8

Multidimensional Calculus: Basics

We've finished the basic "linear algebra" part of the course; now we start a major part on "multidimensional calculus". This will include discussions of field theory, differential geometry and a little tensor analysis. Here is a thumbnail sketch of the topics:

1. *Basic Concepts:* Spaces of positions, curves, variations of "Euclidean-ness", coordinate systems, derivatives of position, partial derivatives and tangent vectors, "metric tensors", the chain rule, general change of coordinates, etc.

2. *The Differential Theory:* Velocity, acceleration and the Christoffel symbols, scalar and vector fields and their derivatives, gradients, divergence, curls and Laplacians, along with related topics.

3. *The Integral Theory:* Path integrals, potential theory, multidimensional integration by parts, and the classical integral theorems such as Gauss's (divergence) theorem, generalized notions of the divergence and curl, etc.

4. *A Little More About "Tensors":* About co- and contravariance, the metric tensor, "tensors" defined. (I might not cover this material as much as would be nice, so that more directly relevant material can be covered this term.)

Warning:

1. I might not cover the above topics in quite the order I just listed. For example, our discussion of "coordinate systems" will be spread out all through this chapter on "Basics".

2. Many authors and speakers tend to be especially sloppy with the notation, terminology and basic concepts of the material in this part of the course.[1] If you already are very knowledgable about the material, you can often compensate for this sloppiness, but if you are just learning it, this sloppiness can be confusing. We will try to be as un-sloppy as practical. After all, if you were really very knowledgable about this material, you probably would not be reading this.

[1] Arfken, Weber and Harris certainly are.

8.1 Fundamental Geometric Entities and Concepts

These are things we are assumed to "know" and do not really need to define.

Spaces of Points/Positions/Locations

The basic underlying geometric entity will always be some sort of "(multidimensional) space of points (with 'points' meaning 'locations' or 'positions')" on which we can visualize ourselves as (possibly very tiny) denizens with the ability to measure distances along curves, and, at any given location in the space, point in directions along curves through that location, and measure angles between directions being pointed out. Our interest is in developing the appropriate calculus for various types of functions on these spaces of positions.

For now, we will leave the concept of just what is a "space" of positions somewhat nebulous, and use the rest of this section to refine this concept.[2] I can, however, at least give you some examples of the sort of spaces we will be discussing:

1. A straight line.
2. A circle.
3. A (flat) plane.
4. A *portion* of a plane — possibly all of a plane except for one point.
5. A warped sheet in space, such as we visualize as the graph of $z = x^2 - y^2$.
6. A sphere (i.e., the *surface* of a ball).
7. The regular three-dimensional space we think we live in.
8. The regular three-dimensional space we think we live in, but with a mysterious hidden force warping space so that apparently parallel lines can intersect.

The first two examples (a straight line and a circle) are actually one-dimensional spaces, and not truly "multi-dimensional".

As we develop our notion of a space of positions, keep in mind that a "point" or "position" is not a vector, even if we can use vector to describe it (and that will not always be possible, and often not advisable when possible). We will have both "vector spaces" and "spaces of points". Rather than constantly repeating the phrase "space of points/positions", let us agree that anything referred to as simply "a space" is a space of points/positions/etc. and not a space of vectors (unless otherwise obviously indicated).

It will be helpful to have notation distinguishing between points in space, vectors and scalars. To denote a position, I will attempt to use either "boldface italics" $\boldsymbol{x}$ (in print) or a "rug tilde" $\underset{\sim}{x}$ (in handwriting). Observe how this notation differs from the corresponding notation for a vector:

$$\boldsymbol{x} \text{ or } \underset{\sim}{x} \text{ is a position.}$$

$$\mathbf{x} \text{ or } \overrightarrow{x} \text{ is a vector.}$$

[2] In some courses, what we will be calling "spaces" are officially known as "manifolds".

Be warned, this notation is not standard. In practice, points in space are often denoted by the same symbols used for scalars or for vectors. The notation I am using is a little more cumbersome, but will help us in our bookkeeping.

Curves and Angles

Geometrically, a curve in a space of points is just some "continuous string of locations" in that space. Visualize a curve as a possible path for a moving dot in that space. It may have one or two endpoints (or even none if it is a loop). We will assume that at each point $\boldsymbol{p}$ of a curve there are one or two unambiguous "directions along the curve" that can be pointed out. These are directions determined by the shape of the curve "right around $\boldsymbol{p}$", and not by where the curve ultimately leads.[3,4]

Let $\boldsymbol{p}$ be a point on a curve C other than an endpoint. We will say that the curve C is *smooth* at $\boldsymbol{p}$ if we can point out two directions along the curve at $\boldsymbol{p}$ with the angle between the directions being π (180^{o}). This means that, if we look closely at the curve around this point, then a small enough piece of C around $\boldsymbol{p}$ looks "fairly straight".

Naturally, if a curve is smooth at every (non-endpoint) point, then we will call it a *smooth* curve.

?►*Exercise 8.1:*

a: *Sketch what you consider to be a smooth curve (never mind our definition), pick a point on the curve and convince yourself that, at that point, there are exactly two "directions along the curve", and that the angle between them is* π *.*

b: *Sketch a curve which is not smooth at a point. What is the angle between the two directions along the curve at that point? Can you even make it so that that angle is* 0 *?*

Suppose now that we have two curves intersecting at a point $\boldsymbol{p}$ and that each is smooth at $\boldsymbol{p}$. Go ahead and sketch this. An angle of intersection here is an angle between directions of travel for one curve and a direction of travel for the other. Indicate those angles in your sketch. We will assume that, in general as in your sketch, there will be at most two values for these angles and that, if θ is the smallest such angle, then $\pi - \theta$ will be the other.

The maximum number N of curves in the space that can intersect at a point $\boldsymbol{p}$ at right angles to each other (and be smooth at $\boldsymbol{p}$) will be assumed to be a finite, positive integer N , and to not depend on the choice of $\boldsymbol{p}$. This integer N is the *dimension* of our space.

Curves and Distance

We will take for granted that we have a reasonable and well-defined notion of "arclength". In particular, we will assume:

1. We (rather, the denizens of the space) have tape measures. To be more precise, if we are given any two points $\boldsymbol{a}$ and $\boldsymbol{b}$ on a some curve C , then we can measure the distance (arclength) along the curve between the two points.

[3] Combined with magnitudes, these "directions" will become tangent vectors.

[4] Our insistence that these "directions along a curve at a point" exist means that we are limiting ourselves to relatively nice curves.

Moreover, if $\boldsymbol{x}$ is any point on the curve C between $\boldsymbol{a}$ and $\boldsymbol{b}$, then

$$\text{“distance along } C \text{ between } \boldsymbol{a} \text{ and } \boldsymbol{x}\text{”} + \text{“distance along } C \text{ between } \boldsymbol{x} \text{ and } \boldsymbol{b}\text{”} = \text{“distance along } C \text{ between } \boldsymbol{a} \text{ and } \boldsymbol{b}\text{”}$$

2. We all have the same tape measure and can agree on distances along curves. (Some of us may also have tape measures using different units, but we know how to convert arclength measurements taken using one to corresponding measurements taken with any other tape measure.)

3. Between every pair of points $\boldsymbol{a}$ and $\boldsymbol{b}$ is a curve of finite length.

Intuitively, the “distance between two points $\boldsymbol{a}$ and $\boldsymbol{b}$” should be the length of the shortest curve between $\boldsymbol{a}$ and $\boldsymbol{b}$. Technically, this curve might not exist (say, if it had to go through a point that was removed from the space). To get around this technicality, we will define the *distance between* $\boldsymbol{a}$ *and* $\boldsymbol{b}$ — denoted $\text{dist}(\boldsymbol{a}, \boldsymbol{b})$ — to be the largest real number that is less than or equal to the length of every curve between $\boldsymbol{a}$ and $\boldsymbol{b}$. If there truly is a shortest curve, then its length gives the distance. Typically, we also consider this “shortest curve” between $\boldsymbol{a}$ and $\boldsymbol{b}$ as as being the “straight line” between the two points. Between some points in some spaces (e.g., points on opposite sides of a sphere), this straight line might not be unique.

Notice that the distance between two points may depend on the space we are considering. For example, if $\boldsymbol{a}$ and $\boldsymbol{b}$ are two points on a circle in a plane, then

$$\text{“distance between } \boldsymbol{a} \text{ and } \boldsymbol{b} \text{ as points in the circle”} > \text{“distance between } \boldsymbol{a} \text{ and } \boldsymbol{b} \text{ as points in the plane”}$$

This becomes quite relevant when the motion of an object is constrained to staying in one subspace of another. It also means that we may occasionally want to denote the distance by $\text{dist}_{\mathcal{S}}(\boldsymbol{a}, \boldsymbol{b})$ to indicate that it is the distance as measured in the space $\mathcal{S}$. On the whole, though, we will try to avoid this.

Finally, let us simply observe the following:

1. For any two points $\boldsymbol{a}$ and $\boldsymbol{b}$ in a space, $\text{dist}(\boldsymbol{a}, \boldsymbol{b}) = \text{dist}(\boldsymbol{b}, \boldsymbol{a})$.

2. $\text{dist}(\boldsymbol{a}, \boldsymbol{b}) = 0 \iff \boldsymbol{a} = \boldsymbol{b}$.

?►*Exercise 8.2:* *Note the similarities and the differences between the basic geometric entities we have here, and those we had for traditional vector spaces.*

Later (section 8.4) we will greatly extend our discussion of curves and explore more deeply the relation between “arclength” and “distance”.

8.2 Euclidean-ness

In our basic assumptions involving curves, we assumed that smooth curves look pretty much like pieces of straight lines if you look at a small enough pieces. We will also assume that portions of our N-dimensional general spaces look pretty much like pieces of N-dimensional Euclidean spaces if we look at small enough portions. As discussed below, our spaces may be actually be Euclidean spaces, or they may merely be "nearly Euclidean".

In all the following, we will let $\mathcal{S}$ be some space of points in which we have interest. We will also assume that $\mathcal{S}$ is truly multidimensional, that is, its dimension is two or greater. That is because "Euclidean-ness" concerns triangles, and there are no nontrivial triangles in a one-dimensional space.

Before actually starting, let me apologize for some of the terminology. It's cumbersome, somewhat nonstandard, and we won't really use all of it. We just need it now to identify the different gradations of "Euclidean-ness".

Euclidean Spaces

Basic Requirements

For convenience, let's first define what it means to be "exactly Euclidean in a neighborhood" of a point $\boldsymbol{p}$ in a space $\mathcal{S}$. This means that there is a positive distance R such that, if we let $\mathcal{S}_{\boldsymbol{p},R}$ be the set of all points in $\mathcal{S}$ within a distance of R from $\boldsymbol{p}$,

$$\mathcal{S}_{\boldsymbol{p},R} = \{ \boldsymbol{x} \in \mathcal{S} \text{ with } \operatorname{dist}(\boldsymbol{p}, \boldsymbol{x}) < R \} \quad ,$$

then we can assume "Euclidean geometry" in $\mathcal{S}_{\boldsymbol{p},R}$. More precisely:

1. In $\mathcal{S}_{\boldsymbol{p},R}$, straight lines are well defined, unique, and yield the distance between pairs of points.

   By all this I mean that, for every $\boldsymbol{a}$ and $\boldsymbol{b}$ in $\mathcal{S}_{\boldsymbol{p},R}$:

   (a) There is exactly one curve in $\mathcal{S}_{\boldsymbol{p},R}$ between those points whose arclength equals $\operatorname{dist}(\boldsymbol{a}, \boldsymbol{b})$. This curve will be called the straight line between $\boldsymbol{a}$ and $\boldsymbol{b}$ and, for now, will be denoted by $\overline{\boldsymbol{ab}}$

   (b) If $0 < \gamma < 1$, there is a unique point $\boldsymbol{x}$ in $\overline{\boldsymbol{ab}}$ such that

   $$\begin{aligned}\text{"distance between } \boldsymbol{a} \text{ and } \boldsymbol{x}\text{"} &= \text{"distance between } \boldsymbol{a} \text{ and } \boldsymbol{x} \text{ along } \overline{\boldsymbol{ab}}\text{"} \\ &= \gamma \times \text{"distance between } \boldsymbol{a} \text{ and } \boldsymbol{b}\text{"}\end{aligned}$$

2. The angles of each triangle in $\mathcal{S}_{\boldsymbol{p},R}$ add up to π (180 degrees).

3. For these triangles, the laws of similar triangles hold. In particular, suppose $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ are the corners of a triangle in $\mathcal{S}_{\boldsymbol{p},R}$, and $0 < \gamma < 1$. Let $\boldsymbol{b}'$ and $\boldsymbol{c}'$ be the points on $\overline{\boldsymbol{ab}}$ and $\overline{\boldsymbol{ac}}$, respectively, satisfying

   $$\operatorname{dist}(\boldsymbol{a}, \boldsymbol{b}') = \gamma \operatorname{dist}(\boldsymbol{a}, \boldsymbol{b}) \qquad \text{and} \qquad \operatorname{dist}(\boldsymbol{a}, \boldsymbol{c}') = \gamma \operatorname{dist}(\boldsymbol{a}, \boldsymbol{c}) \quad ,$$

   as in figure 8.1. Then

   $$\operatorname{dist}(\boldsymbol{b}, \boldsymbol{c}) = \gamma \operatorname{dist}(\boldsymbol{b}', \boldsymbol{c}') \quad .$$

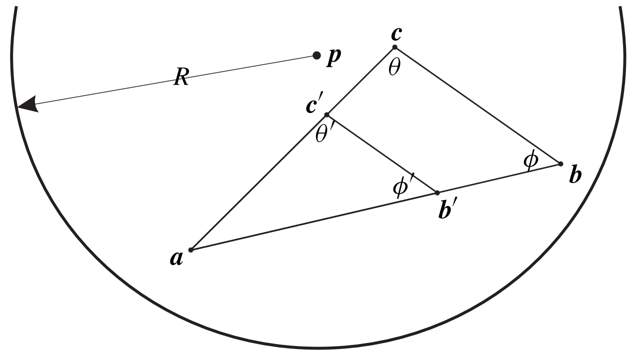

Figure 8.1: Similar triangles in an exactly Euclidean neighborhood of $\boldsymbol{p}$. Here $\text{dist}(\boldsymbol{a}, \boldsymbol{b}') = \gamma\,\text{dist}(\boldsymbol{a}, \boldsymbol{b})$ and $\text{dist}(\boldsymbol{a}, \boldsymbol{c}') = \gamma\,\text{dist}(\boldsymbol{a}, \boldsymbol{c})$ with $\gamma \approx 2/3$.

Moreover, if ϕ and θ are the angles of triangle $\boldsymbol{abc}$ at points $\boldsymbol{b}$ and $\boldsymbol{c}$, respectively, and ϕ' and θ' are the angles of triangle $\boldsymbol{ab'c'}$ at points $\boldsymbol{b}'$ and $\boldsymbol{c}'$, respectively, then

$$\phi' = \phi \qquad \text{and} \qquad \theta = \theta' \quad .$$

4. Parallelograms are well defined in $\mathcal{S}_{\boldsymbol{p},R}$. (This may really follow from the above two rules about triangles.)

