

Calculus Typed

Taylor Hummon

Contents

I	Numbers	7
1	Functions	8
1.1	Components	9
1.2	Dependence	12
1.3	Functions	15
1.4	Inverse Functions	17
1.5	Restrictions	20
2	Height	22
2.1	Height	23
2.2	Five Graphs: Height	25
2.3	Vertical Transformers	27
2.4	Horizontal Transformers	29
3	Differential Laws	32
3.1	Differential Laws	33
3.2	Subtraction and Constants	36
3.3	Squaring and Cubing	39
3.4	Division	41
3.5	Square Root and Cube Root	43
3.6	Review	45
4	Slope	46
4.1	Derivatives	47
4.2	Slope	50
4.3	Level Points	52
4.4	Five Graphs: Slope	54
4.5	Analyzing Slope	56
4.6	Linear Functions	59

5	Bend	62
5.1	Tangent Lines	63
5.2	The Second Derivative	65
5.3	Bend	67
5.4	Five Graphs: Bend	69
5.5	Low Points and High Points	71
5.6	Taylor Polynomials	73
6	Powers	75
6.1	Powers	76
6.2	Base Functions	79
6.3	The Natural Exponential	82
6.4	The Natural Logarithm	84
6.5	Simple Differential Laws	86
6.6	Complicated Differential Laws	88
II	Metrics	90
7	Surfaces	91
7.1	Surfaces as Graphs	92
7.2	Vertical Slices	94
7.3	Horizontal Slices	96
7.4	A Bowl	98
7.5	A Parabolic Saddle	101
7.6	A Warp Saddle	103
7.7	A Thick Parabola	105
8	Partials	107
8.1	Partials	108
8.2	Slope	111
8.3	Planar Functions	113
8.4	Tangent Planes	115
8.5	Bend	117
8.6	Warp	119
9	Differentials as Metrics	121
9.1	The Metric Differential	122
9.2	Parts	125
9.3	Polymetrics	127
9.4	Higher Differentials	129

9.5	The Saddle Discriminant	132
9.6	Low Points and High Points	135
III	Rulers	139
10	Vectors	140
10.1	Vectors	141
10.2	Standard Vectors	143
10.3	Position Vectors	145
10.4	Magnitude	147
10.5	Scaling	150
11	Rulers	153
11.1	Standard Rulers	154
11.2	Rulers	157
11.3	Markings	160
11.4	Measurement	162
11.5	Magnitude	164
11.6	Transpose	166
12	Differentials as Rulers	169
12.1	The Ruler Differential	170
12.2	Directional Slope	172
12.3	The Gradient	174
12.4	Polyvectors	176
12.5	Polyrulers	178
12.6	Higher Differentials	180
12.7	The Second Differential	183
13	Boxes	185
13.1	Integrals of Blips	186
13.2	Intervals	188
13.3	Integrals of Intervals	190
13.4	Anti-Differentiation	193
13.5	Boxes	196
13.6	Boundary	199
13.7	Integrals of Boxes	202
14	Paths	205
14.1	Paths	206

14.2 Motion	208
14.3 Path Integrals	211
14.4 Tame and Wild Rulers	213
14.5 Path Integrals, Revisited	216
15 Angles	219
15.1 Angles	220
15.2 Unit Circle	223
15.3 Differential Laws	226
15.4 Polar Coordinates	228
15.5 Polar Differentials	231
15.6 Polar Area	233
Appendices	237
A Are Variables Dummies?	237
B The Natural Exponential	239
C Multi-Components	244
D Independence	249
E Applying a Metric	251
F Ruler Explanations	254
G Rulers in Space	258
Summaries	262
Summary of Abstract Differential Laws	262
Summary of Vectors, Rulers, Metrics	264
Summary of Notation	265

Introduction

Calculus is a branch of mathematics that, simply enough, lets people make calculations! Most of the time the calculations we'll make using Calculus will have something geometric to say. We'll learn how to calculate whether a curve is steep or shallow. We'll learn to analyze how bent and warped a surface is. And we'll learn to measure length and area in powerful new ways. Remarkably, all of these geometric ideas can be linked to a single algebraic idea: the differential.

In Part I, we'll learn to calculate the differential, and we'll see how we can apply the differential to study curves. Before we can define the differential, however, we'll need to know how to decompose functions. By learning about dependence, we'll see how complicated functions can be built from simpler functions.

In Part II, we'll look at metrics as a way study surfaces. Metrics will help us understand the shape of a surface by comparing the surface to a plane. The differential lives comfortably in the world of dependent metrics.

In Part III, we'll look at a few more types of object: rulers, vectors, and boxes. We'll use rulers to measure vectors, and we'll use dependent rulers to measure boxes. The differential also lives comfortably in the world of dependent rulers.

To get the most out of this text, you'll need to keep pen and paper nearby as you read. Following another person's calculations can get tedious quite quickly. Math is much more fun when you're sorting things out for yourself! Whether you have some background in Calculus already or you're entirely new to the subject, I hope that you find a few ideas that spark your interest and make you want to continue on exploring.

Good luck!

- Taylor

Part I

Numbers

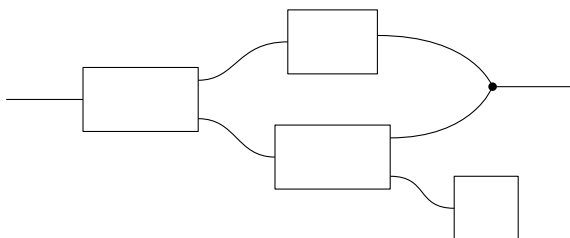
Chapter 1

Functions

A function tells us how one quantity depends on other quantities. You've no doubt already seen plenty examples of functions, and you'll encounter many more as you continue your studies.

- The area of a circle *depends* on its radius.
- The height of a projectile *depends* on the time since its launch.
- The pressure of an ideal gas *depends* on its volume, its temperature, and the number of molecules.
- etc.

Our approach to functions will emphasize the ways in which we can combine simple functions to get more complicated examples. In particular we'll borrow the idea of a *schematic* from engineering to depict how functions fit together.



Schematics are indispensable for explaining subtle issues of dependence, but they are cumbersome for day-to-day calculations. Most of the time we'll stick to using a concise algebraic notation. But schematics will always be nearby: there is no Calculus without dependence!

1.1 Components

Components are the basic building blocks that we'll use to construct functions. A component has an output wire on its left side and a fixed number of input wires on its right side. Here are the *addition*, *subtraction*,



multiplication, and *division* components.



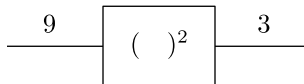
The *squaring* and *cubing* components each have one input wire,



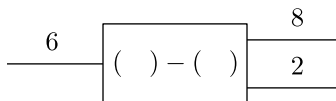
as do the *square root* and *cube root* components.



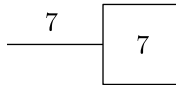
Given values for each of its input wires, a component produces a single value for its output wire in a reproducible way. When we give the squaring component an input three, it will produce the output $3^2 = 9$.



When a component has more than one input wire, we read its inputs from top to bottom. For example, when we give the subtraction component the inputs $(8, 2)$, it will output the result $8 - 2 = 6$.



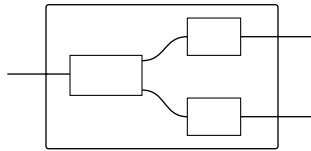
A *constant component* takes no input and will always produce the same output.



Every number gets its own constant component.

Schematics

By connecting up the wires between several components, we can create a *schematic*. We can think of a schematic as a large component made up of smaller components.



A schematic may have multiple input wires and multiple internal wires, but we expect it to have a single output wire, only. We'll always attach wires compatibly: one component's output wire is another component's input wire.

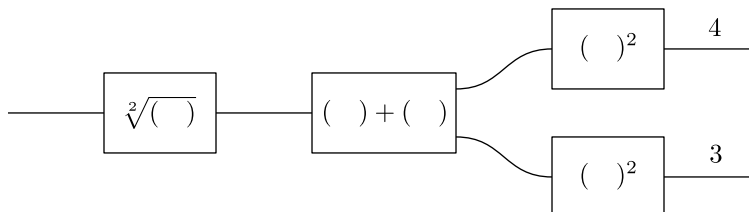
Important.

We will never wire output to output nor input to input.

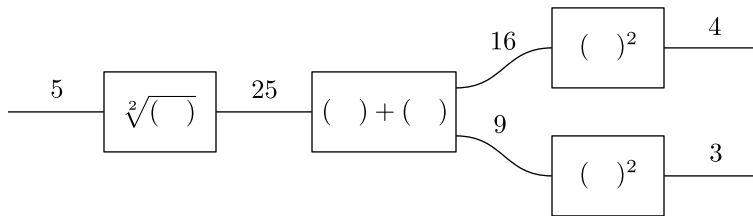
We evaluate a schematic by pushing the given input values across components from right to left. Let's see how this works with an example.

Example.

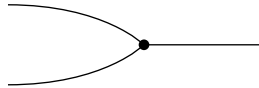
Let's evaluate the following schematic for the pair of inputs (4, 3). We start by placing the inputs on the right-most wires.



Pushing the values across the components, we find that schematic produces the output value five on its left-most wire.



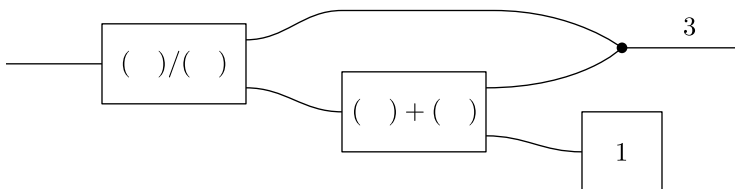
A schematic can include components, wires, and one other kind of object: a splitter.



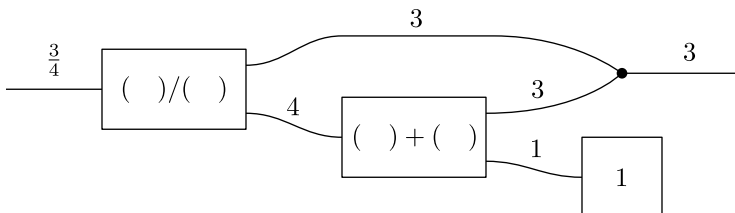
A splitter duplicates its input by producing the same value on each of its output wires. Since a splitter does not have a single output wire, it is not considered a component.

Example.

Let's evaluate the following schematic for the input three.



Our input is duplicated by the splitter, and the constant component introduces its value, one.



The schematic's output is three-quarters.

1.2 Dependence

Schematics are great for illustrating relationships between variables! We'll write

$x : \text{Number}, \quad y : \text{Number}, \quad z : \text{Number}$

to say that x , y , and z are variables that hold numbers. If a schematic has the variable z on its output wire and the variables x and y on its input wires, then we'll write

$z \dot{\leftarrow} (x, y)$

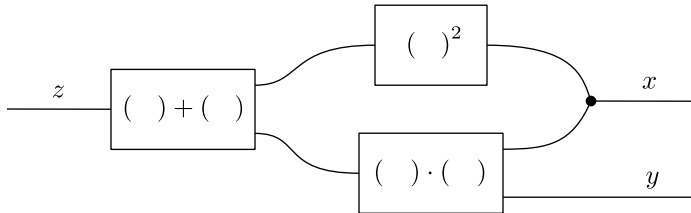
to say that z *depends on* x and y . Similarly, we'll write

$z \dot{\leftarrow} x$

to say that z *depends only on* x . While we are free to use any variable name for any purpose, we'll usually reserve z for output and x and y for input.

Example.

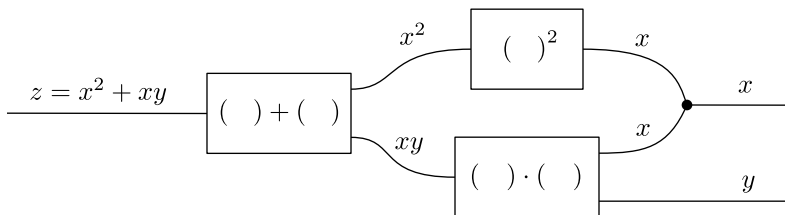
Let's take a look at the following schematic. We've placed variables on some of the wires.



Overall we have an output variable, z , and two input variables, x and y .

$z \dot{\leftarrow} (x, y)$

By thinking of our input variables as values, we can push them from right to left through the schematic.



When two expressions z and $x^2 + xy$ share the same wire, we expect them to be equal! So we've learned not just *that* z depends on x and y , but *how*.

$$z = x^2 + xy$$

If we know how a variable depends on other variables, we can draw its schematic. We'll start on the left with the output wire and work to the right.

Example.

Suppose the variables z : Number and x : Number are related as follows.

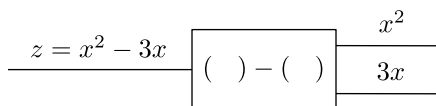
$$z \leftarrow x$$

$$z = x^2 - 3x$$

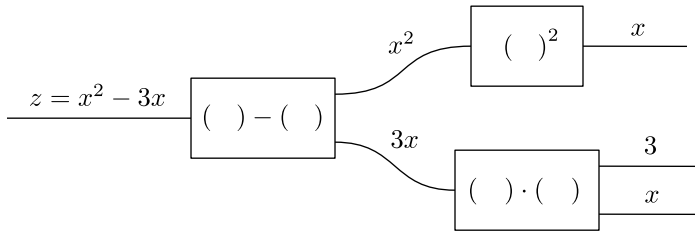
That is, z depends on x , and the equation tells us how. Let's draw this relationship as a schematic. We begin with our output wire z .

$$\underline{z = x^2 - 3x}$$

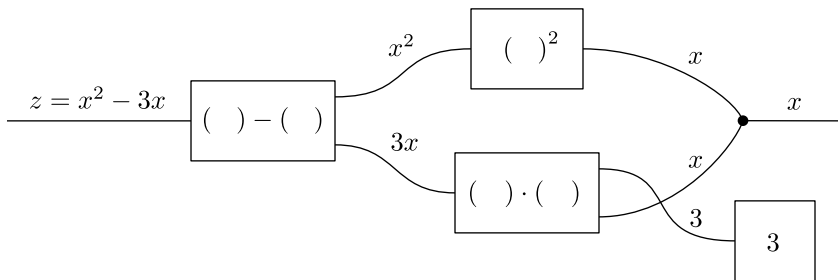
Overall, our expression, $x^2 - 3x$, is made by subtracting. We draw the subtraction component and write the expressions x^2 and $3x$ on its input wires.



Let's further decompose each of these! The expression x^2 is made by squaring, and the expression $3x$ is made by multiplying.



We now have two wires labeled with the variable x . Let's join these together so that that this shared input comes from a single wire. And let's cap off the constant 3 with a constant component.



We've found our schematic!

See if you can draw schematics for each of the following.

$$z = x^3 - 1$$

$$z = \sqrt[2]{x + y}$$

$$z = 5 \cdot (x^2 + x)$$

$$z = \frac{3y}{x - y}$$

It's good to get practice drawing schematics from left to right, as we did in the previous example. At each step decomposing, try to identify what left-most component was used to build the expression.

1.3 Functions

A *function* is a rule that describes how one number variable depends on other number variables. When we define a function,

- we declare the output variable and the input variables,
- we give a defining rule,
- we optionally name the function, and
- we optionally declare restrictions.

We'll focus on the first three aspects of functions, and we'll come back to discuss [restrictions](#) by the end of this chapter.

Example.

Let's see an example of a function definition.

$$z \stackrel{f}{\leftarrow} x$$
$$z = x^2 + x + 3$$

The first line declares that our function has z as its output variable and x as its only input variable. The second line, $z = x^2 + x + 3$, is the defining rule. The defining rule tells us how z depends on x . We gave our function the name f , and we didn't declare any restrictions.

Just as we may evaluate a schematic by placing values on the input wires, we *apply* a function by substituting input values into the defining rule.

Example.

Let's make a table of values by applying the above function to various inputs.

$$f(-2) = (-2)^2 + (-2) + 3 = 5$$

$$f(-1) = (-1)^2 + (-1) + 3 = 3$$

$$f 0 = 0^2 + 0 + 3 = 3$$

$$f 1 = 1^2 + 1 + 3 = 5$$

$$f 2 = 2^2 + 2 + 3 = 9$$

x	z
-2	5
-1	3
0	3
1	5
2	9

Notice that some z -values appear multiple times in the table of values. For example, our function gives the same output $z = 5$ to both $x = -2$ and $x = 1$.

Let's look at a function that has two input variables.

Example.

Consider the function

$$z \stackrel{g}{\leftarrow} (x, y)$$
$$z = 5x + 4xy - 2.$$

To apply this function, we'll give a value for each of the input variables.

$$g(0, 0) = 5 \cdot 0 + 4 \cdot 0 \cdot 0 - 2 = -2$$

$$g(1, 0) = 5 \cdot 1 + 4 \cdot 1 \cdot 0 - 2 = 3$$

$$g(0, 1) = 5 \cdot 0 + 4 \cdot 0 \cdot 1 - 2 = -2$$

We substitute the first input value for x and the second for y .

We'll end this section with a discussion of notation for functions and operations. Here are three equivalent ways to write a function application.

$$f\ 7 = f(7) = (f\ 7)$$

We won't require parentheses when we apply a function that has one input variable, but they may be included for clarity. If a function expects two inputs, we will place parentheses around the pair of input values.

$$g(3, 1)$$

When we perform arithmetic, we'll use the usual order of operations. For example, division binds before addition, and so

$$8 + x/4$$

should be read as $8 + (x/4)$ and not $(8 + x)/4$. Function application binds before any arithmetic operation: addition, subtraction, multiplication, division. As a result,

$$f\ 4 + 6$$

means $(f\ 4) + 6$ and not $f(4 + 6)$. We may always group expressions with parentheses to override the order of operations.

1.4 Inverse Functions

We use a function $z \xleftarrow{\cdot} x$ to describe how one variable, z , depends on another variable, x . By inverting the function, we reverse the direction of dependence.

Definition.

Suppose $b \xleftarrow{f} a$ is a function, and suppose we can solve for the input variable, a . Then

$$a \xleftarrow{\cdot} b$$

is the *inverse function* for f .

We'll avoid using the variables x , y , and, z when inverting a function, since these each come with a bias towards being used either for input or for output. See [Appendix A. Are Variables Dummies?](#) for a brief discussion of how we choose variable names.

Example.

Consider the function

$$\begin{aligned} b &\xleftarrow{f} a \\ b &= 2a - 6. \end{aligned}$$

Let's solve for a .

$$\begin{aligned} b &= 2a - 6 \\ b + 6 &= 2a \\ \frac{b + 6}{2} &= a \end{aligned}$$

We've found the inverse of f .

$$\begin{aligned} a &\xleftarrow{\cdot} b \\ a &= \frac{b + 6}{2} \end{aligned}$$

We may cancel a function by applying its inverse.

Important.

If two functions

$$b \xleftarrow{f} a \quad \text{and} \quad a \xleftarrow{g} b$$

are inverse, then we can make cancellations,

$$g f a = a$$

$$f g b = b.$$

The cube root function is defined as the inverse of the cubing function.

$$\begin{array}{ll} b \xleftarrow{\cdot} a & a \xleftarrow{\cdot} b \\ b = a^3 & \text{and} \quad a = \sqrt[3]{b} \end{array}$$

Let's see how the cube root cancels with cubing.

$$\begin{aligned} \sqrt[3]{3^3} &= \sqrt[3]{27} = 3 \\ (\sqrt[3]{-8})^3 &= (-2)^3 = -8 \end{aligned}$$

If we cube a number and then take the cube root, then we get back the number we started with. And the same is true working in the other order.

Example.

Let's find an inverse for the function

$$\begin{aligned} b &\xleftarrow{\cdot} a \\ b &= (a + 1)^3 - 4. \end{aligned}$$

It is our job to solve for a .

$$\begin{aligned} b &= (a + 1)^3 - 4 \\ b + 4 &= (a + 1)^3 \end{aligned}$$

We'll take cube roots of both sides of this equation so that we may make a cancellation.

$$\sqrt[3]{b+4} = \sqrt[3]{(a+1)^3}$$

$$\sqrt[3]{b+4} = a+1$$

$$\sqrt[3]{b+4} - 1 = a$$

We've found the inverse.

$$a \leftarrow b$$

$$a = \sqrt[3]{b+4} - 1$$

We would like to define the square root as the inverse of the squaring function,

$$b \leftarrow a$$

$$a \leftarrow b$$

$$b = a^2$$

and

$$a = \sqrt[2]{b},$$

but we run into trouble!

- Any positive b -value, say $b = 9$, corresponds to two a -values: $a = 3$ and $a = -3$.
- Any negative b -value, say $b = -9$, corresponds to zero a -values.

For the square root to make any sense, we must have a single a -value for each b -value. To remedy these issues, we'll use restrictions:

$$a \geq 0 \quad \text{and} \quad b \geq 0.$$

We may now define the square root as the inverse to the squaring function. These restrictions have real consequences, however. We'll refuse to take the square root of a negative number. And the cancellation rules

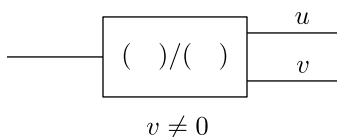
$$\sqrt[2]{a^2} = a$$

$$(\sqrt[2]{b})^2 = b$$

only hold for positive a -values and positive b -values.

1.5 Restrictions

When we define a function, we may specify *restrictions* on its input variables by stating what values are *allowed* or *disallowed*. The division component comes with a restriction: we must avoid dividing by zero.



Whenever we write a function that uses division, we'll include restrictions to ensure that the denominator is non-zero.

Example.

Consider the function

$$z \leftarrow x$$

$$z = \frac{3}{x^2 - 4} + 6.$$

What restrictions do we need on this function's x -values? For division to be defined, we'll need the denominator to be non-zero.

$$x^2 - 4 \neq 0$$

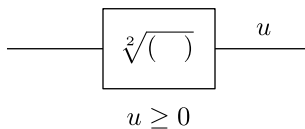
Let's see what x -values must be disallowed.

$$(x + 2) \cdot (x - 2) \neq 0$$

$$x \neq -2 \quad \text{and} \quad x \neq 2$$

All inputs except $x = -2$ and $x = 2$ may be safely allowed.

Apart from division, there is only one other component that needs a restriction at this time.



As we saw in the preceding section, the square root requires its input to be positive.

Example.

Let's see what restrictions we need for the function

$$z \leftarrow x$$

$$z = \sqrt[2]{x^2 + 3} - 5.$$

In order for our function to be defined, we'll need the input to the square root to be positive.

$$x^2 + 3 \geq 0$$

But wait a second! When we square a number x , the result, x^2 , is always positive. And adding three will only make it larger. Since every x -value will make $x^2 + 3$ positive, we don't need any restrictions on x for this function.

We'll say that a function is *undefined* at each of its disallowed inputs. So in particular,

$$\frac{5}{0} \quad \text{and} \quad \sqrt[2]{-4}$$

are both undefined.

Division By Zero

As we study Calculus, there will be times when we will be tempted to divide by zero. Let's take a moment to think through why this is a bad idea.

Question.

What's so wrong with dividing by zero?

For the sake of argument, suppose that we defined the fraction $1/0$ to be some number. Then we could write

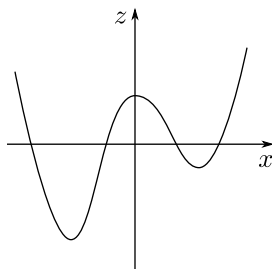
$$0 \cdot \frac{1}{0} = 0 \quad \text{and} \quad 0 \cdot \frac{1}{0} = 1.$$

The first equality must hold because multiplying zero and any number gives zero. The second equality must hold because multiplying any number with its reciprocal gives one. We are forced to conclude that zero and one are equal, but this collapses our mathematical world! We must therefore be vigilant: our computations can never include a division by zero.

Chapter 2

Height

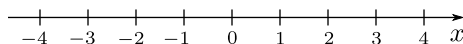
Let's continue our study of functions by looking at their graphs. To graph a function $z \leftarrow x$, we use a vertical z -axis for output and a horizontal x -axis for input. Our approach to graphing will emphasize the height of the graph at any point. This is a bit of foreshadowing. We'll have chapters on [slope](#) and [bend](#) that look at other quantities found on a graph.



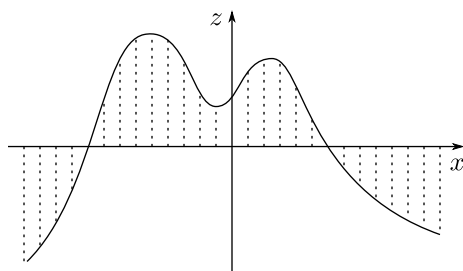
Let's take a short digression to think about our everyday experience with directions and their impact on graphing. In day to day life, we tend to believe that *up* and *down* are absolute directions in a way that *left* and *right* are not. This belief is not some deep mathematical truth, but rather, it comes from a simple physical fact. The Earth's gravity always tells us which way is up and which way is down! We'll take advantage of this asymmetry when graphing. As we'll soon see, it is much easier to work with a function's output than its input. By matching output, z , to the preferred direction, up, we bring significant visual intuition to our graphs.

2.1 Height

To graph a function $z \leftarrow x$ that has one input variable, we'll begin by drawing a horizontal line of input values, the x -axis.



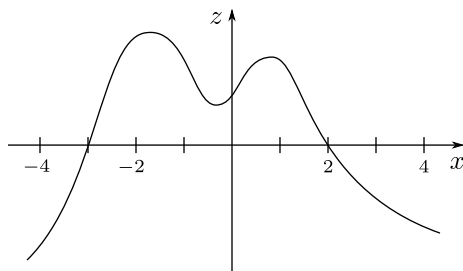
For each input x on the x -axis, our function determines a single point on the *graph* by giving its *height* z .



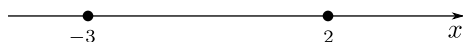
We draw the graph above or below the x -axis according to whether the height is positive or negative.

local quantity	local quality
$z > 0$	above the x -axis
$z < 0$	below the x -axis

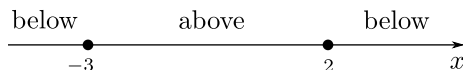
A graph has zero height wherever it intersects the x -axis. Our graph has two points with zero height: one at $x = -3$ and one at $x = 2$.



A *height transition diagram* catalogs the graph's height by splitting up the x -axis at points with zero height.



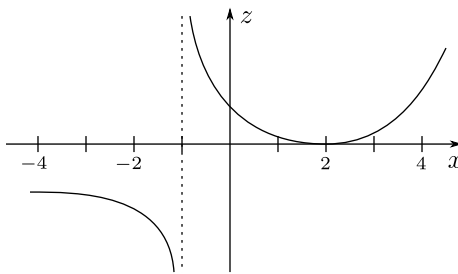
On each region in the diagram, we record whether the height is positive or negative. That is, we'll label a region by stating whether the graph is above or below the x -axis.



The height *transitions* at a point if the graph is above the axis on one side of the point, and below the axis on the other. Looking at our diagram, we see that the height transitions at both $x = -3$ and $x = 2$.

Example.

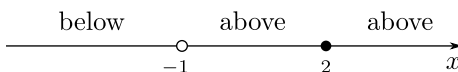
Let's analyze the height of this graph by drawing a transition diagram.



This function has zero height at $x = 2$, and the height is undefined at $x = -1$. We should mark both of these points on the transition diagram. We'll draw the point at $x = -1$ as an open dot to indicate that the height is undefined.



By consulting the graph, we can label each region in the transition diagram according to whether the graph is above or below the x -axis.



The graph's height transitions at $x = -1$. At the point $x = 2$, the story is different. Looking to either side of this point, the graph stays above the x -axis. Although the height is zero at $x = 2$ the graph does not transition here.

In Calculus we'll be working with smooth functions, only! For a graph to "jump," there must be a point that serves as a "gap." In particular, this means that the height of a graph can only transition at points where the height is zero or undefined.

2.2 Five Graphs: Height

Let's take a look at five functions that will make many appearances throughout this text. We'll review their graphs and get some practice drawing their transition diagrams.

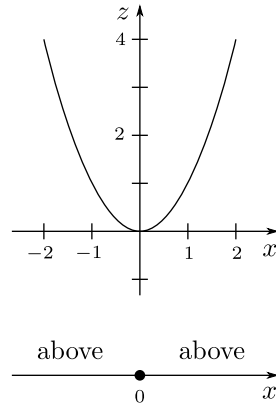
Squaring

First let's consider the squaring function,

$$z \leftarrow x$$

$$z = x^2.$$

The graph of this function is, no doubt, quite familiar: a *parabola*. Although the height is zero at $x = 0$, the height does not transition.



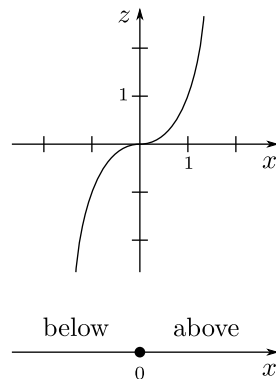
Cubing

Next we'll look at the cubing function,

$$z \leftarrow x$$

$$z = x^3.$$

We can plot some points to produce a graph. Again we find height zero at $x = 0$, and this time the height does transition.



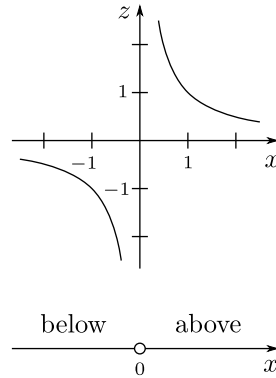
Reciprocal

Let's take a look at the reciprocal function,

$$z \leftarrow x$$

$$z = \frac{1}{x}.$$

The graph of the reciprocal function is called a *hyperbola*. To avoid dividing by zero, we must disallow $x = 0$ as input to the reciprocal. We see that the height transitions at $x = 0$.



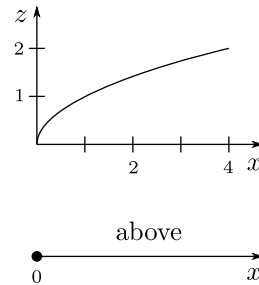
Square Root

The square root function,

$$z \leftarrow x$$

$$z = \sqrt[2]{x},$$

is defined for positive inputs, $x \geq 0$, only. The height is zero at $x = 0$ and is positive for positive inputs.



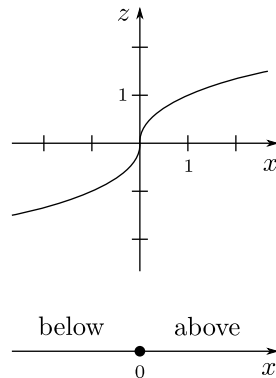
Cube Root

The cube root function,

$$z \leftarrow x$$

$$z = \sqrt[3]{x},$$

is defined for all inputs. The height transitions at $x = 0$.



2.3 Vertical Transformers

Transformers

A transformer is a function that is used to produce geometric effects on another function's graph. For any constant $c : \text{Number}$, we have a *constant-adder transformer*



and a *constant-multiplier transformer*.



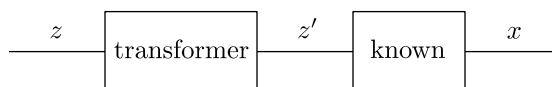
The *negation transformer* is the constant-multiplier that multiplies by negative one.



We'll look at using transformers on a function's output in the remainder of this section, and we'll look at using transformers on a function's input in the next.

Vertical Transformers

Applying a transformer to the output of a function produces a vertical geometric effect on the graph. The following schematic depicts a vertical transformation!



Suppose $z \leftarrow z'$ is a transformer, and $z' \leftarrow x$ is a function with a known graph. Then we can understand the graph of the function $z \leftarrow x$ according to the following table.

algebraic relationship	geometric relationship
$z = z' + 3$	translate up by three
$z = z' - 3$	translate down by three
$z = 2z'$	vertically stretch by a factor of two
$z = z'/2$	vertically shrink by a factor of two
$z = -z'$	vertically reflect

Stretching, shrinking, and reflecting happen with respect to the x -axis.

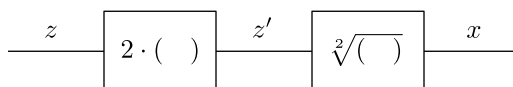
Example.

Let's graph the function

$$z \stackrel{f}{\leftarrow} x$$

$$z = 2 \cdot \sqrt[2]{x}.$$

We'll start by drawing a schematic for f .

