
$\phi^*\theta$	pullback of a differential k-form
$\phi_* \frac{\partial}{\partial u_\alpha} = \sum_{i=1}^k g'_{i,\alpha}(\mathbf{u}) \frac{\partial}{\partial x_i}$	pushforward of a basis vector field
$\phi_* Y$	pushforward of a vector field
χ	chain
$\partial\chi$	boundary of chain
$\int_\chi \theta$	integral of form over chain

The Metric Tensor

$\mathbf{g} = \sum_{i=1}^n \sum_{j=1}^n g_{ij}\, dx_i\, dx_j$	metric tensor
g_{ij}	matrix entries of metric tensor (inner product on vectors)
G	matrix of metric tensor
g^{ij}	matrix entries of inverse matrix (inner product on forms)
G^{-1}	inverse of matrix tensor matrix
$\sqrt{g} = \sqrt{\det G}$	volume factor
$\mathrm{vol} = \sqrt{g}\, dx_1 \cdots dx_n$	volume form

$X \rfloor \text{vol}$	flux form
$\text{element} = \phi_1^* \text{vol}$	hypersurface element
$\phi^*(X \rfloor \text{vol}) = X \rfloor \text{element}$	flux through a hypersurface
$\nabla \cdot X = \text{div } X$	divergence
$\nabla s = \text{grad } s$	gradient
$\nabla^2 s = \text{div grad } s$	Laplacian
$\mathbf{g}^* = \sum_{\alpha=1}^k \sum_{\beta=1}^k g^*_{\alpha\beta}\, du_\alpha\, du_\beta$	metric tensor on surface
$\phi_* \frac{\partial}{\partial u_\alpha} = \sum_{i=1}^k \frac{\partial x_i}{\partial u_\alpha} \frac{\partial}{\partial x_i}$	αth tangent vector
$X^i_\alpha = \frac{\partial x_i}{\partial u_\alpha}$	components of αth tangent vector
$g^*_{\alpha\beta} = X^T_\alpha G X_\beta$	matrix entries of surface metric tensor
X	matrix of tangent vector components
$G^* = X^T G X$	Gram matrix of surface metric tensor
$\sqrt{g^*} = \sqrt{\det G^*}$	area factor
$\text{area} = \sqrt{g^*}\, du_1 \cdots du_k$	surface area form
$\text{area} = \lvert\text{element}\rvert$	hypersurface area form

Measure Zero

$m(A)$	content of Jordan measurable A
$\bar{m}(A)$	outer content of A
$\bar{\mu}(A)$	outer measure of A
$v_n = \frac{1}{n} \frac{2\pi^{n/2}}{\Gamma(n/2)}$	volume coefficient
$B_n(\mathbf{a}, r)$	open n ball of volume $v_n r^n$
$a_n = n v_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}$	area coefficient
$S_{n-1}(\mathbf{a}, r)$	$n-1$ sphere of area $a_{n-1} r^{n-1}$

Index