Now, let me introduce some overly precise terminology:

- If the above is true for each position $\boldsymbol{p}$ in $\mathcal{S}$, then we can say that $\mathcal{S}$ is *locally exactly Euclidean*. Examples of locally exactly Euclidean spaces include a standard two-dimensional plane, a plane with a disk cut out, and that disk.
- If we can let $R \to \infty$ and all the above remains true, then $\mathcal{S}$ is a *complete Euclidean space*. For example, the two-dimensional plane is a complete Euclidean space, but a plane with a disk cut out is not. Neither is that disk.
- If $\mathcal{S}$ is locally exactly Euclidean, but not "complete", then it is an *incomplete Euclidean space*. Both a plane with a disk cut out and that disk are incomplete Euclidean space. Basically, an incomplete Euclidean space is simply a Euclidean space with holes.

Note that a "locally exactly Euclidean space" will be either a complete Euclidean space or an incomplete Euclidean space.

In practice, almost no one uses the terms just defined. Nor should you worry that much about memorizing these terms. When someone says "$\mathcal{S}$ is a Euclidean space", they may mean either that it is a complete Euclidean space (as just defined) or that it is a locally exactly Euclidean space. It depends on the person, the situation, and how careful that person is being. It's up to you to guess which they mean.

?►Exercise 8.3: *Let $\mathcal{S}$ be a sphere (surface only).*

a: *Convince yourself that the great circles on $\mathcal{S}$ are the "straight lines" in $\mathcal{S}$.*

b: *Verify that $\mathcal{S}$ is not locally exactly Euclidean.*

By the way, in practice, I will often denote a Euclidean space by $\mathcal{E}$ or, to indicate the dimension, by $\mathcal{E}^N$ where N is the dimension. Thus, $\mathcal{E}^2$ can be viewed as a "flat plane", and $\mathcal{E}^3$ can be thought of as the three-dimensional space we normally perceive about us.

An Important Observation

If $\mathcal{E}$ is a complete Euclidean space, then (as we saw when we discussed traditional vectors) we can take each pair of points $\boldsymbol{a}$ and $\boldsymbol{b}$, and let $\overrightarrow{\boldsymbol{ab}}$ be the vector from $\boldsymbol{a}$ to $\boldsymbol{b}$. That is, $\overrightarrow{\boldsymbol{ab}}$ is the vector of length

$$\left\|\overrightarrow{\boldsymbol{ab}}\right\| = \operatorname{dist}(\boldsymbol{a}, \boldsymbol{b})$$

pointing in the same direction as a person at $\boldsymbol{a}$ would point if that person were pointing at $\boldsymbol{b}$.

The set of all these $\overrightarrow{\boldsymbol{ab}}$'s generates a traditional vector space $\mathcal{V}$. With a little thought, you'll realize that the dimension of $\mathcal{V}$ will have to be the dimension of $\mathcal{E}$.

In particular, we can choose any point in $\mathcal{E}$, call it "the origin" $\boldsymbol{O}$, and associate each point $\boldsymbol{x}$ in $\mathcal{E}$ with the vector $\overrightarrow{\boldsymbol{Ox}}$ in $\mathcal{V}$. This vector, $\overrightarrow{\boldsymbol{Ox}}$, is properly called the position vector for $\boldsymbol{x}$ (with respect to the given origin). This, of course, can be very useful since it allows us to use everything developed for traditional vector spaces when doing stuff with complete Euclidean spaces.

Later on, we'll do the natural thing and define a coordinate system to match our favorite orthonormal basis.

?►*Exercise 8.4:* *What changes do we have to make to the above discussion if $\mathcal{E}$ is merely an incomplete Euclidean space?*

Warnings:

1. In practice, most authors grossly abuse the idea of matching a Euclidean space of positions with the corresponding vector space, and treat $\boldsymbol{x}$ and $\overrightarrow{\boldsymbol{Ox}}$ as being the same. For elementary applications this not particularly harmful, but if you want to do something advanced, say, use polar coordinates or use two different Cartesian coordinate systems or do calculus with functions on a sphere, then you must break the bad habits developed by pretending $\boldsymbol{x}$ and $\overrightarrow{\boldsymbol{Ox}}$ are the same.

2. In physics texts, it is common to use $\boldsymbol{r}$ or $\mathbf{r}$ to denote an arbitrary position or position-valued function or the corresponding position vector. I prefer using $\boldsymbol{x}$ or $\boldsymbol{p}$ instead of $\boldsymbol{r}$ (unless we are dealing with something like "central-force problems"), but expect to see $\boldsymbol{r}$ in Arfken, Weber and Harris. And at some point, I will probably have to break down and start using $\boldsymbol{r}$ to match some common integral notation.

Non-Euclidean Spaces

If a space $\mathcal{S}$ is not Euclidean, then it must be nonEuclidean. Typically, when we say a space is nonEuclidean, we do not mean merely that is isn't a complete Euclidean space — we mean that $\mathcal{S}$ is not even "locally exactly Euclidean".

However, we want things to be "nearly Euclidean", at least if we look at small enough regions. So we will assume that, for each and every point $\boldsymbol{p}$ in $\mathcal{S}$, there is a corresponding distance R

such that, if we restrict ourselves to

$$\mathcal{S}_{\boldsymbol{p},R} = \{ \boldsymbol{x} \in \mathcal{S} \text{ with } \operatorname{dist}(\boldsymbol{p}, \boldsymbol{x}) < R \} \quad ,$$

then:

1. "Straight" lines are well defined, unique, and yield the distance between pairs of points. That is, for every $\boldsymbol{a}$ and $\boldsymbol{b}$ in $\mathcal{S}_{\boldsymbol{p},R}$:
   - *(a)* There is exactly one curve in $\mathcal{S}_{\boldsymbol{p},R}$ between those points whose arclength equals $\operatorname{dist}(\boldsymbol{a}, \boldsymbol{b})$. This curve will be called the straight line between $\boldsymbol{a}$ and $\boldsymbol{b}$ and, for now, will be denoted by $\overline{\boldsymbol{ab}}$
   - *(b)* If $0 < \gamma < 1$, then there is a unique point $\boldsymbol{x}$ in $\overline{\boldsymbol{ab}}$ such that

$$\begin{aligned}\text{"distance between } \boldsymbol{a} \text{ and } \boldsymbol{x}\text{"} &= \text{"distance between } \boldsymbol{a} \text{ and } \boldsymbol{x} \text{ along } \overline{\boldsymbol{ab}}\text{"} \\ &= \gamma \times \text{"distance between } \boldsymbol{a} \text{ and } \boldsymbol{b}\text{"}\end{aligned}$$

2. The angles of each triangle in $\mathcal{S}_{\boldsymbol{p},R}$ add up to approximately π , with the deviation from π decreasing to zero as $R \to 0$.

3. Likewise, the laws of similar triangles hold approximately for these triangles, with the deviations from those laws decreasing to zero as $R \to 0$.

It would be appropriate to call such a space, *locally approximately Euclidean*, but no one does.[5] The phrase "locally Euclidean" is sometimes used for these spaces, despite the possible confusion with spaces that are truly Euclidean on the local level (e.g., "locally exactly Euclidean" spaces).

Do observe that the apparent errors mentioned above can be zero. So being "locally approximately Euclidean" does not exclude the possibility of actually being Euclidean.

So What Is a "Space"?

Henceforth, when we refer to $\mathcal{S}$ as being a "space" (and not a vector space), then all the assumptions and comments described from the start of this section (on page 8–2) are to be assumed except for the requirements listed for $\mathcal{S}$ being Euclidean. We will automatically assume the space is "nearly Euclidean" in the sense just described (i.e., the requirements hold for it being "locally approximately Euclidean"). Of course, if it is said that $\mathcal{S}$ is Euclidean, then also assume that the corresponding requirements hold. Whether the indicated space is a complete or an incomplete Euclidean space will usually not be a major issue. If it isn't "complete", we can at least view it as part of a complete Euclidean space.

Often, in fact, we may have two spaces $\mathcal{S}_1$ and $\mathcal{S}_2$ (not necessarily Euclidean) with each point in $\mathcal{S}_1$ also being a point in $\mathcal{S}_2$, and with the same basic notions of arclength and angular measurement being used for both. (Think of a warped two-dimensional surface floating in standard three-dimensional Euclidean space.) Then we say $\mathcal{S}_1$ is a subspace of $\mathcal{S}_2$. Some (mainly mathematicians) also say that $\mathcal{S}_1$ is "embedded" in $\mathcal{S}_2$. When we have such a situation, we can either view the physics/calculus occurring in subspace $\mathcal{S}_1$ from the viewpoint of those living in $\mathcal{S}_1$, who unable to see beyond their (sub)space; or we can view the physics/calculus occurring in subspace $\mathcal{S}_1$ from the more godlike viewpoint of those living in the larger space $\mathcal{S}_2$. Exactly which viewpoint we should take will depend on the situation.

[5] except your instructor

By the way, it can be shown that every N-dimensional space can be embedded in a $2N$-dimensional Euclidean space.[6] Thus, any two-dimensional space, no matter how odd the geometry may seem, can be viewed as a two-dimensional object in a four-dimensional Euclidean space. In particular, the Klein bottle which you may have heard of and which cannot be properly sketched in ordinary three-space, can be draw easily by inhabitants of ordinary four-space.

8.3 Coordinate Systems

Basics

To describe position in an N-dimensional space $\mathcal{S}$, we impose an N-dimensional *coordinate system*. That is, to each position $\boldsymbol{x}$ we assign an ordered N-tuple of real numbers called the *coordinates* of $\boldsymbol{x}$ with respect to the given system. If the coordinate system is one of the traditional ones — Cartesian, polar, cylindrical or spherical — then we will usually use the more traditional symbology used in Arfken, Weber and Harris.[7]— (x, y, z), (ρ, ϕ), (ρ, ϕ, z) or (r, θ, ϕ), respectively, with these traditional systems related by

$$\rho = \sqrt{x^2 + y^2} \quad , \quad r = x = \sqrt{x^2 + y^2 + z^2} = \sqrt{\rho^2 + z^2} \quad ,$$

$$x = \rho\cos(\phi) \quad , \quad y = \rho\sin(\phi) \quad \text{and} \quad z = r\cos(\theta) \quad .$$

For an arbitrary coordinate system, we will index the coordinates with *super*scripts,

$$(x^1, x^2, \ldots, x^N)$$

(Using superscripts instead of subscripts here is part of a set of conventions that will simplify bookkeeping later. This convention is standard when studying "coordinates, field theory and tensor analysis", even though the superscripts can be confused with exponentials.[8])

We will require that the assignment of coordinates is done in a reasonable, systematic manner so that nearby points have similar — but different — coordinates. We will denote a coordinate system (as opposed to just coordinates) by $\{(x^1, x^2, \ldots, x^N)\}$. Thus, if our space is a simple two-dimensional plane, then $\{(x, y)\}$ usually denotes the use of a standard Cartesian coordinate system, while the use of a normal polar coordinate system is typically indicated by $\{(\rho, \phi)\}$.

Strictly speaking, a coordinate system $\{(x^1, x^2, \ldots, x^N)\}$ is a mapping between the N-dimensional space $\mathcal{S}$ and $\mathbb{R}^N$. This means that we can view the coordinates as functions of position, writing, say,

$$\left(x^1, x^2, \ldots, x^N\right) = \left(x^1(\boldsymbol{x}), x^2(\boldsymbol{x}), \ldots, x^N(\boldsymbol{x})\right)$$

to indicate that the value of each x^k depends on the position $\boldsymbol{x}$. Alternatively, we can view position as a function of the coordinates, writing, say,

$$\boldsymbol{x} = \boldsymbol{x}\left(x^1, x^2, \ldots, x^N\right)$$

[6] More precisely, any N-dimensional "differentiable manifold" can be embedded in a $2N$-dimensional Euclidean space. The spaces we are considering are basically differentiable manifolds, though I may not have defined things completely and rigorously enough to actually prove it.

[7] See the figures on pages 188 and 182. Be warned, other texts use slightly different conventions. Indeed, you may be used to a slightly different convention.

[8] Even Arfken, Weber and Harris use it, sporatically.

to indicate that each position $\boldsymbol{x}$ is determined by the values of the coordinates $(x^1, x^2, \ldots, x^N)$. In practice, neither of these views is usually explicitly written out. Instead, people just abuse notation by writing

$$\boldsymbol{x} = \left(x^1, x^2, \ldots, x^N\right) \quad ,$$

identifying a position with the coordinates of that position using whatever coordinate system is at hand. Since this notation reinforces the erroneous idea that "position" equals "a description of position", we will use

$$\boldsymbol{x} \sim \left(x^1, x^2, \ldots, x^N\right)$$

to indicate that $\boldsymbol{x}$ is the position corresponding to coordinates $(x^1, x^2, \ldots, x^N)$. That way we may be less likely to derive

$$(1, 1) = \left(\sqrt{2}, \frac{\pi}{4}\right)$$

when dealing with Cartesian and polar coordinates for a position having Cartesian coordinatate $(1, 1)$.

Keep in mind that a coordinate system is not a natural geometric entity. It is a human imposed structure placed on our space so we can do computations. However, the geometry of the space will, itself, impose limits on the types of coordinate systems that can be imposed on it. Consider, for example, the different coordinate systems that can be imposed on a flat plane as opposed to what can be imposed on a sphere.

Basis-Based Coordinate Systems in Euclidean Spaces

The following only makes sense in Euclidean spaces!

Let $\mathcal{E}$ be an N-dimensional Euclidean space. As we discussed long ago, there is the corresponding N-dimensional vector space

$$\mathcal{V} = \left\{ \overrightarrow{\boldsymbol{ab}} \text{ where } \boldsymbol{a} \text{ and } \boldsymbol{b} \text{ are points in } \mathcal{E} \right\} \quad .$$

To construct a "basis-based" coordinate system for $\mathcal{E}$, pick any basis

$$\mathcal{B} = \{ \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N \}$$

for $\mathcal{V}$ you like, along with some point in $\mathcal{E}$. Call that point the *origin* and label as $\boldsymbol{O}$. We can then define the corresponding coordinate system $\{(x^1, x^2, \ldots, x^N)\}$ by

$$\boldsymbol{x} \sim \left(x^1, x^2, \ldots, x^N\right) \quad \Longleftrightarrow \quad \overrightarrow{\boldsymbol{Ox}} = \sum_{k=1}^{N} x^k \mathbf{b}_k \quad .$$

Equivalently,

$$\boldsymbol{x} \sim \left(x^1, x^2, \ldots, x^N\right) \quad \Longleftrightarrow \quad \left|\overrightarrow{\boldsymbol{Ox}}\right\rangle_{\mathcal{B}} = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^N \end{bmatrix} \quad .$$

The coordinate system is said to be *Cartesian* or *rectangular* if $\mathcal{B}$, the basis used, is orthonormal. When this is the case, I will probably denote the basis vectors by $\mathbf{e}_k$'s instead of $\mathbf{b}_k$'s.

The vector $\overrightarrow{\boldsymbol{Ox}}$ is the *position vector* (corresponding to position $\boldsymbol{x}$). This vector depends as much on the choice of the origin $\boldsymbol{O}$ as it does $\boldsymbol{x}$. Unfortunately, many authors will simply use

$\mathbf{x}$ to denote this position vector (or $\mathbf{r}$ to denote a generic position vector), completely ignoring the importance of the origin in determining this vector.

For the rest of this section, let us assume we have a Euclidean space with a basis-based coordinate system based on some basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N\}$ with position $\boldsymbol{O}$ as the origin, and let's see what we can immediately apply from what we developed for traditional vectors. For now, let us not assume $\mathcal{B}$ is orthonormal or even orthogonal.