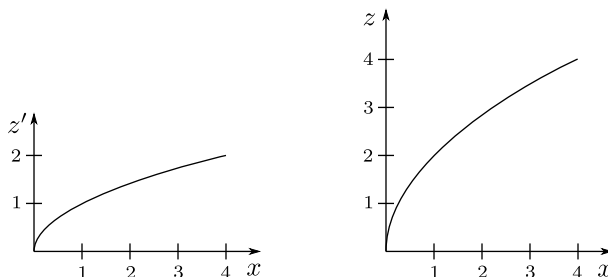


By introducing the variable z' on the intermediate wire, we can decompose f .

$$z \stackrel{\cdot}{\leftarrow} z' \quad \text{and} \quad z' \stackrel{\cdot}{\leftarrow} x$$

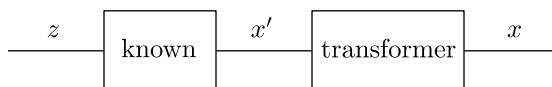
$$z = 2 \cdot z' \quad \text{and} \quad z' = \sqrt[2]{x}$$

We know how to graph the square root function, $z' \stackrel{\cdot}{\leftarrow} x$. To graph the function $z \stackrel{\cdot}{\leftarrow} x$, each height z is twice the height z' .



2.4 Horizontal Transformers

Transforming the input of a function produces a horizontal geometric effect. Here's a schematic depicting a horizontal transformation.



Suppose $z \leftarrow x'$ is a function with a known graph, and $x' \leftarrow x$ is a [transformer](#). Then we can understand the graph of the function $z \leftarrow x$.

algebraic relationship	geometric relationship
$x = x' + 3$	translate right by three
$x = x' - 3$	translate left by three
$x = 2x'$	horizontally stretch by a factor of two
$x = x'/2$	horizontally shrink by a factor of two
$x = -x'$	horizontally reflect

Horizontal transformers looks a lot like vertical transformers, but there is an important subtle distinction to working with inputs!

Important.

We apply a horizontal transformer in the form

$$x' \leftarrow x,$$

but we understand the geometry of a horizontal transformer in the form

$$x \leftarrow x'.$$

This is best understood with examples.

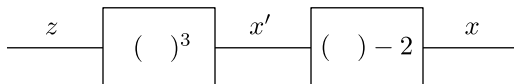
Example.

Consider the function

$$z \stackrel{f}{\leftarrow} x$$

$$z = (x - 2)^3.$$

Let's draw a schematic to understand how to graph this function.



By placing the variable x' on the intermediate wire, we can decompose f as a known function and a transformer.

$$z \stackrel{\cdot}{\leftarrow} x' \qquad x' \stackrel{\cdot}{\leftarrow} x$$

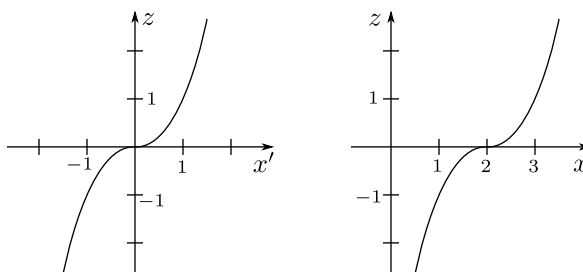
$$z = x'^3 \qquad \text{and} \qquad x' = x - 2$$

We'll make sense of this horizontal transformer by solving for x .

$$x \stackrel{\cdot}{\leftarrow} x'$$

$$x = x' + 2$$

Each x -value is two more than the corresponding x' -value. To graph f , we translate the graph of the cubing function two to the right.



Let's also see an example that includes multiple horizontal transformers.

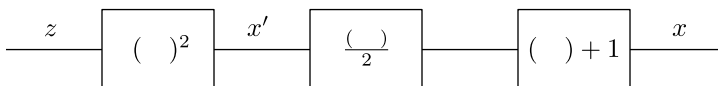
Example.

Consider the function

$$z \stackrel{f}{\leftarrow} x$$

$$z = \left(\frac{x+1}{2} \right)^2.$$

We'd like to graph this function by understanding its transformers. We'll start by drawing a schematic.



We can decompose f as a known function and transformers.

$$z \stackrel{\cdot}{\leftarrow} x' \qquad x' \stackrel{\cdot}{\leftarrow} x$$

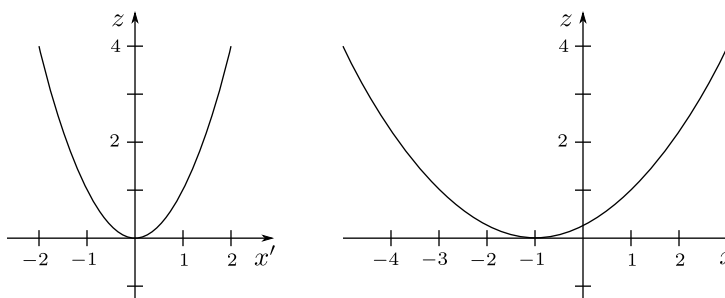
$$z = x'^2 \qquad \text{and} \qquad x' = \frac{x+1}{2}$$

Let's invert the transformers by solving for x .

$$x \stackrel{\cdot}{\leftarrow} x'$$

$$x = 2x' - 1$$

To produce an x -value, we double an x' -value and then subtract one.



To graph f , we take our known parabola, horizontally stretch by a factor of two, and then translate one to the left.

Chapter 3

Differential Laws

The differential is *the* central idea in Calculus. This text is devoted to exploring interpretations and consequences of differential laws. We'll see in due time that the differential allows us to make sense of many geometric ideas including slope, bend, warp, gradient, length, and area. Before we get to all of that, however, it's helpful to get some practice calculating.

$$\partial(\quad)$$

As you work through this chapter, you'll be seeing many laws and many examples of calculations. Don't skip the examples! It's your job to become familiar with these differential laws, and this takes some time and energy. You should be able to justify each step in a calculation with one or more laws. That said, we'll be covering too many laws for you to try to memorize every last one. I would recommend making a hand-written list of the differential laws as you encounter them. This will help you learn the easy laws, and will serve as a reference for the laws that are too unpleasant to commit to memory.

3.1 Differential Laws

The *differential* is an operator ∂ that interacts with functions. Given a function

$$z \leftarrow (x, y)$$

we'll introduce new *differential variables*:

$$\partial z, \quad \partial x, \quad \partial y.$$

We'll pronounce a differential variable ∂z as, "the differential of z," or just, "d of z." We can make differential calculations by applying the following three *fundamental differential laws*.

Law.

The differential law for addition states

$$\partial(u + v) = \partial u + \partial v$$

where $u : \text{Number}$ and $v : \text{Number}$ are numbers.

Law.

The differential law for multiplication states

$$\partial(u \cdot v) = v \cdot \partial u + u \cdot \partial v$$

where $u : \text{Number}$ and $v : \text{Number}$ are numbers.

Law.

The differential law for constants states

$$\partial(c) = 0$$

where $c : \text{Number}$ is constant.

Let's see how we can use these laws to take the differential of a simple function.

Example.

Let's calculate the differential of the function

$$z \leftarrow x$$

$$z = 2x + 5.$$

To calculate ∂z , we start by substituting $2x + 5$ for z . We'll then apply the differential laws.

$$\begin{aligned}\partial z &= \partial(2x + 5) \\ &= \partial(2x) + \partial(5) \\ &= \partial(2x) + 0 \\ &= \partial(2x) \\ &= x \cdot \partial(2) + 2 \cdot \partial(x) \\ &= x \cdot 0 + 2 \cdot \partial x \\ &= 2 \cdot \partial x\end{aligned}$$

That's as far as we can go!

In our example you might have been surprised to see that the result, $2 \cdot \partial x$, contains the differential variable ∂x . We'll be satisfied with a differential calculation when we've found an expression in standard form. Here are the *standard forms for the differential* when working with functions that have one or two input variables.

function	standard form
$z \leftarrow x$	$\partial z = a \cdot \partial x$
$z \leftarrow (x, y)$	$\partial z = a \cdot \partial x + b \cdot \partial y$

Let's look at an example where we work with a function that has two input variables.

Example.

Consider the function

$$z \leftarrow (x, y)$$

$$z = 2x + 3y + 8.$$

Let's calculate the differential of z .

$$\begin{aligned}\partial z &= \partial(2x + 3y + 8) \\ &= \partial(2x) + \partial(3y) + \partial(8) \\ &= x \cdot \partial(2) + 2 \cdot \partial(x) + y \cdot \partial(3) + 3 \cdot \partial(y) + 0 \\ &= x \cdot 0 + 2 \cdot \partial x + y \cdot 0 + 3 \cdot \partial y \\ &= 2 \cdot \partial x + 3 \cdot \partial y\end{aligned}$$

We've found our answer.

The differential probably looks fairly mysterious at this point. As we progress through this chapter, we will see many more differential laws, and we will get lots of experience calculating the differential. And that's just the beginning! The remainder of this text is devoted to studying the differential and learning what it has to teach us about geometry.

3.2 Subtraction and Constants

In the preceding section we learned three fundamental differential laws, and we saw our first examples calculating the differential. In order to make more calculations, we'll need more differential laws! For example, we do not (yet) have a differential law for subtraction. If we try to take the differential of a function that uses subtraction, we'll get stuck.

But don't give up on subtraction just yet! By playing with the fundamental differential laws, we can deduce many new differential laws. In this section, we'll see several differential laws, including one for subtraction. Our goal for the remainder of this chapter is to find a differential law for each of the components we've studied.

Subtraction

To get at a differential law for subtraction, we'll start with something slightly simpler: a differential law for [negation](#).

Law.

The differential law for negation states

$$\partial(-u) = -\partial u$$

where u : Number is a number.

Why?

We can rewrite negation as a product:

$$-u = -1 \cdot u.$$

We can then calculate.

$$\partial(-u) = \partial(-1 \cdot u) = u \cdot \partial(-1) + (-1) \cdot \partial(u) = u \cdot 0 + (-1) \cdot \partial u = -\partial u$$

The differential law for negation says that we may pull a negative sign out of a differential. We're now in a position to state and explain the differential law for subtraction.

Law.

The differential law for subtraction states

$$\partial(u - v) = \partial u - \partial v$$

where u : Number and v : Number are numbers.

Why?

We can rewrite a subtraction as an addition with a negation.

$$u - v = u + (-v)$$

This allows us to calculate,

$$\partial(u - v) = \partial(u + (-v)) = \partial u + \partial(-v) = \partial u + (-\partial v) = \partial u - \partial v.$$

Let's see how we can use the differential law for subtraction.

Example.

Consider the function

$$z \leftarrow (x, y)$$

$$z = 2x - 5y.$$

Let's compute the differential!

$$\begin{aligned} \partial z &= \partial(2x - 5y) \\ &= \partial(2x) - \partial(5y) \\ &= (x \cdot \partial(2) + 2 \cdot \partial(x)) - (y \cdot \partial(5) + 5 \cdot \partial(y)) \\ &= 2 \cdot \partial x - 5 \cdot \partial y \end{aligned}$$

Constants

We've already seen a differential law for constants: the differential of any constant is zero. We use constants all of the time, and so it's worth the effort to learn a couple more differential laws that deal specifically with constants.

Laws.

Suppose u : Number is a number and c : Number is constant. The differential law for constant-adders states

$$\partial(u + c) = \partial u,$$

and the differential law for constant-multipliers states

$$\partial(c \cdot u) = c \cdot \partial u.$$

Why?

We'll check the constant-adder law first.

$$\partial(u + c) = \partial u + \partial c = \partial u + 0 = \partial u$$

And now let's check the constant-multiplier law.

$$\partial(c \cdot u) = u \cdot \partial c + c \cdot \partial u = u \cdot 0 + c \cdot \partial u = c \cdot \partial u$$

These laws are quite useful! They prune a few steps from a differential calculation that weren't actually going anywhere.

Example.

Let's calculate the differential of the function

$$\begin{aligned} z &\leftarrow x \\ z &= 7x + 4. \end{aligned}$$

We'll calculate using our new differential laws.

$$\begin{aligned} \partial z &= \partial(7x + 4) \\ &= \partial(7x) \\ &= 7 \cdot \partial x \end{aligned}$$

We first apply the constant-adder law, and then we apply the constant-multiplier law.

3.3 Squaring and Cubing

Let's learn how to take the differential when squaring is involved!

Law.

The differential law for squaring states

$$\partial(u^2) = 2u \cdot \partial u$$

where u : Number is a number.

Why?

We'll check our law by rewriting the squaring function as a product: $u^2 = u \cdot u$.

$$\partial(u^2) = \partial(u \cdot u) = u \cdot \partial u + u \cdot \partial u = 2 \cdot (u \cdot \partial u) = 2u \cdot \partial u$$

We can now take the differential of any polynomial with degree two.

Example.

Consider the function

$$\begin{aligned} z &\leftarrow x \\ z &= 5x^2 + 3x - 1. \end{aligned}$$

Let's compute the differential.

$$\begin{aligned} \partial z &= \partial(5x^2 + 3x - 1) \\ &= \partial(5x^2 + 3x) \\ &= \partial(5x^2) + \partial(3x) \\ &= 5 \cdot \partial(x^2) + 3 \cdot \partial x \\ &= 5 \cdot 2x \cdot \partial x + 3 \cdot \partial x \\ &= (10x + 3) \cdot \partial x \end{aligned}$$

We factored out the common differential term, ∂x , in order to get our answer in standard form.

And now let's learn to take the differential of the cubing function.

Law.

The differential law for cubing states

$$\partial(u^3) = 3u^2 \cdot \partial u$$

where u : Number is a number.

Why?

We'll write u^3 as a product:

$$u^3 = u \cdot u^2.$$

We can then calculate.

$$\partial(u^3) = \partial(u \cdot u^2) = u^2 \cdot \partial u + u \cdot \partial(u^2) = u^2 \cdot \partial u + u \cdot 2u \cdot \partial u = 3u^2 \cdot \partial u$$

And again, let's have an example!

Example.

Let's take the differential of the function

$$z \leftarrow x$$

$$z = (2x + 1)^3.$$

We calculate.

$$\begin{aligned} \partial z &= \partial((2x + 1)^3) \\ &= 3 \cdot (2x + 1)^2 \cdot \partial(2x + 1) \\ &= 3 \cdot (2x + 1)^2 \cdot \partial(2x) \\ &= 6 \cdot (2x + 1)^2 \cdot \partial x \end{aligned}$$

Notice how we matched

$$u = 2x + 1$$

when we applied the differential law for cubing.

3.4 Division

Let's see how to take the differential of a reciprocal.

Law.

The differential law for reciprocals states

$$\partial\left(\frac{1}{v}\right) = \frac{-1}{v^2} \cdot \partial v$$

where v : Number is a non-zero number, $v \neq 0$.

Why?

We can make sense of the reciprocal using the equation

$$\frac{1}{v} \cdot v = 1.$$

Let's take the differential of both sides of this equation.

$$\partial\left(\frac{1}{v} \cdot v\right) = \partial(1)$$

$$v \cdot \partial\left(\frac{1}{v}\right) + \frac{1}{v} \cdot \partial v = 0$$

$$v \cdot \partial\left(\frac{1}{v}\right) = \frac{-1}{v} \cdot \partial v$$

$$\partial\left(\frac{1}{v}\right) = \frac{-1}{v^2} \cdot \partial v$$

Voila!

We'll use the differential law for the reciprocal to explain the differential law for division.

Law.

The differential law for division states

$$\partial\left(\frac{u}{v}\right) = \frac{1}{v} \cdot \partial u - \frac{u}{v^2} \cdot \partial v$$

where u : Number is a number and v : Number is non-zero, $v \neq 0$.

Why?

We can rewrite any division as a multiplication with a reciprocal.

$$\begin{aligned}\partial\left(\frac{u}{v}\right) &= \partial\left(u \cdot \frac{1}{v}\right) \\ &= \frac{1}{v} \cdot \partial u + u \cdot \partial\left(\frac{1}{v}\right) \\ &= \frac{1}{v} \cdot \partial u + u \cdot \frac{-1}{v^2} \cdot \partial v \\ &= \frac{1}{v} \cdot \partial u - \frac{u}{v^2} \cdot \partial v\end{aligned}$$

Let's have an example that uses the differential law for division.

Example.

Consider the function

$$\begin{aligned}z &\leftarrow x \\ z &= \frac{x^2}{x+5}.\end{aligned}$$

We disallow the input $x = -5$ to avoid dividing by zero. Let's compute the differential.

$$\begin{aligned}\partial z &= \partial\left(\frac{x^2}{x+5}\right) \\ &= \frac{1}{x+5} \cdot \partial(x^2) - \frac{x^2}{(x+5)^2} \cdot \partial(x+5) \\ &= \frac{1}{x+5} \cdot 2x \cdot \partial x - \frac{x^2}{(x+5)^2} \cdot \partial x \\ &= \left(\frac{2x}{x+5} - \frac{x^2}{(x+5)^2}\right) \cdot \partial x\end{aligned}$$

We've written our answer in standard form.

3.5 Square Root and Cube Root

We've made good progress with our differential laws! Let's just look at two more: one for the square root and one for the cube root.

Law.

The differential law for square roots states

$$\partial(\sqrt[2]{u}) = \frac{1}{2 \cdot \sqrt[2]{u}} \cdot \partial u$$

where u : Number is a positive number, $u > 0$.

Why?

The squaring function and the square root function,

$$\begin{array}{ll} b \leftarrow a & a \leftarrow b \\ b = a^2 & \text{and} \quad a = \sqrt[2]{b}, \end{array}$$

are inverse for positive numbers a : Number and b : Number. We'll look for our differential law by rewriting the differential law for squaring.

$$\partial(a^2) = 2a \cdot \partial a$$

$$\partial b = 2a \cdot \partial a$$

$$\frac{1}{2a} \cdot \partial b = \partial a$$

$$\frac{1}{2 \cdot \sqrt[2]{b}} \cdot \partial b = \partial(\sqrt[2]{b})$$

We've found it!

Law.

The differential law for cube roots states

$$\partial(\sqrt[3]{u}) = \frac{1}{3 \cdot (\sqrt[3]{u})^2} \cdot \partial u$$

where u : Number is a non-zero number, $u \neq 0$.

Why?

The cubing function and the cube root function are inverse.

$$\begin{array}{lcl} b \stackrel{\cdot}{\leftarrow} a & & a \stackrel{\cdot}{\leftarrow} b \\ b = a^3 & \text{and} & a = \sqrt[3]{b} \end{array}$$

Starting with the differential law for cubing, we'll try to find the differential law for the cube root.

$$\begin{aligned} \partial(a^3) &= 3a^2 \cdot \partial a \\ \partial b &= 3a^2 \cdot \partial a \\ \frac{1}{3a^2} \cdot \partial b &= \partial a \\ \frac{1}{3 \cdot (\sqrt[3]{b})^2} \cdot \partial b &= \partial(\sqrt[3]{b}) \end{aligned}$$

Got it.

Let's have one more example.

Example.

Consider the function

$$\begin{array}{l} z \stackrel{\cdot}{\leftarrow} x \\ z = \sqrt[2]{x^2 + 1}. \end{array}$$

Let's calculate the differential.

$$\begin{aligned} \partial z &= \partial\left(\sqrt[2]{x^2 + 1}\right) \\ &= \frac{1}{2 \cdot \sqrt[2]{x^2 + 1}} \cdot \partial(x^2 + 1) \\ &= \frac{1}{2 \cdot \sqrt[2]{x^2 + 1}} \cdot \partial(x^2) \\ &= \frac{1}{2 \cdot \sqrt[2]{x^2 + 1}} \cdot 2x \cdot \partial x \\ &= \frac{x}{\sqrt[2]{x^2 + 1}} \cdot \partial x \end{aligned}$$

3.6 Review

Let's review our differential laws by writing them down all in one place. In each of the following laws,

$u : \text{Number}$, $v : \text{Number}$, $c : \text{Number}$

are numbers, c is constant, and any necessary restrictions are present.

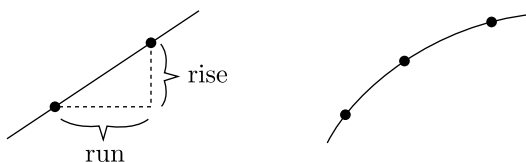
function	differential law
constant	$\partial(c) = 0$
constant-adder	$\partial(u + c) = \partial u$
constant-multiplier	$\partial(c \cdot u) = c \cdot \partial u$
addition	$\partial(u + v) = \partial u + \partial v$
negation	$\partial(-v) = -\partial v$
subtraction	$\partial(u - v) = \partial u - \partial v$
multiplication	$\partial(u \cdot v) = v \cdot \partial u + u \cdot \partial v$
reciprocal	$\partial\left(\frac{1}{v}\right) = \frac{-1}{v^2} \cdot \partial v$
division	$\partial\left(\frac{u}{v}\right) = \frac{1}{v} \cdot \partial u - \frac{u}{v^2} \cdot \partial v$
squaring	$\partial(u^2) = 2u \cdot \partial u$
cubing	$\partial(u^3) = 3u^2 \cdot \partial u$
square root	$\partial(\sqrt[2]{u}) = \frac{1}{2 \cdot \sqrt[2]{u}} \cdot \partial u$
cube root	$\partial(\sqrt[3]{u}) = \frac{1}{3 \cdot (\sqrt[3]{u})^2} \cdot \partial u$

We'll uncover a few more differential laws as we study new functions. At the end of this text, you can find [Summary of Abstract Differential Laws](#) which contains an extended list of differential laws.

Chapter 4

Slope

This chapter is devoted to the studying the slope of a graph. Defining slope can be a little slippery: perhaps you've learned to calculate the slope of a line as the rise over the run.



While this definition works perfectly well for lines, it leaves something to be desired for curved graphs. On a curved graph the slope will change from point to point, and so we'll need a better way to calculate it. Luckily for us, an operator called the *derivative* does just that! By taking the derivative of a function, we learn the slope at every point on the graph. The derivative allows us to become much more discerning graphers of functions.

4.1 Derivatives

In [Chapter 3. Differential Laws](#), we studied the differential, ∂ , as an abstract operator. We'll be using these differential laws to define several operators, the first of which is the derivative. The derivative is defined for functions that have a single input variable x . We'll return to studying functions with two input variables in Part II and Part III of this text.

Definition.

The x -*derivative* is an operator $\overset{x}{\partial}$ that satisfies all [abstract differential laws](#). Further, the x -derivative of the variable x is one.

$$\overset{x}{\partial}x = 1$$

This is known as the *unit law* for the derivative.

We'll pronounce the notation $\overset{x}{\partial}z$ as the “ x -derivative of z .”

Example.

Let's take the derivative of the function

$$\begin{aligned} z &\leftarrow x \\ z &= x^3 - 2x. \end{aligned}$$

We can use all of the differential laws we've studied when calculating the derivative.

$$\begin{aligned} \overset{x}{\partial}z &= \overset{x}{\partial}(x^3 - 2x) \\ &= \overset{x}{\partial}(x^3) - \overset{x}{\partial}(2x) \\ &= 3x^2 \cdot \overset{x}{\partial}x - 2 \cdot \overset{x}{\partial}x \\ &= 3x^2 \cdot 1 - 2 \cdot 1 \\ &= 3x^2 - 2 \end{aligned}$$

We've found the derivative!

The unit law allows us to continue our calculation slightly further with the derivative than we could with the differential. As a consequence, our answer does not include the differential variable ∂x .

Example.

Consider the function

$$z \leftarrow x$$

$$z = \sqrt[2]{x^2 + 1}.$$

Let's compute the x -derivative of z .

$$\begin{aligned} \frac{x}{\partial} z &= \frac{x}{\partial} \left(\sqrt[2]{x^2 + 1} \right) \\ &= \frac{1}{2 \cdot \sqrt[2]{x^2 + 1}} \cdot \frac{x}{\partial} (x^2 + 1) \\ &= \frac{1}{2 \cdot \sqrt[2]{x^2 + 1}} \cdot \frac{x}{\partial} (x^2) \\ &= \frac{1}{2 \cdot \sqrt[2]{x^2 + 1}} \cdot 2x \cdot \frac{x}{\partial} x \\ &= \frac{x}{\sqrt[2]{x^2 + 1}} \end{aligned}$$

We applied the unit law in the last step of this calculation.

The Derivative and the Differential

It is important that we maintain a distinction between the derivative and the differential: there is no “unit law for the differential.”

$$\partial x = 1 \quad \text{No!}$$

$$\frac{x}{\partial} x = 1 \quad \text{Yes!}$$

That being said, it is easy to convert between the derivative and the differential.

Important.

A function $z \leftarrow x$ has x -derivative

$$\frac{x}{\partial} z = s$$

if and only if the function has differential

$$\partial z = s \cdot \partial x.$$

If we know a function's derivative, we can recover its differential.

Example.

If a function $z \leftarrow x$ has x -derivative

$$\frac{\partial z}{\partial x} = x^3 - \frac{4}{x}$$

then the differential of z must be

$$\partial z = \left(x^3 - \frac{4}{x} \right) \cdot \partial x.$$

Converting a differential to a derivative is just as easy.

4.2 Slope

We define the x -slope of a function $z \leftarrow x$ to be equal to the x -derivative.

$$\text{slope } z = \frac{\partial z}{\partial x}$$

We'll write $\frac{\partial}{\partial x}$ when making calculations and we'll write $\text{slope } z$ to emphasize geometry. Looking at a function's slope will help us better understand the shape of its graph. A positive slope gives direct variance, and a negative slope gives indirect variance.

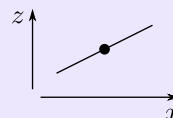
local quantity	local quality
$\text{slope } z > 0$	direct variance
$\text{slope } z < 0$	indirect variance

So wait a second, what are direct and indirect variance?

Important.

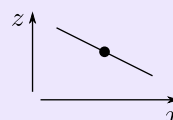
A function $z \leftarrow x$ has *direct variance* near a point if

- z increases as we increase x , and
- z decreases as we decrease x .

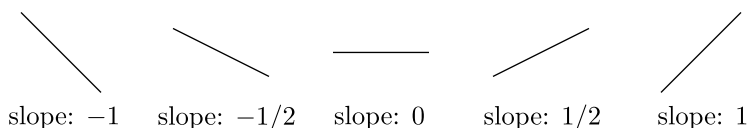


A function $z \leftarrow x$ has *indirect variance* near a point if

- z decreases as we increase x , and
- z increases as we decrease x .



Not only does the slope tell us whether the variance is direct or indirect, it also tells us how steep the graph is.



Shallow graphs have slopes near zero. Steeper graphs have larger positive slopes or smaller negative slopes.

Example.

Let's look at the slope for the squaring function!

$$z \leftarrow x$$
$$z = x^2$$

We find the slope by calculating the derivative.

$$\text{slope } z = 2x$$

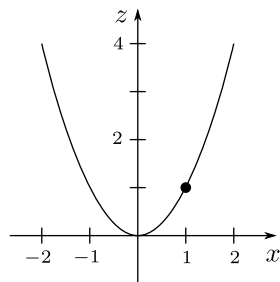
This equation tells us the slope for every x -value. We may *localize* the slope by choosing particular x -values.

At $x = 1$.

Setting x equal to one, we find that the slope is two.

$$\text{slope } z = 2$$

Looking at the graph, we see that near our point the parabola has direct variance and is fairly steep.

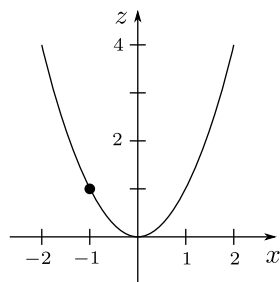


At $x = -1$.

Here we find the slope is negative two.

$$\text{slope } z = -2$$

The parabola is just as steep as at $x = 1$, but the graph has indirect variance.

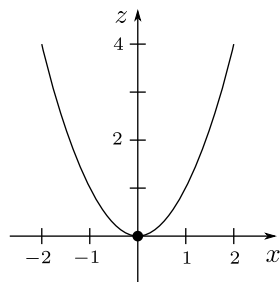


At $x = 0$.

By localizing at $x = 0$ we find that the slope is zero.

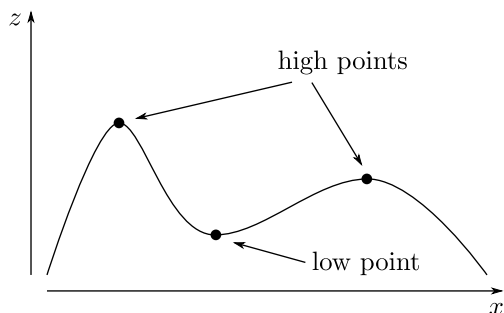
$$\text{slope } z = 0$$

This is an example of a *level point*: a point with zero slope. We'll learn more about level points in the next section.



4.3 Level Points

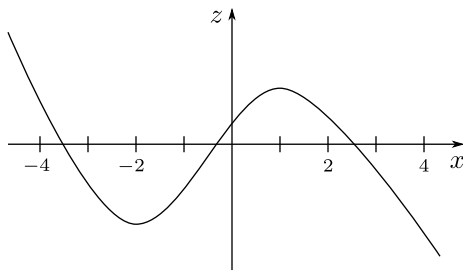
We'll say that a graph has a *level point* wherever the slope is zero. Most examples of level points are either low points or high points. This graph has three level points: one low point and two high points.



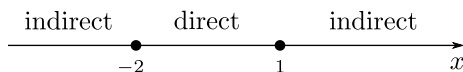
A *high point* does not need to be the highest point on the graph, but rather just higher than other points on the graph nearby. Similarly, a *low point* just needs to be lower than any nearby points on the graph. Transition diagrams allowed us to analyze the [height of a graph](#), and they will be helpful for analyzing slope, too!

Example.

Let's analyze the slope of the following graph.



We'll divide the x -axis up into regions according to whether the graph has direct variance or indirect variance. This graph has two level points: a low point at $x = -2$ and a high point at $x = 1$.

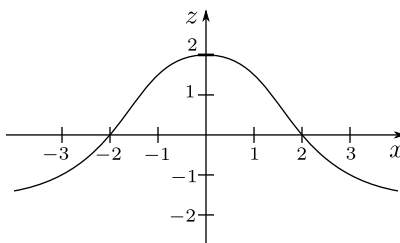


The variance transitions at both the low point and the high point.

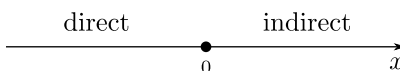
Let's see one more example.

Example.

Now let's take a look at the following graph. We'll draw a transition diagram for slope.



Our graph has a high point at $x = 0$ and no other level points. The graph has direct variance for negative x -values, and the graph has indirect variance for positive x -values.



We see that the slope transitions at the level point.

At the start of this section, we said, “most examples of level points are either low points or high points.” Low points and high points are always level points, but not every level point is either a low point or a high point! In this chapter and the next we'll see many more examples of level points. We'll learn to use the derivative to both locate and classify level points.

4.4 Five Graphs: Slope

Let's revisit our [five graphs](#) but this time wearing glasses that see slope!

Squaring

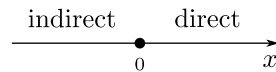
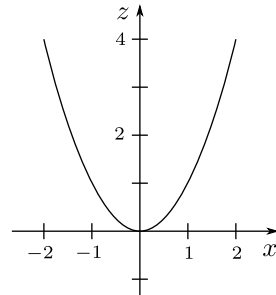
We've already spent a little time looking at the squaring function,

$$\begin{aligned} z &\leftarrow x \\ z &= x^2, \end{aligned}$$

and its slope

$$\text{slope } z = 2x.$$

Let's draw a transition diagram for the slope. The parabola has a level point at $x = 0$. In particular, this is a low point for the graph. The low point appears in the transition diagram: the graph transitions from indirect variance to direct variance at $x = 0$.



Cubing

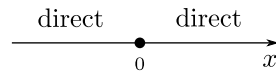
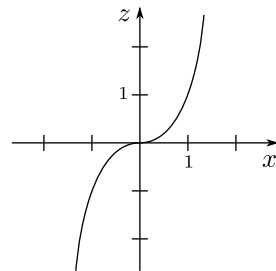
Recall the graph of the cubing function,

$$\begin{aligned} z &\leftarrow x \\ z &= x^3. \end{aligned}$$

The cubing function has slope

$$\text{slope } z = 3x^2.$$

We find a level point at $x = 0$, and the slope is positive at all other points. The cubing function gives us an interesting example of a level point! We find larger heights when moving to right, and we find smaller heights by moving to the left. So this level point at is neither a low point nor a high point for the graph.



Reciprocal

The reciprocal function,

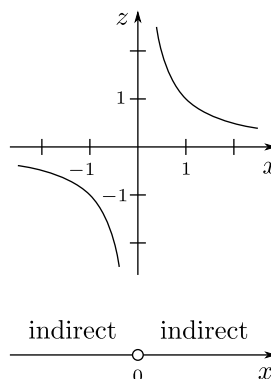
$$z \leftarrow x$$

$$z = \frac{1}{x},$$

is defined at all non-zero inputs, $x \neq 0$. Here's the slope.

$$\text{slope } z = \frac{-1}{x^2}$$

The slope is negative wherever it is defined. In other words, the graph has indirect variance.