First, let $\boldsymbol{p}$ and $\boldsymbol{q}$ be any two points in our Euclidean space, with

$$\boldsymbol{p} \sim \left(p^1, p^2, \ldots, p^N\right) \qquad \text{and} \qquad \boldsymbol{q} \sim \left(q^1, q^2, \ldots, q^N\right) \quad .$$

using our basis-based coordinate system. From the basic definition of addition for traditional vectors, we have

$$\overrightarrow{\boldsymbol{O}\boldsymbol{p}} + \overrightarrow{\boldsymbol{p}\boldsymbol{q}} = \overrightarrow{\boldsymbol{O}\boldsymbol{q}} \quad .$$

So the relation between the coordinates of positions $\boldsymbol{p}$ and $\boldsymbol{q}$, and the scalar components of the vector $\overrightarrow{\boldsymbol{p}\boldsymbol{q}}$ is given by

$$\overrightarrow{\boldsymbol{p}\boldsymbol{q}} = \overrightarrow{\boldsymbol{O}\boldsymbol{q}} - \overrightarrow{\boldsymbol{O}\boldsymbol{p}} = \sum_{k=1}^{N} q^k \mathbf{b}_k - \sum_{k=1}^{N} p^k \mathbf{b}_k = \sum_{k=1}^{N} \left(q^k - p^k\right) \mathbf{b}_k \quad .$$

Equivalently,

$$\left|\overrightarrow{\boldsymbol{p}\boldsymbol{q}}\right\rangle_{\mathcal{B}} = \begin{bmatrix} q^1 - p^1 \\ q^2 - p^2 \\ \vdots \\ q^N - p^N \end{bmatrix} \quad .$$

From the above, it immediately follows that

$$\begin{aligned}
\operatorname{dist}(\boldsymbol{p}, \boldsymbol{q}) &= \left\|\overrightarrow{\boldsymbol{p}\boldsymbol{q}}\right\| \\
&= \sqrt{\overrightarrow{\boldsymbol{p}\boldsymbol{q}} \cdot \overrightarrow{\boldsymbol{p}\boldsymbol{q}}} \\
&= \sqrt{\left[\sum_{j=1}^{N} \left(q^j - p^j\right) \mathbf{b}_j\right] \cdot \left[\sum_{k=1}^{N} \left(q^k - p^k\right) \mathbf{b}_k\right]} \\
&= \sqrt{\sum_{j=1}^{N} \sum_{k=1}^{N} \left(q^j - p^j\right) \left(q^k - p^k\right) \mathbf{b}_j \cdot \mathbf{b}_k} \quad .
\end{aligned}$$

Of course, if the coordinate system is Cartesian (so the basis $\mathcal{B}$ is orthonormal) then this reduces to the classic Pythagorean formula

$$\operatorname{dist}(\boldsymbol{p}, \boldsymbol{q}) = \sqrt{\sum_{k=1}^{N} \left(q^k - p^k\right)^2} \quad .$$

WARNING: Do NOT blindly apply vector operations to positions even when the space is Euclidean. You will usually end up with an operation that is not "well defined" in that the resulting position/vector will depend on the choice of the origin $\boldsymbol{O}$. An exception that will be of value to us is that the *difference* of two positions $\boldsymbol{p}$ and $\boldsymbol{q}$ is well defined provided you view this difference " $\boldsymbol{q} - \boldsymbol{p}$ " as the *vector* from $\boldsymbol{p}$ to $\boldsymbol{q}$. That is, we define the difference of two positions by

$$\text{“}\boldsymbol{q} - \boldsymbol{p}\text{”} = \overrightarrow{\boldsymbol{q} - \boldsymbol{p}} = \overrightarrow{\boldsymbol{pq}} \quad .$$

(This will be relevant in discussing derivatives.)

Along the same lines, you can specify a point in $\mathcal{E}$ by "adding" a vector $\mathbf{v}$ to another point $\boldsymbol{p}$. The point obtained $\boldsymbol{q}$ is simply the point such that $\overrightarrow{\boldsymbol{pq}} = \mathbf{v}$ (equivalently, $\overrightarrow{\boldsymbol{q} - \boldsymbol{p}} = \mathbf{v}$).

On the other hand, "adding" two positions using their coordinates makes no sense.

?►*Exercise 8.5:* *Show that "addition" is not well defined for positions. In particular, show that for two points*

$$\boldsymbol{p} \sim \left(p^1, p^2, \ldots, p^N\right) \qquad \text{and} \qquad \boldsymbol{q} \sim \left(q^1, q^2, \ldots, q^N\right)$$

in some Euclidean space with a basis-based coordinate system, the position and vector

$$\boldsymbol{s} \sim \left(p^1 + q^1, p^2 + q^2, \ldots, p^N + q^N\right) \qquad \text{and} \qquad \mathbf{s} = \sum_{k=1}^{N} \left(p^k + q^k\right) \mathbf{b}_k$$

will depend on the choice of $\boldsymbol{O}$ *. (And explain why that means "addition is not well defined for position.")*

?►*Exercise 8.6:* *Show that a polar coordinate system in the plane is* not *basis based.*

Hint: Pick two convenient points $\boldsymbol{p}$ *and* $\boldsymbol{q}$ *with polar coordinates* (r, θ) *and* (ρ, ϕ) *, respectively (actually, it's the polar coordinate values that should be 'conveniently chosen'), and show that there is no basis* $\{\mathbf{b}_1, \mathbf{b}_2\}$ *such that we have both*

$$\boldsymbol{p} = r\mathbf{b}_1 + \theta\mathbf{b}_2 \qquad \text{and} \qquad \boldsymbol{q} = \rho\mathbf{b}_1 + \phi\mathbf{b}_2 \quad .$$

Coordinate systems that are not basis based also have associated basis vectors, but these vectors vary from point to point. We will develop these associated basis sets later, after discussing curves and the notion of differentiating position, because, in general, our basis vectors at each point will come from differentiating position with respect to the different coordinates.

8.4 Describing Curves

In the following, all points/positions are in some space $\mathcal{S}$, not necessarily Euclidean.

Parametrizations

Now, let C be some curve in $\mathcal{S}$. A *parametrization* of C is a (continuous) position-valued function $\boldsymbol{x}(\tau)$ that traces out the curve as the *parameter* τ varies over some interval (a, b) (or $[a, b]$ or $(a, b]$ or $[a, b)$). There are always many different possible parametrizations for any given curve. If one has been chosen, then that curve is said to be a *parametrized* *curve*. In applications, of course, we like to use parametrized curves with particularly relevant parametrizations . For example, if C is the path of a missile, then you would probably like to parametrize it with

$$\mathbf{x}(\tau) \;=\; \text{position of the missile at time } \tau \qquad \text{where} \quad \tau_{\text{launch}} \le \tau \le \tau_{\text{hit}} \quad .$$

Also, in applications such as this, we are likely to replace the symbol τ with one more indicative of the actual variable, say, t for "time".

In practice, a position-valued function $\mathbf{x}(\tau)$ is usually described in coordinate form using whatever coordinate system is being used,

$$\mathbf{x}(\tau) \;\sim\; \left(x^1(\tau),\, x^2(\tau),\, \dots,\, x^N(\tau)\right)$$

Each $x^k(\tau)$ is a real-valued function which gives the k^{th} coordinate of $\mathbf{x}(\tau)$ for each value of τ .

It's worth noting that the curve C is, itself, a one-dimensional space of positions. Since you can use a parametrization $\mathbf{x}(\tau)$ to identify each point in the curve with a value of τ , you can view a parametrization as defining a coordinate system for that curve.

!►Example 8.1: *Suppose C is a straight line in a three-dimensional Euclidean space $\mathcal{E}$ passing through two points $\mathbf{a}$ and $\mathbf{b}$. Assume we have already chosen some Cartesian coordinate system $\{(x, y, z)\}$ for our Euclidean space, and, using this coordinate system,*

$$\mathbf{a} \;\sim\; (1, 0, 3) \qquad \text{and} \qquad \mathbf{b} \;\sim\; (6, 5, 4) \quad .$$

For each real number τ , let $\mathbf{x}(\tau)$ be given by $\overrightarrow{\mathbf{Ox}}(\tau) = \overrightarrow{\mathbf{Oa}} + \tau\overrightarrow{\mathbf{ab}}$. That is,

$$\mathbf{x} \;=\; \mathbf{x}(\tau) \;\sim\; (x(\tau),\, y(\tau),\, z(\tau))$$

is the point on this line with

$$\overrightarrow{\mathbf{ax}} \;=\; \tau\,\overrightarrow{\mathbf{ab}} \quad .$$

That is

$$\left|\overrightarrow{\mathbf{ax}}\right\rangle \;=\; \tau\left|\overrightarrow{\mathbf{ab}}\right\rangle \quad ,$$

which, in terms of the coordinates of the points $\mathbf{x}$, $\mathbf{a}$ and $\mathbf{b}$ is

$$\begin{bmatrix} x(\tau) - 1 \\ y(\tau) - 0 \\ z(\tau) - 3 \end{bmatrix} \;=\; \tau \begin{bmatrix} 6 - 1 \\ 5 - 0 \\ 4 - 3 \end{bmatrix} \;=\; \begin{bmatrix} 5\tau \\ 5\tau \\ \tau \end{bmatrix} \quad ,$$

and which we can also write as

$$\begin{aligned} (\,x(\tau) - 1,\; y(\tau) - 0,\; z(\tau) - 3\,) \;&=\; (\,\tau[6 - 1],\; \tau[5 - 0],\; \tau[4 - 3]\,) \\ &=\; (\,5\tau,\; 5\tau,\; \tau\,) \quad . \end{aligned}$$

Solving for the $x(\tau)$, $y(\tau)$ and $z(\tau)$, we see that the curve's parametrization is

$$\mathbf{x}(\tau) \;\sim\; (\,x(\tau),\, y(\tau),\, z(\tau)\,) \;=\; (\,1 + 5\tau,\; 5\tau,\; 3 + \tau\,) \quad .$$

Of particular interest are the coordinate curves with the obvious parametrizations. For example, the " X^1 coordinate curve through the point with coordinates $(0, 0, \dots, 0)$ " is naturally parametrized by

$$\mathbf{x}(\tau) \;\sim\; (\tau, 0, 0, \dots, 0) \quad ,$$

and, naturally, in practice, we would use the symbol x^1 instead of τ .

Orientation for Curves

A curve C is said to be *oriented* if, at each point on C, a "direction along C at that point" has been chosen. It is required that all these directions have been chosen so that, given any two points $\boldsymbol{a}$ and $\boldsymbol{b}$ on C, a particle will get from one point to the other by traveling along the curve and following the "direction of travel" given at each point. Implicit in this is that the chosen "direction of travel" never has the particle reversing course along the curve.

The general "direction of flow along the curve" indicated by all these chosen directions at points is called the *orientation* of the curve. It's also called the *positive direction* along the curve. Obviously, a curve can only have two possible orientations. If the curve has two endpoints, then choosing an orientation for the curve is the same as choosing one of the endpoints as the "starting point" and the other as the "ending point" for the curve.

If a curve C is oriented, then we will insist that any parametrization $\boldsymbol{x}(\tau)$ traces out the curve in the positive direction as τ increases.

Arclength Parametrization for Curves

It is often convenient to assume that a given curve C has been parametrized by its arclength. This implicitly requires that the curve has been oriented. We take a point $\boldsymbol{x}_0$ on the curve and, for each possible $s \geq 0$, we let

$$\boldsymbol{x}(s) = \text{the point on } C \text{ reached by traveling a distance of } s \text{ along the curve in the positive direction from } \boldsymbol{x}_0$$

Naturally, for each possible $s < 0$, we let

$$\boldsymbol{x}(s) = \text{the point on } C \text{ reached by traveling a distance of } s \text{ along the curve in the negative direction from } \boldsymbol{x}_0$$

The possible values of s will be limited only by the distances along the curve from $\boldsymbol{x}_0$ to any endpoint.

The symbol "s" used above is traditional. It is worth noting that we have two "distances" between the points $\boldsymbol{x}_0$ and $\boldsymbol{x}(s)$:

1. the distance as measured in the space, $\text{dist}(\boldsymbol{x}_0, \boldsymbol{x}(s))$, and
2. the distance as measured along the curve, which will simply be s (or $|s|$) by the definition of $\boldsymbol{x}(s)$.

The requirement that "small pieces of smooth curves look like straight lines" can now be refined to the following: For small positive values of arclength $s = \Delta s$,

$$\frac{\text{dist}(\boldsymbol{x}_0, \boldsymbol{x}(\Delta s))}{\Delta s} = \frac{\text{distance from } \boldsymbol{x}_0 \text{ to } \boldsymbol{x}(\Delta s)}{\text{arclength along } C \text{ from } \boldsymbol{x}_0 \text{ to } \boldsymbol{x}(\Delta s)} \approx 1 \quad ,$$

with the deviation from 1 decreasing as $\Delta s \to 0$. Letting $\Delta s \to 0$, the above becomes

$$\frac{d}{ds} \text{dist}(\boldsymbol{x}_0, \boldsymbol{x}(s)) = 1 \quad .$$

What this tells us is that, for all practical purposes, any derivative involving the distance between points will equal the same derivative with that distance replaced with arclength s along as suitable curve. In particular, if $\boldsymbol{x}_0$ is a point on a smooth curve with any parametrization $\boldsymbol{x}(\tau)$, then, applying the elementary calculus chain rule, we have

$$\frac{d}{d\tau}\operatorname{dist}(\boldsymbol{x}_0, \boldsymbol{x}) = \frac{ds}{d\tau}\frac{d}{ds}\operatorname{dist}(\boldsymbol{x}_0, \boldsymbol{x}) = \frac{ds}{d\tau}\cdot 1 = \frac{ds}{d\tau} \quad . \tag{8.1}$$

Note that, in writing the above derivative, we are also treating s as a function of τ. It should be clear that, in fact, $^{ds}/_{d\tau}$ is the *speed* at which $\boldsymbol{x}(\tau)$ is tracing out the curve. Moreover, to find the arclength of the curve from $\boldsymbol{x}(\tau_1)$ to $\boldsymbol{x}(\tau_2)$, we simply evaluate the integral

$$\text{arclength} = \int_{\tau_1}^{\tau_2} \frac{ds}{d\tau}\,d\tau \quad .$$

This speed and arclength are quantities naturally measured by those on the curve using speedometers and tape measures. We will need to develop coordinate formulas to compute them. In general, using these coordinate formulas will not be nearly as simple as using tape measures or speedometers.

8.5 Parametrized Surfaces in $\mathcal{E}^3$

Since we will later have applications involving "surfaces in a three-dimensional Euclidean space", and since such surfaces can serve as simple examples of nonEuclidean spaces, let us say a few words about such things.

Parametrizations

Let $\mathcal{S}$ be a surface in $\mathcal{E}^3$ (remember $\mathcal{E}^3$ is simply three-dimensional Euclidean space), and let $\{(x^1, x^2, x^3)\}$ be any coordinate system for $\mathcal{E}^3$. To parametrize $\mathcal{S}$ we need a region $\mathcal{R}$ of $\mathbb{R}^2$ and a (continuous) position-valued function of two variables

$$\boldsymbol{x}(\sigma, \tau) \sim \left(x^1(\sigma, \tau),\ x^2(\sigma, \tau),\ x^3(\sigma, \tau)\right)$$

such that each point in the surface is given by $\boldsymbol{x}(\sigma, \tau)$ for some (σ, τ) in $\mathcal{R}$. This function $\boldsymbol{x}(\sigma, \tau)$, along with its domain $\mathcal{R}$ in $\mathbb{R}^2$, is a *parametrization* of the surface $\mathcal{S}$.

A few simple observations are in order, here:

1. If we fix σ at some single value σ_0 and let τ vary, then

$$\boldsymbol{p}(\tau) = \boldsymbol{x}(\sigma_0, \tau) \sim \left(x^1(\sigma_0, \tau),\ x^2(\sigma_0, \tau),\ x^3(\sigma_0, \tau)\right)$$

traces out a curve in the surface. Changing the value of σ_0 changes the curve. With a little thought, you can see that this surface, $\mathcal{S}$, can be viewed as the surface swept out by the curves $\boldsymbol{p}(\tau) = \boldsymbol{x}(\sigma_0, \tau)$ as we vary the parameter σ_0.

2. If we fix τ at some single value τ_0 and let σ vary, then

$$\boldsymbol{q}(\sigma) = \boldsymbol{x}(\sigma, \tau_0) \sim \left(x^1(\sigma, \tau_0),\ x^2(\sigma, \tau_0),\ x^3(\sigma, \tau_0)\right)$$

traces out a curve in the surface. Changing the value of τ_0 changes the curve. With a little thought, you can see that this surface, $\mathcal{S}$, can be viewed as the surface swept out by the curves $\boldsymbol{q}(\sigma) = \boldsymbol{x}(\sigma, \tau_0)$ as we vary the parameter τ_0.

3. The parameters of the parametrization σ and τ can serve as a natural coordinate system $\{(\sigma,\tau)\}$ for the surface.

!► ***Example 8.2 (A sphere):*** *Consider a sphere of radius 3 about some point. Let's call that sphere $\mathcal{S}_3$. Using a spherical coordinate system $\{(r,\theta,\phi)\}$ centered at the sphere's center, we can describe each point $\boldsymbol{x}$ in this sphere by*

$$\boldsymbol{x} \sim (r,\theta,\phi) \quad \text{with} \quad r = 3 \quad , \quad 0 \le \theta \le \pi \quad \text{and} \quad 0 \le \phi \le 2\pi \quad .$$

So

$$\boldsymbol{x}(\theta,\phi) \sim (3,\theta,\phi) \quad \text{for} \quad 0 \le \theta \le \pi \quad \text{and} \quad 0 \le \phi \le 2\pi$$

is a natural parameterization of this sphere, and $\{(\theta,\phi)\}$ is a natural coordinate system for this two-dimensional space, with

$$\boldsymbol{x} \sim (\theta,\phi) \quad \text{in the sphere}$$

being the point given by

$$\boldsymbol{x} \sim (3,\theta,\phi) \quad \text{in} \quad \mathcal{E}^3$$

using the given spherical coordinates.

Surfaces Given by Graphs

One easy way to generate a surface $\mathcal{S}$ is as the graph of a real-valued function of two variables,

$$z = f(\sigma,\tau) \quad \text{for} \quad (\sigma,\tau) \text{ in some region } \mathcal{R} \text{ of } \mathbb{R}^2$$

assuming a standard "XYZ" Cartesian coordinate system with $\{(\sigma,\tau)\}$ being a convenient coordinate system for the XY–plane (usually Cartesian, $\{(x,y)\}$, or polar, $\{(\rho,\phi)\}$). This gives us $\mathcal{S}$ as "surface over the XY–plane" (though, in fact, this surface can dip below the XY–plane). This surface is parameterized by

$$\boldsymbol{x}(\sigma,\tau) \sim (\sigma,\ \tau,\ z) \quad \text{with} \quad z = f(\sigma,\tau) \quad \text{and} \quad (\sigma,\tau) \in \mathcal{R} \quad ,$$

which we can write more concisely as

$$\boldsymbol{x}(\sigma,\tau) \sim (\sigma,\ \tau,\ f(x,y)) \quad \text{for} \quad (\sigma,\tau) \in \mathcal{R} \quad .$$

!► ***Example 8.3:*** *Consider the surface $\mathcal{S}$ given by*

$$z = x^2 + y^2 \quad \text{for} \quad -\infty < x < \infty \quad \text{and} \quad -\infty < y < \infty$$

using the standard Cartesian coordinate system on the XY–plane. You should recognize this surface $\mathcal{S}$ as a circular paraboloid about the Z–axis. The corresponding parametrization (in terms of the corresponding Cartesian coordinate system for $\mathcal{E}^3$) is

$$\boldsymbol{x}(x,y) \sim \left(x,\ y,\ x^2+y^2\right) \quad \text{for} \quad -\infty < x < \infty \quad \text{and} \quad -\infty < y < \infty \quad .$$

This automatically defines a "$\{(x,y)\}$" coordinate system for $\mathcal{S}$, with each

$$\boldsymbol{x} \sim (x,y) \quad \text{in the surface } \mathcal{S}$$

being the point

$$\boldsymbol{x} \sim \left(x,\ y,\ x^2+y^2\right) \quad \text{in the thee-dimensional space } \mathcal{E}^3 \quad .$$

!►Example 8.4: *A possibly better way to describe that paraboloid from the last example is by*

$$z = \rho^2 \qquad \text{for} \quad 0 \leq \rho < \infty$$

using the standard polar coordinate system $\{(\rho, \phi)\}$ on the XY–plane. (Convince yourself that the graph of this is the same as the surface given in the previous example). The corresponding parametrization (in terms of the corresponding cylindrical coordinate system for $\mathcal{E}^3$) is

$$\boldsymbol{x}(\rho, \phi) \sim \left(\rho,\, \phi,\, \rho^2\right) \qquad \text{for} \quad 0 \leq \rho < \infty \quad \text{and} \quad 0 \leq \phi < 2\pi \quad .$$

This automatically defines a "$\{(\rho, \phi)\}$" coordinate system for $\mathcal{S}$, with each

$$\boldsymbol{x} \sim (\rho, \phi) \quad \text{in the surface } \mathcal{S}$$

being the point

$$\boldsymbol{x} \sim \left(\rho, \phi\, \rho^2\right) \quad \text{in the thee-dimensional space } \mathcal{E}^3 \quad .$$

?►Exercise 8.7: *Sketch, as well as you can, the circular parabolic surface described in examples 8.3 and 8.4. Sketch it twice and:*

a: *On one, sketch curves representing some of the $\{(x, y)\}$ coordinate curves (say, $x = 0$, $x = 1$, $x = 2$, $y = 0$, $y = -1$ and $y = 1$).*

b: *On the other, sketch curves representing some of the $\{(\rho, \phi)\}$ coordinate curves (say, $\rho = 1/2$, $\rho = 1$, $\rho = 3/2$, $\rho = 2$, $\phi = 0$, $\phi = \pi/4$, $\phi = \pi/2$, etc.)*

?►Exercise 8.8: *Using standard Cartesian coordinates, sketch the plane in space given by*

$$z = 1 + 2x + y \quad .$$

(Suggestion: Just sketch the piece in the first octant by sketching the intersections of this plane with the XZ–plane and YZ–plane.) Then sketch in some of the corresponding $\{(x, y)\}$ coordinate curves on this plane.

Generalizations

Mathematically, we can have a parametrized surface in any three-dimensional space, whether that space is Euclidean or not, though visualizing a surface in a nonEuclidean three-dimensional space may be difficult. And, of course, we can extend all this to a discussion of "parametrizations of K-dimensional spaces in arbitrary N-dimensional spaces" for any pair of positive integers K and N with $K < N$.

But, unless great need drives us, we won't.

Oh, by the way, in an N-dimensional space, the best analog of a "surface in three-space" is an $(N - 1)$-dimensional subspace. These are sometimes called *hyper-surfaces*. On a plane (or any two-dimensional space), a "hyper-surface" is just a curve. Hyper-surfaces may arise later, just for fun.

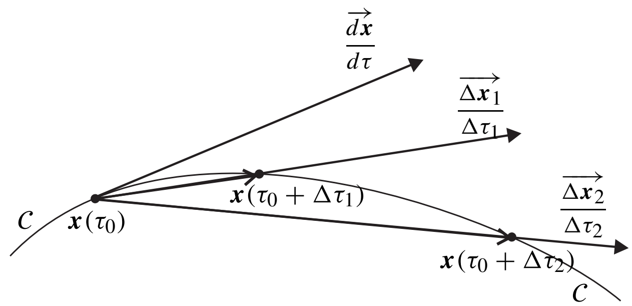

Figure 8.2: Derivative of position at a point $\mathbf{x}(\tau_0)$, along with approximations in which Δx_1 is the vector from $\mathbf{x}(\tau_0)$ to $\mathbf{x}(\tau_1)$, Δx_2 is the vector from $\mathbf{x}(\tau_0)$ to $\mathbf{x}(\tau_2)$, and $\Delta\tau_1 < \Delta\tau_2 < 1$.

8.6 Differentiation of Position (Along a Curve) In Euclidean Space

For the moment, let us restrict ourselves to an N-dimensional Euclidean space $\mathcal{E}$, and let $\mathbf{x}(\tau)$ be a position-valued function (hence $\mathbf{x}(\tau)$ traces out some curve C in $\mathcal{E}$ as τ varies over some interval (a, b)). The *derivative* of $\mathbf{x}$ with respect to τ is defined and denoted in a manner analogous to the way we defined and denoted derivatives in elementary calculus. Making use of the "difference of positions" discussed earlier, we have

$$\mathbf{x}'(\tau) = \frac{d\mathbf{x}}{d\tau} = \lim_{\Delta\tau\to 0} \frac{\text{“}\mathbf{x}(\tau+\Delta\tau) - \mathbf{x}(\tau)\text{”}}{\Delta\tau} = \lim_{\Delta\tau\to 0} \frac{\overrightarrow{\Delta\mathbf{x}}}{\Delta\tau}$$

where

$$\overrightarrow{\Delta\mathbf{x}} = \text{“}\mathbf{x}(\tau+\Delta\tau) - \mathbf{x}(\tau)\text{”} = \text{vector from } \mathbf{x}(\tau) \text{ to } \mathbf{x}(\tau+\Delta\tau) = \overrightarrow{\mathbf{x}(\tau)\,\mathbf{x}(\tau+\Delta\tau)}$$

(see figure 8.2). Notice that, being the limit of a vector, the derivative of a position is a *vector*. We probably should denote it by

$$\overrightarrow{\mathbf{x}'(\tau)} \quad\text{and}\quad \frac{\overrightarrow{d\mathbf{x}}}{d\tau} \qquad \text{instead of} \qquad \mathbf{x}'(\tau) \quad\text{and}\quad \frac{d\mathbf{x}}{d\tau} \quad .$$

This is rarely done since typing it out takes so long, but I may attempt to include the arrows to help us remember. Just remember that, even if the arrows aren't there, *this derivative is a vector*.

All this assumes the derivative exists. If it exists for a given value $\tau = \tau_0$ then we say $\mathbf{x}(\tau)$ is *differentiable at* τ_0. If it exists for all values of τ over some interval, then we say $\mathbf{x}(\tau)$ is *differentiable over that interval*. And if we are using $\mathbf{x}(\tau)$ to parametrize a curve, then we say that it is a *differentiable parametrization*.

Let's look more closely at this derivative at the point $\mathbf{x}_0 = \mathbf{x}(\tau_0)$. As illustrated in figure 8.2, this vector is tangent to the curve and points in one of the "directions along the curve C" at $\mathbf{x}_0$. For the magnitude:

$$\left\| \frac{\overrightarrow{d\mathbf{x}}}{d\tau}\bigg|_{\tau_0} \right\| = \left\| \lim_{\Delta\tau\to 0} \frac{\text{“}\mathbf{x}(\tau_0+\Delta\tau) - \mathbf{x}(\tau_0)\text{”}}{\Delta\tau} \right\|$$

$$= \left\| \lim_{\Delta\tau\to 0} \frac{\overrightarrow{\mathbf{x}(\tau_0)\mathbf{x}(\tau_0+\Delta\tau)}}{\Delta\tau} \right\|$$

$$= \lim_{\Delta\tau\to 0^+} \frac{\left\| \overrightarrow{\boldsymbol{x}(\tau_0)\boldsymbol{x}(\tau_0+\Delta\tau)} \right\|}{\Delta\tau}$$

$$= \lim_{\Delta\tau\to 0} \frac{\operatorname{dist}(\boldsymbol{x}(\tau_0), \boldsymbol{x}(\tau_0+\Delta\tau))}{\Delta\tau} = \frac{d}{d\tau} \operatorname{dist}(\boldsymbol{x}(\tau_0), \boldsymbol{x}(\tau))\Big|_{\tau_0}$$

So the magnitude of $\boldsymbol{x}'(\tau_0)$ is the "instantaneous ratio" at τ_0 of how the distance between $\boldsymbol{x}(\tau_0)$ and $\boldsymbol{x}(\tau)$ varies as τ varies. Recalling the relation between the distance between $\boldsymbol{x}(\tau_0)$ and $\boldsymbol{x}(\tau)$ and the arclength s between $\boldsymbol{x}(\tau_0)$ and $\boldsymbol{x}(\tau)$ (i.e., equation (8.1) on page 8–15), we see that we can write this equation a bit more simply as

$$\left\| \frac{\overrightarrow{d\boldsymbol{x}}}{d\tau}\Big|_{\tau_0} \right\| = \frac{ds}{d\tau}\Big|_{\tau_0} \quad .$$

Thus, at each point $\boldsymbol{x}(\tau)$ on the curve,

$$\boldsymbol{x}'(\tau) = \frac{\overrightarrow{d\boldsymbol{x}}}{d\tau} = \left\| \frac{\overrightarrow{d\boldsymbol{x}}}{d\tau} \right\| \times \big(\text{unit vector in the direction of } \boldsymbol{x}'(\tau)\big) = \frac{ds}{d\tau}\mathbf{T}(\tau) \tag{8.2}$$

where $\mathbf{T}(\tau)$ is the unit vector in the appropriate direction along the curve C at $\boldsymbol{x}(\tau)$. The vector $\mathbf{T}$ is called a *unit tangent vector* to the curve at that point. If $\boldsymbol{x}(\tau)$ is a parametrization of an oriented curve, then, at each point $\boldsymbol{x}(\tau)$, $\mathbf{T}(\tau)$ points in the direction of the orientation at that point.

If $\boldsymbol{x}(\tau)$ is the position of an object at time τ, then we know the magnitude ${ds}/{d\tau} = \left\| {d\boldsymbol{x}}/{d\tau} \right\|$ as the *speed* of the object at time τ. The direction of $\boldsymbol{x}'(\tau)$ is the *direction of travel* of the object at time τ (with $\mathbf{T}(\tau)$ being the unit vector in the direction of travel). And the entire vector, $\boldsymbol{x}'(\tau)$, is the *velocity* of the object at time τ.

?►*Exercise 8.9:* *Verify that, in the above,*

$$\mathbf{T} = \frac{\overrightarrow{d\boldsymbol{x}}}{ds} \quad .$$

(Hint: What if $\tau = s$?)

(Hence ${d\boldsymbol{x}}/{ds}$ is a unit vector at each point of C, and the above formula can be written as

$$\frac{\overrightarrow{d\boldsymbol{x}}}{d\tau} = \frac{ds}{d\tau}\frac{\overrightarrow{d\boldsymbol{x}}}{ds} \quad ,$$

which should remind you of a chain rule.)