Square Root

Next let's look at the square root function,

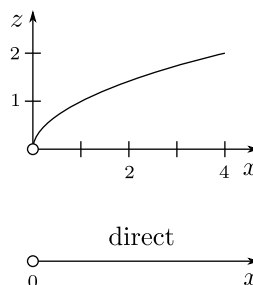
$$z \leftarrow x$$

$$z = \sqrt[2]{x}.$$

The square root has slope

$$\text{slope } z = \frac{1}{2 \cdot \sqrt[2]{x}}.$$

The graph has direct variance at all allowed inputs. We must disallow $x = 0$, however, for the slope to be defined.



Cube Root

Last let's check in with the cube root function,

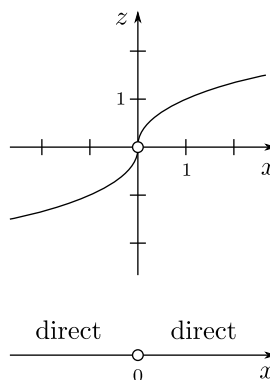
$$z \leftarrow x$$

$$z = \sqrt[3]{x}.$$

We've calculated the slope to be

$$\text{slope } z = \frac{1}{3 \cdot (\sqrt[3]{x})^2}.$$

Once again, we must disallow the input $x = 0$ for the slope to be defined. The graph has direct variance at all allowed points.



4.5 Analyzing Slope

Let's get a little more practice analyzing slope using transition diagrams.

Example.

Consider the function

$$z \leftarrow x$$

$$z = x^3 - 3x.$$

By taking the derivative

$$\overset{x}{\partial} z = \overset{x}{\partial} (x^3 - 3x) = \overset{x}{\partial} (x^3) - \overset{x}{\partial} (3x) = 3x^2 \cdot \overset{x}{\partial} x - 3 \cdot \overset{x}{\partial} x = 3x^2 - 3,$$

we find our function's slope.

$$\overset{x}{\text{slope}} z = 3x^2 - 3$$

Let's draw a transition diagram for slope. We find level points by setting the slope equal to zero.

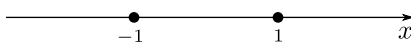
$$\overset{x}{\text{slope}} z = 0$$

$$3x^2 - 3 = 0$$

$$x^2 = 1$$

$$x = -1 \quad \text{or} \quad x = 1$$

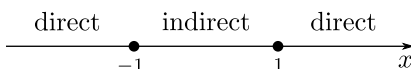
We'll mark both level points on an x -axis to begin drawing a transition diagram.



Between negative one and one, the slope is negative. At $x = 0$ for example, the slope is negative three.

$$\overset{x}{\text{slope}} z = 3 \cdot 0^2 - 3 = -3$$

As you can check, the slope is positive in each of the other regions.



We've found our transition diagram for slope!

Let's have one more example.

Example.

Now consider the function

$$z \leftarrow x$$

$$z = \frac{1}{x^2 - 1}.$$

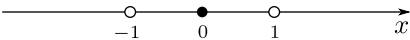
Let's take the derivative of z .

$$\begin{aligned} \overset{x}{\partial} z &= \overset{x}{\partial} \left(\frac{1}{x^2 - 1} \right) = \frac{-1}{(x^2 - 1)^2} \cdot \overset{x}{\partial} (x^2 - 1) = \frac{-1}{(x^2 - 1)^2} \cdot \overset{x}{\partial} (x^2) \\ &= \frac{-1}{(x^2 - 1)^2} \cdot 2x \cdot \overset{x}{\partial} x = \frac{-2x}{(x^2 - 1)^2} \end{aligned}$$

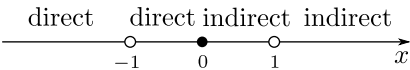
We've found the slope.

$$\overset{x}{\text{slope}} z = \frac{-2x}{(x^2 - 1)^2}$$

The slope is zero at $x = 0$. The slope is undefined at $x = -1$ and at $x = 1$.



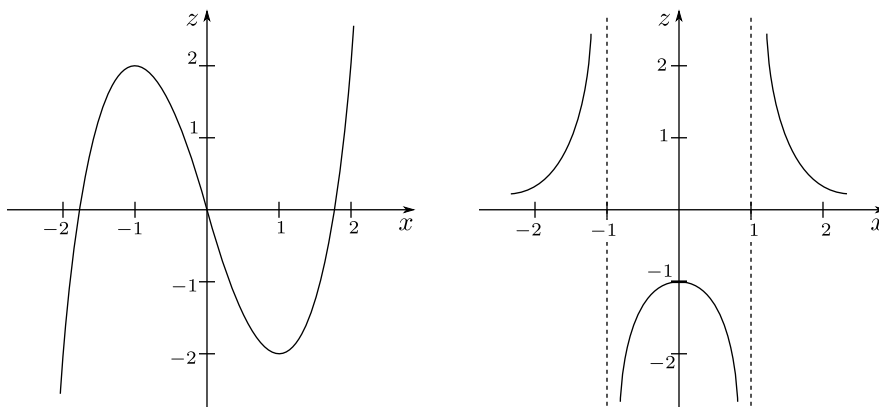
We can figure out where the slope is positive or negative by testing an x -value from each region.



x	$\overset{x}{\text{slope}} z$
-2	4/9
-1	
-1/2	16/9
0	0
1/2	-16/9
1	
2	-4/9

The slope transitions at the level point, only.

For those that are curious, here are the graphs of the functions in our examples.



The first has a high point at $x = -1$ and a low point at $x = 1$. The second graph has a high point at $x = 0$ and no other level points.

4.6 Linear Functions

Let's end our chapter on slope by taking a look at linear functions. That is, we'll look at functions that have lines as graphs.

Definition.

A function $z \stackrel{f}{\leftarrow} x$ is *linear* if we can write it in the form

$$z - z_0 = s \cdot (x - x_0)$$

where

$$x_0 : \text{Number}, \quad z_0 : \text{Number}, \quad s : \text{Number}$$

are constants.

We may always easily solve for the output variable, z , but we'll often prefer to write a linear function in the above form.

Important.

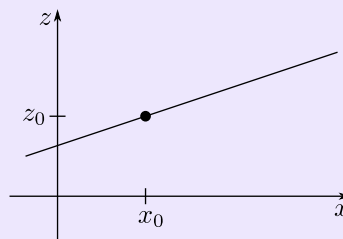
The graph of a linear function

$$z - z_0 = s \cdot (x - x_0)$$

is the line that

- has height $z = z_0$ at $x = x_0$, and
- has slope s .

The line has the same slope s at every point on the graph.



Why?

Let's solve for height z .

$$z = z_0 + s \cdot (x - x_0)$$

By evaluating our equation at $x = x_0$, we find the height.

$$z = z_0 + s \cdot (x_0 - x_0) = z_0 + s \cdot 0 = z_0$$

Now let's verify the slope by calculating the derivative.

$$\overset{x}{\partial} z = \overset{x}{\partial} (z_0 + s \cdot (x - x_0)) = \overset{x}{\partial} (s \cdot (x - x_0)) = s \cdot \overset{x}{\partial} (x - x_0) = s \cdot \overset{x}{\partial} x = s$$

Sure enough, we've found the slope to be the constant s .

Let's have a quick example of a linear function and its graph.

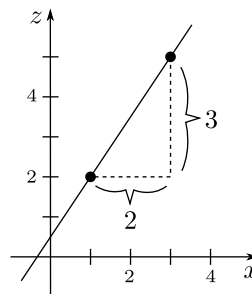
Example.

Consider the linear function

$$z - 2 = \frac{3}{2} \cdot (x - 1).$$

According to our linear form,

- the graph has height $z = 2$ at $x = 1$, and
- the graph has constant slope $s = 3/2$.



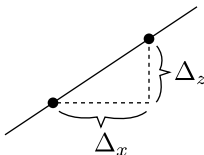
You should double check both of these claims! For the first, solve for the height z and then localize it at $x = 1$. For the second, calculate the derivative, $\overset{x}{\partial} z$.

Slope as Rise Over Run

Because a line has a constant slope, it is easy to calculate the slope as

$$\text{slope } z = \frac{\Delta z}{\Delta x}$$

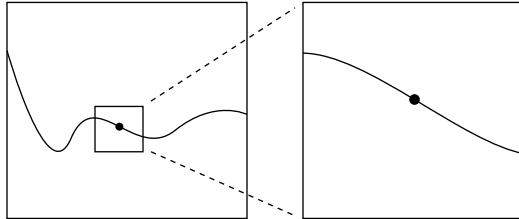
where Δ_z and Δ_x are the *change in* z and the *change in* x for any two points on the line.



When we work with a curved graph, the slope changes from point to point. We should only expect this “rise over run” formula to be an approximation of the slope.

$$\text{slope } z^x \approx \frac{\Delta z}{\Delta x}$$

Even though this formula is just an approximation, it is still useful! If we zoom in towards a point on a curved graph, the graph tends to look more like a straight line.



In fact, the more we zoom in on our graph, the more confidence we should have that the “rise over run” formula gives a value close to the actual slope.

It takes hard work to make this idea of zooming fully rigorous: you can study *limits* in Real Analysis if you are interested in seeing the details. But let’s at least take a moment to think about what makes slope so difficult. As we zoom in towards a point, our concept of slope becomes dangerously close to containing a division by zero!

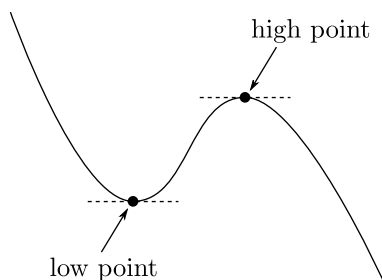
$$\text{slope } z^x \approx \frac{\Delta z}{\Delta x} \approx \frac{0}{0}$$

We saw that defining the fraction $1/0$ immediately [leads to a contradiction](#). The fraction $0/0$ has a different problem: it could arguably be defined to be any number! We won’t ever officially define $0/0$ but it lurks in the shadows whenever we use the derivative.

Chapter 5

Bend

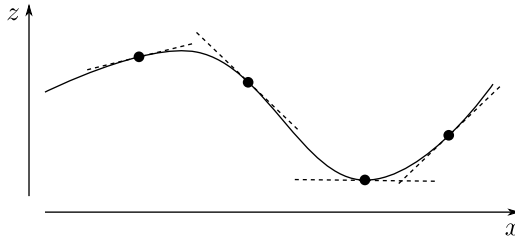
Already we have learned about two important quantities for analyzing the shape of a graph: height and slope. The bend is a third quantity that helps us go further. A line has no bend. Instead we find bend when looking at curved graphs. Nonetheless, we'll see that certain lines play an important role for understanding a graph's bend.



By learning to calculate the bend, we'll be able to classify level points. The bend is just the right tool to tell whether a level point is a low point or a high point.

5.1 Tangent Lines

The tangent line is a terrific tool for better understanding the graph of a function $z \leftarrow^f x$. We may center a tangent line at any point on the graph. The tangent line matches both the height and the slope of the graph at the center point. Here's a graph with four tangents drawn in as dotted lines.



Near its center, the tangent is the line that most closely resembles the graph.

Definition.

Suppose $z \leftarrow^f x$ is a function. The *tangent to f* is the linear function $z_T \leftarrow^x x_T$ given by

$$z_T - z = \text{slope } z \cdot (x_T - x)$$

where

$$x : \text{Number}, \quad z : \text{Number}, \quad \text{slope } z : \text{Number}$$

are the constants found at the center.

Let's make sense of this definition by working through an example.

Example.

Consider the cubing function,

$$z \leftarrow^f x$$

$$z = x^3,$$

and its slope

$$\text{slope } z = 3x^2.$$

Let's find three tangent lines by centering at $x = 1$, at $x = -1$, and at $x = 0$.

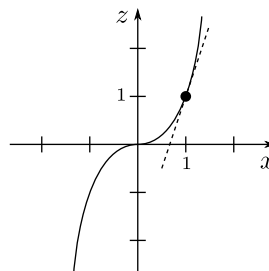
Centering at $x = 1$.

By centering at $x = 1$, we find constants

$$x = 1, \quad z = 1, \quad \text{slope } z = 3.$$

The equation for the tangent line reads

$$z_T - 1 = 3 \cdot (x_T - 1).$$



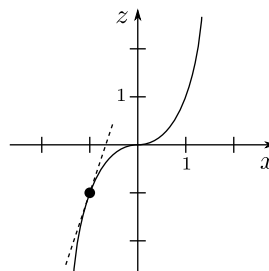
Centering at $x = -1$.

Here the constants for the tangent line are

$$x = -1, \quad z = -1, \quad \text{slope } z = 3.$$

The equation for the tangent line reads

$$z_T + 1 = 3 \cdot (x_T + 1).$$



Centering at $x = 0$.

At $x = 0$ we find

$$x = 0, \quad z = 0, \quad \text{slope } z = 0.$$

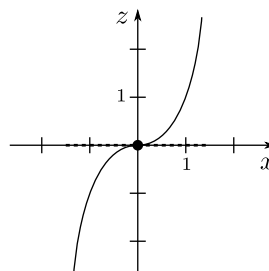
The tangent line can be written as

$$z_T - 0 = 0 \cdot (x_T - 0)$$

or more simply

$$z_T = 0.$$

This is an equation for the x -axis: all points with height zero.



The tangent line gives us a new geometric way to understand level points.

Important.

A level point is any point on a graph that has a horizontal tangent line.

We see, once again, that the cubing function has a level point at $x = 0$. If we zoomed in around the origin, the cubing function would look more and more like a horizontal line.

5.2 The Second Derivative

When we take the derivative of a function $z \leftarrow x$, the result, $\overset{x}{\partial} z \leftarrow x$, is also a function. We find the *second derivative of z* by taking another derivative.

$$\overset{x}{\partial} \overset{x}{\partial} z \leftarrow x$$

And we could keep going, taking the derivative again and again!

Example.

Let's compute the first two derivatives of the function

$$\begin{aligned} z &\leftarrow x \\ z &= x^3 + 5x^2. \end{aligned}$$

Here's the first x -derivative.

$$\begin{aligned} \overset{x}{\partial} z &= \overset{x}{\partial} (x^3 + 5x^2) \\ &= \overset{x}{\partial} (x^3) + \overset{x}{\partial} (5x^2) \\ &= 3x^2 + 5 \cdot \overset{x}{\partial} (x^2) \\ &= 3x^2 + 5 \cdot 2x \cdot \overset{x}{\partial} x \\ &= 3x^2 + 10x \end{aligned}$$

Now let's take the x -derivative, again.

$$\begin{aligned} \overset{x}{\partial} \overset{x}{\partial} z &= \overset{x}{\partial} (3x^2 + 10x) \\ &= \overset{x}{\partial} (3x^2) + \overset{x}{\partial} (10x) \\ &= 3 \cdot \overset{x}{\partial} (x^2) + 10 \cdot \overset{x}{\partial} x \\ &= 3 \cdot 2x \cdot \overset{x}{\partial} x + 10 \\ &= 6x + 10 \end{aligned}$$

We've found our function's second derivative!

Let's take the second derivative of the square root function.

Example.

The square root function,

$$z \leftarrow x$$

$$z = \sqrt[2]{x},$$

is defined for positive inputs, $x \geq 0$. We've seen that the square root has first derivative

$$\frac{\partial z}{\partial x} = \frac{1}{2 \cdot \sqrt[2]{x}}.$$

Let's calculate the second derivative.

$$\begin{aligned} \frac{\partial^2 z}{\partial x^2} &= \frac{\partial}{\partial x} \left(\frac{1}{2 \cdot \sqrt[2]{x}} \right) \\ &= \frac{1}{2} \cdot \frac{\partial}{\partial x} \left(\frac{1}{\sqrt[2]{x}} \right) \\ &= \frac{1}{2} \cdot \frac{-1}{(\sqrt[2]{x})^2} \cdot \frac{\partial}{\partial x} (\sqrt[2]{x}) \\ &= \frac{1}{2} \cdot \frac{-1}{(\sqrt[2]{x})^2} \cdot \frac{1}{2 \cdot \sqrt[2]{x}} \cdot \frac{\partial x}{\partial x} \\ &= \frac{-1}{4x \cdot \sqrt[2]{x}} \end{aligned}$$

Like the first derivative, the second derivative is defined only for $x > 0$.

For an easy exercise, you can calculate the second derivative of the cubing function.

$$\frac{\partial^2}{\partial x^2} (x^3) = 6x$$

For a much more challenging exercise, calculate the second derivative of the cube root function.

$$\frac{\partial^2}{\partial x^2} (\sqrt[3]{x}) = \frac{-2}{9 \cdot (\sqrt[3]{x})^5}$$

We'll find an easier way to calculate this cube root exercise when we study [base functions](#).

5.3 Bend

The x -*bend* of a function $z \leftarrow x$ is the second x -derivative of z , halved.

$$\text{bend } z = \frac{\frac{x}{\partial} \frac{x}{\partial} z}{2}$$

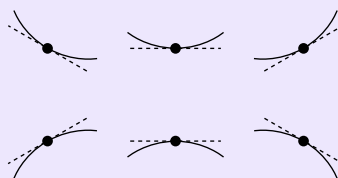
A positive bend gives a concave up graph, and a negative bend gives a concave down graph.

local quantity	local quality
$\text{bend } z > 0$	concave up
$\text{bend } z < 0$	concave down

Important.

A *concave up* graph peels away from the tangent line towards larger z -values.

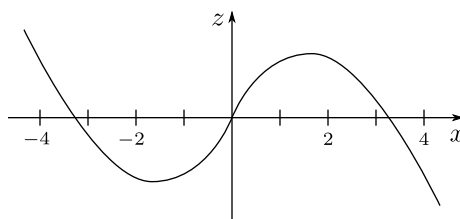
A *concave down* graph peels away from the tangent line towards smaller z -values.



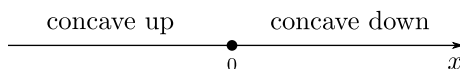
We have transition diagrams for bend just like we had for height and for slope.

Example.

Let's analyze the bend for the following graph.



This graph has positive bend for negative x -values and negative bend for positive x -values.



The bend transitions at $x = 0$. We'll call such a transition for bend an *inflection point*.

A line has constant slope and zero bend.

Example.

Let's calculate the first two derivatives for the linear function

$$z - 2 = 3 \cdot (x - 5).$$

We find

$$\overset{x}{\partial} z = \overset{x}{\partial} (3 \cdot (x - 5) + 2) = \overset{x}{\partial} (3 \cdot (x - 5)) = 3 \cdot \overset{x}{\partial} (x - 5) = 3 \cdot \overset{x}{\partial} x = 3,$$

$$\overset{x}{\partial} \overset{x}{\partial} z = \overset{x}{\partial} (3) = 0.$$

Our line has constant slope three and zero bend.

$$\overset{x}{\text{slope}} z = 3 \quad \overset{x}{\text{bend}} z = 0$$

You should be able to visually discern whether a graph is concave up or concave down. But you shouldn't expect to be able to estimate the size of the bend from a graph. It might seem strange, then, that we don't just define the bend as the second derivative.

Question.

Why do we halve the second derivative to get the bend?

Our favorite parabola,

$$z \leftarrow x$$

$$z = x^2,$$

has second x -derivative equal to two.

$$\overset{x}{\partial} \overset{x}{\partial} z = \overset{x}{\partial} \overset{x}{\partial} (x^2) = \overset{x}{\partial} (2x) = 2$$

Our theory will work best if this parabola has a bend of one.

$$\overset{x}{\text{bend}} z = \frac{\overset{x}{\partial} \overset{x}{\partial} z}{2} = 1$$

In short, we've defined the bend so that our favorite parabola is the standard against which we compare all other bends.

5.4 Five Graphs: Bend

We've analyzed our five graphs [for height](#) and [for slope](#) in previous chapters. Let's revisit these graphs one more time, this time looking at the bend!

Squaring

As we just discussed in the preceding section, the squaring function,

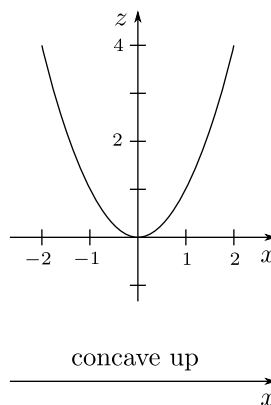
$$z \leftarrow x$$

$$z = x^2,$$

has a constant bend of one.

$$\text{bend } z = 1$$

The graph is concave up everywhere.



Cubing

Now let's analyze the cubing function,

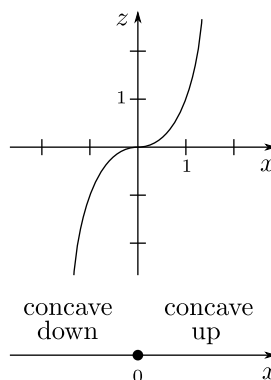
$$z \leftarrow x$$

$$z = x^3.$$

The cubing function has bend

$$\text{bend } z = 3x.$$

The bend is positive for positive x -values, is negative for negative x -values, and is zero at $x = 0$. The cubing function illustrates something that is rather rare! Not only is $x = 0$ a level point, but it is also an inflection point.



Reciprocal

Let's take a look at the reciprocal function,

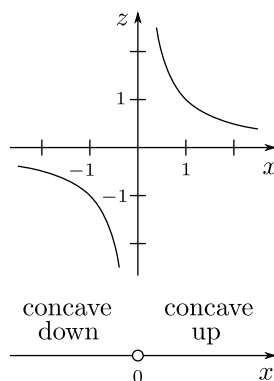
$$z \leftarrow x$$

$$z = \frac{1}{x}.$$

We can calculate the bend.

$$\text{bend } z = \frac{1}{x^3}$$

Although the bend transitions $x = 0$, we won't call this an inflection point! We'll want the bend to be defined at inflection points.



Square Root

Let's consider the square root function,

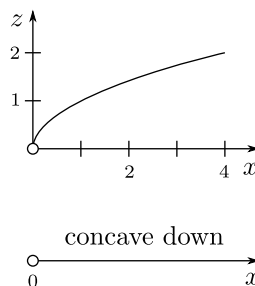
$$z \leftarrow x$$

$$z = \sqrt[2]{x}.$$

This function has bend

$$\text{bend } z = \frac{-1}{8x \cdot \sqrt[2]{x}}.$$

The graph is concave down at all allowed inputs, $x > 0$.



Cube Root

Last but not least, let's look at the graph of the cube root,

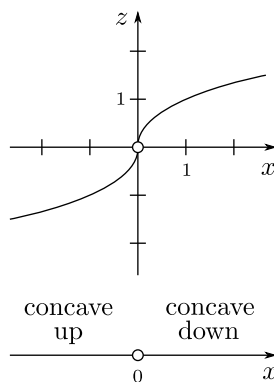
$$z \leftarrow x$$

$$z = \sqrt[3]{x}.$$

We can calculate the bend to be

$$\text{bend } z = \frac{-1}{9 \cdot (\sqrt[3]{x})^5}.$$

The graph is concave down for positive x -values and concave up for negative x -values.



5.5 Low Points and High Points

Let's return to the problem of locating a graph's low points and high points. First, we should locate the graph's [level points](#).

$$\text{slope } z = 0$$

Now that we can calculate the bend, we can use it to help decide whether each level point is a low point or a high point.

Important.

Suppose we've found a level point for a function $z \leftarrow x$.

If the graph has positive bend at the level point,

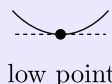
$$\text{bend } z > 0,$$

then that is a low point for the graph.

If the graph has negative bend at the level point,

$$\text{bend } z < 0,$$

then that is a high point for the graph.



low point



high point

If a level point has zero bend, then the level point might be a low point, a high point, or an inflection point! In this case, we might draw a transition diagram for slope to better understand the shape of the graph.

Example.

Let's analyze the level points for the function

$$z \leftarrow x$$

$$z = x^3 - 3x.$$

We looked at this function once before in the section [4.5 Analyzing Slope](#). By calculating the slope

$$\text{slope } z = 3x^2 - 3,$$

we found that this function has two level points: $x = -1$ and $x = 1$. Let's use the bend

$$\text{bend } z = 3x$$

to classify the level points.

At $x = -1$.

Localizing the bend, we find

$$\text{bend } z = -3.$$

A negative bend means the graph is concave down, and so $x = -1$ is a high point.

At $x = 1$.

Here the bend is positive.

$$\text{bend } z = 3$$

The graph is concave up, and so $x = 1$ must be a low point.

Let's conclude by making a table of the height, slope, and bend at each of the level points.

x	z	$\text{slope } z$	$\text{bend } z$
-1	2	0	-3
1	-2	0	3

The graph has height $z = 2$ at the high point, and height $z = -2$ at the low point.

5.6 Taylor Polynomials

We began this chapter by finding [tangent lines](#) for a graph by matching the graph's height and slope. Now that we know how to calculate bend, we can use height, slope, and bend together to find parabolas that hug our graph.

Definition.

Consider the graph of a function $z \leftarrow x$. The *Taylor polynomial* is the function $z_T \leftarrow x_T$ described by the equation

$$z_T - z = \overset{x}{\text{slope } z} \cdot (x_T - x) + \overset{x}{\text{bend } z} \cdot (x_T - x)^2$$

where

$$x : \text{Number}, \quad z : \text{Number}, \quad \overset{x}{\text{slope } z} : \text{Number}, \quad \overset{x}{\text{bend } z} : \text{Number}$$

are the constants found at the center.

Let's find a Taylor polynomial!

Example.

Consider the square root function

$$z \leftarrow x$$

$$z = \sqrt[2]{x}.$$

Let's find the Taylor polynomial centered at $x = 1$. We've calculated the slope and bend.

$$\overset{x}{\text{slope } z} = \frac{1}{2 \cdot \sqrt[2]{x}}$$

$$\overset{x}{\text{bend } z} = \frac{-1}{8x \cdot \sqrt[2]{x}}$$

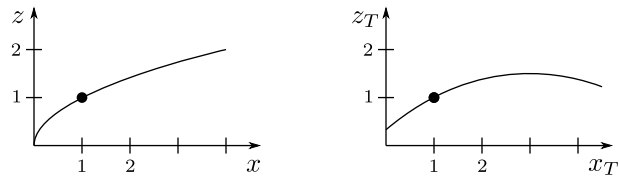
So at the center $x = 1$ our function has height, slope, and bend:

$$z = 1, \quad \overset{x}{\text{slope } z} = \frac{1}{2}, \quad \overset{x}{\text{bend } z} = \frac{-1}{8}.$$

We can use these constants to write down the Taylor polynomial.

$$z_T - 1 = \frac{1}{2} \cdot (x_T - 1) - \frac{1}{8} \cdot (x_T - 1)^2$$

Here are the graphs of the functions from the example. The square root function is on the left, and its Taylor polynomial is on the right.



The graph of the Taylor polynomial is the *best fit parabola*. The Taylor polynomial, $z_T \leftarrow x_T$, is defined so that its height, slope, and bend match the function $z \leftarrow x$ at the center.

Chapter 6

Powers

It's time for us to turn our attention to powers so that we may confront some impressive differential laws. There are two kinds of functions that can be built as powers: base functions and exponential functions. A base function allows its base b to vary. These include familiar functions like the squaring function

$$p = b^2$$

and the cubing function

$$p = b^3.$$

We'll see that reciprocals, square roots, and cube roots can also be written as base functions. And we'll learn a single differential law for base functions that unifies many of the differential laws we've already studied!

An exponential function, in contrast, allows its exponent e to vary. People use different exponential functions for different purposes. Engineers often like base ten exponentials

$$p = 10^e$$

since these can be handy when estimating the size of decimal numbers. Computer scientists using binary numbers prefer to work with base two exponentials

$$p = 2^e.$$

It turns out that another base is best for Calculus.

$$p = \text{nat}^e$$

Mathematicians have discovered that the *natural base*, $\text{nat} \approx 2.718$, gives the simplest differential law. Moreover, the natural base turns up when we take the differential of any exponential, natural or otherwise!

6.1 Powers

A *power*

$$p = b^e$$

is a number p : Number that is built from a base b : Number and an exponent e : Number. We know how powers are defined for whole number exponents: repeated multiplication.

$$b^1 = b$$

$$b^2 = b \cdot b$$

$$b^3 = b \cdot b \cdot b$$

$\vdots$

We wish to define powers for other exponents including negative numbers and fractions. The naive power laws

$$b^{e_1+e_2} = b^{e_1} \cdot b^{e_2}$$

$$b^{e_1 \cdot e_2} = (b^{e_1})^{e_2}$$

are true when we use whole number exponents, and we'd like them to be true for all exponents. Let's see what definitions we'll need for these laws to hold more generally.

Definition.

An exponent of zero gives

$$b^0 = 1$$

where b : Number is a non-zero number, $b \neq 0$.

Why?

By playing with our power laws, we can solve for b^0 .

$$b^0 \cdot b^1 = b^1$$

$$b^0 \cdot b = b$$

$$b^0 = 1$$

We used our assumption, $b \neq 0$, to avoid dividing by zero.

Definition.

An exponent of negative one gives

$$b^{-1} = \frac{1}{b}$$

where b : Number is non-zero, $b \neq 0$.

Why?

Let's use our laws to solve for b^{-1} .

$$b^{-1} \cdot b^1 = b^0$$

$$b^{-1} \cdot b = 1$$

$$b^{-1} = \frac{1}{b}$$

Definition.

A fractional exponent gives

$$b^{1/2} = \sqrt[2]{b}$$

$$b^{1/3} = \sqrt[3]{b}$$

where b : Number is a positive number, $b > 0$.

Why?

We'll take a look at the exponent one-half.

$$(b^{1/2})^2 = b$$

$$\sqrt[2]{(b^{1/2})^2} = \sqrt[2]{b}$$

$$b^{1/2} = \sqrt[2]{b}$$

A similar argument works for the exponent one-third.

With these definitions in place, we can revise our naive power laws!

Laws.

The power laws state

$$b^{e_1+e_2} = b^{e_1} \cdot b^{e_2}$$

$$b^{e_1 \cdot e_2} = (b^{e_1})^{e_2}$$

where

$$b : \text{Number} \quad e_1 : \text{Number} \quad e_2 : \text{Number}$$

are numbers and the base is positive, $b > 0$.

The naive power laws are true for all exponents once we've included an assumption that the base be positive. We can run into trouble if we don't include this assumption. The following calculations appear to contradict each other.

$$((-1)^2)^{1/2} = (-1)^1 = -1$$

$$((-1)^2)^{1/2} = (1)^{1/2} = 1$$

Which calculation contains the mistake?

6.2 Base Functions

Now that we are familiar with negative exponents and fractional exponents, let's take another look at some familiar functions.

function	power	exponent
$p = b^2$	$p = b^2$	2
$p = b^3$	$p = b^3$	3
$p = \frac{1}{b}$	$p = b^{-1}$	-1
$p = \sqrt[2]{b}$	$p = b^{1/2}$	1/2
$p = \sqrt[3]{b}$	$p = b^{1/3}$	1/3

Each of these functions is an example of a base function!

Definition.

A *base function* is a power

$$p \leftarrow b$$

$$p = b^e$$

where the base, b , is allowed to vary and the exponent, e , is constant.

Let's take a look at the differential law for base functions.

Law.

The differential law for base functions states

$$\partial(v^c) = c \cdot v^{c-1} \cdot \partial v$$

where c : Number is a constant number and the base v : Number is a positive number, $v > 0$.

We won't check the differential base law, but we have already seen significant evidence for it! The differential laws for squaring, cubing, reciprocals, square roots, and cube roots are all special cases of the differential law for base functions.

differential law	differential base law
$\partial(v^2) = 2v \cdot \partial v$	$\partial(v^2) = 2v^1 \cdot \partial v$
$\partial(v^3) = 3v^2 \cdot \partial v$	$\partial(v^3) = 3v^2 \cdot \partial v$
$\partial\left(\frac{1}{v}\right) = \frac{-1}{v^2} \cdot \partial v$	$\partial(v^{-1}) = -v^{-2} \cdot \partial v$
$\partial(\sqrt[2]{v}) = \frac{1}{2 \cdot \sqrt[2]{v}} \cdot \partial v$	$\partial(v^{1/2}) = \frac{1}{2} \cdot v^{-1/2} \cdot \partial v$
$\partial(\sqrt[3]{v}) = \frac{1}{3 \cdot (\sqrt[3]{v})^2} \cdot \partial v$	$\partial(v^{1/3}) = \frac{1}{3} \cdot v^{-2/3} \cdot \partial v$

The differential law for base functions packs quite a punch!