A formula for $\boldsymbol{x}'(\tau)$ is easily obtained from a coordinate formula for $\boldsymbol{x}(\tau)$,

$$\boldsymbol{x}(\tau) \sim (x^1(\tau), x^2(\tau), \ldots, x^N(\tau)) \quad ,$$

when we are using a basis-based coordinate system. So assume $\{(x^1, x^2, \ldots, x^N)\}$ is such a system based on a basis $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N\}$ and origin $\boldsymbol{O}$. Then, recalling about describing the "difference in positions" with such a system,

$$\boldsymbol{x}'(\tau) = \frac{d\boldsymbol{x}}{d\tau} = \lim_{\Delta\tau\to 0} \frac{\text{“}\boldsymbol{x}(\tau+\Delta\tau) - \boldsymbol{x}(\tau)\text{”}}{\Delta\tau}$$

$$= \lim_{\Delta\tau\to 0} \frac{\sum_{k=1}^N \left[x^k(\tau+\Delta\tau) - x^k(\tau)\right]\mathbf{b}_k}{\Delta\tau}$$

$$= \sum_{k=1}^N \left[\lim_{\Delta\tau\to 0} \frac{x^k(\tau+\Delta\tau) - x^k(\tau)}{\Delta\tau}\right]\mathbf{b}_k \quad .$$

But the limits in the last formula just give the ordinary "Calc. I" derivatives of the individual coordinate formulas, so this reduces to

$$\boldsymbol{x}'(\tau) \;=\; \frac{d\boldsymbol{x}}{d\tau} \;=\; \sum_{k=1}^{N} \frac{dx^k}{d\tau}\,\mathbf{b}_k \quad . \tag{8.3}$$

The magnitude, then, can be computed by traditional vector methods

$$\begin{aligned}\frac{ds}{d\tau} \;&=\; \left\|\frac{d\boldsymbol{x}}{d\tau}\right\| \;=\; \left(\frac{d\boldsymbol{x}}{d\tau}\cdot\frac{d\boldsymbol{x}}{d\tau}\right)^{1/2} \\ &=\; \left(\sum_{j=1}^{N}\frac{dx^j}{d\tau}\mathbf{b}_j\cdot\sum_{k=1}^{N}\frac{dx^k}{d\tau}\mathbf{b}_k\right)^{1/2} \;=\; \left(\sum_{j=1}^{N}\sum_{k=1}^{N}\frac{dx^j}{d\tau}\frac{dx^k}{d\tau}\mathbf{b}_j\cdot\mathbf{b}_k\right)^{1/2} \quad .\end{aligned}$$

If the coordinate system is Cartesian (so the basis is orthonormal), this simplifies to

$$\frac{ds}{d\tau} \;=\; \left\|\frac{d\boldsymbol{x}}{d\tau}\right\| \;=\; \sqrt{\sum_{k=1}^{N}\left(\frac{dx^k}{d\tau}\right)^2} \quad .$$

In particular, if we are using a standard Cartesian system in Euclidean three-space with

$$\{(x^1, x^2, x^3)\} \;=\; \{x,\ y,\ z\} \qquad \text{and} \qquad \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\} \;=\; \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$$

Then

$$\boldsymbol{x}(\tau) \;\sim\; (\,x(\tau),\ y(\tau),\ z(\tau)\,) \quad ,$$

$$\boldsymbol{x}'(\tau) \;=\; \frac{d\boldsymbol{x}}{d\tau} \;=\; \frac{dx}{d\tau}\mathbf{i} \;+\; \frac{dy}{d\tau}\mathbf{j} \;+\; \frac{dz}{d\tau}\mathbf{k}$$

and

$$\frac{ds}{d\tau} \;=\; \left\|\frac{d\boldsymbol{x}}{d\tau}\right\| \;=\; \sqrt{\left(\frac{dx}{d\tau}\right)^2 + \left(\frac{dy}{d\tau}\right)^2 + \left(\frac{dz}{d\tau}\right)^2} \quad .$$

In General

Now let's suppose our position-valued function $\boldsymbol{x}(\tau)$ is tracing out a curve C in some space $\mathcal{S}$ which might not be Euclidean. In this case, we can no longer use the relation between vectors and positions to define the "$\overrightarrow{\Delta\boldsymbol{x}}$" used to define ${}^{d\boldsymbol{x}}/_{d\tau}$ in the previous discussion. We could try to appeal to the local approximate Euclidean-ness and pretend that there are "infinitesimally small" vectors between "infinitesimally close points", but I never liked that approach. Instead, let's just observe that formula (8.2) tells us exactly how to generalize our definition of $\boldsymbol{x}'(\tau)$; namely, at each point $\boldsymbol{x}(\tau)$ of the curve,

$$\boldsymbol{x}'(\tau) \;=\; \frac{\overrightarrow{d\boldsymbol{x}}}{d\tau} \;=\; \frac{ds}{d\tau}\,\mathbf{T}(\tau) \quad . \tag{8.4}$$

Where ${}^{ds}/_{d\tau}$ is the rate at which the arclength along the curve varies as τ varies, and $\mathbf{T}(\tau)$ is the unit tangent vector at $\boldsymbol{x}(\tau)$ pointing in the direction in the appropriate "direction of travel" there.

It is worth noting that the results of exercise 8.9 on page 8–19 did not require the space to be Euclidean. Thus, those results apply here, telling us that

$$\mathbf{T} = \frac{\overrightarrow{d\boldsymbol{x}}}{ds} \qquad \text{and} \qquad \frac{d\boldsymbol{x}}{d\tau} = \frac{ds}{d\tau}\frac{\overrightarrow{d\boldsymbol{x}}}{ds} \quad .$$

Partial Derivatives

If we have a position-valued function that depends on two or more real variables $\boldsymbol{x}(\tau, \mu, \ldots)$ then we refer to the partial derivatives with respect to each variable, instead of "the derivative". As in elementary calculus,

$$\frac{\partial \boldsymbol{x}}{\partial \tau} = \text{derivative of } \boldsymbol{x}(\tau, \mu, \ldots) \text{ assuming } \tau \text{ is the only variable}$$

$$\frac{\partial \boldsymbol{x}}{\partial \mu} = \text{derivative of } \boldsymbol{x}(\tau, \mu, \ldots) \text{ assuming } \mu \text{ is the only variable}$$

$$\vdots$$

Of particular interest will be the partial derivatives of position with respect to the coordinates of any given coordinate system $\{(x^1, x^2, \ldots, x^N)\}$ for our space $\mathcal{S}$.

Let us just consider a basis-based coordinate system $\{(x^1, x^2, \ldots, x^N)\}$ with corresponding basis $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N\}$. It should be clear that we have the partial derivative analog to formula (8.3),

$$\frac{\partial \boldsymbol{x}}{\partial \tau} = \sum_{k=1}^{N} \frac{\partial x^k}{\partial \tau} \mathbf{b}_k \quad . \tag{8.5}$$

In particular, this formula applies when the variable τ is actually any one of the coordinates x^j. But because individual coordinates do not depend on each other,

$$\frac{\partial x^k}{\partial x^j} = \begin{Bmatrix} 0 & \text{if} & k \neq j \\ 1 & \text{if} & k = j \end{Bmatrix} = \delta_{jk} \quad ,$$

and so

$$\frac{\partial \boldsymbol{x}}{\partial x^j} = \sum_{k=1}^{N} \frac{\partial x^k}{\partial x_j} \mathbf{b}_k = \sum_{k=1}^{N} \delta_{jk} \mathbf{b}_k = \mathbf{b}_j \quad .$$

In particular, if $\{(x, y, z)\}$ is a standard Cartesian coordinate system for $\mathcal{E}^3$, then

$$\frac{\partial \boldsymbol{x}}{\partial x} = \mathbf{i} \quad , \quad \frac{\partial \boldsymbol{x}}{\partial y} = \mathbf{j} \quad \text{and} \quad \frac{\partial \boldsymbol{x}}{\partial z} = \mathbf{k} \quad .$$

Note that, for a basis-based coordinate system in Euclidean space, the set of partial derivatives of position with respect to the different coordinates

$$\left\{ \frac{\partial \boldsymbol{x}}{\partial x^1}, \frac{\partial \boldsymbol{x}}{\partial x^2}, \ldots, \frac{\partial \boldsymbol{x}}{\partial x^N} \right\}$$

is a basis, $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N\}$, for the corresponding vector space, and differential formulas (8.3) and (8.5) can be written as

$$\frac{d\boldsymbol{x}}{d\tau} = \sum_{k=1}^{N} \frac{dx^k}{d\tau} \frac{\partial \boldsymbol{x}}{\partial x^k} \quad . \tag{8.6}$$

and

$$\frac{\partial \boldsymbol{x}}{\partial \tau} = \sum_{k=1}^{N} \frac{\partial x^k}{\partial \tau} \frac{\partial \boldsymbol{x}}{\partial x^k} \quad . \tag{8.7}$$

If these formulas remind you of the "chain rule", good.

If you think about it, even if the space is not Euclidean, you should expect the last two equations to always hold at each point $\boldsymbol{x}$ of the space. That is, at each point $\boldsymbol{x}$,

$$\left\{ \frac{\partial \boldsymbol{x}}{\partial x^1}, \frac{\partial \boldsymbol{x}}{\partial x^2}, \ldots, \frac{\partial \boldsymbol{x}}{\partial x^N} \right\}$$

should be a basis for the vector space of all vectors tangent to curves passing through that point.

8.7 Tangent Spaces of Vectors

Think of a point $\boldsymbol{p}$ on a surface $\mathcal{S}$ in ordinary three-dimensional space, and visualize the plane tangent to this surface at $\boldsymbol{p}$ (we assume the surface is sufficiently smooth for this plane to exist). Note that each point $\boldsymbol{q}$ on that plane corresponds to a vector from $\boldsymbol{p}$ to that point. That vector $\overrightarrow{\boldsymbol{pq}}$ will, of course, be a vector "tangent to $\mathcal{S}$ at $\boldsymbol{p}$", and we can view the tangent plane as the vector space of all tangent vectors.

Now keep that example in mind as we formally define tangent vectors and tangent spaces in general.

Let $\boldsymbol{p}$ be a point in an N-dimensional space $\mathcal{S}$ (not necessarily the surface mentioned above), and assume $\mathcal{S}$ be "sufficiently smooth" at $\boldsymbol{p}$.[9] A *tangent vector to $\mathcal{S}$ at $\boldsymbol{p}$* is any vector that can be written as the derivative at $\boldsymbol{p}$ of some curve in $\mathcal{S}$ passing through $\boldsymbol{p}$,

$$\boldsymbol{x}'(\tau_0) \qquad \text{with} \quad \boldsymbol{x}(\tau_0) = \boldsymbol{p} \quad .$$

The set of all tangent vectors to $\mathcal{S}$ at $\boldsymbol{p}$ forms an N-dimensional vector space $\mathcal{V}_{\boldsymbol{p}}$ called *tangent space for $\mathcal{S}$ at $\boldsymbol{p}$*.[10]

!►Example 8.5: *Consider the point $\boldsymbol{p} \sim (1, 2)$ in the circular paraboloid $\mathcal{S}$ from example 8.3 on page 8–16 using the coordinate system $\{(x, y)\}$ described in that example. Remember, this is the surface $\mathcal{S}$ given by the graph of*

$$z = x^2 + y^2 \qquad \textit{for} \quad -\infty < x < \infty \quad \textit{and} \quad -\infty < y < \infty$$

using a standard Cartesian coordinate system, and with each

$$\boldsymbol{x} \sim (x, y) \quad \textit{in the surface } \mathcal{S}$$

being the point

$$\boldsymbol{x} \sim \left(x,\ y,\ x^2 + y^2\right) \quad \textit{in the thee-dimensional space } \mathcal{E}^3 \quad .$$

[9] I.e., that we can have N smooth curves passing through $\boldsymbol{p}$ at right angles.

[10] In a sense, $\mathcal{V}_{\boldsymbol{p}}$ can be thought of as the vector space of all possible velocities of objects moving on $\mathcal{S}$ through position $\boldsymbol{p}$.

Now, let $\boldsymbol{x}(\tau) \sim (x(\tau), y(\tau))$ be any (smooth) curve on our surface $\mathcal{S}$. Using the fact that this is also a curve in $\mathcal{E}^3$, we have

$$\boldsymbol{x}(\tau) \sim \left(x\,,\, y\,,\, x^2 + y^2\right) \qquad \text{with} \quad x = x(\tau) \quad \text{and} \quad y = y(\tau) \quad .$$

Thus, with $x = x(\tau)$ and $y = y(\tau)$,

$$\begin{aligned}\frac{\overrightarrow{d\boldsymbol{x}}}{d\tau} &= \frac{dx}{d\tau}\mathbf{i} + \frac{dy}{d\tau}\mathbf{j} + \left[2x\frac{dx}{d\tau} + 2y\frac{dy}{d\tau}\right]\mathbf{k} \\ &= \frac{dx}{d\tau}\big[\mathbf{i} + 2x\mathbf{k}\big] + \frac{dy}{d\tau}\big[\mathbf{j} + 2y\mathbf{k}\big] \quad .\end{aligned}$$

In particular, if $\boldsymbol{x}(\tau_0) = \boldsymbol{p}$ with $\boldsymbol{p} \sim (1, 2)$, then the corresponding tangent vector to $\mathcal{S}$ at $\boldsymbol{p}$ is

$$\begin{aligned}\frac{\overrightarrow{d\boldsymbol{x}}}{d\tau} &= \left.\frac{dx}{d\tau}\right|_{\tau_0}\big[\mathbf{i} + 2\cdot 1\mathbf{k}\big] + \left.\frac{dy}{d\tau}\right|_{\tau_0}\big[\mathbf{j} + 2\cdot 2\mathbf{k}\big] \\ &= x'(\tau_0)\big[\mathbf{i} + 2\mathbf{k}\big] + y'(\tau_0)\big[\mathbf{j} + 4\mathbf{k}\big] \quad .\end{aligned}$$

If you think about it, you will realize that we can choose functions $x(\tau)$ and $y(\tau)$ so that their derivatives equal anything we wish. Thus, the derivatives of x and y in the last expression above are arbitrary. This means that the set of all tangent vectors to $\mathcal{S}$ at this $\boldsymbol{p}$ — i.e., the tangent space to $\mathcal{S}$ at $\boldsymbol{p}$ — is the set of all vectors of the form

$$\mathbf{v} = \alpha\,[\mathbf{i} + 2\mathbf{k}] + \beta\,\big[\mathbf{j} + 4\mathbf{k}\big]$$

where α and β are arbitrary real numbers. (Though we still try to visualize these tangent vectors as "starting" at position $\boldsymbol{p}$.)

Note that $\{\mathbf{i} + 2\mathbf{k}\,,\, \mathbf{j} + 4\mathbf{k}\}$ is a basis for this tangent space. If you check, you'll see that this set is $\{{}^{\partial\boldsymbol{x}}/_{\partial x}\,,\, {}^{\partial\boldsymbol{x}}/_{\partial y}\}$

?►*Exercise 8.10:* *Let $\mathcal{S}$ be as in the last example with the same coordinate system as in that example. Find the tangent space at the point $\boldsymbol{q} \sim (3, 4)$.*

?►*Exercise 8.11:* *Convince yourself that if $\mathcal{S}$ is a three-dimensional Euclidean space, then the tangent space $\mathcal{V}_{\boldsymbol{p}}$ is simply the set of all vectors from $\boldsymbol{p}$ to other points in $\mathcal{S}$.*

Strictly speaking, $\mathcal{V}_{\boldsymbol{p}}$ and $\mathcal{V}_{\boldsymbol{q}}$, the tangent spaces of a space $\mathcal{S}$ at two different points $\boldsymbol{p}$ and $\boldsymbol{q}$, are two *different* vector spaces. This is obvious if $\mathcal{S}$ is a warped surface in space, and less obvious if, say, $\mathcal{S}$ is the three-dimensional Euclidean space in which we think we live. One way to reinforce our awareness of this fact is to always draw a tangent vector at $\boldsymbol{p}$ with its 'tail' at $\boldsymbol{p}$.