Example.

Consider the function

$$z \leftarrow x$$

$$z = \frac{x^4}{2} + \sqrt[2]{3x}$$

defined for positive inputs, $x > 0$. Let's take the derivative.

$$\begin{aligned}
 \overset{x}{\partial} z &= \overset{x}{\partial} \left(\frac{x^4}{2} + \sqrt[2]{3x} \right) \\
 &= \overset{x}{\partial} \left(\frac{x^4}{2} \right) + \overset{x}{\partial} (\sqrt[2]{3x}) \\
 &= \frac{1}{2} \cdot \overset{x}{\partial} (x^4) + \overset{x}{\partial} ((3x)^{1/2}) \\
 &= \frac{1}{2} \cdot 4x^3 \cdot \overset{x}{\partial} x + \frac{1}{2} \cdot (3x)^{-1/2} \cdot \overset{x}{\partial} (3x) \\
 &= 2x^3 + \frac{1}{2} \cdot \frac{1}{\sqrt[2]{3x}} \cdot 3 \cdot \overset{x}{\partial} x \\
 &= 2x^3 + \frac{3}{2 \cdot \sqrt[2]{3x}}
 \end{aligned}$$

Let's use the differential law for base functions to calculate two derivatives of the cube root.

Example.

We can write the cube root as a base function.

$$z \leftarrow x$$

$$z = x^{1/3}$$

Here's the first derivative

$$\frac{\partial}{\partial z} z = \frac{\partial}{\partial} (x^{1/3}) = \frac{1}{3} \cdot x^{-2/3}$$

and here's the second.

$$\begin{aligned} \frac{\partial}{\partial} \frac{\partial}{\partial} z &= \frac{\partial}{\partial} \left(\frac{1}{3} \cdot x^{-2/3} \right) \\ &= \frac{1}{3} \cdot \frac{\partial}{\partial} (x^{-2/3}) \\ &= \frac{1}{3} \cdot \frac{-2}{3} \cdot x^{-5/3} \cdot \frac{\partial}{\partial} x \\ &= \frac{-2}{9} \cdot x^{-5/3} \end{aligned}$$

Using the differential law for base functions made this calculation much easier.

As wonderful as the differential base law is, we still have a long way to go to understand powers! We would like to be able to take the differential of any power

$$\partial(v^u)$$

where both the exponent u : Number and the base v : Number are allowed to vary. In order to get there, we'll need to study exponential functions.

6.3 The Natural Exponential

An exponential function allows its exponent to vary.

Definition.

An *exponential function* is a power

$$p \leftarrow e$$

$$p = b^e$$

where the exponent, e , is allowed to vary and the base, b , is constant.

The *natural exponential function* is the exponential function

$$p \stackrel{\text{exp}}{\leftarrow} e$$

$$p = \text{nat}^e$$

where the constant

$$\text{nat} : \text{Number}$$

$$\text{nat} = 1 + \frac{1}{1} + \frac{1}{1 \cdot 2} + \frac{1}{1 \cdot 2 \cdot 3} + \frac{1}{1 \cdot 2 \cdot 3 \cdot 4} + \dots$$

$$\text{nat} = 2.718\dots$$

is an irrational number is known as the *natural base*. Don't worry about the natural base's strange definition: it was carefully chosen to give us simple differential laws. The constant `nat` is just some number between two and three that we call "natural."

Theorem.

The natural exponential function satisfies the laws

$$\text{exp } 0 = 1$$

$$\text{exp } 1 = \text{nat}$$

$$\text{exp}(e_1 + e_2) = \text{exp } e_1 \cdot \text{exp } e_2$$

where $e_1 : \text{Number}$ and $e_2 : \text{Number}$ are numbers.

Why?

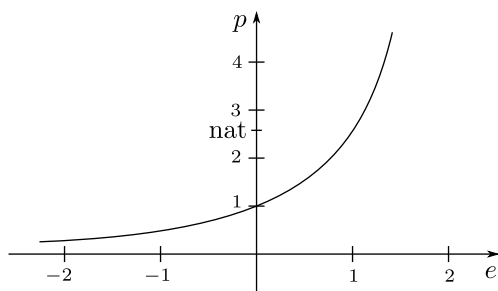
These follow quickly from our definition of the natural exponential.

$$\exp 0 = \text{nat}^0 = 1$$

$$\exp 1 = \text{nat}^1 = \text{nat}$$

$$\exp(e_1 + e_2) = \text{nat}^{e_1 + e_2} = \text{nat}^{e_1} \cdot \text{nat}^{e_2} = \exp e_1 \cdot \exp e_2$$

Graphing the natural exponential function by hand takes some real effort. We can quickly plot a few points if we use a calculator or computer.



e	p
-2	≈ 0.135
-1	≈ 0.367
0	1
1	≈ 2.718
2	≈ 7.398

To gain more familiarity with exponential functions, you can graph the base two exponential, 2^e . This function is much easier to graph, and it has a very similar shape. In fact, the graphs of exponential functions are all quite similar. We'll soon see a change-of-base formula which reveals an important geometric fact: exponential functions only differ by [horizontal transformers](#).

6.4 The Natural Logarithm

In the last section we studied the natural exponential function.

$$p \stackrel{\text{exp}}{\leftarrow} e$$
$$p = \exp e$$

The natural exponential takes exponents, e , as input and produces powers, p , as output. The *natural logarithm* inverts the relationship between p and e .

$$e \stackrel{\text{log}}{\leftarrow} p$$
$$e = \log p$$

Since the natural logarithm is an [inverse](#), it cancels with the natural exponential.

$$\log \exp e = e \quad \text{and} \quad \exp \log p = p$$

We need one restriction: the natural exponential function produces positive output, $p > 0$, only, so the natural logarithm will only accept positive input.

Theorem.

The natural logarithm satisfies the laws

$$\log 1 = 0$$

$$\log \text{nat} = 1$$

$$\log(p_1 \cdot p_2) = \log p_1 + \log p_2$$

where $p_1 : \text{Number}$ and $p_2 : \text{Number}$ are positive numbers.

Why?

The first two laws are easy enough cancellations.

$$\log 1 = \log \exp 0 = 0$$

$$\log \text{nat} = \log \exp 1 = 1$$

For the third we define

$$e_1 = \log p_1 \quad \text{and} \quad e_2 = \log p_2$$

so that we can rewrite the analogous exponential law.

$$\exp(e_1 + e_2) = \exp e_1 \cdot \exp e_2$$

$$\exp(\log p_1 + \log p_2) = \exp \log p_1 \cdot \exp \log p_2$$

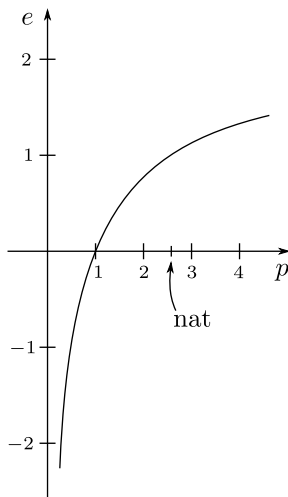
$$\exp(\log p_1 + \log p_2) = p_1 \cdot p_2$$

$$\log \exp(\log p_1 + \log p_2) = \log(p_1 \cdot p_2)$$

$$\log p_1 + \log p_2 = \log(p_1 \cdot p_2)$$

We've found it!

We can graph the natural logarithmic function by plotting points. Each point (e, p) on the natural exponential's graph, is a point (p, e) on the natural logarithm's graph.



p	e
nat^{-2}	-2
nat^{-1}	-1
nat^0	0
nat^1	1
nat^2	2

See if you can draw transition diagrams for the graph's height, slope, and bend!

6.5 Simple Differential Laws

The natural exponential and the natural logarithmic functions have simple differential laws.

Law.

The differential law for the natural exponential function states

$$\partial(\exp u) = \exp u \cdot \partial u$$

where u : Number is a number.

If you'd like to see why this differential rule holds, you can read [Appendix B. The Natural Exponential](#) which explores the natural exponential in much more detail. This appendix also helps explain our curious definition for the natural base, `nat`.

Law.

The differential law for the natural logarithm states

$$\partial(\log v) = \frac{1}{v} \cdot \partial v$$

where v : Number is any positive number, $v > 0$.

Why?

We'll use the following variables to solve for our differential law.

$$\begin{array}{ll} u \stackrel{\cdot}{\leftarrow} v & v \stackrel{\cdot}{\leftarrow} u \\ u = \log v & \text{and} \quad v = \exp u \end{array}$$

We find the logarithmic law by rewriting the corresponding law for the natural exponential.

$$\begin{aligned} \partial(\exp u) &= \exp u \cdot \partial u \\ \partial v &= v \cdot \partial(\log v) \\ \frac{1}{v} \cdot \partial v &= \partial(\log v) \end{aligned}$$

We've found it!

Let's use our differential law for the natural exponential to make a calculation.

Example.

Consider the function

$$\begin{aligned} z &\stackrel{\cdot}{\leftarrow} x \\ z &= \exp(3x - 1). \end{aligned}$$

We'll take the x -derivative.

$$\begin{aligned} \overset{x}{\partial} z &= \overset{x}{\partial} (\exp(3x - 1)) \\ &= \exp(3x - 1) \cdot \overset{x}{\partial} (3x - 1) \\ &= \exp(3x - 1) \cdot \overset{x}{\partial} (3x) \\ &= \exp(3x - 1) \cdot 3 \cdot \overset{x}{\partial} x \\ &= 3 \cdot \exp(3x - 1) \end{aligned}$$

Let's also see an example that uses a natural logarithm.

Example.

Consider the function

$$\begin{aligned} z &\stackrel{\cdot}{\leftarrow} x \\ z &= \log(x^2 + 1). \end{aligned}$$

Let's compute the derivative.

$$\begin{aligned} \overset{x}{\partial} z &= \overset{x}{\partial} (\log(x^2 + 1)) \\ &= \frac{1}{x^2 + 1} \cdot \overset{x}{\partial} (x^2 + 1) \\ &= \frac{1}{x^2 + 1} \cdot \overset{x}{\partial} (x^2) \\ &= \frac{1}{x^2 + 1} \cdot 2x \cdot \overset{x}{\partial} x \\ &= \frac{2x}{x^2 + 1} \end{aligned}$$

6.6 Complicated Differential Laws

We're now ready to see the differential law for powers in full, glorious detail.

Law.

The differential law for powers states

$$\partial(v^u) = v^u \cdot \log v \cdot \partial u + u \cdot v^{u-1} \cdot \partial v$$

where u : Number and v : Number are numbers, and the base is positive, $v > 0$.

Why?

Our differential law relies on the *change-of-base formula*.

$$v^u = \exp(u \cdot \log v)$$

Let's check that this formula holds.

$$\exp(u \cdot \log v) = \text{nat}^{u \cdot \log v} = (\text{nat}^{\log v})^u = (\exp \log v)^u = v^u$$

We'll use the change-of-base formula to rewrite the power v^u and thereby calculate its differential law!

$$\begin{aligned} \partial(v^u) &= \partial(\exp(u \cdot \log v)) \\ &= \exp(u \cdot \log v) \cdot \partial(u \cdot \log v) \\ &= v^u \cdot \partial(u \cdot \log v) \\ &= v^u \cdot (\log v \cdot \partial u + u \cdot \partial(\log v)) \\ &= v^u \cdot \left(\log v \cdot \partial u + u \cdot \frac{1}{v} \cdot \partial v \right) \\ &= v^u \cdot \log v \cdot \partial u + u \cdot v^{u-1} \cdot \partial v \end{aligned}$$

All that remains is to discuss the differential law for exponential functions.

Law.

The differential law for exponential functions states

$$\partial(c^u) = c^u \cdot \log c \cdot \partial u$$

where $u : \text{Number}$ is a number and the base, $c : \text{Number}$, is a positive constant, $c > 0$.

Why?

This follows directly from the differential law for powers.

$$\partial(c^u) = c^u \cdot \log c \cdot \partial u + u \cdot c^{u-1} \cdot \partial c = c^u \cdot \log c \cdot \partial u$$

Let's take the derivative of an exponential function with a non-natural base.

Example.

The exponential function with base two

$$\begin{aligned} p &\leftarrow e \\ p &= 2^e \end{aligned}$$

has e -derivative

$$\begin{aligned} \overset{e}{\partial} p &= \overset{e}{\partial} (2^e) \\ &= 2^e \cdot \log 2 \cdot \overset{e}{\partial} e \\ &= 2^e \cdot \log 2. \end{aligned}$$

Even when we work with a non-natural base, the natural logarithm has a way of getting involved!

Part II

Metrics

Chapter 7

Surfaces

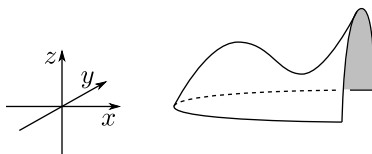
It's time for us to improve our graphing game! Up to this point, we've only needed two axes for graphing:

- a vertical z -axis for output, and
- a horizontal x -axis for input.

This is enough when we look at functions $z \leftarrow x$ that only have one input variable. But we'd like to start graphing functions $z \leftarrow (x, y)$ with two input variables, and for this, we'll need another axis:

- a horizontal y -axis for input.

Because two of our three axes are for input variables, a graph will be two-dimensional, a surface, sitting in three-dimensional space!



It's our job, then, to learn how to visualize surfaces. But surfaces can be quite complicated to draw on paper or a screen. Even with years experience, many of our drawings won't quite capture their subject. So we offer a compromise: we'll study surfaces by slicing them. Slices are relatively easy to draw because they ignore some of the surface's complexity. By looking at multiple slices, we can go a long way towards reconstructing a surface!

7.1 Surfaces as Graphs

Points

To start graphing functions of the form $z \leftarrow (x, y)$, we'll need some new notation for points.

Definition.

A *point* in the (x, y) -coordinate plane

$$p : \text{Point}_{(x,y)}$$

$$p = (x_p, y_p)$$

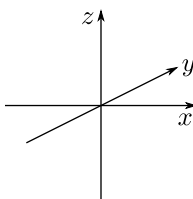
has two coordinates, and a *point* in (x, y, z) -coordinate space

$$q : \text{Point}_{(x,y,z)}$$

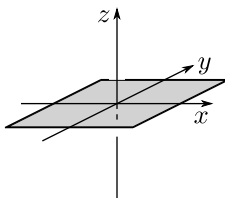
$$q = (x_q, y_q, z_q)$$

has three coordinates.

We'll locate points in (x, y, z) -coordinate space using three axes: the x -axis points right, the y -axis points back, and the z -axis points up.



In space, we may picture the (x, y) -coordinate plane as the horizontal plane with height zero, $\{z = 0\}$.



Each point (x_p, y_p) in the coordinate plane can be thought of as the point $(x_p, y_p, 0)$ in space.

Graphing

Let's think through how we will graph a function $z \leftarrow^f (x, y)$ that has two input variables. The *graph* is the surface whose *height* z is determined by f . In detail, suppose

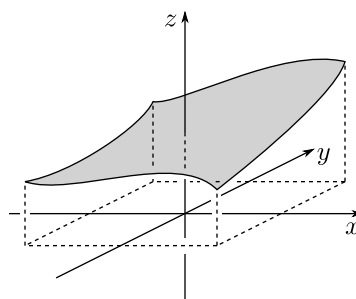
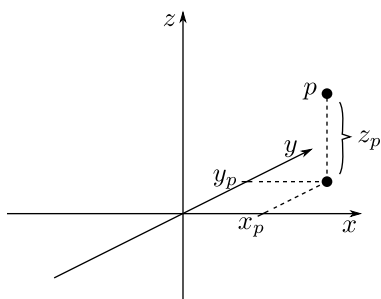
$$p : \text{Point}_{(x,y)}$$

$$p = (x_p, y_p)$$

is a point in the (x, y) -coordinate plane. Instead of drawing p at height zero, we'll draw it at height

$$z_p = f(x_p, y_p).$$

Locating each point p in this way, we'll find a surface!



In practice, graphing a surface point by point is far too much work to do by hand. And it is often difficult to draw a surface using perspective, in any case. To better understand a surface, we'll find it much more helpful to picture the curves found by slicing the surface.

7.2 Vertical Slices

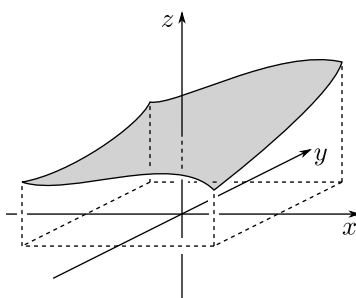
We can take a *vertical slice* of a function $z \leftarrow (x, y)$ by fixing either of its input variables.

Important.

The vertical slices of a function $z \leftarrow (x, y)$ are graphs.

- By fixing a y -value, we find a $(z \leftarrow x)$ -graph.
- By fixing an x -value, we find a $(z \leftarrow y)$ -graph.

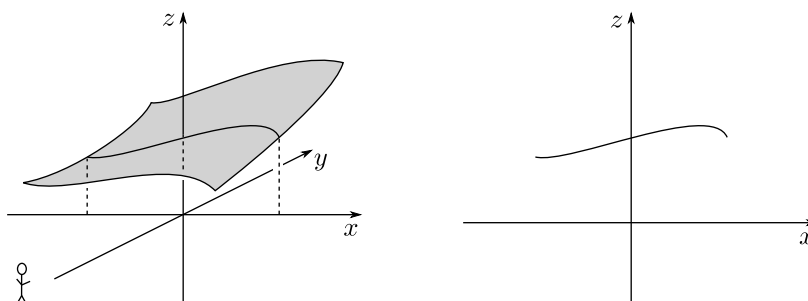
Let's take a look at how we vertically slice a graph.



We'll look at a slice $\{y = 0\}$ that fixes y , and then we'll look at a slice $\{x = 0\}$ that fixes x .

$\{y = 0\}$ -slice.

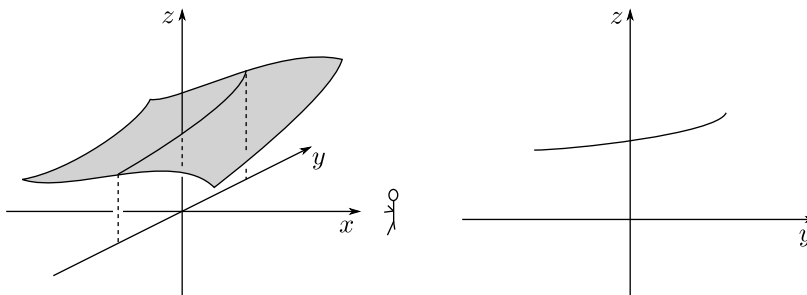
By fixing y , we see how z depends on x . The slice lives in an (x, z) -coordinate plane.



We see the slice by standing in front of the surface and looking along the y -axis.

$\{x = 0\}$ -slice.

By fixing x , we see how z depends on y . This slice lives in a (y, z) -coordinate plane.



We see the slice by standing to the right of the surface and looking along the x -axis.

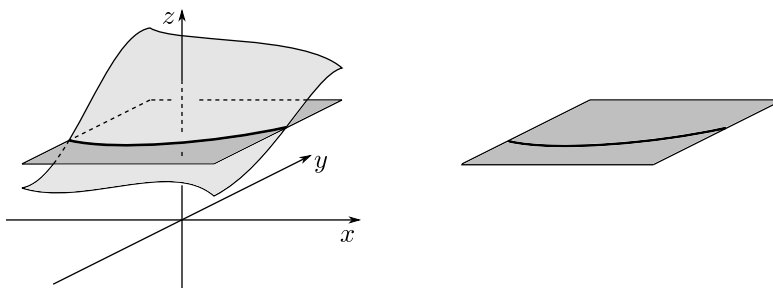
We'll often take $\{x = 0\}$ and $\{y = 0\}$ slices of a surface because they help us understand what's happening for inputs near the origin $(0, 0)$. We can take other vertical slices by fixing other x or y -values. We'll take lots of slices later in this chapter when we study the [bowl](#), the [parabolic saddle](#), the [warp saddle](#), and the [thick parabola](#). As you may have guessed, it is also possible to horizontally slice a surface. Horizontal slices are important enough that they get a section of their own!

7.3 Horizontal Slices

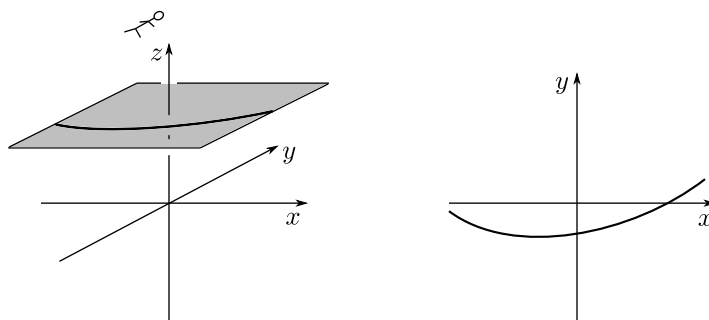
We just learned how to take vertical slices of a function $z \leftarrow (x, y)$ with two input variables: we fix either of the input variables. We'll take a *horizontal slice* of the graph by fixing the output variable. Given any constant z_0 : Number, the horizontal slice

$$\{z = z_0\}$$

consists of all points on the graph that have height z_0 .



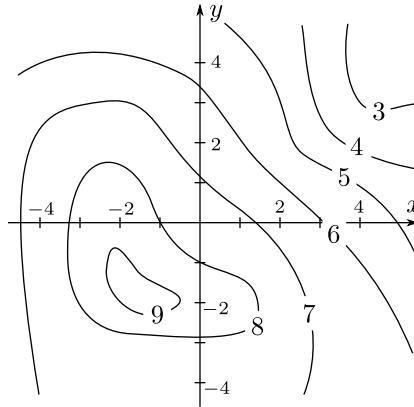
We see horizontal slices in (x, y, z) -coordinate space by looking down from above.



Unlike vertical slices, we do not expect a horizontal slice to be the graph of a function! In general, we have no reason to think that y depends on x or vice versa.

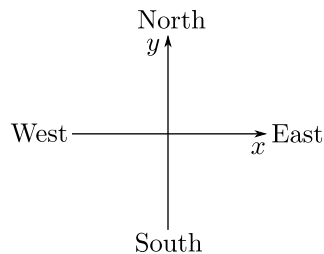
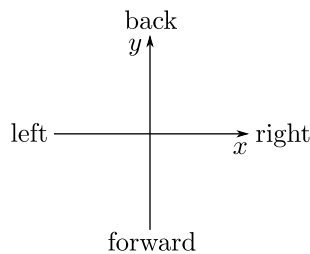
Contour Maps

A *contour map* for a function $z \leftarrow (x, y)$ gives a top-down picture of the function's graph. The contour map consists of several horizontal slices all drawn in the same (x, y) -coordinate plane. We'll call each slice a *contour*, and we'll label each contour with its z -value height.



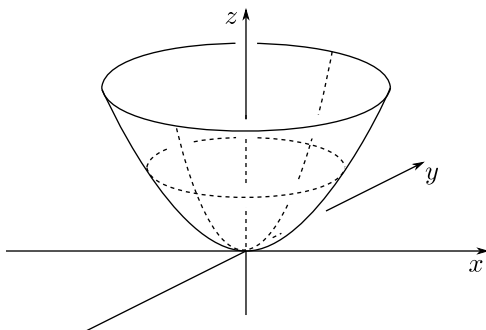
Contour maps are often used in cartography to display the elevation of a landscape. If we imagine ourselves hiking around on a surface, we can gain or lose elevation by hiking from one contour to another. To maintain our elevation, we may hike along a contour.

When we work in the (x, y) -coordinate plane, we should be careful how we describe movement in the y -direction. While it can be tempting to use the words “up” and “down” for the y -direction, it’s better to reserve these words for the z -direction. If we keep our picture of [space](#) in mind, we could use the words “back” and “forward” for the y -direction. Cartography offers another solution: we can use cardinal directions for the (x, y) -coordinate plane.



7.4 A Bowl

In the remainder of this chapter we'll be looking at several examples of surfaces. Let's start by studying the *bowl*



which is the graph of the function

$$z \leftarrow (x, y)$$

$$z = x^2 + y^2.$$

We'd like to understand this surface by taking some vertical slices and by looking at some contours.

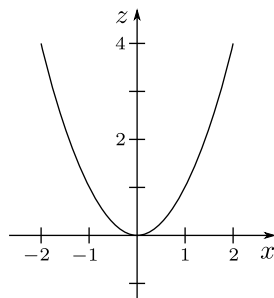
$\{y = 0\}$ -slice.

Setting y equal to zero gives a vertical slice.

$$z \leftarrow x$$

$$z = x^2$$

This is a parabola in the (x, z) -plane. This slice can be seen in the surface by looking from the front.



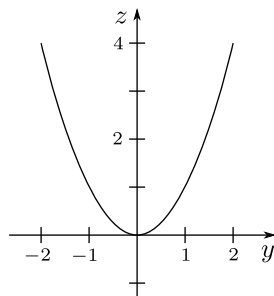
$\{x = 0\}$ -slice.

Setting x equal to zero also gives a vertical slice.

$$z \leftarrow y$$

$$z = y^2$$

This is the same parabola, except it sits in the (y, z) -plane. This slice can be seen in the surface by looking from the right side.



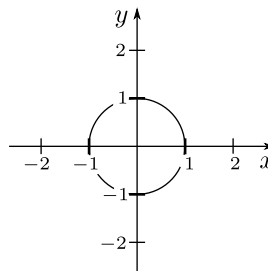
The $\{y = 0\}$ and $\{x = 0\}$ vertical slices go a long way in helping us understand our surface. They do not, however, capture the fact that our bowl is round. We'll turn to horizontal slices, or contours, to see this aspect of our surface.

$\{z = 1\}$ -contour.

When we set our output variable, z , equal to one

$$\{x^2 + y^2 = 1\}$$

we find the equation for a circle with radius one. We can locate this contour in our surface as the horizontal slice at height one. This slice should be visible in the surface by looking down from above.

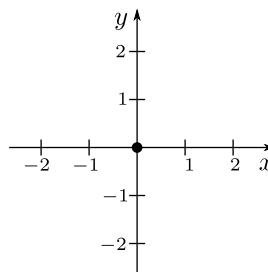


$\{z = 0\}$ -contour.

Now let's set our output variable, z , equal to zero.

$$\{x^2 + y^2 = 0\}$$

To satisfy this equation both x and y must be zero. The only point (x, y) that satisfies this equation is $(0, 0)$. In other words, the horizontal slice at height zero only intersects the bowl at one point, the origin.

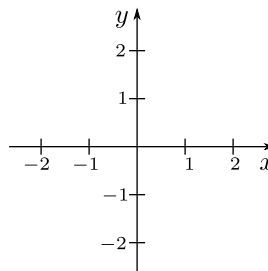


$\{z = -1\}$ -contour.

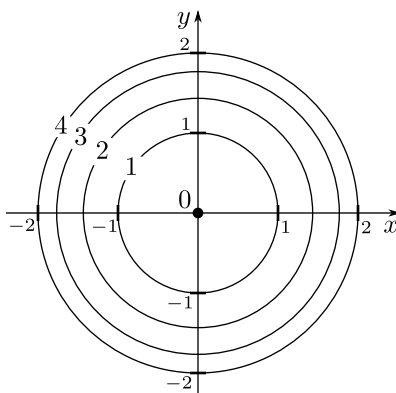
Setting z equal to negative one gives an equation that cannot be satisfied!

$$\{x^2 + y^2 = -1\}$$

Both x^2 and y^2 must be at least zero, and so their sum cannot be negative. What does this say about our graph? If we attempt to slice our surface at height negative one, we won't intersect the bowl.



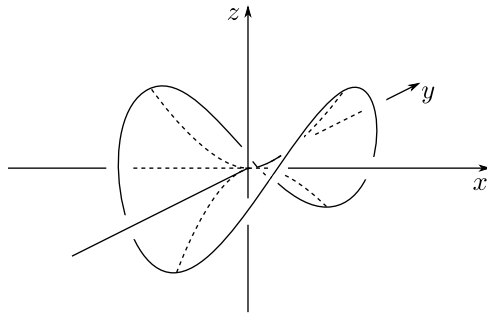
Great! These contours tell us a good deal of information about the shape of the bowl beyond the vertical slices. Let's wrap up our analysis of the bowl by looking at a contour map.



Our $\{z = 0\}$ and $\{z = 1\}$ contours can be seen on this map. We've gone ahead and added a few more contours to give a better sense of the bowl. Notice how the spacing between contour lines varies across the contour map. The bowl is a shallow surface near the origin, but it becomes steeper as we move away from the origin. A shallow graph has sparsely packed contours, and a steep graph has densely packed contours.

7.5 A Parabolic Saddle

Let's consider the *saddle*



that is made by graphing the function

$$z \leftarrow (x, y)$$

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

We'll call this saddle *parabolic* because its vertical slices are parabolas.

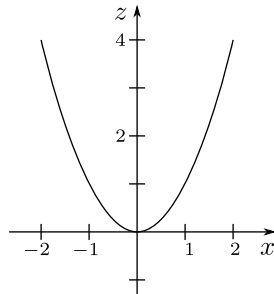
$\{y = 0\}$ -slice.

By setting y equal to zero, we find the classic upward facing parabola in the (x, z) -plane.

$$z \leftarrow x$$

$$z = x^2$$

This slice is visible in our surface when looking from the front.



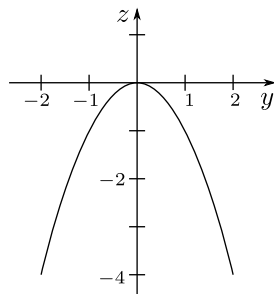
$\{x = 0\}$ -slice.

By fixing x equal to zero, we find a parabola that opens downward.

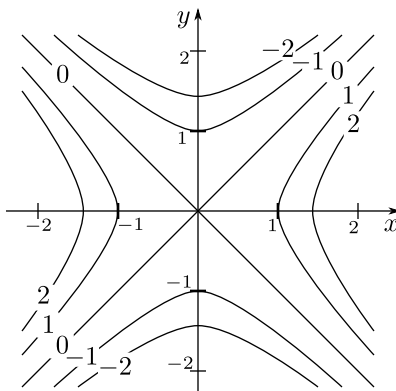
$$z \leftarrow y$$

$$z = -y^2$$

We can see this slice in the saddle when looking from the right.



Now let's look at a contour map for our surface.



Most of the contours for our saddle are hyperbolas, but there is one contour that should stand out. The $\{z = 0\}$ -contour consists of two lines that intersect at the origin. Let's take a moment to see how the algebra works out for this contour.

$\{z = 0\}$ -contour.

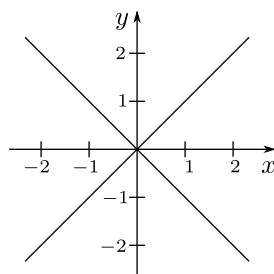
Let's take a horizontal slice at height zero.

$$\{x^2 - y^2 = 0\}$$

$$\{(x - y) \cdot (x + y) = 0\}$$

$$\{x - y = 0 \quad \text{or} \quad x + y = 0\}$$

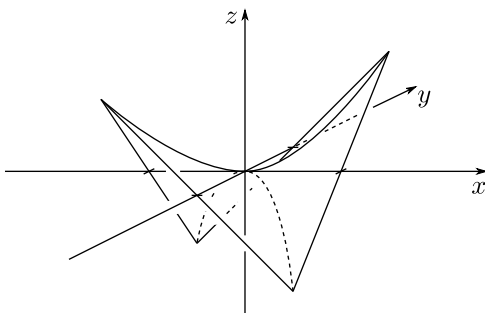
$$\{y = x \quad \text{or} \quad y = -x\}$$



We've found the two lines that make up the $\{z = 0\}$ -contour.

7.6 A Warp Saddle

Let's look at the *warp saddle*



made by graphing the function

$$z \leftarrow (x, y)$$

$$z = xy.$$

We'll take a few vertical slices by fixing y -values. Each of these slices can be seen by looking down the y -axis when standing in front of the surface.

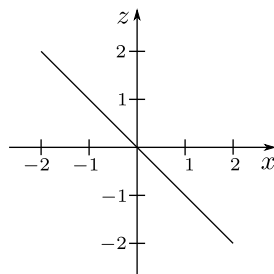
$\{y = -1\}$ -slice.

Setting y equal to negative one,

$$z \leftarrow x$$

$$z = -x$$

we find a line with slope negative one.



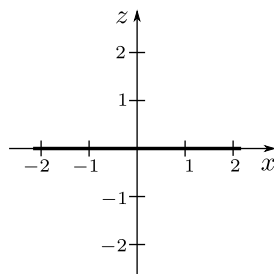
$\{y = 0\}$ -slice.