Now we can state a basic requirement linking coordinate systems and tangent spaces; namely that, for a coordinate system $\{(x^1, x^2, \ldots, x^N)\}$ to be a "reasonable" coordinate system for a space $\mathcal{S}$,

$$\left\{\frac{\partial\boldsymbol{x}}{\partial x^1}\,,\, \frac{\partial\boldsymbol{x}}{\partial x^2}\,,\, \ldots\,,\, \frac{\partial\boldsymbol{x}}{\partial x^N}\right\}$$

must be a basis for the tangent space at $\boldsymbol{p}$ for each point $\boldsymbol{p}$ in $\mathcal{S}$. If this basis is orthogonal at each point, then we say the coordinate system *orthogonal*. In practice, "orthonormality" is too much to ask for unless we have a Cartesian system.

In fact, we can have coordinate systems that are not reasonable at some points, but are reasonable everywhere where they might be useful.

?►***Exercise 8.12:*** *Why would we not consider a polar coordinate system as "reasonable" at the origin?*

?►***Exercise 8.13:*** *These problems all concern surfaces in Euclidean three-space $\mathcal{E}^3$, with each being the graph of a function $z = f(x, y)$ using a standard Cartesian system, and with $\{(x, y)\}$ being the corresponding coordinate system on the surface.*

a: *Describe all the tangent vectors in the tangent space at $\boldsymbol{p} \sim (2, 3)$ to the (flat) surface given by*

$$z = 1 + 2x + y \quad .$$

Also, try to sketch the surface and this tangent plane.

b: *Describe all the tangent vectors in the tangent space at $\boldsymbol{p} \sim (3, 4)$ to the surface given by*

$$z = \sqrt{x^2 + y^2} \quad .$$

Also, try to sketch the surface and this tangent plane.

c: *Describe all the tangent vectors in the tangent space at $\boldsymbol{p} \sim (2, 1)$ to the surface given by*

$$z = x^2 - y^2 \quad .$$

Also, try to sketch the surface and this tangent plane.

d: *Describe all the tangent vectors in the tangent space at $\boldsymbol{p} \sim (2, 4)$ to the surface given by*

$$z = 27 \quad .$$

Also, try to sketch the surface and this tangent plane.

8.8 Scaling Factors and Vectors Associated with Coordinate Systems

Scaling Factors and Associated Tangent Vectors

Suppose we have some N-dimensional space, not necessarily Euclidean, and let

$$\left\{\left(x^1, x^2, \ldots, x^N\right)\right\}$$

be a coordinate system for this space. So each position $\boldsymbol{x}$ can be described in terms of these coordinates,

$$\boldsymbol{x} \sim \left(x^1, x^2, \ldots, x^N\right) \quad .$$

As each coordinate x^k varies (with the other coordinates held fixed) the corresponding position $\boldsymbol{x}$ traces out a corresponding (coordinate) curve, and, at each point of that coordinate curve, we have the *associated tangent vector*

$$\boldsymbol{\varepsilon}_k = \frac{\partial \boldsymbol{x}}{\partial x^k} = \text{rate position varies as } x^k \text{ varies} \quad .$$

We then define the *associated scaling factor* h_k by

$$h_k = \|\boldsymbol{\varepsilon}_k\| = \left\| \frac{\partial \boldsymbol{x}}{\partial x^k} \right\| = \frac{\partial s}{\partial x^k}$$
$$= \text{rate distance traveled varies as } x^k \text{ varies} \quad ,$$

and the *associated normalized tangent vector* $\mathbf{e}_k$ by

$$\mathbf{e}_k = \text{the unit vector in the direction of } \boldsymbol{\varepsilon}_k$$
$$= \frac{\boldsymbol{\varepsilon}_k}{\|\boldsymbol{\varepsilon}_k\|} = \frac{1}{h_k}\boldsymbol{\varepsilon}_k = \frac{1}{h_k}\frac{\partial \boldsymbol{x}}{\partial x^k} \quad .$$

By definition,

$$\frac{\partial \boldsymbol{x}}{\partial x^k} = \boldsymbol{\varepsilon}_k = h_k \mathbf{e}_k \quad . \tag{8.8}$$

We will be using this along with the notation $\boldsymbol{\varepsilon}_k$, h_k and $\mathbf{e}_k$ (or immediately obvious variations of this notation) for the rest of the term. Often though, I may just use ${}^{\partial \boldsymbol{x}}/_{\partial x^k}$ instead of $\boldsymbol{\varepsilon}_k$. Do observe that the scaling factor h_k and the vector $\mathbf{e}_k$ are the same as the ${}^{\partial s}/_{\partial x^k}$ and unit tangent $\mathbf{T}$ mentioned earlier. We've just introduced special terminology and notation for the cases where the parameter is a coordinate.

!►*Example 8.6:* *If our space is just a two-dimensional Euclidean plane, and the coordinate system is a Cartesian system* $\{(x, y)\}$ *with corresponding orthonormal basis* $\{\mathbf{i}, \mathbf{j}\}$ *, then we already know that*

$$\frac{\partial \boldsymbol{x}}{\partial x} = \mathbf{i} \qquad \textit{and} \qquad \frac{\partial \boldsymbol{x}}{\partial y} = \mathbf{j} \quad .$$

Letting

$$\left(x^1, x^2\right) = (x, y)$$

in the above, we have

$$\boldsymbol{\varepsilon}_1 = h_1\mathbf{e}_1 = \frac{\partial \boldsymbol{x}}{\partial x^1} = \frac{\partial \boldsymbol{x}}{\partial x} = \mathbf{i} \qquad \textit{and} \qquad \boldsymbol{\varepsilon}_2 = h_2\mathbf{e}_2 = \frac{\partial \boldsymbol{x}}{\partial x^2} = \frac{\partial \boldsymbol{x}}{\partial y} = \mathbf{j} \quad .$$

Clearly then

$$\boldsymbol{\varepsilon}_1 = \mathbf{i} \quad , \quad \boldsymbol{\varepsilon}_2 = \mathbf{j} \quad ,$$

$$h_1 = \|\boldsymbol{\varepsilon}_1\| = \|\mathbf{i}\| = 1 \quad , \quad h_2 = \|\boldsymbol{\varepsilon}_2\| = \|\mathbf{j}\| = 1 \quad ,$$

$$\mathbf{e}_1 = \frac{1}{h_1}\frac{\partial \boldsymbol{x}}{\partial x^1} = \mathbf{i} \qquad \textit{and} \qquad \mathbf{e}_2 = \frac{1}{h_2}\frac{\partial \boldsymbol{x}}{\partial x^2} = \mathbf{j} \quad .$$

?►*Exercise 8.14:* *Assume our space is an* N*-dimensional Euclidean space and we have a basis-based coordinate system* $\{(x^1, x^2, \ldots, x^N)\}$ *with corresponding basis* $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N\}$ *. Remember, this means that, letting* $\boldsymbol{O}$ *be the origin,*

$$\boldsymbol{x} \sim \left(x^1, x^2, \ldots, x^N\right) \iff \overrightarrow{\boldsymbol{Ox}} = \sum_{k=1}^{N} x^k \mathbf{b}_k \quad .$$

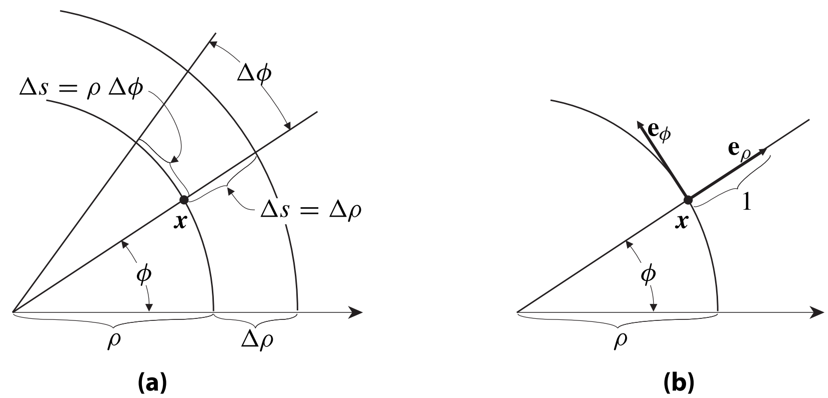

Figure 8.3: (a) Change in position corresponding to change in the values of the polar coordinates, and (b) the unit vectors in the directions of increasing r and ϕ (For exercise 8.7).

Show that, for $k = 1, 2, \ldots, N$,

$$\boldsymbol{\varepsilon}_k = \mathbf{b}_k \quad ,$$

and, thus,

$$h_k = \|\mathbf{b}_k\| \qquad \textit{and} \qquad \mathbf{e}_k = \frac{\mathbf{b}_k}{\|\mathbf{b}_k\|} \quad .$$

As demonstrated above, the scaling factors and the associated tangent vectors are essentially constants when the coordinate system is basis based. In general, though, the scaling factors and the associated tangent vectors will vary with position.

Before looking at a simple example in which the h_k's and $\mathbf{e}_k$'s vary with position, take another brief look at the definition of the scaling factor,

$$h_k = \frac{\partial s}{\partial x^k} = \left\| \frac{\partial \boldsymbol{x}}{\partial x^k} \right\| = \text{rate distance traveled varies as } x^k \text{ varies} \quad .$$

If x^k happens to exactly correspond to the distance along the x^k coordinate curve, then the above reduces to $h_k = 1$. More commonly, the change Δs in distance traveled is related to a corresponding change Δx^k in the k^{th} coordinate by some constant C of proportionality,

$$\Delta s = C \, \Delta x^k \quad .$$

Comparing this to the definition of h_k , we see that the scaling factor is this constant,

$$h_k = \frac{\partial s}{\partial x^k} = \lim_{\Delta x^k \to 0} \frac{\Delta s}{\Delta x^k} = \lim_{\Delta x^k \to 0} \frac{C \, \Delta x^k}{\Delta x^k} = C \quad .$$

Now, let's get the scaling factors and corresponding tangent vectors for our second favorite coordinate system for the plane.

!►*Example 8.7 (polar coordinates):* *Let us determine the scaling factors and unit tangent vectors at a point* $\boldsymbol{x} \sim (\rho, \phi)$ *associated with a standard polar coordinate system* $\{(\rho, \phi)\}$ *in the plane as indicated in figure 8.3.*

Since ρ is the distance from the chosen origin, changing ρ by $\Delta\rho$ changes the distance (from the origin) by $\Delta s = \Delta\rho$. So the scaling factor corresponding to the ρ coordinate is

$$h_\rho = \frac{\partial s}{\partial \rho} = \lim_{\Delta\rho\to 0} \frac{\Delta s}{\Delta \rho} = \lim_{\Delta\rho\to 0} \frac{\Delta r}{\Delta \rho} = 1 \quad .$$

Increasing ρ also moves position in the radial direction away from the origin. So $\mathbf{e}_\rho$, the unit tangent associated with the ρ coordinate at $\boldsymbol{x}$, will point in the radial direction away from the origin, as sketched in figure 8.3b.

On the other hand, increasing ϕ by $\Delta\phi$ moves position counterclockwise along a circle of radius ρ . So the unit tangent associated with this coordinate, $\mathbf{e}_\phi$, is tangent to the circle and pointing in the direction indicated in figure 8.3b. Assuming radian measure for ϕ , the distance traveled as ϕ increases by $\Delta\phi$ is $\Delta s = \rho\,\Delta\phi$. So the scaling factor corresponding to the ϕ coordinate is

$$h_\phi = \frac{\partial s}{\partial \phi} = \lim_{\Delta\phi\to 0} \frac{\Delta s}{\Delta \phi} = \lim_{\Delta\phi\to 0} \frac{\rho\,\Delta \phi}{\Delta \phi} = \rho \quad .$$

Finally, from basic geometric considerations, it should be quite apparent that $\{\mathbf{e}_\rho, \mathbf{e}_\phi\}$ is an orthonormal pair of vectors at each point of the plane (other than the origin).

?►*Exercise 8.15 (a sphere):* *Consider a sphere of radius 3 (as in example 8.2). Find the scaling factors associated with the two-dimensional spherical coordinate system $\{(\theta, \phi)\}$ on this sphere where θ is the angle with the "north pole" ("latitude") and ϕ is the "longitude" (as in the spherical coordinate system described at the start of section 8.3 on page 8–9). Also, sketch the associated unit tangents $\mathbf{e}_\theta$ and $\mathbf{e}_\phi$ for different points on the sphere. Do you expect $\{\mathbf{e}_\theta, \mathbf{e}_\phi\}$ to be orthonormal?*

?►*Exercise 8.16 (any sphere):* *What are the scaling factors h_θ and h_ϕ using the same sort of coordinate system as in the last exercise, but on a sphere of radius R ?*

?►*Exercise 8.17 (spherical coordinates):* *Find the scaling factors h_r , h_θ and h_ϕ for the spherical coordinate system $\{(r, \theta, \phi)\}$ described at the start of section 8.3. (Make use of what you obtained above!). Also, sketch the vectors $\mathbf{e}_r$, $\mathbf{e}_\theta$ and $\mathbf{e}_\phi$ at various points in space. Do you expect $\{\mathbf{e}_r, \mathbf{e}_\phi, \mathbf{e}_\theta\}$ to be orthonormal?*

All of the above are examples and exercises involving coordinate systems in Euclidean spaces. Let's look at an example and some exercises involving nonEuclidean spaces (but still inside a larger Euclidean space).

!►*Example 8.8:* *Again, consider the circular paraboloid from example 8.3 on page 8–16 and example 8.5 on page 8–22. Using a standard Cartesian coordinate system, this surface $\mathcal{S}$ is the graph of*

$$z = x^2 + y^2 \qquad \text{for} \quad -\infty < x < \infty \quad \text{and} \quad -\infty < y < \infty \quad ,$$

with each

$$\boldsymbol{x} \sim (x, y) \quad \text{in the surface } \mathcal{S}$$

being the point

$$\boldsymbol{x} \sim \left(x,\ y,\ x^2+y^2\right) \quad \text{in the thee-dimensional space } \mathcal{E}^3 \quad .$$

Using this fact, we have, for each $\boldsymbol{p} \sim (x, y)$ *on the surface,*

$$\boldsymbol{\varepsilon}_1 = h_1\mathbf{e}_1 = \frac{\partial \boldsymbol{x}}{\partial x} = \frac{\partial x}{\partial x}\mathbf{i} + \frac{\partial y}{\partial x}\mathbf{j} + \frac{\partial}{\partial x}\left[x^2+y^2\right]\mathbf{k} = 1\mathbf{i} + 2x\mathbf{k}$$

and

$$\boldsymbol{\varepsilon}_2 = h_2\mathbf{e}_2 = \frac{\partial \boldsymbol{x}}{\partial y} = \frac{\partial x}{\partial y}\mathbf{i} + \frac{\partial y}{\partial y}\mathbf{j} + \frac{\partial}{\partial y}\left[x^2+y^2\right]\mathbf{k} = 1\mathbf{j} + 2y\mathbf{k} \quad .$$

The corresponding scaling factors are

$$h_1 = \|\boldsymbol{\varepsilon}_1\| = \|1\mathbf{i}+2x\mathbf{k}\| = \sqrt{1+4x^2}$$

and

$$h_2 = \|\boldsymbol{\varepsilon}_2\| = \|1\mathbf{j}+2y\mathbf{k}\| = \sqrt{1+4y^2} \quad ,$$

and the corresponding unit tangents are

$$\mathbf{e}_1 = \frac{\boldsymbol{\varepsilon}_1}{h_1} = \frac{1}{\sqrt{1+4x^2}}\left[1\mathbf{i}+2x\mathbf{k}\right]$$

and

$$\mathbf{e}_2 = \frac{\boldsymbol{\varepsilon}_2}{h_2} = \frac{1}{\sqrt{1+4y^2}}\left[1\mathbf{j}+2y\mathbf{k}\right] \quad .$$

In particular, at the point on $\mathcal{S}$ *where* $(x, y) = (1, 2)$ *,*

$$\{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2\} = \{\mathbf{i}+2\mathbf{k}\,,\ \mathbf{j}+4\mathbf{k}\} \quad ,$$

$$\{h_1, h_2\} = \left\{\sqrt{5}, \sqrt{17}\right\} \quad ,$$

and

$$\{\mathbf{e}_1, \mathbf{e}_2\} = \left\{\frac{1}{\sqrt{5}}[\mathbf{i}+2\mathbf{k}]\,,\ \frac{1}{\sqrt{17}}[\mathbf{j}+4\mathbf{k}]\right\} \quad .$$

?►*Exercise 8.18:* *Several surfaces in Euclidean three-space* $\mathcal{E}^3$ *are given below as graphs of functions using a standard Cartesian system* $\{x, y, z)\}$ *, and with* $\{(x, y)\}$ *being the corresponding coordinate system on the surface. For each, find the corresponding* $\boldsymbol{\varepsilon}_k$*'s ,* h_k*'s and* $\mathbf{e}_k$*'s for the tangent plane at* $\boldsymbol{p} \sim (x, y)$ *.*

a: $\mathcal{S}$ *is the flat surface given by*

$$z = 1 + 2x + y \quad .$$

b: $\mathcal{S}$ *is the graph of*

$$z = \sqrt{x^2+y^2} \quad .$$

c: $\mathcal{S}$ *is the graph of*

$$z = x^2 - y^2 \quad .$$

Local Bases and Orthogonality

Remember that the tangent vector space at a point $\boldsymbol{x}$ is the vector space of all vectors tangent at $\boldsymbol{x}$ to curves passing through $\boldsymbol{x}$. If $\{(x^1, x^2, \ldots, x^N)\}$ is a reasonable coordinate system at $\boldsymbol{x}$, then each of the two sets of associated tangent vectors at this point

$$\{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_N\} \qquad \text{and} \qquad \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N\}$$

(with all vectors evaluated at $\boldsymbol{x}$) will be a basis for the tangent vector space at that point. We will call these the *associated local bases*. In practice, we usually prefer using the $\mathbf{e}_k$'s since they are unit vectors. Still there are occasions, especially when differentiating vector fields, when can be convenient to use the $\boldsymbol{\varepsilon}_k$'s.

"Orthogonality", naturally, is highly desired. We refer to a coordinate system as being *orthogonal* if and only if the associated bases are orthogonal at every point (excluding isolated points where the coordinate system is not well defined). Remember, each $\mathbf{e}_k$ is just a normalization of $\boldsymbol{\varepsilon}_k$, so

$$\{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_N\} \text{ is orthogonal} \quad \Longleftrightarrow \quad \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N\} \text{ is orthonormal} \quad .$$

Hence, a coordinate system $\{(x^1, x^2, \ldots, x^N)\}$ is orthogonal if and only if the associated local basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N\}$ is orthonormal everywhere. This does not necessarily mean that the coordinate curves are straight lines! As you should have already noted (and as we will rigorously confirm later), the polar and spherical coordinate systems are orthogonal systems.

When computing derivatives of entities involving the $\boldsymbol{\varepsilon}_k$'s and $\mathbf{e}_k$'s, it will be important to remember that these basis vector fields can no longer be assumed constant. They may vary from point to point. This will become especially relevant later when discussing acceleration and differentiating vector fields while using non-Cartesian coordinate systems.

The Metric

In many of the formulas that we will derive, we will see one of the following equivalent expressions:

$$\boldsymbol{\varepsilon}_j \cdot \boldsymbol{\varepsilon}_k \quad , \quad (h_j\,\mathbf{e}_j) \cdot (h_k \mathbf{e}_k) \quad \text{or} \quad h_j h_k\,\mathbf{e}_j \cdot \mathbf{e}_k \quad .$$

Of these $\boldsymbol{\varepsilon}_j \cdot \boldsymbol{\varepsilon}_k$ is the most concise. An even more concise notation is given by setting

$$g_{jk} = \boldsymbol{\varepsilon}_j \cdot \boldsymbol{\varepsilon}_k = h_j h_k\,\mathbf{e}_j \cdot \mathbf{e}_k \quad .$$

This is fairly standard notation. The g_{ij}'s encode the information about obtaining distances using the coordinate system, and are refered to as the components of the *metric* associated with the coordinate system. Later we will see that they are, in fact, the "covariant components of the metric tensor" for the space.

For now, though, we will just treat them as convenient shorthand.

From the above definition, we have the following immediate observations:

1. The g_{jk}'s are symmetric, $g_{jk} = g_{kj}$.

2. The coordinate system $\{(x^1, \ldots, x^N)\}$ is orthogonal if and only if

$$g_{jk} = \begin{cases} (h_j)^2 & \text{if} \quad j = k \\ 0 & \text{if} \quad j \neq k \end{cases} \quad .$$

?►Exercise 8.19: *Find the components of the metric (as described above) for each of the following:*

a: *Cartesian coordinates*

b: *Polar coordinates (Assume orthogonality.)*

c: *Spherical coordinates (Assume orthogonality.)*

d: *The $\{(x, y)\}$ coordinate system for the circular paraboloid discussed in example 8.8.*

e: *Each of the $\{(x, y)\}$ coordinate systems for the surfaces discussed in exercises 8.18.*

8.9 The Chain Rule In Euclidean Spaces

Let us recall something derived few pages ago (page 8–21), assuming our space is Euclidean and our coordinate system $\{(x^1, x^2, \ldots, x^N)\}$ is basis based. In this case, if $\boldsymbol{x}(\tau)$ is a position-valued function of a real variable τ, then

$$\frac{d\boldsymbol{x}}{d\tau} = \sum_{k=1}^{N} \frac{dx^k}{d\tau} \frac{\partial \boldsymbol{x}}{\partial x^k} \quad . \tag{8.9}$$

And if $\boldsymbol{x}$ depends on several variables, then for any one of these variables τ,

$$\frac{\partial \boldsymbol{x}}{\partial \tau} = \sum_{k=1}^{N} \frac{\partial x^k}{\partial \tau} \frac{\partial \boldsymbol{x}}{\partial x^k} \quad , \tag{8.10}$$

These are the *chain rule formulas* for derivatives of position. We've derived them assuming $\{(x^1, x^2, \ldots, x^N)\}$ is a basis-based coordinate system for a Euclidean space. However, these formulas always hold, even if the coordinate system is not basis-based. The verification of that will be left as an exercise for the adventurous.

?►Exercise 8.20 (optional, and mainly for math majors): *Let $\boldsymbol{x}(\tau)$ be any smooth position-valued function in an N-dimensional Euclidean space $\mathcal{E}$. Show that, for any "reasonable" coordinate system $\{(x^1, x^2, \ldots, x^N)\}$,*

$$\frac{d\boldsymbol{x}}{d\tau} = \sum_{k=1}^{N} \frac{dx^k}{d\tau} \frac{\partial \boldsymbol{x}}{\partial x^k} \quad .$$

(Suggestions: Fix $\boldsymbol{x}_0 = \boldsymbol{x}(\tau_0)$ and consider the basis-based coordinate system with origin $\boldsymbol{x}_0$ and basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N\}$ where

$$\mathbf{b}_k = \left. \frac{\partial \boldsymbol{x}}{\partial x^k} \right|_{\boldsymbol{x}_0} \quad .$$

You might even be clever and think about using a 'straight line' position-valued function $\boldsymbol{y}(\tau)$ such that $\boldsymbol{y}(\tau_0) = \boldsymbol{x}_0$ and $\boldsymbol{y}'(\tau_0) = \boldsymbol{x}'(\tau_0)$.)

The Chain Rule in Non-Euclidean Spaces

Chain rules (8.9) and (8.10) also hold when the space is non-Euclidean. This follows from the "local near-Euclideaness" of our spaces (you may want to glance back at our discussion of non-Euclidean spaces starting on page 8–7). It can also be verified without too much difficulty when the non-Euclidean space is contained in a larger Euclidean space (e.g., a sphere in a three-dimensional Euclidean space).

I hope it seems reasonable for these chain rules to hold even when the space is not Euclidean (just think of what "local near-Euclideaness" means). We will need them, but we don't have time to rigorously derive/prove them beyond what we've already done.

The General Result

Theorem 8.1 (Chain Rule)
Let $\{(x^1, x^2, \ldots, x^N)\}$ *be a (reasonable) coordinate system for some space* $\mathcal{S}$ *, and let*

$$\boldsymbol{x}(\tau) \sim \left(x^1(\tau),\, x^2(\tau),\, \ldots,\, x^N(\tau)\right)$$

be any (differentiable) position-valued function. Then

$$\frac{d\boldsymbol{x}}{d\tau} = \sum_{k=1}^{N} \frac{dx^k}{d\tau} \frac{\partial \boldsymbol{x}}{\partial x^k} \quad . \tag{8.11}$$

If, instead, $\boldsymbol{x}$ *depends on several variables, then for any one of these variables* τ *,*

$$\frac{\partial \boldsymbol{x}}{\partial \tau} = \sum_{k=1}^{N} \frac{\partial x^k}{\partial \tau} \frac{\partial \boldsymbol{x}}{\partial x^k} \quad . \tag{8.12}$$

We should observe that, using the scaling factors and bases associated with the coordinate system, the two equations above can also be written as

$$\frac{d\boldsymbol{x}}{d\tau} = \sum_{k=1}^{N} \frac{dx^k}{d\tau} \boldsymbol{\varepsilon}_k = \sum_{k=1}^{N} \frac{dx^k}{d\tau} h_k \mathbf{e}_k \tag{8.11$'$}$$

and

$$\frac{\partial \boldsymbol{x}}{\partial \tau} = \sum_{k=1}^{N} \frac{\partial x^k}{\partial \tau} \boldsymbol{\varepsilon}_k = \sum_{k=1}^{N} \frac{\partial x^k}{\partial \tau} h_k \mathbf{e}_k \quad . \tag{8.12$'$}$$

8.10 Converting Between Two Coordinate Systems

Often, we wish to convert formulas for vector fields, differential operators, path and surface integrals, etc. that we've derived in terms of one coordinate system, say, a two-dimensional Cartesian system $\{(x, y)\}$, to the corresponding formulas in terms of a second system, say, polar coordinates $\{(\rho, \phi)\}$. Typically, in such cases, we have a set of "change of coordinates formulas" for at least converting the basic coordinates. For the Cartesian/polar conversion, for example, we have the change of coordinates formulas

$$x = x(\rho,\phi) = \rho\cos(\phi) \qquad \text{and} \qquad y = y(\rho,\phi) = \rho\sin(\phi) \quad .$$

Getting this set of formulas is the first step in the conversion process. The second step is finding the scaling factors and tangent vectors for the second coordinate system.

Let us consider the second step in general.

Assume we have two coordinate systems with associated scaling factors and unit tangent vectors, with the first one being

$$\left\{\left(x^1, x^2, \ldots, x^N\right)\right\} \quad , \quad \{h_1, h_2, \ldots, h_N\} \quad \text{and} \quad \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N\} \quad ,$$

and the second one being[11]

$$\left\{\left(x^{1'}, x^{2'}, \ldots, x^{N'}\right)\right\} \quad , \quad \{h_1{}', h_2{}', \ldots, h_N{}'\} \quad \text{and} \quad \{\mathbf{e}_1{}', \mathbf{e}_2{}', \ldots, \mathbf{e}_N{}'\} \quad .$$

Remember that

$$\boldsymbol{\varepsilon}_i = h_i\mathbf{e}_i = \frac{\partial \boldsymbol{x}}{\partial x^i} \qquad \text{and} \qquad \boldsymbol{\varepsilon}_i{}' = h_i{}'\mathbf{e}_i{}' = \frac{\partial \boldsymbol{x}}{\partial x^{i'}} \quad .$$

Suppose, further, that we have a set of change of coordinates formulas for computing the x^k coordinates of any position from the position's $x^{k'}$ coordinates:

$$\begin{aligned} x^1 &= x^1(x^{1'}, x^{2'}, \ldots, x^{N'}) \quad , \\ x^2 &= x^2(x^{1'}, x^{2'}, \ldots, x^{N'}) \quad , \\ &\vdots \\ x^N &= x^N(x^{1'}, x^{2'}, \ldots, x^{N'}) \quad . \end{aligned}$$

Then, using the definitions and the chain rule,

$$h_i{}'\mathbf{e}_i{}' = \frac{\partial \boldsymbol{x}}{\partial x^{i'}} = \sum_{k=1}^{N} \frac{\partial \boldsymbol{x}}{\partial x^k}\frac{\partial x^k}{\partial x^{i'}} = \sum_{k=1}^{N} h_k\mathbf{e}_k \frac{\partial x^k}{\partial x^{i'}} = \sum_{k=1}^{N} h_k \frac{\partial x^k}{\partial x^{i'}}\,\mathbf{e}_k \quad . \tag{8.13}$$

We could cut out the middle and use the $\boldsymbol{\varepsilon}_i$'s and $\boldsymbol{\varepsilon}_i{}'$'s to write the above more concisely as

$$\boldsymbol{\varepsilon}_i{}' = \sum_{k=1}^{N} \frac{\partial x^k}{\partial x^{i'}}\,\boldsymbol{\varepsilon}_k \quad ,$$

but, frankly, it is better to simply remember the derivation via the chain rule expressed in (8.13) than to memorize the result.