Setting y equal to zero has an interesting effect.

$$z \leftarrow x$$

$$z = 0$$

This slice is allowed to depend on x , but is just the constant, zero. The graph consists of all points on the x -axis.



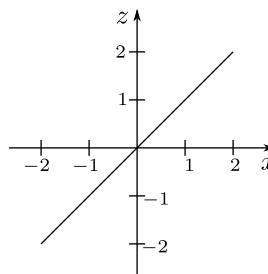
$\{y = 1\}$ -slice.

By fixing y equal to one,

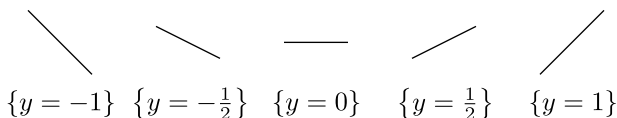
$$z \leftarrow x$$

$$z = x$$

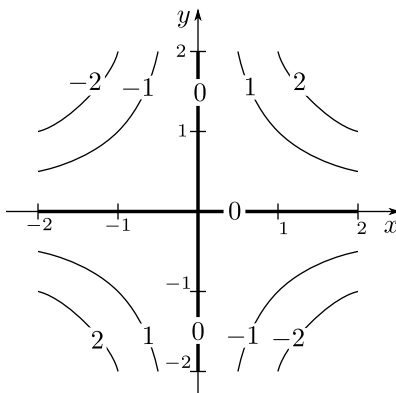
we find a line with slope one.



Notice how the vertical slices twist as we progress y -values. The $\{y = 0\}$ -slice is level: it has zero slope. Increasing y increases the slope, and decreasing y decreases the slope.



Let's finish our exploration of the warp saddle by taking a look at its contour map.



Most contours are hyperbolas, with the notable exception of the $\{z = 0\}$ -contour.

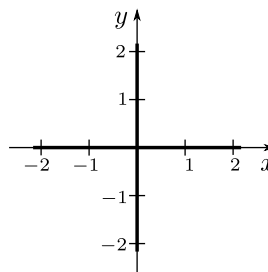
$\{z = 0\}$ -contour.

Setting z equal to zero gives us

$$\{xy = 0\}$$

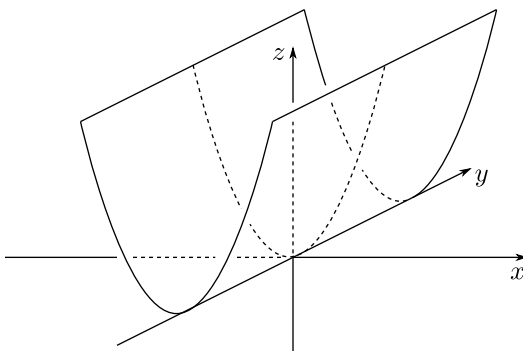
$$\{x = 0 \text{ or } y = 0\}.$$

In the (x, y) -coordinate plane, these are the equations for the y -axis and the x -axis, respectively.



7.7 A Thick Parabola

The *thick parabola* is the last in our series of example surfaces.

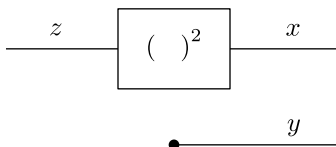


This is the graph of the function

$$z \leftarrow (x, y)$$

$$z = x^2.$$

Notice that one of the input variables y does not even appear in the definition of z . We'll say that the variable y has been *forgotten*. We can draw this function as a schematic by using a forgetter.



A forgetter is a splitter that creates zero copies of its input. For more on forgetters, see [Appendix C. Multi-Components](#). Let's take some vertical slices by fixing y -values.

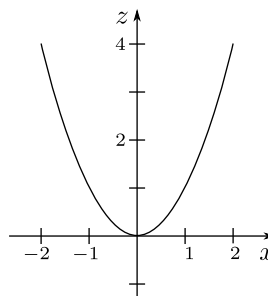
$\{y = 1\}$ -slice.

We take the $\{y = 1\}$ -slice by substituting the number, one, for every occurrence of y .

$$z \leftarrow x$$

$$z = x^2$$

Since there are no occurrences of y , no substitutions are needed.



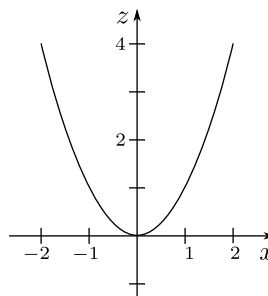
$\{y = -1\}$ -slice.

Fixing a different y -value gives an identical looking slice.

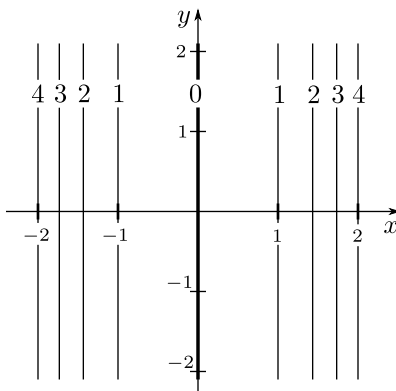
$$z \leftarrow x$$

$$z = x^2$$

Once again, there was no substitution needed because there is no dependence on y .



These vertical slices help explain our surface: every $(z \leftarrow x)$ -slice of this function is just the standard parabola! Let's finish up by taking a quick look at the contour map for the thick parabola.



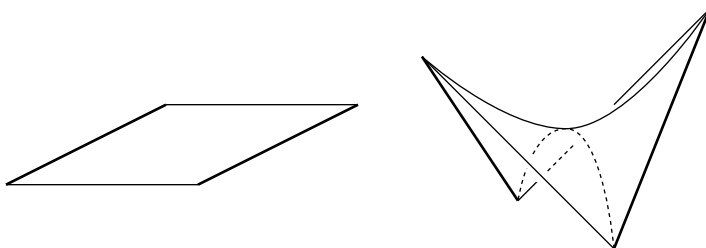
Each contour is a line. If we ever travel North or South on this contour map, we'll stay at the same height. To change our height, we should instead move East or West. Notice that the contours become more densely packed as we move away from the y -axis. Our surface is shallow near the y -axis and it becomes steeper moving to the East or West.

Chapter 8

Partials

By studying derivatives we learned to calculate slope for functions that have one input variable. Partials are a generalization of derivatives that work for functions with more input variables. We'll soon see that partials are slopes that are found in a surface's vertical slices.

Moving to higher partials, we'll find that a surface has both bend and warp. We've studied bend with derivatives, and bend will generalize in a simple way. Warp, however, is something new for surfaces! The [warp saddle](#) gives an excellent illustration of warp.



Imagine holding a square of some flexible material on its left and right edges. We create a warped surface by twisting those edges in opposite directions, one clockwise and the other counterclockwise. Using partials, we'll be able to calculate a surface's warp at any point. In [Chapter 9. Differentials as Metrics](#), we'll learn how to decide whether a surface is saddle-shaped. Warp will be an important, but not exclusive, factor in making this decision!

8.1 Partials

It's time that we study derivatives for functions with two input variables x and y .

Definition.

The x -partial, $\overset{x}{\partial}$, and the y -partial, $\overset{y}{\partial}$, both satisfy all [abstract differential laws](#). Further, partials satisfy *unit laws*

$$\overset{x}{\partial}x = 1 \quad \text{and} \quad \overset{y}{\partial}y = 1$$

and *independence laws*

$$\overset{x}{\partial}y = 0 \quad \text{and} \quad \overset{y}{\partial}x = 0.$$

We'll reserve the word, "[derivative](#)," for the operator $\overset{x}{\partial}$ that operates on functions with a single input variable, and we'll use the word, "partial," when more input variables are in play. That being said, derivatives and partials are so similar that they'll share the same notation.

Example.

Let's find both partials for the function

$$\begin{aligned} z &\leftarrow (x, y) \\ z &= x^2y + 4. \end{aligned}$$

We'll start with the x -partial of z .

$$\begin{aligned} \overset{x}{\partial}z &= \overset{x}{\partial}(x^2y + 4) \\ &= \overset{x}{\partial}(x^2y) \\ &= y \cdot \overset{x}{\partial}(x^2) + x^2 \cdot \overset{x}{\partial}y \\ &= y \cdot 2x \cdot \overset{x}{\partial}x + x^2 \cdot 0 \\ &= y \cdot 2x \cdot 1 \\ &= 2xy \end{aligned}$$

Now let's find the y -partial of z .

$$\begin{aligned}\frac{y}{\partial} z &= \frac{y}{\partial} (x^2 y + 4) \\ &= \frac{y}{\partial} (x^2 y) \\ &= y \cdot \frac{y}{\partial} (x^2) + x^2 \cdot \frac{y}{\partial} y \\ &= y \cdot 2x \cdot \frac{y}{\partial} x + x^2 \cdot 1 \\ &= y \cdot 2x \cdot 0 + x^2 \\ &= x^2\end{aligned}$$

Independence Shortcuts

The independence laws are new to partials: when a partial finds a non-matching input variable, the result is zero. The independence laws allow us to take shortcuts!

Important.

- If an expression u does not depend on x , then the x -partial treats u as a constant.
- If an expression v does not depend on y , then the y -partial treats v as a constant.

These shortcuts are probably best understood with an example.

Example.

Let's compute the partials for the function

$$\begin{aligned}z &\leftarrow (x, y) \\ z &= x^2 y^3 + 7y.\end{aligned}$$

Here's the x -partial of z .

$$\begin{aligned}\frac{x}{\partial} z &= \frac{x}{\partial} (x^2 y^3 + 7y) \\ &= \frac{x}{\partial} (x^2 y^3) \\ &= y^3 \cdot \frac{x}{\partial} (x^2) \\ &= y^3 \cdot 2x \cdot \frac{x}{\partial} x \\ &= 2xy^3\end{aligned}$$

Notice that the x -partial treats the expression $7y$ as an added constant, and it treats the expression y^3 as a multiplied constant. Now let's find the y -partial of z .

$$\begin{aligned}\frac{\partial}{\partial y} z &= \frac{\partial}{\partial y} (x^2 y^3 + 7y) \\ &= \frac{\partial}{\partial y} (x^2 y^3) + \frac{\partial}{\partial y} (7y) \\ &= x^2 \cdot \frac{\partial}{\partial y} (y^3) + 7 \cdot \frac{\partial}{\partial y} y \\ &= x^2 \cdot 3y^2 \cdot \frac{\partial}{\partial y} y + 7 \\ &= 3x^2 y^2 + 7\end{aligned}$$

We pulled the multiplier x^2 out of the y -partial.

You should take advantage of these independence shortcuts whenever you calculate a partial. For a deeper discussion of the independence laws, see [Appendix D. Independence](#).

8.2 Slope

The x -slope and the y -slope of a function $z \leftarrow (x, y)$ are the partials:

$\overset{x}{\text{slope } z} : \text{Number}$

$\overset{y}{\text{slope } z} : \text{Number}$

$$\overset{x}{\text{slope } z} = \overset{x}{\partial} z$$

and

$$\overset{y}{\text{slope } z} = \overset{y}{\partial} z.$$

We can picture these slopes in the vertical slices of the graph.

Example.

Let's investigate the slopes for the bowl,

$$z \leftarrow (x, y)$$

$$z = x^2 + y^2,$$

at the point $p = (1, -1)$. We calculate slopes as partials.

$$\overset{x}{\text{slope } z} = 2x$$

$$\overset{y}{\text{slope } z} = 2y$$

We can find the slopes at p by setting $x = 1$ and $y = -1$.

$$\overset{x}{\text{slope } z} = 2$$

$$\overset{y}{\text{slope } z} = -2$$

What do these slopes say about our surface, the bowl? Let's take some vertical slices through p to find out!

$\{y = -1\}$ -slice.

In the $\{y = -1\}$ -slice, we see how z depends on x .

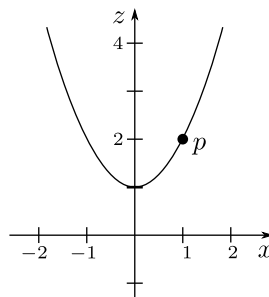
$$z \leftarrow x$$

$$z = x^2 + 1$$

This is the correct slice to find the x -slope.

$$\overset{x}{\text{slope } z} = 2$$

We locate p in this slice at $x = 1$. At p , we find our slope!



$\{x = 1\}$ -slice.

Now let's see how z depends on y .

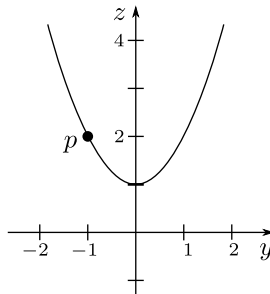
$$z \leftarrow y$$

$$z = y^2 + 1$$

We should be able to find the y -slope in this slice.

$$\overset{y}{\text{slope}} z = -2$$

By looking at $y = -1$, we find our point p and the y -slope.



When we calculate partials we learn the slopes for every point in the (x, y) -coordinate plane.

$$\overset{x}{\partial} z \leftarrow (x, y) \quad \overset{y}{\partial} z \leftarrow (x, y)$$

We may localize the slope at any particular point $p = (x_p, y_p)$.

Example.

Let's find the x and y -slopes at the point $p = (0, -1)$ for the function

$$z \leftarrow (x, y)$$

$$z = y^3 \cdot \exp(x^2).$$

The slopes are

$$\overset{x}{\text{slope}} z = 2xy^3 \cdot \exp(x^2)$$

$$\overset{y}{\text{slope}} z = 3y^2 \cdot \exp(x^2).$$

Localizing at the point $p = (0, -1)$, we find

$$\overset{x}{\text{slope}} z = 0$$

$$\overset{y}{\text{slope}} z = 3.$$

At p , the x -slope of z is zero, and the y -slope of z is three.

8.3 Planar Functions

A function $z \stackrel{f}{\leftarrow} (x, y)$ is *planar* if it can be written in the form

$$z - z_p = s_x \cdot (x - x_p) + s_y \cdot (y - y_p)$$

where

$$\begin{aligned} x_p &: \text{Number}, & y_p &: \text{Number}, & z_p &: \text{Number}, \\ s_x &: \text{Number}, & s_y &: \text{Number} \end{aligned}$$

are all constants.

Important.

The graph of a planar function

$$z - z_p = s_x \cdot (x - x_p) + s_y \cdot (y - y_p)$$

is the plane that

- has height z_p at the point $p = (x_p, y_p)$,
- has constant x -slope s_x , and
- has constant y -slope s_y .

Why?

We can solve for height z .

$$z = z_p + s_x \cdot (x - x_p) + s_y \cdot (y - y_p)$$

By localizing at $p = (x_p, y_p)$, we find

$$z = z_p.$$

Taking partials of z , we find the constants

$$\frac{x}{\partial} z = s_x \quad \frac{y}{\partial} z = s_y.$$

You may find it helpful to compare the definition of a planar function to that of a [linear function](#). Let's see an example of a planar function.

Example.

Consider the planar function $z \leftarrow (x, y)$ given by the equation

$$z - 2 = -3 \cdot (x - 1) + \frac{1}{2} \cdot (y - 4).$$

We can understand this plane by reading off the constants:

$$\begin{aligned} x_p &= 1, & y_p &= 4, & z_p &= 2, \\ s_x &= -3, & s_y &= \frac{1}{2}. \end{aligned}$$

At the point $(1, 4)$, our plane has height two. The x -slope is negative three everywhere, and the y -slope is one-half everywhere.

Horizontal Planes

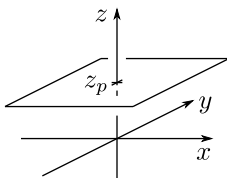
We can describe a horizontal plane as a planar function by setting both slopes s_x and s_y equal to zero. We'll write

$$z - z_p = 0 \cdot (x - x_p) + 0 \cdot (y - y_p)$$

or more simply

$$z = z_p.$$

This is an equation we've seen before: we set the height z equal to a constant to find [horizontal slices](#).



Unlike horizontal planes, a vertical plane is never the graph of a function $z \leftarrow (x, y)$. After all, a function must produce only a single height z for each point $p = (x_p, y_p)$.

8.4 Tangent Planes

We can find a tangent plane at any point on the graph of a function $z \leftarrow (x, y)$. Like [tangent lines](#), the tangent plane will help us better understand the shapes of graphs.

Definition.

Suppose $z \leftarrow (x, y)$ is a function. The *tangent to f* is the planar function $z_T \leftarrow (x_T, y_T)$ described by the equation

$$z_T - z = \text{slope } z^x \cdot (x_T - x) + \text{slope } z^y \cdot (y_T - y)$$

where

$$\begin{array}{lll} x : \text{Number}, & y : \text{Number}, & z : \text{Number}, \\ \text{slope } z^x : \text{Number}, & \text{slope } z^y : \text{Number} & \end{array}$$

are constants found by localizing at a point $p : \text{Point}_{(x,y)}$ called the *center*.

The tangent plane is the plane that best matches the graph of f near the center p .

Example.

We'll find the tangent plane for the [bowl](#)

$$\begin{aligned} z &\leftarrow (x, y) \\ z &= x^2 + y^2 \end{aligned}$$

centered at the point $p = (1, -1)$. We start by taking partials.

$$\text{slope } z^x = 2x \quad \text{slope } z^y = 2y$$

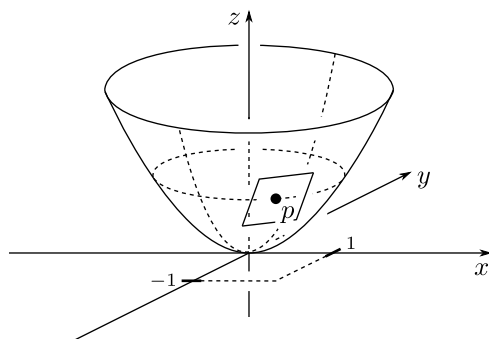
Localizing at the point $p = (1, -1)$, we find constants:

$$\begin{array}{lll} x = 1, & y = -1, & z = 2, \\ \text{slope } z^x = 2, & \text{slope } z^y = -2. & \end{array}$$

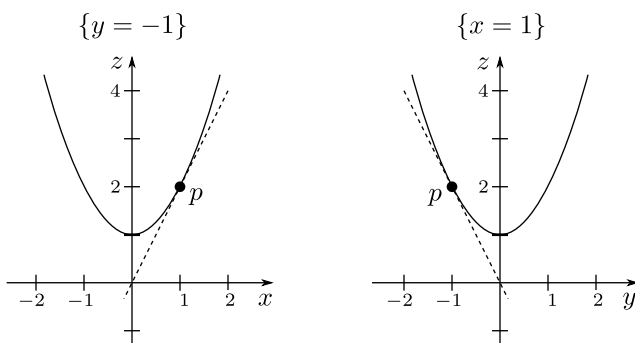
This is everything we need to write down an equation for the tangent plane!

$$z_T - 2 = 2 \cdot (x_T - 1) - 2 \cdot (y_T + 1)$$

Here's a sketch of what the tangent plane from the example actually looks like.



In each vertical slice the tangent plane shows up as a tangent line!



We see the $\{y = -1\}$ -slice when viewing the bowl from the front, and we see the $\{x = 1\}$ -slice when viewing the bowl from the right.

8.5 Bend

A function $z \leftarrow (x, y)$ has two bends.

$$\overset{x}{\text{bend}} z : \text{Number}$$

$$\overset{y}{\text{bend}} z : \text{Number}$$

$$\overset{x}{\text{bend}} z = \frac{\overset{x x}{\partial \partial} z}{2}$$

and

$$\overset{y}{\text{bend}} z = \frac{\overset{y y}{\partial \partial} z}{2}$$

The x -*bend* is the second x -partial, halved. And the y -*bend* is the second y -partial, halved.

Example.

Consider the function

$$z \leftarrow (x, y)$$

$$z = xy^3 - 5x^2.$$

Let's compute the x -bend and the y -bend.

$$\overset{x}{\partial} z = y^3 - 10x$$

$$\overset{y}{\partial} z = 3xy^2$$

$$\overset{x x}{\partial \partial} z = -10$$

$$\overset{y y}{\partial \partial} z = 6xy$$

$$\overset{x}{\text{bend}} z = -5$$

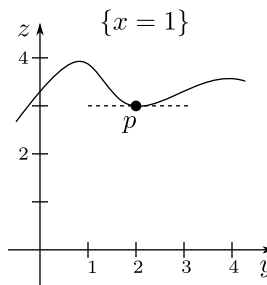
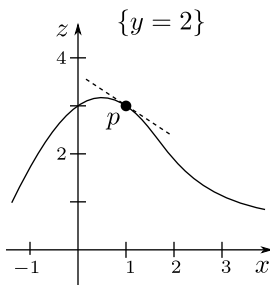
$$\overset{y}{\text{bend}} z = 3xy$$

This function has a constant x -bend of negative five, but the y -bend varies from point to point.

Like we saw with slopes, a graph's x -bend and y -bend can be found in its vertical slices.

Example.

Suppose a function $z \leftarrow (x, y)$ has vertical slices through the point $p = (1, 2)$ as pictured.



We've drawn in the tangents at p as dotted lines. Let's see what we can say about slope and bend at p .

$\{y = 2\}$ -slice.

In the $\{y = 2\}$ -slice, we see how z depends on x . We can learn about the x -slope and the x -bend in this slice. At p , we see a graph that has indirect variance and is concave down.

$$\overset{x}{\text{slope } z} < 0 \quad \overset{x}{\text{bend } z} < 0$$

The slope is negative, and the bend is negative, also.

$\{x = 1\}$ -slice.

In the $\{x = 1\}$ -slice, we can see how z depends on y .

$$\overset{y}{\text{slope } z} = 0 \quad \overset{y}{\text{bend } z} > 0$$

Here we see zero slope and a positive bend.

8.6 Warp

The yx -warp and xy -warp of a function $z \leftarrow (x, y)$ are the mixed second partials:

$$\begin{array}{ll} \overset{yx}{\text{warp}} z : \text{Number} & \overset{xy}{\text{warp}} z : \text{Number} \\ \overset{yx}{\text{warp}} z = \overset{yx}{\partial \partial} z & \text{and} \quad \overset{xy}{\text{warp}} z = \overset{xy}{\partial \partial} z. \end{array}$$

The first thing to know about warp is that we are free to take the partials in either order.

Theorem.

For any function $z \leftarrow (x, y)$, the yx -warp and the xy -warp are equal.

$$\overset{yx}{\partial \partial} z = \overset{xy}{\partial \partial} z$$

We'll usually just write $\text{warp } z$ for the warp as long as there are only two input variables.

Example.

Let's compute the warp of the function

$$\begin{aligned} z &\leftarrow (x, y) \\ z &= xy^3 - 5x^2 \end{aligned}$$

in both possible orders.

$$\begin{aligned} \overset{x}{\partial} z &= \overset{x}{\partial} (xy^3 - 5x^2) = y^3 - 10x \\ \overset{yx}{\partial \partial} z &= \overset{y}{\partial} (y^3 - 10x) = 3y^2 \end{aligned}$$

Here we took the y -partial of the x -partial of z . Now let's take the x -partial of the y -partial of z .

$$\begin{aligned} \overset{y}{\partial} z &= \overset{y}{\partial} (xy^3 - 5x^2) = 3xy^2 \\ \overset{xy}{\partial \partial} z &= \overset{x}{\partial} (3xy^2) = 3y^2 \end{aligned}$$

We found the same result.

$$\text{warp } z = 3y^2$$

Let's take a look at the warp for an important example.

Example.

Recall the [warp saddle](#),

$$z \leftarrow (x, y)$$

$$z = xy.$$

Let's compute the warp! We'll start by taking an x -partial.

$$\frac{x}{\partial} z = \frac{x}{\partial} (xy) = y \cdot \frac{x}{\partial} x = y$$

And now we'll take the y -partial of the x -partial.

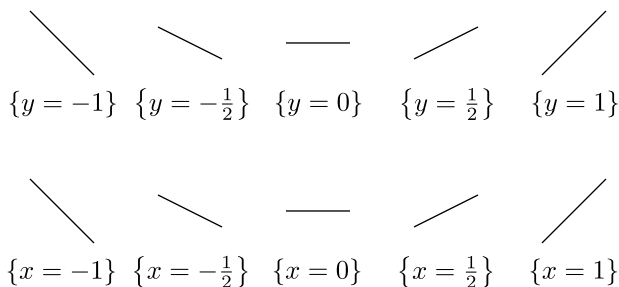
$$\frac{y}{\partial} \frac{x}{\partial} z = \frac{y}{\partial} y = 1$$

We've found our result.

$$\text{warp } z = 1$$

The warp saddle has a constant warp of one! Just as the parabola, $z = x^2$, is our ideal example of bend, the warp saddle is our ideal example of warp.

You may have noticed that we did not provide any explanation for our theorem about warp. Unfortunately, a full explanation requires advanced topics: integrals and limits. But by looking at the warp saddle, we see pretty good evidence for the theorem. A surface's warp tells us how quickly the vertical slices twist.



The $(z \leftarrow x)$ -slices and the $(z \leftarrow y)$ -slices of the warp saddle twist at the same rate in the same counterclockwise direction.

Chapter 9

Differentials as Metrics

In this chapter we'll continue our work analyzing surfaces. We've learned that a function's partials give us a surface's slopes, and higher partials let us calculate bend and warp. Our goal is to apply this material towards identifying low points and high points on a surface. To get there, however, we'll need to introduce a new type of object: metrics.

$$\hat{\partial}(\quad)$$

By studying metrics and their differentials, we'll learn how to identify the shape of a surface by comparing the surface to its tangent plane. A surface will typically peel away from its tangent plane in either a saddle-shaped fashion or a bowl-shaped fashion. Making this distinction is crucial for locating low points and high points on a surface.

9.1 The Metric Differential

Metrics

There are two standard metrics for the (x, y) -coordinate plane:

- the *standard x-metric*, $\hat{e}_x : \text{Metric}_{(x,y)}$, and
- the *standard y-metric*, $\hat{e}_y : \text{Metric}_{(x,y)}$.

Definition.

Any *metric* $m : \text{Metric}_{(x,y)}$ can be written in *standard form* as

$$m = a \cdot \hat{e}_x + b \cdot \hat{e}_y$$

where the numbers $a : \text{Number}$ and $b : \text{Number}$ are called the metric's *parts*.

In standard form, we'll say that m is written as a *combination* of $\hat{e}_x$ and $\hat{e}_y$.

Example.

The metric

$$m : \text{Metric}_{(x,y)}$$

$$m = 2 \cdot \hat{e}_x + 5 \cdot \hat{e}_y$$

has x -part two and y -part five.

The Differential

Let's cut right to the chase: we can interpret the differential as a metric.

Definition.

For any function $z \leftarrow (x, y)$ there is a metric

$$\hat{\partial}z : \text{Metric}_{(x,y)}$$

known as the *metric differential of z* . The operator $\hat{\partial}$ satisfies all [abstract differential laws](#). We also have the following *metric interpretation laws*.

$$\hat{\partial}x = \hat{e}_x \quad \text{and} \quad \hat{\partial}y = \hat{e}_y$$

The metric differential does not assume the unit laws or the independence laws that we saw

with **partials**. So we must give up our independence shortcuts when calculating the metric differential.

Example.

Consider the function

$$\begin{aligned} z &\leftarrow (x, y) \\ z &= \sqrt[2]{x^2 + y^2}. \end{aligned}$$

Let's take the metric differential of z .

$$\begin{aligned} \hat{\partial}z &= \hat{\partial}\left(\sqrt[2]{x^2 + y^2}\right) \\ &= \frac{1}{2 \cdot \sqrt[2]{x^2 + y^2}} \cdot \hat{\partial}(x^2 + y^2) \\ &= \frac{1}{2 \cdot \sqrt[2]{x^2 + y^2}} \cdot (\hat{\partial}(x^2) + \hat{\partial}(y^2)) \\ &= \frac{1}{2 \cdot \sqrt[2]{x^2 + y^2}} \cdot (2x \cdot \hat{\partial}x + 2y \cdot \hat{\partial}y) \\ &= \frac{x}{\sqrt[2]{x^2 + y^2}} \cdot \hat{\partial}x + \frac{y}{\sqrt[2]{x^2 + y^2}} \cdot \hat{\partial}y \\ &= \frac{x}{\sqrt[2]{x^2 + y^2}} \cdot \hat{\varepsilon}_x + \frac{y}{\sqrt[2]{x^2 + y^2}} \cdot \hat{\varepsilon}_y \end{aligned}$$

We've found our answer.

To calculate the metric differential, we apply our abstract differential laws as usual. When we are left with $\hat{\partial}x$ and $\hat{\partial}y$ terms, we interpret these as the standard metrics, $\hat{\varepsilon}_x$ and $\hat{\varepsilon}_y$.

Example.

Let's find the metric differential for the function

$$\begin{aligned} z &\leftarrow (x, y) \\ z &= x^2y - xy^2. \end{aligned}$$

We compute.

$$\begin{aligned}\hat{\partial}z &= \hat{\partial}(x^2y - xy^2) \\ &= \hat{\partial}(x^2y) - \hat{\partial}(xy^2) \\ &= (y \cdot \hat{\partial}(x^2) + x^2 \cdot \hat{\partial}y) - (y^2 \cdot \hat{\partial}x + x \cdot \hat{\partial}(y^2)) \\ &= y \cdot 2x \cdot \hat{\partial}x + x^2 \cdot \hat{\partial}y - y^2 \cdot \hat{\partial}x - x \cdot 2y \cdot \hat{\partial}y \\ &= 2xy \cdot \hat{\varepsilon}_x + x^2 \cdot \hat{\varepsilon}_y - y^2 \cdot \hat{\varepsilon}_x - 2xy \cdot \hat{\varepsilon}_y \\ &= (2xy - y^2) \cdot \hat{\varepsilon}_x + (x^2 - 2xy) \cdot \hat{\varepsilon}_y\end{aligned}$$

We've written our result in standard form.

9.2 Parts

The metric differential, $\hat{\partial}$, is closely related to the partials, $\overset{x}{\partial}$ and $\overset{y}{\partial}$. In fact, we may think of partials as “part differentials.”

Formula.

The metric differential of a function $z \leftarrow (x, y)$ is the metric

$$\begin{aligned}\hat{\partial}z &: \text{Metric}_{(x,y)} \\ \hat{\partial}z &= a \cdot \hat{e}_x + b \cdot \hat{e}_y\end{aligned}$$

whose parts

$$\begin{array}{ll}a : \text{Number} & b : \text{Number} \\ a = \overset{x}{\partial}z & \text{and} \quad b = \overset{y}{\partial}z\end{array}$$

are the partials of z .

Let's see what this looks like with an example.

Example.

The function

$$\begin{aligned}z &\leftarrow (x, y) \\ z &= (x - y)^3\end{aligned}$$

has differential

$$\hat{\partial}z = 3 \cdot (x - y)^2 \cdot \hat{e}_x - 3 \cdot (x - y)^2 \cdot \hat{e}_y.$$

The partials of z are the parts of $\hat{\partial}z$.

$$\begin{aligned}\overset{x}{\partial}z &= 3 \cdot (x - y)^2 \\ \overset{y}{\partial}z &= -3 \cdot (x - y)^2\end{aligned}$$

The metric differential can be read as a concise description of the [tangent plane](#): at any point, the differential records the graph's x -slope and y -slope. To emphasize a geometric point of view, we'll call the differential the *tangent metric*.

tangent z : Metric

$$\text{tangent } z = \hat{\partial} z$$

Let's rewrite our formula for the metric differential using geometric language.

Important.

The parts of the tangent metric are the x -slope and the y -slope. That is,

$$\text{tangent } z = \overset{x}{\text{slope } z} \cdot \hat{\epsilon}_x + \overset{y}{\text{slope } z} \cdot \hat{\epsilon}_y$$

where $z \leftarrow (x, y)$ is any function.

We can localize the tangent metric at any point $p : \text{Point}_{(x,y)}$ of interest.

Example.

Consider the [parabolic saddle](#).

$$\begin{aligned} z &\leftarrow (x, y) \\ z &= x^2 - y^2 \end{aligned}$$

Let's compute the metric differential.

$$\begin{aligned} \hat{\partial} z &= \hat{\partial}(x^2 - y^2) \\ &= \hat{\partial}(x^2) - \hat{\partial}(y^2) \\ &= 2x \cdot \hat{\epsilon}_x - 2y \cdot \hat{\epsilon}_y \end{aligned}$$

We've found the tangent metric.

$$\text{tangent } z = 2x \cdot \hat{\epsilon}_x - 2y \cdot \hat{\epsilon}_y$$

Let's see what the tangent metric looks like at the point $(3, 1) : \text{Point}_{(x,y)}$.

$$\text{tangent } z = 6 \cdot \hat{\epsilon}_x - 2 \cdot \hat{\epsilon}_y$$

At $(3, 1)$, the tangent plane has a positive x -slope and a negative y -slope. Can you find a point where both the x -slope and the y -slope are positive? Or a point where the tangent plane is horizontal?

For more the tangent metric, see [Appendix E. Applying a Metric](#).

9.3 Polymetrics

In order to go further with the differential we'll need to understand how to work with polymetrics. The standard metrics are *polymetrics*,

$$\hat{\varepsilon}_x : \text{Polymetric}_{(x,y)}$$

$$\hat{\varepsilon}_y : \text{Polymetric}_{(x,y)}$$

numbers are *polymetrics*,

$$7 : \text{Polymetric}_{(x,y)}$$

$$-4 : \text{Polymetric}_{(x,y)}$$

and we build all other *polymetrics* from these by adding and multiplying.

$$5 + 3 \cdot \hat{\varepsilon}_x - \hat{\varepsilon}_y + \hat{\varepsilon}_x \cdot \hat{\varepsilon}_y : \text{Polymetric}_{(x,y)}$$

Law.

The multiplication of polymetrics is commutative. We can reorder multiplication

$$k \cdot j = j \cdot k$$

where $j : \text{Polymetric}_{(x,y)}$ and $k : \text{Polymetric}_{(x,y)}$ are any two polymetrics.

Polymetrics generalize metrics: every metric is a polymetric, but there are many polymetrics which are not metrics.

Example.

Let's multiply the metrics

$$m : \text{Metric}_{(x,y)}$$

$$m = 3 \cdot \hat{\varepsilon}_x + 4 \cdot \hat{\varepsilon}_y$$

and

$$n : \text{Metric}_{(x,y)}$$

$$n = 2 \cdot \hat{\varepsilon}_x - \hat{\varepsilon}_y.$$

We calculate.

$$\begin{aligned} m \cdot n &= (3 \cdot \hat{\varepsilon}_x + 4 \cdot \hat{\varepsilon}_y) \cdot (2 \cdot \hat{\varepsilon}_x - \hat{\varepsilon}_y) \\ &= 6 \cdot \hat{\varepsilon}_x^2 - 3 \cdot \hat{\varepsilon}_x \cdot \hat{\varepsilon}_y + 8 \cdot \hat{\varepsilon}_y \cdot \hat{\varepsilon}_x - 4 \cdot \hat{\varepsilon}_y^2 \\ &= 6 \cdot \hat{\varepsilon}_x^2 + 5 \cdot \hat{\varepsilon}_x \cdot \hat{\varepsilon}_y - 4 \cdot \hat{\varepsilon}_y^2 \end{aligned}$$

Multiplying two metrics does not result in a metric. Rather, the product of metrics is a polymetric!

In the previous example, we used the fact

$$\hat{\varepsilon}_y \cdot \hat{\varepsilon}_x = \hat{\varepsilon}_x \cdot \hat{\varepsilon}_y.$$

This holds because our multiplication is commutative.

Degree of a Polymetric

Just like working with polynomials, we can always simplify a polymetric until it becomes a sum of terms. The *degree of a term* is the number of standard metrics it contains. So,

$$3 \cdot \hat{\varepsilon}_x : \text{Polymetric}_{(x,y)}^1$$

$$4 \cdot \hat{\varepsilon}_y : \text{Polymetric}_{(x,y)}^1$$

are degree one terms, and

$$6 \cdot \hat{\varepsilon}_x^2 : \text{Polymetric}_{(x,y)}^2$$

$$5 \cdot \hat{\varepsilon}_x \cdot \hat{\varepsilon}_y : \text{Polymetric}_{(x,y)}^2$$

$$-4 \cdot \hat{\varepsilon}_y^2 : \text{Polymetric}_{(x,y)}^2$$

are degree two terms. Unlike polynomials, we expect a polymetric to be *homogeneous* to have a degree: each term should have the same degree.

degree	polymetric standard form
0	a
1	$a \cdot \hat{\varepsilon}_x + b \cdot \hat{\varepsilon}_y$
2	$a \cdot \hat{\varepsilon}_x^2 + b \cdot \hat{\varepsilon}_x \cdot \hat{\varepsilon}_y + c \cdot \hat{\varepsilon}_y^2$

In degrees zero and one, polymetrics are familiar objects.

Important.

A degree zero polymetric is a number and a degree one polymetric is a metric.

We'll be most interested in polymetrics with degrees zero, one, or two, but you can find polymetrics in higher degrees, too.

9.4 Higher Differentials

Let's learn to take the differential of a polymetric. Using polymetrics, we can iterate the differential to our hearts content!

Laws.

The differential law for addition states

$$\hat{\partial}(j + k) = \hat{\partial}j + \hat{\partial}k$$

and the differential law for multiplication states

$$\hat{\partial}(j \cdot k) = k \cdot \hat{\partial}j + j \cdot \hat{\partial}k$$

where $j : \text{Polymetric}_{(x,y)}$ and $k : \text{Polymetric}_{(x,y)}$ are polymetrics.

Both of the standard metrics, $\hat{\varepsilon}_x$ and $\hat{\varepsilon}_y$, have zero differential.

$$\hat{\partial}(\hat{\varepsilon}_x) = 0 \quad \text{and} \quad \hat{\partial}(\hat{\varepsilon}_y) = 0$$

As such, they look a bit like constants when taking the differential.

Laws.

The standard metrics pull out of the differential

$$\hat{\partial}(j \cdot \hat{\varepsilon}_x) = \hat{\partial}j \cdot \hat{\varepsilon}_x$$

$$\hat{\partial}(j \cdot \hat{\varepsilon}_y) = \hat{\partial}j \cdot \hat{\varepsilon}_y$$

where $j : \text{Polymetric}_{(x,y)}$ is any polymetric.

Let's see how we take the differential of a metric.

Example.

Consider the metric

$$m : \text{Metric}_{(x,y)}$$

$$m = 2xy \cdot \hat{e}_x + x^3 \cdot \hat{e}_y.$$

Let's calculate the differential.

$$\begin{aligned} \hat{\partial}m &= \hat{\partial}(2xy \cdot \hat{e}_x + x^3 \cdot \hat{e}_y) \\ &= \hat{\partial}(2xy \cdot \hat{e}_x) + \hat{\partial}(x^3 \cdot \hat{e}_y) \\ &= \hat{\partial}(2xy) \cdot \hat{e}_x + \hat{\partial}(x^3) \cdot \hat{e}_y \\ &= 2 \cdot \hat{\partial}(xy) \cdot \hat{e}_x + 3x^2 \cdot \hat{\partial}x \cdot \hat{e}_y \\ &= 2 \cdot (y \cdot \hat{\partial}x + x \cdot \hat{\partial}y) \cdot \hat{e}_x + 3x^2 \cdot \hat{e}_x \cdot \hat{e}_y \\ &= 2y \cdot \hat{e}_x \cdot \hat{e}_x + 2x \cdot \hat{e}_y \cdot \hat{e}_x + 3x^2 \cdot \hat{e}_x \cdot \hat{e}_y \\ &= 2y \cdot \hat{e}_x \cdot \hat{e}_x + (2x + 3x^2) \cdot \hat{e}_x \cdot \hat{e}_y \end{aligned}$$

The result is a polymetric with degree two.

Taking the differential of a polymetric produces a polymetric with degree one larger. So starting with a function,

$$z \leftarrow (x, y)$$

we find

$$\begin{aligned} \hat{\partial}z &: \text{Metric}_{(x,y)} \\ \hat{\partial}\hat{\partial}z &: \text{Polymetric}_{(x,y)}^2 \\ \hat{\partial}\hat{\partial}\hat{\partial}z &: \text{Polymetric}_{(x,y)}^3 \end{aligned}$$

and so on.

Example.

Consider the function

$$z \leftarrow (x, y)$$

$$z = xy^3 + 1.$$

Let's take the differential twice. The first differential is a metric

$$\begin{aligned}\hat{\partial}z &= \hat{\partial}(xy^3 + 1) \\ &= \hat{\partial}(xy^3) \\ &= y^3 \cdot \hat{\partial}x + x \cdot \hat{\partial}(y^3) \\ &= y^3 \cdot \hat{e}_x + x \cdot 3y^2 \cdot \hat{\partial}y \\ &= y^3 \cdot \hat{e}_x + 3xy^2 \cdot \hat{e}_y\end{aligned}$$

and the second differential is a polymetric with degree two.

$$\begin{aligned}\hat{\partial}\hat{\partial}z &= \hat{\partial}(y^3 \cdot \hat{e}_x + 3xy^2 \cdot \hat{e}_y) \\ &= \hat{\partial}(y^3 \cdot \hat{e}_x) + \hat{\partial}(3xy^2 \cdot \hat{e}_y) \\ &= \hat{\partial}(y^3) \cdot \hat{e}_x + \hat{\partial}(3xy^2) \cdot \hat{e}_y \\ &= 3y^2 \cdot \hat{\partial}y \cdot \hat{e}_x + 3 \cdot \hat{\partial}(xy^2) \cdot \hat{e}_y \\ &= 3y^2 \cdot \hat{e}_y \cdot \hat{e}_x + 3 \cdot (y^2 \cdot \hat{\partial}x + x \cdot \hat{\partial}(y^2)) \cdot \hat{e}_y \\ &= 3y^2 \cdot \hat{e}_y \cdot \hat{e}_x + 3 \cdot (y^2 \cdot \hat{\partial}x + x \cdot 2y \cdot \hat{\partial}y) \cdot \hat{e}_y \\ &= 3y^2 \cdot \hat{e}_y \cdot \hat{e}_x + 3y^2 \cdot \hat{e}_x \cdot \hat{e}_y + 6xy \cdot \hat{e}_y \cdot \hat{e}_y \\ &= 6y^2 \cdot \hat{e}_x \cdot \hat{e}_y + 6xy \cdot \hat{e}_y \cdot \hat{e}_y\end{aligned}$$

9.5 The Saddle Discriminant

The *Hessian polymetric* of a function $z \leftarrow (x, y)$ is the second differential, halved.

$$\text{hess } z : \text{Polymetric}_{(x,y)}^2$$

$$\text{hess } z = \frac{\hat{\partial}\hat{\partial}z}{2}$$

The Hessian describes how a surface peels away from its tangent plane by taking the bend and warp into account.

Formula.

The Hessian

$$\text{hess } z = a \cdot \hat{\varepsilon}_x^2 + b \cdot \hat{\varepsilon}_x \cdot \hat{\varepsilon}_y + c \cdot \hat{\varepsilon}_y^2$$

records the x -bend, the warp, and the y -bend.

$$a = \overset{x}{\text{bend}} z \quad b = \text{warp } z \quad c = \overset{y}{\text{bend}} z$$

Why?

Let's first find a formula for the second differential.

$$\begin{aligned} \hat{\partial}\hat{\partial}z &= \hat{\partial}\left(\overset{x}{\partial}z \cdot \hat{\varepsilon}_x + \overset{y}{\partial}z \cdot \hat{\varepsilon}_y\right) \\ &= \hat{\partial}\left(\overset{x}{\partial}z \cdot \hat{\varepsilon}_x\right) + \hat{\partial}\left(\overset{y}{\partial}z \cdot \hat{\varepsilon}_y\right) \\ &= \hat{\partial}\overset{x}{\partial}z \cdot \hat{\varepsilon}_x + \hat{\partial}\overset{y}{\partial}z \cdot \hat{\varepsilon}_y \\ &= \left(\overset{x x}{\partial\partial}z \cdot \hat{\varepsilon}_x + \overset{y x}{\partial\partial}z \cdot \hat{\varepsilon}_y\right) \cdot \hat{\varepsilon}_x + \left(\overset{x y}{\partial\partial}z \cdot \hat{\varepsilon}_x + \overset{y y}{\partial\partial}z \cdot \hat{\varepsilon}_y\right) \cdot \hat{\varepsilon}_y \\ &= \overset{x x}{\partial\partial}z \cdot \hat{\varepsilon}_x^2 + 2 \cdot \overset{y x}{\partial\partial}z \cdot \hat{\varepsilon}_y \cdot \hat{\varepsilon}_x + \overset{y y}{\partial\partial}z \cdot \hat{\varepsilon}_y^2 \end{aligned}$$

We get our formula for the Hessian by dividing by two.

$$\text{hess } z = \frac{\overset{x x}{\partial\partial}z}{2} \cdot \hat{\varepsilon}_x^2 + \overset{y x}{\partial\partial}z \cdot \hat{\varepsilon}_x \cdot \hat{\varepsilon}_y + \frac{\overset{y y}{\partial\partial}z}{2} \cdot \hat{\varepsilon}_y^2$$

The *saddle discriminant* of a function $z \leftarrow (x, y)$ is the number

$\text{disc } z : \text{Number}$

$$\text{disc } z = (\text{warp } z)^2 - 4 \cdot \text{bend } z^x \cdot \text{bend } z^y.$$

The discriminant helps us describe the shape of the graph.

Theorem.

Suppose we localize the saddle discriminant at a point $p : \text{Point}_{(x,y)}$.

- If the discriminant is positive, $\text{disc } z > 0$, then the graph of z is *saddle-shaped* near p .
- If the discriminant is negative, $\text{disc } z < 0$, then the graph of z is *bowl-shaped* near p .

We'll say that the discriminant is *inconclusive at p* if it is equal to zero.

Why?

We learn how the graph peels away from its tangent plane by looking at the Hessian

$$\text{hess } z = a \cdot \hat{\epsilon}_x^2 + b \cdot \hat{\epsilon}_x \cdot \hat{\epsilon}_y + c \cdot \hat{\epsilon}_y^2$$

where

$$a = \text{bend } z^x, \quad b = \text{warp } z, \quad c = \text{bend } z^y.$$

By using a bit of algebraic dark magic—completing the square—we can rewrite the Hessian polymetric as

$$\text{hess } z = a \cdot \left(\left(\hat{\epsilon}_x + \frac{b}{2a} \cdot \hat{\epsilon}_y \right)^2 - \frac{b^2 - 4ac}{(2a)^2} \cdot \hat{\epsilon}_y^2 \right).$$

The discriminant

$$\text{disc } z = b^2 - 4ac$$

determines whether our polymetric factors, and thereby determines the shape of the graph. You might recognize the discriminant: it also shows up in the quadratic formula!

Let's take a look at saddle discriminants for some familiar surfaces.

shape	function	discriminant
bowl	$z = x^2 + y^2$	$\text{disc } z = -4$
parabolic saddle	$z = x^2 - y^2$	$\text{disc } z = 4$
warp saddle	$z = xy$	$\text{disc } z = 1$
thick parabola	$z = x^2$	$\text{disc } z = 0$
plane	$z - z_p = s_x \cdot (x - x_p) + s_y \cdot (y - y_p)$	$\text{disc } z = 0$

Each of these functions has a constant saddle discriminant. This will actually be uncommon! For most functions, the saddle discriminant will vary from point to point, depending on whether the graph is saddle-shaped or bowl-shaped. If the discriminant is zero at a point, the graph may be saddle-shaped, bowl-shaped, or some other shape. In this case, we might look at a contour map to better understand the graph.

9.6 Low Points and High Points

We're now ready to look for low points and high points on the graph of a function $z \leftarrow (x, y)$. We'll say that a point $p : \text{Point}_{(x,y)}$ is a *level point* for the function if both partials are zero at p .

$$\frac{\partial z}{\partial x} = 0 \quad \frac{\partial z}{\partial y} = 0$$

Level points are points that have a horizontal tangent plane. Finding level points is the first step to finding low points and high points. Once we've found a level point, we'll need some criteria to decide whether it is a low point, a high point, or something else.

Important.

Suppose $p : \text{Point}_{(x,y)}$ is a level point. We may *classify* p by computing the [saddle discriminant](#).

The level point is a *low point* if the graph is bowl-shaped,

$$\text{disc } z < 0,$$

and the bends are positive,

$$\text{bend } z > 0 \quad \text{and} \quad \text{bend } z > 0.$$



low point

The level point is a *high point* if the graph is bowl-shaped

$$\text{disc } z < 0,$$

and the bends are negative

$$\text{bend } z < 0 \quad \text{and} \quad \text{bend } z < 0.$$



high point

A level point with a saddle-shaped graph, $\text{disc } z > 0$, is neither a low point nor a high point.

Let's see a couple examples!

Example.

Consider the function

$$z \leftarrow (x, y)$$

$$z = x^3 - y^2 - 3x.$$

Let's use our tools to look for low points and high points! We'll start by finding slopes

$$\overset{x}{\text{slope } z} = 3x^2 - 3$$

$$\overset{y}{\text{slope } z} = -2y$$

so that we can find level points. Setting the x -slope equal to zero, we find

$$3x^2 - 3 = 0$$

$$3x^2 = 3$$

$$x^2 = 1$$

$$x = 1 \quad \text{or} \quad x = -1.$$

Setting the y -slope equal to zero gives

$$-2y = 0$$

$$y = 0.$$

We conclude that there are only two level points on our graph, $(1, 0)$ and $(-1, 0)$. Let's classify these level points! To this end, we've computed the x -bend, the y -bend, the warp, and the saddle discriminant.

$$\overset{x}{\text{bend } z} = 3x \quad \overset{y}{\text{bend } z} = -1 \quad \text{warp } z = 0 \quad \text{disc } z = 12x$$

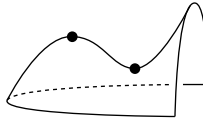
At $(1, 0)$.

At the point $(1, 0)$, the discriminant is positive, $\text{disc } z = 12$, and so we've found a saddle. This level point is neither a low point nor a high point.

At $(-1, 0)$.

At the point $(-1, 0)$, the discriminant is negative, $\text{disc } z = -12$, so our graph is bowl-shaped here. Both bends are negative, so our level point $(-1, 0)$ must be a high point!

In the previous example we found a saddle to the right of a high point. If we graph the function, we'll find the following surface.



If you'd like to get a better intuition for this surface, you can take some vertical slices through the level points.

Example.

Let's look for low points and high points on the graph of the function

$$z \leftarrow (x, y)$$

$$z = x^2 - 3xy + y^2 + 3.$$

We'll start by calculating slopes

$$\text{slope } z = 2x - 3y \quad \text{slope } z = 2y - 3x$$

so that we can look for level points.

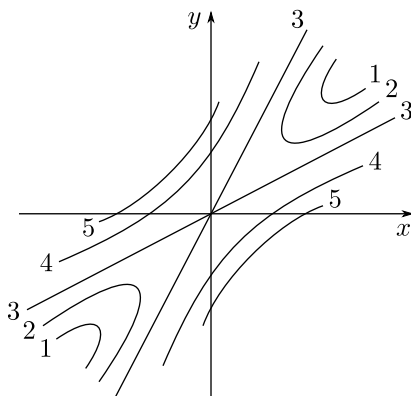
$$2x - 3y = 0 \quad \text{and} \quad 2y - 3x = 0$$

The only way to satisfy both equations is if both x and y are zero. So the origin $(0, 0)$ is the only level point for our function. Let's compute bends, warp, and the discriminant to classify the graph's shape.

$$\text{bend } z = 1 \quad \text{bend } z = 1 \quad \text{warp } z = -3 \quad \text{disc } z = 5$$

A positive discriminant tells us that we've found a saddle. Our level point $(0, 0)$ is neither a low point nor a high point.

Looking at a contour map for the previous example, we see the saddle at the origin.



Notice, however, that we found both a positive x -bend and a positive y -bend. If we had just looked at bend, we might have been fooled into thinking that our level point is a low point! It is important to calculate the saddle discriminant: both bend and warp play a role in determining whether a graph is saddle-shaped.

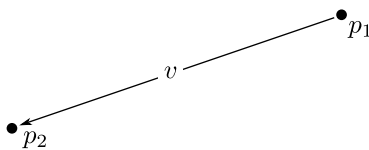
Part III

Rulers

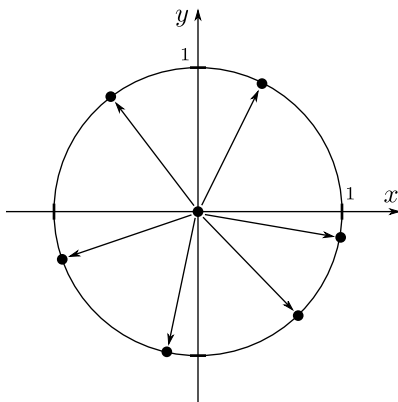
Chapter 10

Vectors

Vectors are used to represent a change in location: given any two points p_1 and p_2 , we can find a vector v that starts at p_1 and ends at p_2 .



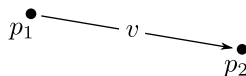
We can think of vectors as having both a length and a direction. The length, or *magnitude*, of a vector is the distance between the endpoints. By looking at vectors with magnitude one, we find vectors that fit as rays on the unit circle.



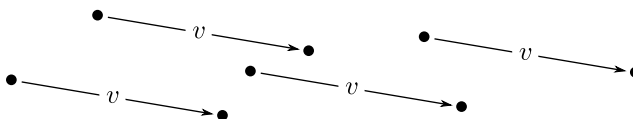
Each such *unit vector* represents a unique direction. Our goal for this chapter is to show that we can produce any vector v by scaling a unit vector up or down to the desired length.

10.1 Vectors

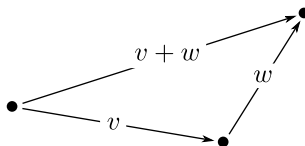
A *vector* is an object that records a change in location. We'll draw vectors as arrows. Here is a vector that travels from a starting point p_1 to an ending point p_2 .



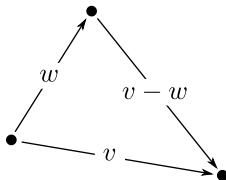
Oddly enough, we do not expect a vector to remember where it is located. We may move a vector for free, as long as we do not change its length or direction. All of the following arrows are portraits of that same vector v .



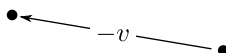
We may perform arithmetic with vectors which looks a bit like the arithmetic of numbers. We add two vectors, v and w , by placing them head-to-tail and drawing a new arrow $v + w$ from the start to the end.



By starting v and w at the same point, the difference, $v - w$, travels from the end of w to the end of v .



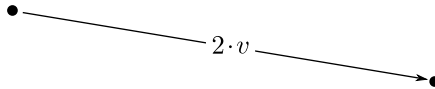
We may reverse the direction of a vector by negating.



If we add a vector v with itself, we get a vector

$$2 \cdot v = v + v$$

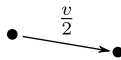
that points in the same direction as v and is twice as long.



Dividing a vector by two

$$\frac{v}{2} = \frac{1}{2} \cdot v$$

gives a vector that is half as long as v .



As you encounter expressions built from numbers and vectors, it is good to be in the habit of deducing what is a number and what is a vector.

- The sum of vectors is a vector.
- The difference of vectors is a vector.
- The product of a vector and a number is a vector.
- The quotient of a vector by a number is a vector.

Each of these operations produces a vector. In this chapter and the next, we'll see two ways to turn vectors into numbers.

- The magnitude of a vector is a number.
- The measurement of a vector by a ruler is a number.

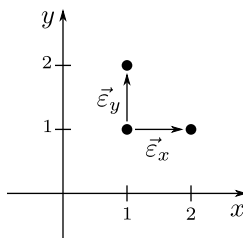
Last, you should know that vectors and division don't always mix. While it's fine to use a vector as the numerator of a fraction, we won't ever allow a vector as the denominator!

10.2 Standard Vectors

We have two *standard vectors* in the (x, y) -coordinate plane:

- the *standard x -vector*, $\vec{\varepsilon}_x : \text{Vector}_{(x,y)}$, travels one East, and
- the *standard y -vector*, $\vec{\varepsilon}_y : \text{Vector}_{(x,y)}$, travels one North.

As before, we'll use **cardinal directions** to discuss movement in the (x, y) -coordinate plane.



Standard vectors give us a uniform way to work with vectors. We may write any vector as a combination of the standard vectors.

Definition.

Any *vector* $v : \text{Vector}_{(x,y)}$ can be written in *standard form* as

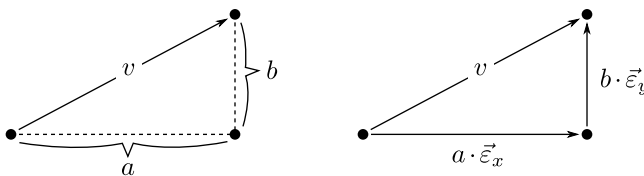
$$v = a \cdot \vec{\varepsilon}_x + b \cdot \vec{\varepsilon}_y$$

where the numbers

$a : \text{Number}$ and $b : \text{Number}$

are called the *parts* of v .

Let's think through how standard form works, geometrically. Suppose $v : \text{Vector}_{(x,y)}$ is a vector that travels some amount $a : \text{Number}$ to the East and some amount $b : \text{Number}$ to the North.



The vector $a \cdot \vec{\varepsilon}_x$ travels a to the East, and the vector $b \cdot \vec{\varepsilon}_y$ travels b to the North. Adding these together gives our vector v .

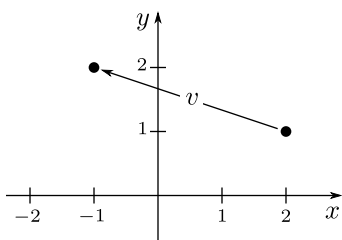
This same process works for the other cardinal directions, too! The vectors

$$-\vec{\varepsilon}_x \quad -\vec{\varepsilon}_y$$

travel one West and one South, respectively. So we'll use a negative x -part a to represent traveling West, and we'll use a negative y -part b to represent traveling South.

Example.

Consider the vector $v : \text{Vector}_{(x,y)}$ as pictured.



This vector travels three to the West and one North.

$$v = -3 \cdot \vec{\varepsilon}_x + \vec{\varepsilon}_y$$

Our vector has x -part negative three and y -part one.

The Zero Vector

The *zero vector*

$$0 : \text{Vector}_{(x,y)}$$

$$0 = 0 \cdot \vec{\varepsilon}_x + 0 \cdot \vec{\varepsilon}_y$$

represents a lack of movement. As such, the zero vector is difficult to draw as an arrow: it has no length nor direction. But it is a useful object for vector arithmetic. Subtracting a vector from itself gives the zero vector.

$$v - v = 0$$

And adding the zero vector to any vector has no effect.

$$v + 0 = v$$

We use the same symbol for the number zero and the vector zero,

$$0 : \text{Number} \quad \text{and} \quad 0 : \text{Vector},$$

but it's best to think of these as distinct objects.

10.3 Position Vectors

Given any point

$$p : \text{Point}_{(x,y)}$$

$$p = (x_p, y_p),$$

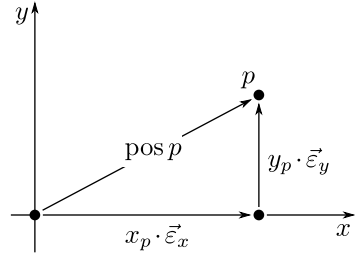
the *position vector for p* is the vector

$$\text{pos } p : \text{Vector}_{(x,y)}$$

$$\text{pos } p = x_p \cdot \vec{\epsilon}_x + y_p \cdot \vec{\epsilon}_y$$

that travels from the origin to p .

Position vectors give us a way to convert any point to a vector. By changing types, we get access to vector arithmetic.



Formula.

The vector $v : \text{Vector}_{(x,y)}$ that starts at a point $p_1 : \text{Point}_{(x,y)}$ and ends at a point $p_2 : \text{Point}_{(x,y)}$ can be calculated as the difference of position vectors.

$$v = \text{pos } p_2 - \text{pos } p_1$$

Whenever we compute a change, we'll use this pattern:

to – from.

Let's see how this works in an example.

Example.

Let's find the vector v that starts at the point $p_1 = (1, 3)$ and ends at the point $p_2 = (5, 2)$. We'll start by converting our points to position vectors.

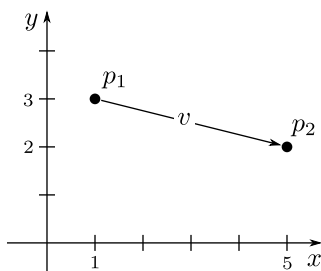
$$\text{pos } p_1 = \vec{\epsilon}_x + 3 \cdot \vec{\epsilon}_y$$

$$\text{pos } p_2 = 5 \cdot \vec{\epsilon}_x + 2 \cdot \vec{\epsilon}_y$$

And now let's subtract to find the vector v .

$$\begin{aligned} v &= \text{pos } p_2 - \text{pos } p_1 \\ &= (5 \cdot \vec{\epsilon}_x + 2 \cdot \vec{\epsilon}_y) - (\vec{\epsilon}_x + 3 \cdot \vec{\epsilon}_y) \\ &= 4 \cdot \vec{\epsilon}_x - \vec{\epsilon}_y \end{aligned}$$

To move from p_1 to p_2 , we travel four East and one South. We can check our work by drawing our vector v .



Sometimes it can be helpful to rearrange the above formula.

Example.

Suppose we locate the vector $v = \vec{\varepsilon}_x - 3 \cdot \vec{\varepsilon}_y$ so that it starts at the point $p_1 = (1, 2)$. At what point p_2 does the vector end?

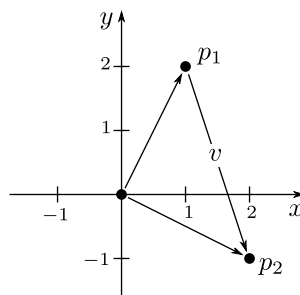
To answer this question, let's find the position vector for p_2 .

$$\begin{aligned} \text{pos } p_2 &= \text{pos } p_1 + v \\ &= (\vec{\varepsilon}_x + 2 \cdot \vec{\varepsilon}_y) + (\vec{\varepsilon}_x - 3 \cdot \vec{\varepsilon}_y) \\ &= 2 \cdot \vec{\varepsilon}_x - \vec{\varepsilon}_y \end{aligned}$$

What point p_2 has this position vector? Easy.

$$p_2 = (2, -1)$$

To sum up, if we move one East and three South from $p_1 = (1, 2)$, we'll end up at $p_2 = (2, -1)$.



10.4 Magnitude

The *magnitude* of a vector $v : \text{Vector}_{(x,y)}$ is the number

$$\text{mag } v : \text{Number}$$

$$\text{mag } v = \sqrt[2]{a^2 + b^2}$$

when we've written

$$v = a \cdot \vec{\varepsilon}_x + b \cdot \vec{\varepsilon}_y$$

in standard form.

Example.

Let's find the magnitude of the vector

$$v = 2 \cdot \vec{\varepsilon}_x - \vec{\varepsilon}_y.$$

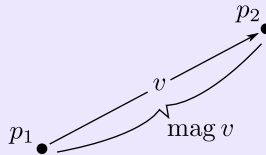
Our vector has x -part two and y -part negative one. We calculate.

$$\text{mag } v = \sqrt[2]{2^2 + (-1)^2} = \sqrt[2]{4 + 1} = \sqrt[2]{5}$$

The magnitude of a vector is a distance. We'll sometimes also refer to a vector's magnitude as its length.

Important.

Suppose a vector $v : \text{Vector}_{(x,y)}$ travels from a point $p_1 : \text{Point}_{(x,y)}$ to a point $p_2 : \text{Point}_{(x,y)}$.



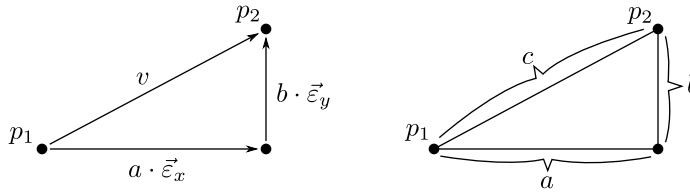
Then the magnitude, $\text{mag } v$, is equal to the distance between p_1 and p_2 .

Why?

By writing our vector in standard form

$$v = a \cdot \vec{\varepsilon}_x + b \cdot \vec{\varepsilon}_y,$$

we can picture its parts as the legs of right triangle.



Let c : Number be the distance between p_1 and p_2 so that we can apply the Pythagorean Theorem.

$$a^2 + b^2 = c^2$$

Taking square roots,

$$c = \sqrt{c^2} = \sqrt{a^2 + b^2} = \text{mag } v$$

we see that c is equal to the magnitude of v .

We can calculate the distance between two points by finding the magnitude of a vector between them.

Example.