Using the just derived fact that

$$h_i{}'\mathbf{e}_i{}' = \sum_{k=1}^{N} h_k \frac{\partial x^k}{\partial x^{i'}}\,\mathbf{e}_k \quad ,$$

[11] and may I be forgiven for using this "prime" notation, but it is 'standard'.

we immediately get

$$h_i{}' \;=\; \left\| h_i{}' \mathbf{e}_i{}' \right\| \;=\; \left\| \sum_{k=1}^{N} h_k \frac{\partial x^k}{\partial x^{i'}} \mathbf{e}_k \right\|$$

and

$$\mathbf{e}_i{}' \;=\; \frac{1}{h_i{}'} h_i{}' \mathbf{e}_i{}' \;=\; \frac{1}{h_i{}'} \sum_{k=1}^{N} h_k \frac{\partial x^k}{\partial x^{i'}} \mathbf{e}_k \quad .$$

The hardest part is computing the scaling factor. In general, the above norm is computed by taking a dot product,

$$h_i{}' \;=\; \sqrt{\left[\sum_{j=1}^{N} h_j \frac{\partial x^j}{\partial x^{i'}} \mathbf{e}_j \right] \cdot \left[\sum_{k=1}^{N} h_k \frac{\partial x^k}{\partial x^{i'}} \mathbf{e}_k \right]} \;=\; \sqrt{\sum_{j=1}^{N} \sum_{k=1}^{N} \frac{\partial x^j}{\partial x^{i'}} \frac{\partial x^k}{\partial x^{i'}} h_j h_k \, \mathbf{e}_j \cdot \mathbf{e}_k} \quad . \tag{8.14}$$

In terms of the metric, this is

$$h_i{}' \;=\; \sqrt{\sum_{j=1}^{N} \sum_{k=1}^{N} \frac{\partial x^j}{\partial x^{i'}} \frac{\partial x^k}{\partial x^{i'}} g_{jk}} \quad .$$

If we are lucky enough for the first coordinate system to be orthogonal (so that $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ is orthonormal), then the above formula for $h_i{}'$ reduces to

$$h_i{}' \;=\; \sqrt{\sum_{j=1}^{N} \left[\frac{\partial x^j}{\partial x^{i'}} h_j \right]^2} \quad . \tag{8.15}$$

!►*Example 8.9 (Cartesian and polar coordinates):* *For the first system, use Cartesian coordinates,*

$$\left\{ \left(x^1, x^2 \right) \right\} \;=\; \{ (x, y) \} \quad ,$$

and for the second, use polar coordinates,

$$\left\{ \left(x^{1'}, x^{2'} \right) \right\} \;=\; \{ (\rho, \phi) \} \quad .$$

Here, the basic change of coordinates formulas are

$$x \;=\; x(\rho, \phi) \;=\; \rho \cos(\phi) \qquad \text{and} \qquad y \;=\; y(\rho, \phi) \;=\; \rho \sin(\phi) \quad .$$

For the tangent vector at each point associated with the radial coordinate, we have

$$\begin{aligned} \boldsymbol{\varepsilon}_\rho \;=\; h_\rho \mathbf{e}_\rho \;=\; \frac{\partial \boldsymbol{x}}{\partial \rho} \;&=\; \frac{\partial \boldsymbol{x}}{\partial x} \frac{\partial x}{\partial \rho} \;+\; \frac{\partial \boldsymbol{x}}{\partial y} \frac{\partial y}{\partial \rho} \\ &=\; \mathbf{i} \frac{\partial}{\partial \rho} [\rho \cos(\phi)] \;+\; \mathbf{j} \frac{\partial}{\partial \rho} [\rho \sin(\phi)] \\ &=\; \cos(\phi) \, \mathbf{i} \;+\; \sin(\phi) \, \mathbf{j} \quad . \end{aligned}$$

So, the associated scaling factor is

$$h_\rho \;=\; \left\| h_\rho \mathbf{e}_\rho \right\| \;=\; \| \cos(\phi) \, \mathbf{i} \;+\; \sin(\phi) \, \mathbf{j} \| \;=\; \sqrt{\cos^2(\phi) + \sin^2(\phi)} \;=\; 1 \quad ,$$

and the associated unit tangent is

$$\mathbf{e}_\rho = \frac{\boldsymbol{\varepsilon}_\rho}{\|\boldsymbol{\varepsilon}_\rho\|} = \frac{\cos(\phi)\,\mathbf{i} + \sin(\phi)\,\mathbf{j}}{1} = \cos(\phi)\,\mathbf{i} + \sin(\phi)\,\mathbf{j} \quad .$$

For the tangent vector at each point associated with the ϕ *coordinate, we have*

$$\begin{aligned}\boldsymbol{\varepsilon}_\phi = h_\phi \mathbf{e}_\phi = \frac{\partial \boldsymbol{x}}{\partial \phi} &= \frac{\partial \boldsymbol{x}}{\partial x}\frac{\partial x}{\partial \phi} + \frac{\partial \boldsymbol{x}}{\partial y}\frac{\partial y}{\partial \phi} \\ &= \mathbf{i}\,\frac{\partial}{\partial \phi}[\rho\cos(\phi)] + \mathbf{j}\,\frac{\partial}{\partial \phi}[\rho\sin(\phi)] \\ &= -\rho\sin(\phi)\,\mathbf{i} + \rho\cos(\phi)\,\mathbf{j} \quad .\end{aligned}$$

So, the associated scaling factor is

$$h_\phi = \|h_\phi \mathbf{e}_\phi\| = \|-\rho\sin(\phi)\,\mathbf{i} + \rho\cos(\phi)\,\mathbf{j}\| = \cdots = \rho \quad ,$$

and the associated unit tangent is

$$\mathbf{e}_\phi = \frac{\boldsymbol{\varepsilon}_\phi}{\|\boldsymbol{\varepsilon}_\phi\|} = \frac{-\rho\sin(\phi)\,\mathbf{i} + \rho\cos(\phi)\,\mathbf{j}}{\rho} = -\sin(\phi)\,\mathbf{i} + \cos(\phi)\,\mathbf{j} \quad .$$

Do observe that

$$\mathbf{e}_\phi \cdot \mathbf{e}_\rho = -\sin(\phi)\cos(\phi) + \cos(\phi)\sin(\phi) = 0 \quad ,$$

re-confirming the observation made earlier that polar coordinates is an orthogonal coordinate system.

Finally, we can use the above to find the corresponding components of the metric:

$$g_{\rho\rho} = \boldsymbol{\varepsilon}_\rho \cdot \boldsymbol{\varepsilon}_\rho = (h_\rho)^2 = 1 \quad ,$$

$$g_{\phi\phi} = \boldsymbol{\varepsilon}_\phi \cdot \boldsymbol{\varepsilon}_\phi = (h_\phi)^2 = \rho^2$$

and

$$g_{\rho\phi} = g_{\phi\rho} = \boldsymbol{\varepsilon}_\phi \cdot \boldsymbol{\varepsilon}_\rho = 0 \quad .$$

?►Exercise 8.21 (spherical coordinates): *Let* $\{(r,\theta,\phi\}$ *be the spherical coordinate system described at the start of section 8.3, page 8–9.*

a: *Find the associated basis vectors* $\boldsymbol{\varepsilon}_r$, $\boldsymbol{\varepsilon}_\theta$, $\boldsymbol{\varepsilon}_\phi$, $\mathbf{e}_r$, $\mathbf{e}_\theta$ *and* $\mathbf{e}_\phi$ *in terms of spherical coordinates and* $\mathbf{i}$, $\mathbf{j}$ *and* $\mathbf{k}$.

b: *Using the results of the above, find the scaling factors* h_r , h_θ *and* h_ϕ *(compare with the results obtained in exercise 8.17 on page 8–27).*

c: *Verify that the spherical coordinate system is orthogonal.*

d: *Find the corresponding components of the metric.*

!►Example 8.10: *Again, consider the circular paraboloid* $\mathcal{S}$ *we first discussed in example 8.3 on page 8–16, and returned our attention to in examples 8.5 and 8.8 (pages 8–22 and 8–27) Using a standard Cartesian coordinate system,* $\mathcal{S}$ *is the graph of*

$$z = x^2 + y^2 \qquad \text{for} \quad -\infty < x < \infty \quad \text{and} \quad -\infty < y < \infty \quad ,$$

with each

$$\boldsymbol{x} \sim (x^1, x^2) = (x, y) \quad \text{in the surface } \mathcal{S}$$

being the point

$$\boldsymbol{x} \sim \left(x,\ y,\ x^2 + y^2\right) \quad \text{in the thee-dimensional space } \mathcal{E}^3 \quad .$$

In example 8.8, we derived that, at each point $\boldsymbol{p} \sim (x, y)$ *in* $\mathcal{S}$,

$$\{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2\} = \{\mathbf{i} + 2x\mathbf{k}\,,\ \mathbf{j} + 2y\mathbf{k}\} \quad ,$$

$$\{h_1, h_2\} = \left\{\sqrt{1+4x^2}, \sqrt{1+4y^2}\right\} \quad ,$$

and

$$\{\mathbf{e}_1, \mathbf{e}_2\} = \left\{ \frac{1}{\sqrt{1+4x^2}}[\mathbf{i} + 2x\mathbf{k}]\,,\ \frac{1}{\sqrt{1+4y^2}}[\mathbf{j} + 2y\mathbf{k}] \right\} \quad .$$

This was using the $\{(x, y)\}$ *coordinate system corresponding to the Cartesian* $\{(x, y)\}$ *coordinate system on the* XY*–plane. But we can also use the* $\{(\rho, \phi)\}$ *coordinate system on* $\mathcal{S}$ *corresponding to the polar* $\{(\rho, \phi)\}$ *coordinate system on the* XY*–plane (see example 8.4 on page 8–17). These two coordinate systems are still related via*

$$x = \rho\cos(\phi) \qquad \text{and} \qquad y = \rho\sin(\phi) \quad .$$

So,

$$\begin{aligned}
\boldsymbol{\varepsilon}_\rho = \frac{\partial \boldsymbol{x}}{\partial \rho} &= \frac{\partial \boldsymbol{x}}{\partial x}\frac{\partial x}{\partial \rho} + \frac{\partial \boldsymbol{x}}{\partial y}\frac{\partial y}{\partial \rho} \\
&= \boldsymbol{\varepsilon}_1 \frac{\partial}{\partial \rho}[\rho\cos(\phi)] + \boldsymbol{\varepsilon}_2 \frac{\partial}{\partial \rho}[\rho\sin(\phi)] \\
&= [\mathbf{i} + 2x\mathbf{k}]\cos(\phi) + [\mathbf{j} + 2y\mathbf{k}]\sin(\phi) \\
&= [\mathbf{i} + 2\rho\cos(\phi)\,\mathbf{k}]\cos(\phi) + [\mathbf{j} + 2\rho\sin(\phi)\,\mathbf{k}]\sin(\phi) \\
&= \cos(\phi)\,\mathbf{i} + \sin(\phi)\,\mathbf{j} + 2\rho[\cos^2(\phi) + \sin^2(\phi)]\mathbf{k} \\
&= \cos(\phi)\,\mathbf{i} + \sin(\phi)\,\mathbf{j} + 2\rho\mathbf{k} \quad ,
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{\varepsilon}_\phi = \frac{\partial \boldsymbol{x}}{\partial \phi} &= \frac{\partial \boldsymbol{x}}{\partial x}\frac{\partial x}{\partial \phi} + \frac{\partial \boldsymbol{x}}{\partial y}\frac{\partial y}{\partial \phi} \\
&= \boldsymbol{\varepsilon}_1 \frac{\partial}{\partial \phi}[\rho\cos(\phi)] + \boldsymbol{\varepsilon}_2 \frac{\partial}{\partial \phi}[\rho\sin(\phi)] \\
&= [\mathbf{i} + 2x\mathbf{k}][-\rho\sin(\phi)] + [\mathbf{j} + 2y\mathbf{k}]\rho\cos(\phi) \\
&= [\mathbf{i} + 2\rho\cos(\phi)\,\mathbf{k}][-\rho\sin(\phi)] + [\mathbf{j} + 2\rho\sin(\phi)\,\mathbf{k}]\rho\cos(\phi) \\
&= -\rho\sin(\phi)\,\mathbf{i} + \rho\cos(\phi)\,\mathbf{j} + 0\mathbf{k} \\
&= -\rho\sin(\phi)\,\mathbf{i} + \rho\cos(\phi)\,\mathbf{j} \quad ,
\end{aligned}$$

$$\begin{aligned}
h_\rho = \left\|\boldsymbol{\varepsilon}_\rho\right\| &= \|\cos(\phi)\,\mathbf{i} + \sin(\phi)\,\mathbf{j} + 2\rho\mathbf{k}\| \\
&= \sqrt{\cos^2(\phi) + \sin^2(\phi) + (2\rho)^2} = \sqrt{1 + 4\rho^2} \quad ,
\end{aligned}$$

$$\begin{aligned}
h_\phi = \left\|\boldsymbol{\varepsilon}_\phi\right\| &= \|-\rho\sin(\phi)\,\mathbf{i} + \rho\cos(\phi)\,\mathbf{j}\| \\
&= \sqrt{\rho^2\cos^2(\phi) + \rho^2\sin^2(\phi)} = \rho \quad ,
\end{aligned}$$

$$\mathbf{e}_\rho \;=\; \frac{1}{h_\rho}\boldsymbol{\varepsilon}_\rho \;=\; \frac{1}{\sqrt{1+4\rho^2}}\big[\cos(\phi)\,\mathbf{i} + \sin(\phi)\,\mathbf{j} \;+\; 2\rho\mathbf{k}\big]$$

and

$$\mathbf{e}_\phi \;=\; \frac{1}{h_\phi}\boldsymbol{\varepsilon}_\phi \;=\; \frac{1}{\rho}\big[-\rho\sin(\phi)\,\mathbf{i} + \rho\cos(\phi)\,\mathbf{j}\big] \;=\; -\sin(\phi)\,\mathbf{i} \;+\; \cos(\phi)\,\mathbf{j} \quad .$$

Next, we compute the components of the metric using the $\{(\rho,\phi)\}$ *system:*

$$g_{\rho\rho} \;=\; \boldsymbol{\varepsilon}_\rho \cdot \boldsymbol{\varepsilon}_\rho \;=\; \left(h_\rho\right)^2 \;=\; 1 \;+\; 4\rho^2 \quad ,$$

$$g_{\phi\phi} \;=\; \boldsymbol{\varepsilon}_\phi \cdot \boldsymbol{\varepsilon}_\phi \;=\; \left(h_\phi\right)^2 \;=\; \rho^2$$

and

$$g_{\rho\phi} \;=\; g_{\phi\rho} \;=\; \boldsymbol{\varepsilon}_\phi \cdot \boldsymbol{\varepsilon}_\rho \;=\; \cdots \;=\; 0 \quad .$$

Note that the $\{(\rho,\phi)\}$ *system is an orthogonal coordinate system on* $\mathcal{S}$ *.*

Now you do the same computations using the surfaces from exercise 8.18 on page 8–28.

?►*Exercise 8.22:* *Several surfaces in Euclidean three-space* $\mathcal{E}^3$ *are given below as graphs of functions using a standard Cartesian system* $\{x, y, z)\}$ *. In exercise 8.18, your found the local bases and scaling factors using the corresponding* $\{(x, y)\}$ *coordinate system. This time, let* $\{(\rho,\phi)\}$ *be the coordinate system on each surface corresponding to the standard polar coordinate system on the* XY*–plane and do the following:*

i. *Rewrite the given formula for* z *in terms of* ρ *and* ϕ *.*

ii. *Find the local bases, scaling factors, and metric components at each point* $\boldsymbol{p} \sim (\rho,\phi)$ *(as just done in the previous example).*

iii. *Determine if the* $\{(\rho,\phi)\}$ *system is an orthogonal coordinate system on the surface.*

a: $\mathcal{S}$ *is the flat surface given by*

$$z \;=\; 1 \;+\; 2x \;+\; y \quad .$$

b: $\mathcal{S}$ *is the graph of*

$$z \;=\; \sqrt{x^2 + y^2} \quad .$$

c: $\mathcal{S}$ *is the graph of*

$$z \;=\; x^2 \;-\; y^2 \quad .$$