Let's find the distance between the points

$$p_1 = (-1, 1) \quad \text{and} \quad p_2 = (4, -2).$$

We'll start by finding the vector from p_1 to p_2 .

$$\begin{aligned} v &= \text{pos } p_2 - \text{pos } p_1 \\ &= (4 \cdot \vec{\epsilon}_x - 2 \cdot \vec{\epsilon}_y) - (-\vec{\epsilon}_x + \vec{\epsilon}_y) \\ &= 5 \cdot \vec{\epsilon}_x - 3 \cdot \vec{\epsilon}_y \end{aligned}$$

Now that we have our vector v , let's find its magnitude.

$$\text{mag } v = \sqrt{5^2 + (-3)^2} = \sqrt{25 + 9} = \sqrt{34}$$

The distance between p_1 and p_2 is $\sqrt{34}$.

Because we've defined magnitude as a square root, magnitudes are always positive.

$$\text{mag } v \geq 0$$

In fact, the only vector to have magnitude zero is the zero vector.

$$\text{mag } 0 = \sqrt{0^2 + 0^2} = 0$$

Every other vector actually travels some distance.

10.5 Scaling

Scaling a vector allows us to change the vector's length without changing its direction. We've seen, for example, that doubling a vector v produces a vector

$$2 \cdot v$$

that is twice as long. Let's look at a theorem that makes this idea precise by comparing magnitudes.

Theorem.

Suppose $v : \text{Vector}$ is a vector. Then

$$\text{mag}(n \cdot v) = n \cdot \text{mag } v$$

where $n : \text{Number}$ is any positive number, $n \geq 0$.

Why?

Writing v in standard form

$$v = a \cdot \vec{\epsilon}_x + b \cdot \vec{\epsilon}_y$$

allows us to calculate:

$$\begin{aligned} \text{mag}(n \cdot v) &= \text{mag}(na \cdot \vec{\epsilon}_x + nb \cdot \vec{\epsilon}_y) \\ &= \sqrt{(na)^2 + (nb)^2} \\ &= \sqrt{n^2 \cdot (a^2 + b^2)} \\ &= \sqrt{n^2} \cdot \sqrt{a^2 + b^2} \\ &= n \cdot \text{mag } v. \end{aligned}$$

A *unit vector* is a vector that has magnitude one. Just as magnitude captures the intuitive idea of length, unit vectors capture the intuitive idea of direction. Every vector is determined by a length and a direction.

Important.

We may write any vector $v : \text{Vector}$ as the product

$$v = \text{mag } v \cdot v'$$

where $v' : \text{Vector}$ is a unit vector.

In fact, we can find the unit vector v' by scaling

$$v' = \frac{v}{\text{mag } v}$$

as long as v is not the zero vector. This formula for the unit vector v' almost looks like it breaks our rule, “no vectors as denominators.” Looking carefully at types, however, we see that we are dividing v , a vector, by $\text{mag } v$, a number.

Example.

Consider the vector

$$v : \text{Vector}_{(x,y)}$$

$$v = \vec{e}_x + 2 \cdot \vec{e}_y.$$

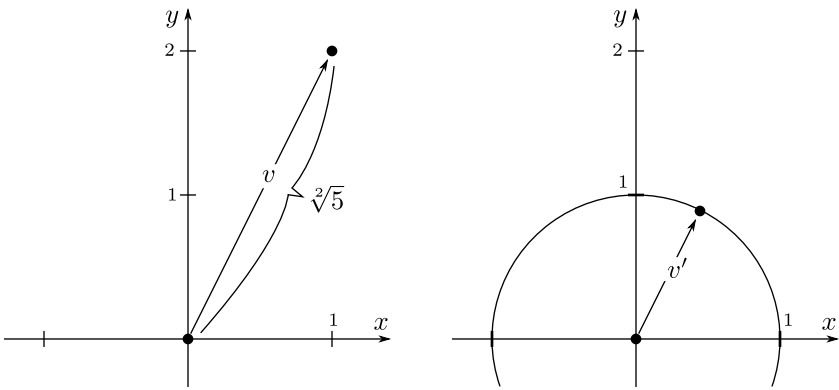
Let’s find the magnitude of v .

$$\text{mag } v = \sqrt{1^2 + 2^2} = \sqrt{5}$$

We’ll scale v to produce a unit vector v' .

$$v' = \frac{v}{\text{mag } v} = \frac{\vec{e}_x + 2 \cdot \vec{e}_y}{\sqrt{5}} = \frac{1}{\sqrt{5}} \cdot \vec{e}_x + \frac{2}{\sqrt{5}} \cdot \vec{e}_y$$

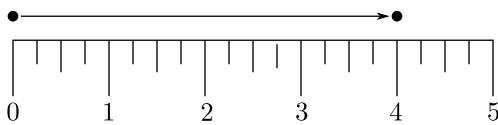
Since v' is a unit vector, it fits as a ray on the unit circle.



Chapter 11

Rulers

All too often rulers get overlooked when people study Calculus. It takes quite a bit of time to draw a ruler. And to be perfectly honest, it takes significant effort to develop an intuition for rulers. But don't let this discourage you!



The good news is that it is no more difficult to make calculations with rulers than with vectors. In fact, the algebra of rulers and vectors are essentially identical. They share the same definitions, laws, theorems, etc. Or perhaps it is better to say that rulers and vectors are dual: they are complimentary types that together give a theory of measurement. Despite their similarities, rulers will be much more useful to us as we move forward. There will be significant benefits to learning to calculate with rulers!

11.1 Standard Rulers

Before we give a general definition of rulers, let's spend a little time working with the standard rulers. Our goal for this section is to learn to make measurements using standard rulers. We'll see how measurement extends to all rulers in the next section.

We have two *standard rulers* for the (x, y) -coordinate plane: a *standard x -ruler* and a *standard y -ruler*.

$$\bar{\epsilon}_x : \text{Ruler}_{(x,y)} \quad \text{and} \quad \bar{\epsilon}_y : \text{Ruler}_{(x,y)}$$

The *measurement product* of a ruler $r : \text{Ruler}$ and a vector $v : \text{Vector}$ is a number

$$\langle r \mid v \rangle : \text{Number}.$$

We write our ruler on the left side of a measurement and our vector on the right side.

Laws.

We enforce vector measurement laws

$$\begin{aligned} \langle r \mid v_1 + v_2 \rangle &= \langle r \mid v_1 \rangle + \langle r \mid v_2 \rangle \\ \langle r \mid n \cdot v \rangle &= n \cdot \langle r \mid v \rangle \end{aligned}$$

where $r : \text{Ruler}$ is a ruler, $n : \text{Number}$ is a number, and

$$v_1 : \text{Vector}, \quad v_2 : \text{Vector}, \quad v : \text{Vector}$$

are vectors.

We also have standard measurements.

$$\langle \bar{\epsilon}_x \mid \vec{\epsilon}_x \rangle = 1 \quad \langle \bar{\epsilon}_x \mid \vec{\epsilon}_y \rangle = 0 \quad \langle \bar{\epsilon}_y \mid \vec{\epsilon}_x \rangle = 0 \quad \langle \bar{\epsilon}_y \mid \vec{\epsilon}_y \rangle = 1$$

That's enough to us get started measuring with standard rulers!

Example.

Let's measure the vector

$$v : \text{Vector}_{(x,y)}$$

$$v = 2 \cdot \vec{\varepsilon}_x - 3 \cdot \vec{\varepsilon}_y$$

with both of the standard rulers. We'll start by measuring with the standard x -ruler.

$$\begin{aligned} \langle \bar{\varepsilon}_x \mid v \rangle &= \langle \bar{\varepsilon}_x \mid 2 \cdot \vec{\varepsilon}_x - 3 \cdot \vec{\varepsilon}_y \rangle \\ &= \langle \bar{\varepsilon}_x \mid 2 \cdot \vec{\varepsilon}_x \rangle - \langle \bar{\varepsilon}_x \mid 3 \cdot \vec{\varepsilon}_y \rangle \\ &= 2 \cdot \langle \bar{\varepsilon}_x \mid \vec{\varepsilon}_x \rangle - 3 \cdot \langle \bar{\varepsilon}_x \mid \vec{\varepsilon}_y \rangle \\ &= 2 \cdot 1 - 3 \cdot 0 \\ &= 2 \end{aligned}$$

And now let's measure using the standard y -ruler.

$$\begin{aligned} \langle \bar{\varepsilon}_y \mid v \rangle &= \langle \bar{\varepsilon}_y \mid 2 \cdot \vec{\varepsilon}_x - 3 \cdot \vec{\varepsilon}_y \rangle \\ &= \langle \bar{\varepsilon}_y \mid 2 \cdot \vec{\varepsilon}_x \rangle - \langle \bar{\varepsilon}_y \mid 3 \cdot \vec{\varepsilon}_y \rangle \\ &= 2 \cdot \langle \bar{\varepsilon}_y \mid \vec{\varepsilon}_x \rangle - 3 \cdot \langle \bar{\varepsilon}_y \mid \vec{\varepsilon}_y \rangle \\ &= 2 \cdot 0 - 3 \cdot 1 \\ &= -3 \end{aligned}$$

In the previous example, the measurement $\langle \bar{\varepsilon}_x \mid v \rangle$ is equal to the vector's x -part, and the measurement $\langle \bar{\varepsilon}_y \mid v \rangle$ is equal to the vector's y -part. The same thing will happen for other vectors, too!

Formulas.

We can solve for the parts of a vector $v : \text{Vector}_{(x,y)}$ by measuring the vector with standard rulers. If

$$v = a \cdot \vec{\varepsilon}_x + b \cdot \vec{\varepsilon}_y$$

then

$$\langle \bar{\varepsilon}_x \mid v \rangle = a \quad \text{and} \quad \langle \bar{\varepsilon}_y \mid v \rangle = b.$$

Why?

Let's check these formulas by measuring.

$$\begin{aligned}\langle \bar{\varepsilon}_x | v \rangle &= \langle \bar{\varepsilon}_x | a \cdot \vec{\varepsilon}_x + b \cdot \vec{\varepsilon}_y \rangle \\ &= \langle \bar{\varepsilon}_x | a \cdot \vec{\varepsilon}_x \rangle + \langle \bar{\varepsilon}_x | b \cdot \vec{\varepsilon}_y \rangle \\ &= a \cdot \langle \bar{\varepsilon}_x | \vec{\varepsilon}_x \rangle + b \cdot \langle \bar{\varepsilon}_x | \vec{\varepsilon}_y \rangle \\ &= a \cdot 1 + b \cdot 0 \\ &= a\end{aligned}$$

$$\begin{aligned}\langle \bar{\varepsilon}_y | v \rangle &= \langle \bar{\varepsilon}_y | a \cdot \vec{\varepsilon}_x + b \cdot \vec{\varepsilon}_y \rangle \\ &= \langle \bar{\varepsilon}_y | a \cdot \vec{\varepsilon}_x \rangle + \langle \bar{\varepsilon}_y | b \cdot \vec{\varepsilon}_y \rangle \\ &= a \cdot \langle \bar{\varepsilon}_y | \vec{\varepsilon}_x \rangle + b \cdot \langle \bar{\varepsilon}_y | \vec{\varepsilon}_y \rangle \\ &= a \cdot 0 + b \cdot 1 \\ &= b\end{aligned}$$

11.2 Rulers

In the preceding section we learned to measure using standard rulers. We're now ready to look at rulers and measurement in full generality.

Definition.

A ruler for the (x, y) -coordinate plane is a combination of the standard rulers

$$r : \text{Ruler}_{(x,y)}$$

$$r = a \cdot \bar{\epsilon}_x + b \cdot \bar{\epsilon}_y$$

where the numbers

$$a : \text{Number} \quad \text{and} \quad b : \text{Number}$$

are called the *parts*.

Let's introduce ruler measurement laws that complement the vector measurement laws we've already seen.

Laws.

We enforce ruler measurement laws

$$\langle r_1 + r_2 \mid v \rangle = \langle r_1 \mid v \rangle + \langle r_2 \mid v \rangle$$

$$\langle n \cdot r \mid v \rangle = n \cdot \langle r \mid v \rangle$$

where $v : \text{Vector}$ is a vector, $n : \text{Number}$ is a number, and

$$r_1 : \text{Ruler}, \quad r_2 : \text{Ruler}, \quad r : \text{Ruler}$$

are rulers.

The measurement laws tell us we can measure by distributing! That said, we'll lose any terms where a standard ruler measures a non-matching standard vector.

Example.

Let's measure the vector

$$v : \text{Vector}_{(x,y)}$$

$$v = \vec{\varepsilon}_x - 2 \cdot \vec{\varepsilon}_y$$

with the ruler

$$r : \text{Ruler}_{(x,y)}$$

$$r = 3 \cdot \vec{\varepsilon}_x + 2 \cdot \vec{\varepsilon}_y.$$

We calculate!

$$\begin{aligned} \langle r | v \rangle &= \langle 3 \cdot \vec{\varepsilon}_x + 2 \cdot \vec{\varepsilon}_y | \vec{\varepsilon}_x - 2 \cdot \vec{\varepsilon}_y \rangle \\ &= \langle 3 \cdot \vec{\varepsilon}_x | \vec{\varepsilon}_x \rangle + \langle 3 \cdot \vec{\varepsilon}_x | -2 \cdot \vec{\varepsilon}_y \rangle + \langle 2 \cdot \vec{\varepsilon}_y | \vec{\varepsilon}_x \rangle + \langle 2 \cdot \vec{\varepsilon}_y | -2 \cdot \vec{\varepsilon}_y \rangle \\ &= 3 \cdot \langle \vec{\varepsilon}_x | \vec{\varepsilon}_x \rangle - 6 \cdot \langle \vec{\varepsilon}_x | \vec{\varepsilon}_y \rangle + 2 \cdot \langle \vec{\varepsilon}_y | \vec{\varepsilon}_x \rangle - 4 \cdot \langle \vec{\varepsilon}_y | \vec{\varepsilon}_y \rangle \\ &= 3 - 4 \\ &= -1 \end{aligned}$$

Let's have another example of measurement.

Example.

Consider the ruler and vector.

$$r : \text{Ruler}_{(x,y)}$$

$$r = 3 \cdot \vec{\varepsilon}_x - 6 \cdot \vec{\varepsilon}_y$$

and

$$v : \text{Vector}_{(x,y)}$$

$$v = 4 \cdot \vec{\varepsilon}_x + 2 \cdot \vec{\varepsilon}_y$$

Let's measure.

$$\begin{aligned} \langle r | v \rangle &= \langle 3 \cdot \vec{\varepsilon}_x - 6 \cdot \vec{\varepsilon}_y | 4 \cdot \vec{\varepsilon}_x + 2 \cdot \vec{\varepsilon}_y \rangle \\ &= 12 \cdot \langle \vec{\varepsilon}_x | \vec{\varepsilon}_x \rangle + 6 \cdot \langle \vec{\varepsilon}_x | \vec{\varepsilon}_y \rangle - 24 \cdot \langle \vec{\varepsilon}_y | \vec{\varepsilon}_x \rangle - 12 \cdot \langle \vec{\varepsilon}_y | \vec{\varepsilon}_y \rangle \\ &= 12 - 12 \\ &= 0 \end{aligned}$$

Curiously, the measurement is zero! We'll see how this can happen when we study the geometry of measurement.

The following formulas allow us to solve for a ruler's parts.

Formulas.

We can solve for the parts of a ruler $r : \text{Ruler}_{(x,y)}$ by measuring the standard vectors. If

$$r = a \cdot \bar{\varepsilon}_x + b \cdot \bar{\varepsilon}_y$$

then

$$\langle r \mid \vec{\varepsilon}_x \rangle = a \quad \text{and} \quad \langle r \mid \vec{\varepsilon}_y \rangle = b.$$

I recommend that you check these with pen and paper!

11.3 Markings

Up until now, all of our work with rulers has been algebraic. We have learned to measure vectors with rulers by calculating. To broaden our understanding of measurement, we'll need to be able to visualize a ruler. Although the algebra of rulers and vectors are essentially the same, the ways we picture rulers and vectors are quite different. We picture a ruler by plotting its markings.

Definition.

Consider a ruler, a point, and a number.

$$\begin{aligned} r : \text{Ruler}_{(x,y)} \quad p : \text{Point}_{(x,y)} \quad m : \text{Number} \\ r = a \cdot \bar{\epsilon}_x + b \cdot \bar{\epsilon}_y \quad p = (x_p, y_p) \end{aligned}$$

The m -marking of r is the line

$$a \cdot (x - x_p) + b \cdot (y - y_p) = m$$

when we've centered our ruler r at the point p .

Just as we are free to choose where to locate a vector, we may choose any point p to sit on our ruler's zero marking. We'll often choose to center a ruler at the origin, $p = (0, 0)$, to make our equations that much simpler.

Example.

Consider the ruler

$$\begin{aligned} r : \text{Ruler}_{(x,y)} \\ r = \bar{\epsilon}_x + \frac{1}{2} \cdot \bar{\epsilon}_y. \end{aligned}$$

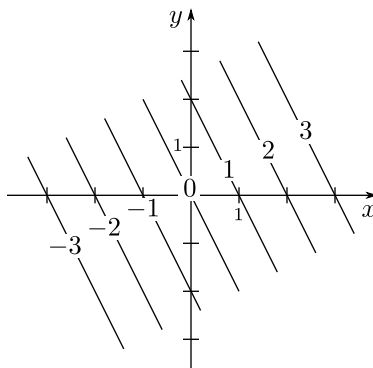
Let's draw markings for r where we've centered r at the origin, $p = (0, 0)$. The m -marking is the line

$$m = 1 \cdot (x - 0) + \frac{1}{2} \cdot (y - 0)$$

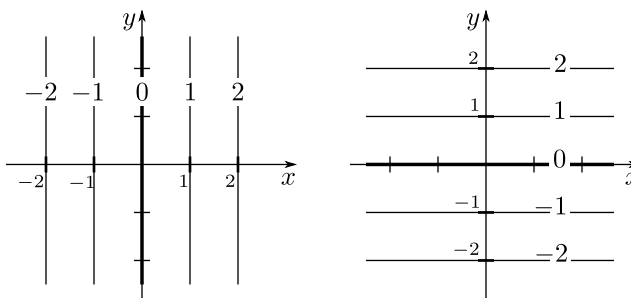
or more simply

$$m = x + \frac{y}{2}.$$

We get a picture of our ruler by plotting several m -markings.



We can think of the standard rulers, $\bar{\varepsilon}_x$ and $\bar{\varepsilon}_y$, as extending the markings we draw on the x and y -axes.



We've drawn the standard x -ruler on the left, and the standard y -ruler on the right. Just like these standard rulers, the markings on any ruler are evenly-spaced parallel lines.

11.4 Measurement

Now that we can draw a ruler's markings, we can visualize measurement.

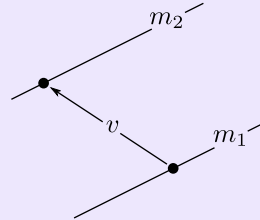
Theorem.

Consider a ruler and a vector.

$$r : \text{Ruler}_{(x,y)} \quad v : \text{Vector}_{(x,y)}$$

If the vector v travels from a marking m_1 to a marking m_2 , then the measurement is equal to the difference of the markings.

$$\langle r \mid v \rangle = m_2 - m_1$$



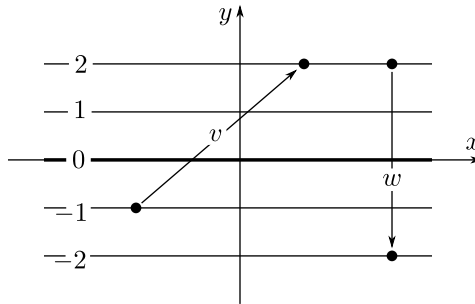
We do not require the vector v to be perpendicular to the markings of r . Let's get some intuition for measurement by looking at examples.

Example.

Consider the standard y -ruler, $\bar{\epsilon}_y$, and the vectors

$$v : \text{Vector}_{(x,y)} \quad \text{and} \quad w : \text{Vector}_{(x,y)}$$

pictured as follows.



We measure each vector as a difference of markings.

$$\langle \bar{\epsilon}_y \mid v \rangle = 2 - (-1) = 3$$

$$\langle \bar{\epsilon}_y \mid w \rangle = -2 - 2 = -4$$

The standard ruler $\bar{\epsilon}_y$ only sees change in height: it ignores how much a vector travels in the x -direction.

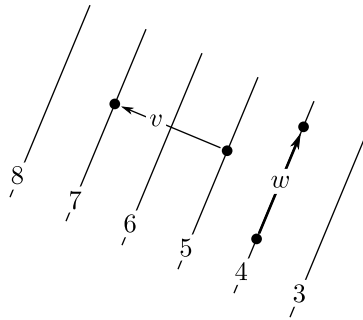
We get the most bang for our buck when measuring vectors that are perpendicular to a ruler's markings.

Example.

Consider the vectors

$$v : \text{Vector}_{(x,y)} \quad \text{and} \quad w : \text{Vector}_{(x,y)}$$

and the ruler $r : \text{Ruler}_{(x,y)}$ pictured as follows.



The vectors v and w have the same length, but their measurements are quite different.

$$\langle r \mid v \rangle = 7 - 5 = 2$$

$$\langle r \mid w \rangle = 4 - 4 = 0$$

The vector w is parallel to the ruler's markings. As a result, measuring w with r gives zero!

11.5 Magnitude

The *magnitude* of a ruler $r : \text{Ruler}_{(x,y)}$ is the number

$$\text{mag } r : \text{Number}$$

$$\text{mag } r = \sqrt[2]{a^2 + b^2}$$

when we've written

$$r = a \cdot \bar{\epsilon}_x + b \cdot \bar{\epsilon}_y$$

in standard form.

Example.

The ruler

$$r : \text{Ruler}_{(x,y)}$$

$$r = -3 \cdot \bar{\epsilon}_x + 4 \cdot \bar{\epsilon}_y$$

has magnitude

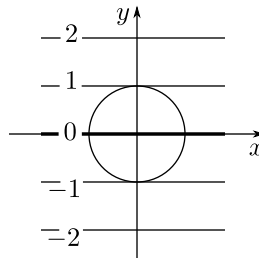
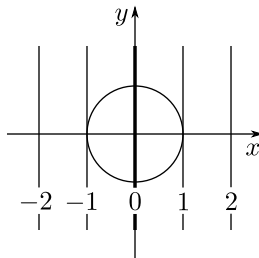
$$\text{mag } r = \sqrt[2]{(-3)^2 + 4^2} = \sqrt[2]{25} = 5.$$

We calculate the magnitude of a ruler just like we calculate the [magnitude of a vector](#). The magnitude of a vector is the vector's length. But unlike a vector, a ruler does not have a length! This raises a question.

Question.

How do we picture the magnitude of a ruler?

We can use the unit circle to help us visualize a ruler's magnitude! By definition, the unit circle has radius one: if we use a standard ruler to measure the radius, we'll get one as the result.



If we look at non-standard rulers, however, we may find other measurements!

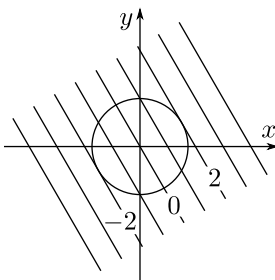
Theorem.

A ruler's magnitude is equal to the measurement the ruler gives for the radius of the unit circle.

Let's see an example.

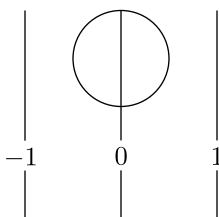
Example.

Consider the ruler $r : \text{Ruler}_{(x,y)}$ pictured with the unit circle.

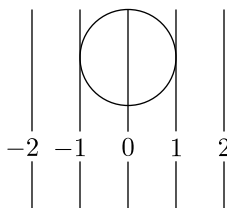


According to this ruler, the radius of the unit circle is two! So r has magnitude two.

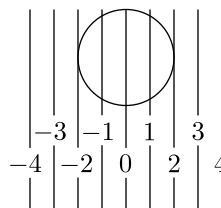
We'll call any ruler that has magnitude, one, a *unit ruler*. All other rulers either over-report or under-report, depending on their magnitude.



magnitude $\frac{1}{2}$



magnitude 1



magnitude 2

When we increase the magnitude of a ruler, the spacing between markings shrinks! This is quite different from what happens with vectors: a larger magnitude gives a longer vector.

11.6 Transpose

The transpose is an operator tran that converts rulers to vectors.

Definition.

The *transpose* of a ruler

$$r : \text{Ruler}_{(x,y)}$$

$$r = a \cdot \bar{\epsilon}_x + b \cdot \bar{\epsilon}_y$$

is the vector

$$\text{tran } r : \text{Vector}_{(x,y)}$$

$$\text{tran } r = a \cdot \vec{\epsilon}_x + b \cdot \vec{\epsilon}_y.$$

The transpose gives a one-to-one correspondence between rulers and vectors.

Example.

The transpose of the ruler

$$r : \text{Ruler}_{(x,y)}$$

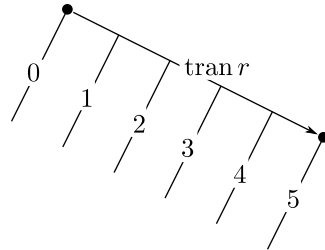
$$r = 2 \cdot \bar{\epsilon}_x - \bar{\epsilon}_y$$

is the vector

$$\text{tran } r : \text{Vector}_{(x,y)}$$

$$\text{tran } r = 2 \cdot \vec{\epsilon}_x - \vec{\epsilon}_y.$$

We just swap out each standard ruler for the corresponding standard vector.



Take a moment to look at the picture of the ruler and its transpose vector from the previous example. Notice how the transpose is perpendicular to the markings of the ruler. If we think of a ruler as having both a straight-edge and markings, the transpose vector gives us that straight-edge!

Theorem.

Suppose $r : \text{Ruler}_{(x,y)}$ is a ruler.

- The markings of r are perpendicular to the transpose vector, $\text{tran } r$.
- The transpose, $\text{tran } r$, points in the direction of increasing markings.

We've verified this theorem and two others in [Appendix F. Ruler Explanations](#).

Important.

The magnitude of a ruler is equal to the magnitude of its transpose.

Why?

Suppose

$$r : \text{Ruler}_{(x,y)}$$

$$r = a \cdot \bar{\epsilon}_x + b \cdot \bar{\epsilon}_y$$

is a ruler. Then we may compute.

$$\text{mag } r = \text{mag}(a \cdot \bar{\epsilon}_x + b \cdot \bar{\epsilon}_y) = \sqrt[2]{a^2 + b^2}$$

$$\text{mag } \text{tran } r = \text{mag}(a \cdot \bar{\epsilon}_x + b \cdot \bar{\epsilon}_y) = \sqrt[2]{a^2 + b^2}$$

Surprisingly, this fact—that a ruler and its transpose have the same magnitude—is not an easy thing to picture! The ways we visualize the magnitude of a ruler and the magnitude of a vector are entirely different. But there is something that is easy to see when picturing a ruler and its transpose: we can use the ruler to measure its transpose vector!

Theorem.

By using a ruler to measure its transpose vector,

$$\langle r \mid \text{tran } r \rangle = (\text{mag } r)^2$$

we get the magnitude of the ruler, squared.

Why?

Writing our ruler in standard form

$$r = a \cdot \bar{\varepsilon}_x + b \cdot \bar{\varepsilon}_y$$

allows us to compute

$$\langle r \mid \text{tran } r \rangle = \langle a \cdot \bar{\varepsilon}_x + b \cdot \bar{\varepsilon}_y \mid a \cdot \vec{\varepsilon}_x + b \cdot \vec{\varepsilon}_y \rangle = a^2 + b^2 = (\text{mag } r)^2.$$

Even after checking this theorem, we still might find it surprising! Why should the measurement be the square of the magnitude?

Well, suppose the ruler r has magnitude m . Then the transpose vector must also have magnitude m .

$$\text{mag } r = m \quad \text{mag } \text{tran } r = m$$

So while the length of our vector is m , when we measure it with our ruler r , the length will be over-reported by a factor of m .

$$\langle r \mid \text{tran } r \rangle = m \cdot m = (\text{mag } r)^2$$

In other words, both the ruler and the vector contribute a factor of m to the measurement!

Chapter 12

Differentials as Rulers

When we studied metrics we learned to interpret the differential as a metric. We're now ready to give an alternative interpretation of the differential: as a ruler.

$$\bar{\partial}(\quad)$$

At first glance this should not actually make that much of a difference. Both metrics and rulers are tools that let us assess change. But rulers will allow us to make more sophisticated measurements. Using vectors and rulers together, we will learn how to definitively tell how steep a surface is at any point.

Moving beyond slope, polymetrics and polyrulers look quite different. Polymetrics allowed us to investigate how a surface bends and warps away from its tangent plane. Polyrulers have an entirely different purpose: they allow us to measure length and area. That said, we won't be measuring any length or area until [Chapter 13. Boxes](#) when begin our study of integrals. For now, our focus will be on taking the differential of a polyruler.

12.1 The Ruler Differential

Let's interpret the differential as a ruler!

Definition.

For any function $z \leftarrow (x, y)$, the *ruler differential* of z is a ruler

$$\bar{\partial}z : \text{Ruler}_{(x,y)}.$$

The ruler differential, $\bar{\partial}$, satisfies all [abstract differential laws](#). We interpret the ruler differential of an input variable as a standard ruler:

$$\bar{\partial}x = \bar{\varepsilon}_x \quad \text{and} \quad \bar{\partial}y = \bar{\varepsilon}_y.$$

We calculate the ruler differential in very much the same way that we calculate the [metric differential](#).

Example.

Let's compute the differential of the function

$$\begin{aligned} z &\leftarrow (x, y) \\ z &= \exp(xy) \end{aligned}$$

as a ruler. We find

$$\begin{aligned} \bar{\partial}z &= \bar{\partial}(\exp(xy)) \\ &= \exp(xy) \cdot \bar{\partial}(xy) \\ &= \exp(xy) \cdot (y \cdot \bar{\partial}x + x \cdot \bar{\partial}y) \\ &= \exp(xy) \cdot (y \cdot \bar{\varepsilon}_x + x \cdot \bar{\varepsilon}_y) \\ &= y \cdot \exp(xy) \cdot \bar{\varepsilon}_x + x \cdot \exp(xy) \cdot \bar{\varepsilon}_y. \end{aligned}$$

This calculation should be familiar. A similar calculation will show that the metric differential of z is

$$\hat{\partial}z = y \cdot \exp(xy) \cdot \hat{\varepsilon}_x + x \cdot \exp(xy) \cdot \hat{\varepsilon}_y.$$

We can localize the ruler differential at any point $p : \text{Point}_{(x,y)}$ in the coordinate plane.

Example.

The bowl

$$z \leftarrow (x, y)$$

$$z = x^2 + y^2$$

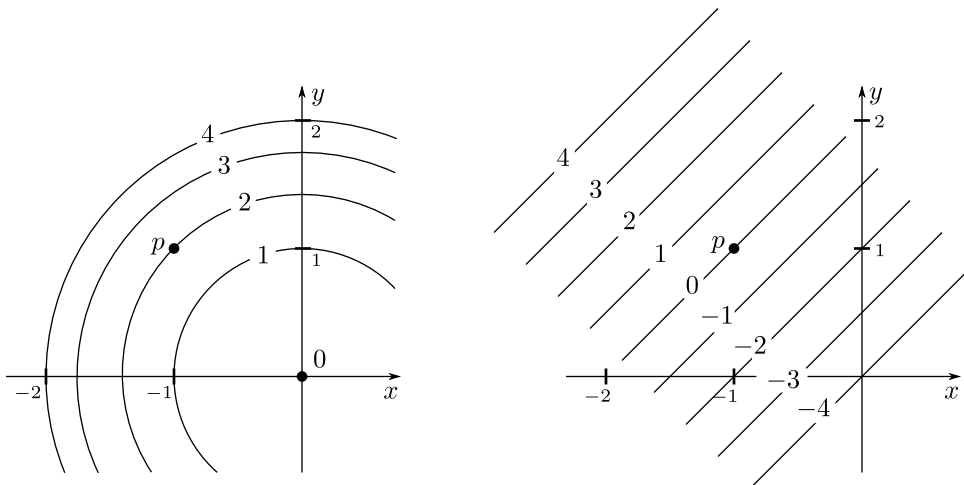
has differential

$$\bar{\partial}z = 2x \cdot \bar{\varepsilon}_x + 2y \cdot \bar{\varepsilon}_y.$$

Localizing at the point $p = (-1, 1)$, we find the ruler

$$\bar{\partial}z = -2 \cdot \bar{\varepsilon}_x + 2 \cdot \bar{\varepsilon}_y.$$

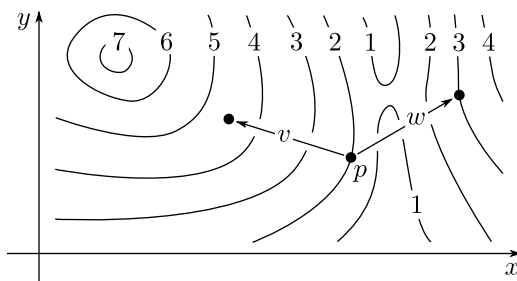
Let's plot the ruler differential from the previous example. We see some similarity when comparing this ruler to a contour map for the bowl. Near p , the markings of the differential resemble the contours of the bowl.



The differential is a best fit ruler!

12.2 Directional Slope

Before we write down a definition of directional slope, let's get some intuition by looking at a contour map. To visualize a directional slope, we'll need both a point and a vector.



We've drawn in two vectors v and w at the point p . Let's think about what the slope looks like as we leave p in traveling in either of these directions.

The direction given by v .

Imagine standing at the point p , ready to embark on a hike in the direction given by the vector v . Our height is two, and our height will increase as we begin to hike. Traveling in this direction, we'll be walking uphill. We should expect the directional slope to be positive.

The direction given by w .

Now let's imagine that we instead leave p going in the direction given by the vector w . Although we'll reach larger z -values if we hike long enough, our hike begins by heading downhill. Directional slopes will only care about what happens near p , so we should expect this directional slope to be negative.

Definition.

Suppose $z \leftarrow (x, y)$ is a function and $v : \text{Vector}_{(x,y)}$ is a non-zero vector, $v \neq 0$. The *slope in the direction given by v* is the number

$$\begin{aligned} \text{direc } z &: \text{Number} \\ \text{direc } z &= \frac{\langle \bar{\partial} z \mid v \rangle}{\text{mag } v}. \end{aligned}$$

If v happens to be a **unit vector**, then the formula for the directional slope becomes that much simpler:

$$\text{direc } z = \langle \bar{\partial} z \mid v \rangle.$$

We just measure v with the differential.

Example.

Let's look at some directional slopes for the **parabolic saddle**,

$$\begin{aligned} z &\leftarrow (x, y) \\ z &= x^2 - y^2, \end{aligned}$$

at the point $p = (1, 1)$. We can calculate the differential as

$$\bar{\partial} z = 2x \cdot \bar{\epsilon}_x - 2y \cdot \bar{\epsilon}_y$$

and by localizing at our point, we find the ruler

$$\bar{\partial} z = 2 \cdot \bar{\epsilon}_x - 2 \cdot \bar{\epsilon}_y.$$

Let's find the directional slopes for the vectors $v = \vec{\epsilon}_x - \vec{\epsilon}_y$ and $w = \vec{\epsilon}_x + \vec{\epsilon}_y$. These are not unit vectors, so we'll work from the definition.

$$\begin{aligned} \text{direc } z &= \frac{\langle \bar{\partial} z \mid v \rangle}{\text{mag } v} = \frac{\langle 2 \cdot \bar{\epsilon}_x - 2 \cdot \bar{\epsilon}_y \mid \vec{\epsilon}_x - \vec{\epsilon}_y \rangle}{\sqrt{2}} = \frac{4}{\sqrt{2}} \\ \text{direc } z &= \frac{\langle \bar{\partial} z \mid w \rangle}{\text{mag } w} = \frac{\langle 2 \cdot \bar{\epsilon}_x - 2 \cdot \bar{\epsilon}_y \mid \vec{\epsilon}_x + \vec{\epsilon}_y \rangle}{\sqrt{2}} = 0 \end{aligned}$$

If we leave our point p traveling in the direction given by v , we'll be heading uphill at a slope of $4/\sqrt{2}$. If we instead travel in the direction of w , we'll find a slope of zero.

By looking at directional slopes for the standard directions, $\vec{\epsilon}_x$ and $\vec{\epsilon}_y$, we'll recover the x -slope and the y -slope.

$$\text{direc } z = \text{slope } x \quad \text{and} \quad \text{direc } z = \text{slope } y$$

So directional slopes generalize the slopes we defined as partials!

12.3 The Gradient

The *gradient* of a function $z \leftarrow (x, y)$ is the magnitude of the differential.

$\text{grad } z : \text{Number}$

$$\text{grad } z = \text{mag } \bar{\partial} z$$

The following theorem tells us that we can think of the gradient as a slope.

Theorem.

Suppose $z \leftarrow (x, y)$ is a function. The gradient can be written as the directional slope

$$\text{grad } z = \text{direc } z^v$$

where the vector

$v : \text{Vector}_{(x,y)}$

$$v = \text{tran } \bar{\partial} z$$

is the **transpose** of the differential.

Why?

To simplify notation, let's use the variable $r : \text{Ruler}_{(x,y)}$ to refer to the differential, $r = \bar{\partial} z$. Calculating the directional slope, we find

$$\text{direc } z^v = \frac{\langle \bar{\partial} z \mid v \rangle}{\text{mag } v} = \frac{\langle r \mid \text{tran } r \rangle}{\text{mag } r} = \frac{(\text{mag } r)^2}{\text{mag } r} = \text{mag } r = \text{mag } \bar{\partial} z = \text{grad } z.$$

The gradient isn't just any slope: the gradient automatically chooses the direction that gives the largest possible slope! Put simply, the gradient is the slope found when traveling directly uphill.

Example.

Let's compute the gradient for the **bowl**

$$z \leftarrow (x, y)$$

$$z = x^2 + y^2.$$

The differential is

$$\bar{\partial}z = 2x \cdot \bar{\varepsilon}_x + 2y \cdot \bar{\varepsilon}_y$$

and so the gradient is

$$\begin{aligned} \text{grad } z &= \text{mag } \bar{\partial}z \\ &= \text{mag}(2x \cdot \bar{\varepsilon}_x + 2y \cdot \bar{\varepsilon}_y) \\ &= \sqrt{(2x)^2 + (2y)^2} \\ &= 2 \cdot \sqrt{x^2 + y^2}. \end{aligned}$$

The gradient tells us how steep the graph is at any point. At the origin, $(0, 0)$, the gradient is zero.

$$\text{grad } z = 0$$

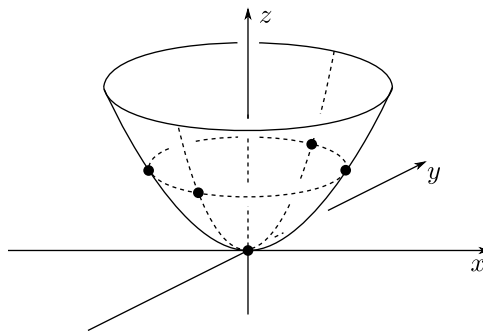
At each of the points

$$\begin{aligned} (1, 0) : \text{Point}_{(x,y)}, & \quad (-1, 0) : \text{Point}_{(x,y)}, \\ (0, 1) : \text{Point}_{(x,y)}, & \quad (0, -1) : \text{Point}_{(x,y)} \end{aligned}$$

the gradient is two.

$$\text{grad } z = 2$$

Let's take a look at the bowl and each of the points discussed in the previous example.



At the origin, the tangent plane is horizontal, and this gives a gradient of zero. At each of the four other points in question, the tangent planes will be equally steep, even if each is angled in a different direction.

12.4 Polyvectors

We've now come to the key definition that splits our theory of vectors and rulers away from our theory of metrics!

Definition.

The product of vectors is *anti-commutative*

$$w \cdot v = -v \cdot w$$

and the square of a vector is zero

$$v \cdot v = 0$$

where $v : \text{Vector}$ and $w : \text{Vector}$ are any vectors.

Applied to the standard vectors, we have:

$$\vec{\epsilon}_x \cdot \vec{\epsilon}_x = 0, \quad \vec{\epsilon}_y \cdot \vec{\epsilon}_x = -\vec{\epsilon}_x \cdot \vec{\epsilon}_y, \quad \vec{\epsilon}_y \cdot \vec{\epsilon}_y = 0.$$

With these multiplication laws in place, we can talk about polyvectors.

Definition.

Numbers are *polyvectors* with degree zero, the standard vectors,

$$\vec{\epsilon}_x : \text{Polyvector}_{(x,y)}^1 \quad \text{and} \quad \vec{\epsilon}_y : \text{Polyvector}_{(x,y)}^1,$$

are *polyvectors* with degree one, and all other *polyvectors* are made from these by adding and multiplying.

Let's get a little practice with polyvectors by multiplying two vectors together.

Example.

Let's multiply the vectors

$$\begin{aligned} v &: \text{Vector}_{(x,y)} & w &: \text{Vector}_{(x,y)} \\ v &= \vec{\epsilon}_x - 4 \cdot \vec{\epsilon}_y & \text{and} & \\ w &= 2 \cdot \vec{\epsilon}_x + \vec{\epsilon}_y. \end{aligned}$$

We compute.

$$\begin{aligned}
 v \cdot w &= (\vec{e}_x - 4 \cdot \vec{e}_y) \cdot (2 \cdot \vec{e}_x + \vec{e}_y) \\
 &= 2 \cdot \vec{e}_x \cdot \vec{e}_x + \vec{e}_x \cdot \vec{e}_y - 8 \cdot \vec{e}_y \cdot \vec{e}_x - 4 \cdot \vec{e}_y \cdot \vec{e}_y \\
 &= \vec{e}_x \cdot \vec{e}_y - 8 \cdot \vec{e}_y \cdot \vec{e}_x \\
 &= \vec{e}_x \cdot \vec{e}_y + 8 \cdot \vec{e}_x \cdot \vec{e}_y \\
 &= 9 \cdot \vec{e}_x \cdot \vec{e}_y
 \end{aligned}$$

The result is a polyvector with degree two!

Polyvector Standard Forms

Using our multiplication laws, we can rewrite polyvectors into standard forms.

degree	polyvector standard form
0	a
1	$a \cdot \vec{e}_x + b \cdot \vec{e}_y$
2	$a \cdot \vec{e}_x \cdot \vec{e}_y$

Like [polymetrics](#), polyvectors in degrees zero and one are familiar objects.

Important.

A degree zero polyvector is a number and a degree one polyvector is a vector.

Unlike polymetrics, polyvectors for the (x, y) -coordinate plane exist only in degrees zero, one, and two. In degree three, each of the possible products

$$\vec{e}_x \cdot \vec{e}_x \cdot \vec{e}_x, \quad \vec{e}_x \cdot \vec{e}_x \cdot \vec{e}_y, \quad \vec{e}_x \cdot \vec{e}_y \cdot \vec{e}_y, \quad \vec{e}_y \cdot \vec{e}_y \cdot \vec{e}_y$$

either contains $\vec{e}_x \cdot \vec{e}_x$ or $\vec{e}_y \cdot \vec{e}_y$ and is therefore zero. To get any degree three polyvectors, we would need to work in a space with more standard vectors.

12.5 Polyrulers

Our theory of polyrulers perfectly mirrors our theory of polyvectors.

Definition.

The product of rulers is anti-commutative

$$s \cdot r = -r \cdot s$$

and the square of a ruler is zero

$$r \cdot r = 0$$

where $r : \text{Ruler}$ and $s : \text{Ruler}$ are any rulers.

Because rulers and vector have the same multiplication, we get the same standard forms for polyrulers.

degree	polyruler standard form
0	a
1	$a \cdot \bar{\epsilon}_x + b \cdot \bar{\epsilon}_y$
2	$a \cdot \bar{\epsilon}_x \cdot \bar{\epsilon}_y$

We can measure a polyvector using a polyruler of the same degree. Using the polyruler $\bar{\epsilon}_x \cdot \bar{\epsilon}_y$ to measure the polyvector $\vec{\epsilon}_x \cdot \vec{\epsilon}_y$ gives a result of one.

$$\langle \bar{\epsilon}_x \cdot \bar{\epsilon}_y \mid \vec{\epsilon}_x \cdot \vec{\epsilon}_y \rangle = 1.$$

Let's make a measurement in degree two.

Example.

Let's measure the polyvector

$$q : \text{Polyvector}_{(x,y)}^2$$

$$q = 3 \cdot \vec{\epsilon}_x \cdot \vec{\epsilon}_y$$

with the polyruler

$$p : \text{Polyruler}_{(x,y)}^2$$

$$p = -2 \cdot \bar{\epsilon}_x \cdot \bar{\epsilon}_y.$$

We compute.

$$\begin{aligned}\langle p \mid q \rangle &= \langle -2 \cdot \bar{\epsilon}_x \cdot \bar{\epsilon}_y \mid 3 \cdot \vec{\epsilon}_x \cdot \vec{\epsilon}_y \rangle \\ &= -2 \cdot 3 \cdot \langle \bar{\epsilon}_x \cdot \bar{\epsilon}_y \mid \vec{\epsilon}_x \cdot \vec{\epsilon}_y \rangle \\ &= -6\end{aligned}$$

That was easy!

We've now seen how to make measurements in degrees one and two. In degree zero, measuring reduces to multiplying numbers. Let's quickly summarize how measurement works with a formula for each degree.

Formulas.

We may measure a polyvector with a polyruler as follows.

$$\begin{aligned}\langle a_r \mid a_v \rangle &= a_r a_v \\ \langle a_r \cdot \bar{\epsilon}_x + b_r \cdot \bar{\epsilon}_y \mid a_v \cdot \vec{\epsilon}_x + b_v \cdot \vec{\epsilon}_y \rangle &= a_r a_v + b_r b_v \\ \langle a_r \cdot \bar{\epsilon}_x \cdot \bar{\epsilon}_y \mid a_v \cdot \vec{\epsilon}_x \cdot \vec{\epsilon}_y \rangle &= a_r a_v\end{aligned}$$

In the above,

$$a_r : \text{Number}, \quad b_r : \text{Number}, \quad a_v : \text{Number}, \quad b_v : \text{Number}$$

are all numbers.

Our theory of vectors and rulers extends to three-dimensional space and beyond. For those interested, we've sketched how this works in [Appendix G. Rulers in Space](#).

12.6 Higher Differentials

Let's learn to take differentials of polyrulers!

Law.

The differential law for addition states

$$\bar{\partial}(p + q) = \bar{\partial}p + \bar{\partial}q$$

where $p : \text{Polyruler}_{(x,y)}$ and $q : \text{Polyruler}_{(x,y)}$ are polyrulers.

The differential law for addition is easy. Unfortunately, things get a bit trickier when we look at multiplication.

Laws.

The differential laws for multiplication state

$$\bar{\partial}(n \cdot p) = \bar{\partial}n \cdot p + n \cdot \bar{\partial}p$$

$$\bar{\partial}(r \cdot p) = \bar{\partial}r \cdot p - r \cdot \bar{\partial}p$$

where

$$n : \text{Number}, \quad r : \text{Ruler}, \quad p : \text{Polyruler}$$

are a number, a ruler, and a polyruler.

These multiplication laws are quite easy to mess up! It's difficult to remember when you need a negative sign and in what order everything goes. Luckily, we have an alternative pair of laws that are much easier to use in practice.

Laws.

The standard rulers pull out of the differential on the right side

$$\bar{\partial}(p \cdot \bar{\varepsilon}_x) = \bar{\partial}p \cdot \bar{\varepsilon}_x$$

$$\bar{\partial}(p \cdot \bar{\varepsilon}_y) = \bar{\partial}p \cdot \bar{\varepsilon}_y$$

where $p : \text{Polyruler}_{(x,y)}$ is any polyruler.

It's good to get some practice taking the differential.

Example.

Let's take the differential of the ruler

$$r : \text{Ruler}_{(x,y)}$$

$$r = xy^2 \cdot \bar{\varepsilon}_x + x^3 \cdot \bar{\varepsilon}_y.$$

We compute.

$$\begin{aligned} \bar{\partial}r &= \bar{\partial}(xy^2 \cdot \bar{\varepsilon}_x + x^3 \cdot \bar{\varepsilon}_y) \\ &= \bar{\partial}(xy^2 \cdot \bar{\varepsilon}_x) + \bar{\partial}(x^3 \cdot \bar{\varepsilon}_y) \\ &= \bar{\partial}(xy^2) \cdot \bar{\varepsilon}_x + \bar{\partial}(x^3) \cdot \bar{\varepsilon}_y \\ &= (y^2 \cdot \bar{\partial}x + 2xy \cdot \bar{\partial}y) \cdot \bar{\varepsilon}_x + 3x^2 \cdot \bar{\partial}x \cdot \bar{\varepsilon}_y \\ &= (y^2 \cdot \bar{\varepsilon}_x + 2xy \cdot \bar{\varepsilon}_y) \cdot \bar{\varepsilon}_x + 3x^2 \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y \\ &= y^2 \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_x + 2xy \cdot \bar{\varepsilon}_y \cdot \bar{\varepsilon}_x + 3x^2 \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y \\ &= -2xy \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y + 3x^2 \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y \\ &= (-2xy + 3x^2) \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y \end{aligned}$$

We've found a polyruler with degree two.

Let's now see an example where we take the differential of a differential.

Example.

The function

$$z \leftarrow (x, y)$$

$$z = x^2y^3$$

has differential

$$\bar{\partial}z : \text{Ruler}_{(x,y)}$$

$$\bar{\partial}z = 2xy^3 \cdot \bar{\varepsilon}_x + 3x^2y^2 \cdot \bar{\varepsilon}_y.$$

Let's take the second differential.

$$\begin{aligned}
\bar{\partial}\bar{\partial}z &= \bar{\partial}(2xy^3 \cdot \bar{\varepsilon}_x + 3x^2y^2 \cdot \bar{\varepsilon}_y) \\
&= \bar{\partial}(2xy^3) \cdot \bar{\varepsilon}_x + \bar{\partial}(3x^2y^2) \cdot \bar{\varepsilon}_y \\
&= 2 \cdot \bar{\partial}(xy^3) \cdot \bar{\varepsilon}_x + 3 \cdot \bar{\partial}(x^2y^2) \cdot \bar{\varepsilon}_y \\
&= 2 \cdot (y^3 \cdot \bar{\partial}x + 3xy^2 \cdot \bar{\partial}y) \cdot \bar{\varepsilon}_x + 3 \cdot (2xy^2 \cdot \bar{\partial}x + 2x^2y \cdot \bar{\partial}y) \cdot \bar{\varepsilon}_y \\
&= 2 \cdot (y^3 \cdot \bar{\varepsilon}_x + 3xy^2 \cdot \bar{\varepsilon}_y) \cdot \bar{\varepsilon}_x + 3 \cdot (2xy^2 \cdot \bar{\varepsilon}_x + 2x^2y \cdot \bar{\varepsilon}_y) \cdot \bar{\varepsilon}_y \\
&= 2y^3 \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_x + 6xy \cdot \bar{\varepsilon}_y \cdot \bar{\varepsilon}_x + 6xy \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y + 6x^2y \cdot \bar{\varepsilon}_y \cdot \bar{\varepsilon}_y \\
&= -6xy \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y + 6xy \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y \\
&= 0
\end{aligned}$$

The second ruler differential of z is identically zero.

You might think it strange that we chose an example that has second differential equal to zero. But remarkably, every function $z \leftarrow (x, y)$ has second ruler differential equal to zero!

$$\bar{\partial}\bar{\partial}z = 0$$

We'll explore this fact in detail in the next section.

12.7 The Second Differential

The goal of this section is to see why the following theorem is true.

Theorem.

The second ruler differential is identically zero

$$\bar{\partial}\bar{\partial}z = 0$$

where $z \leftarrow (x, y)$ is any function.

Before we can explain this theorem, we'll need a couple formulas for differentials. The first formula should be familiar from [our work with metrics](#).

Formula.

The differential of a function $z \leftarrow (x, y)$ is the ruler

$$\bar{\partial}z : \text{Ruler}_{(x,y)}$$

$$\bar{\partial}z = a \cdot \bar{\varepsilon}_x + b \cdot \bar{\varepsilon}_y$$

whose parts

$$a = \overset{x}{\partial}z \quad \text{and} \quad b = \overset{y}{\partial}z$$

are the partials of z .

Partials are the parts of the differential, whether that's the metric differential or the ruler differential. We'll also need a formula for the differential of a ruler.

Formula.

The differential of a ruler $r : \text{Ruler}_{(x,y)}$ is the polyruler

$$\bar{\partial}r : \text{Polyruler}_{(x,y)}^2$$

$$\bar{\partial}r = \left(\overset{x}{\partial}b - \overset{y}{\partial}a \right) \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y$$

when we've written $r = a \cdot \bar{\varepsilon}_x + b \cdot \bar{\varepsilon}_y$ in standard form.

Why?

We compute.

$$\begin{aligned}\bar{\partial}r &= \bar{\partial}(a \cdot \bar{\varepsilon}_x + b \cdot \bar{\varepsilon}_y) \\ &= \bar{\partial}a \cdot \bar{\varepsilon}_x + \bar{\partial}b \cdot \bar{\varepsilon}_y \\ &= \left(\overset{x}{\partial}a \cdot \bar{\varepsilon}_x + \overset{y}{\partial}a \cdot \bar{\varepsilon}_y \right) \cdot \bar{\varepsilon}_x + \left(\overset{x}{\partial}b \cdot \bar{\varepsilon}_x + \overset{y}{\partial}b \cdot \bar{\varepsilon}_y \right) \cdot \bar{\varepsilon}_y \\ &= \overset{x}{\partial}a \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_x + \overset{y}{\partial}a \cdot \bar{\varepsilon}_y \cdot \bar{\varepsilon}_x + \overset{x}{\partial}b \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y + \overset{y}{\partial}b \cdot \bar{\varepsilon}_y \cdot \bar{\varepsilon}_y \\ &= \overset{y}{\partial}a \cdot \bar{\varepsilon}_y \cdot \bar{\varepsilon}_x + \overset{x}{\partial}b \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y \\ &= -\overset{y}{\partial}a \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y + \overset{x}{\partial}b \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y \\ &= \left(\overset{x}{\partial}b - \overset{y}{\partial}a \right) \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y\end{aligned}$$

We're now ready to explain our theorem!

Why?

The first ruler differential

$$\bar{\partial}z = a \cdot \bar{\varepsilon}_x + b \cdot \bar{\varepsilon}_y$$

has partials for parts

$$a = \overset{x}{\partial}z \quad b = \overset{y}{\partial}z.$$

Taking the second differential, we find

$$\bar{\partial}\bar{\partial}z = \left(\overset{x}{\partial}b - \overset{y}{\partial}a \right) \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y = \left(\overset{x}{\partial}\overset{y}{\partial}z - \overset{y}{\partial}\overset{x}{\partial}z \right) \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y.$$

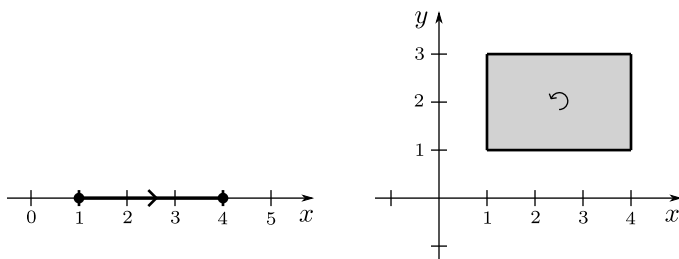
But this is zero because we can [reorder the partials](#)

$$\overset{y}{\partial}\overset{x}{\partial}z = \overset{x}{\partial}\overset{y}{\partial}z.$$

Chapter 13

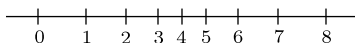
Boxes

Integrals are tough. Integrals make us work hard to do something simple: calculate a length or area. In its simplest form, an integral tells us that the length of the pictured interval is three, or the area of the box is six.



Given that we can already easily find such lengths and areas without using integrals, one might get to wondering, why should we learn about integrals?

Integrals are very powerful tools for measuring: we can switch out a standard ruler $\bar{\varepsilon}_x$ for a non-standard ruler $\bar{\partial}z : \text{Ruler}$ that varies from point to point. A non-standard ruler can give more weight to one location than another. We might picture a non-standard ruler as having a non-uniform density of markings.



Non-standard rulers appear in any number of contexts: by changing coordinates, we'll turn standard rulers into non-standard rulers. We'll find that non-standard rulers can be useful even if our ultimate measurement involves only standard rulers!

13.1 Integrals of Blips

We'll start our theory of integrals somewhere simple by looking at integrals of blips. The *blip*

$$[x_0]_x : \text{Blip}_x$$

selects the point x_0 on the x -axis.



We can think of the blip as representing the equation $x = x_0$.

Definition.

Suppose we have a function and a blip:

$$z \dot{\leftarrow} x, \quad b : \text{Blip}_x.$$

The *integral of b*

$$\left\langle z \int b \right\rangle : \text{Number}$$

is the number obtained by evaluating z at the blip b .

Let's see how this works.

Example.

Let's compute the integral for the following function and blip.

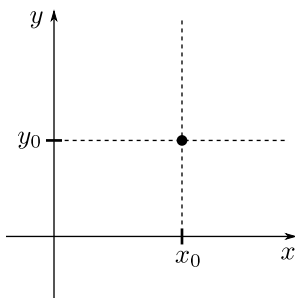
$$\begin{array}{ll} z \dot{\leftarrow} x & b : \text{Blip}_x \\ z = 3x^2 & \text{and} \quad b = [2]_x \end{array}$$

This is a quick calculation.

$$\left\langle z \int b \right\rangle = \left\langle 3x^2 \int [2]_x \right\rangle = \left\langle 12 \int \cdot \right\rangle = 12$$

The blip $[2]_x$ tells us to substitute two for x . Having used up the blip, there is nothing left to be done, and so we've found our result.

A product of blips $[x_0]_x \cdot [y_0]_y$ allows us to select the point (x_0, y_0) in the coordinate plane.



We define an integral of blips for a function $z \leftarrow (x, y)$ in a similar manner.

Example.

Let's compute the following integral of blips.

$$\left\langle \frac{x+y}{x-y} \int [5]_x \cdot [3]_y \right\rangle = \left\langle \frac{5+y}{5-y} \int [3]_y \right\rangle = \left\langle \frac{8}{2} \int \cdot \right\rangle = 4$$

Here we applied the x -blip before the y -blip. We could equally well have applied the y -blip first.

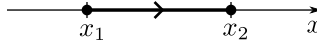
13.2 Intervals

The *interval* from x_1 to x_2 , written as

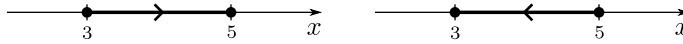
$$i : \text{Interval}_x$$

$$i = [x_1, x_2]_x$$

consists of all x -values from a number $x_1 : \text{Number}$ to a number $x_2 : \text{Number}$.



We draw an interval with an *orientation* to indicate the interval's direction. The interval $[3, 5]_x$ is pictured on the left, and the interval $[5, 3]_x$ is pictured on the right.



Negating an interval reverses the orientation.

$$-[x_1, x_2]_x = [x_2, x_1]_x$$

So

$$-[3, 5]_x = [5, 3]_x$$

for example.

Boundary

Let's define the *boundary operator*, $\ddot{\partial}$, for blips and intervals. The boundary of any blip is zero.

$$\ddot{\partial}[x_0]_x = 0$$

And the boundary of an interval is a difference of blips.

$$\ddot{\partial}[x_1, x_2]_x = [x_2]_x - [x_1]_x$$

We use positive and negative signs on blips as orientations:

- a positive blip, $+[x_2]_x$, represents an *arrival*, and
- a negative blip, $-[x_1]_x$, represents a *departure*.

The orientation of an interval determines the orientations on the boundary.

Example.

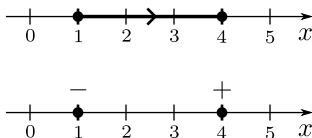
Let's picture the interval

$$[1, 4]_x : \text{Interval}_x$$

and its boundary

$$\partial[1, 4]_x = [4]_x - [1]_x.$$

We've labeled each blip with a positive or negative sign to record its orientation.



13.3 Integrals of Intervals

By computing the integral of an interval, we can measure the interval's length.

Definition.

Suppose we have a ruler and an interval

$$\bar{\partial}z : \text{Ruler}_x \quad i : \text{Interval}_x$$

where $\bar{\partial}z$ is the differential of a function $z \leftarrow x$. The *integral of i*

$$\left\langle \bar{\partial}z \int i \right\rangle : \text{Number}$$

$$\left\langle \bar{\partial}z \int i \right\rangle = \left\langle z \int \ddot{\partial}i \right\rangle$$

is computed by trading the differential, $\bar{\partial}$, for a boundary, $\ddot{\partial}$.

The standard ruler

$$\bar{\varepsilon}_x : \text{Ruler}_x$$

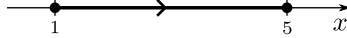
is also known as the *length ruler* for the x -coordinate line.

Example.

Let's use the length ruler, $\bar{\varepsilon}_x$, to measure the interval $[1, 5]_x$. We compute.

$$\begin{aligned} \left\langle \bar{\varepsilon}_x \int [1, 5]_x \right\rangle &= \left\langle \bar{\partial}x \int [1, 5]_x \right\rangle \\ &= \left\langle x \int \ddot{\partial}[1, 5]_x \right\rangle \\ &= \left\langle x \int [5]_x - [1]_x \right\rangle \\ &= \left\langle x \int [5]_x \right\rangle - \left\langle x \int [1]_x \right\rangle \\ &= \left\langle 5 \int \cdot \right\rangle - \left\langle 1 \int \cdot \right\rangle \\ &= 4 \end{aligned}$$

The length integral just had us subtract one from five. That's fair enough!



Integrals become more interesting when our rulers include dependence.

Example.

Consider the function

$$z \leftarrow x$$

$$z = x^2.$$

Let's compute the integral

$$\left\langle \bar{\partial} z \int [1, 2]_x \right\rangle$$

that measures the interval $[1, 2]_x$ with the ruler $\bar{\partial} z$. We find

$$\begin{aligned} \left\langle \bar{\partial} z \int [1, 2]_x \right\rangle &= \left\langle z \int \bar{\partial} [1, 2]_x \right\rangle \\ &= \left\langle z \int [2]_x - [1]_x \right\rangle \\ &= \left\langle z \int [2]_x \right\rangle - \left\langle z \int [1]_x \right\rangle \\ &= \left\langle x^2 \int [2]_x \right\rangle - \left\langle x^2 \int [1]_x \right\rangle \\ &= \left\langle 4 \int \cdot \right\rangle - \left\langle 1 \int \cdot \right\rangle \\ &= 3. \end{aligned}$$

This is not the measurement we would find using the length ruler, $\bar{\varepsilon}_x$.

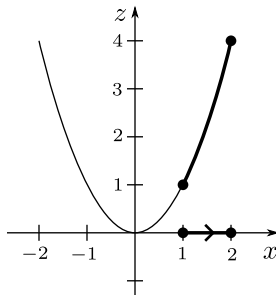
What role did rulers play in previous example's integral? On the x -axis, the ruler

$$\bar{\partial}z = 2x \cdot \bar{\epsilon}_x$$

is non-standard. We would like to measure the interval $[1, 2]_x$ with this ruler. But to make the measurement, we should switch to z -coordinates! On the z -axis

$$\bar{\partial}z = \bar{\epsilon}_z$$

is the standard ruler.



Measuring the corresponding interval $[1, 4]_z$ just computes the change in height.

$$\Delta_z = 4 - 1 = 3$$

What would you expect to find if we measured the interval $[-1, 2]_x$ instead? Or the interval $[-1, 1]_x$?

13.4 Anti-Differentiation

To compute an integral,

$$\left\langle r \int i \right\rangle$$

it is our job to write a ruler $r : \text{Ruler}_x$ as a differential

$$r = \bar{\partial} z.$$

We have lots of practice taking the differential, but for integrals, we need to reverse this process!

Definition.

Suppose $r : \text{Ruler}_x$ is a ruler. An *anti-differential* for r is a function

$$z \leftarrow x$$

whose differential is r .

$$r = \bar{\partial} z$$

Finding an anti-differential is often a difficult process. There is no one method that always delivers anti-differentials. Rather, we must be willing to make guesses, experiment with the differential laws, and learn from our mistakes.

Example.

Let's find an anti-differential for

$$r : \text{Ruler}_x$$

$$r = x^2 \cdot \bar{\epsilon}_x.$$

Taking the differential of x^3 will almost produce our ruler.

$$\bar{\partial}(x^3) = 3x^2 \cdot \bar{\epsilon}_x$$

If we divide both sides of this equation by three, we'll have our ruler.

$$x^2 \cdot \bar{\epsilon}_x = \frac{1}{3} \cdot \bar{\partial}(x^3)$$

$$x^2 \cdot \bar{\epsilon}_x = \bar{\partial}\left(\frac{x^3}{3}\right)$$

We've written our ruler as the differential of a function!

Let's calculate an integral by finding an anti-differential.

Example.

Consider the integral

$$\left\langle \sqrt[2]{x} \cdot \bar{\varepsilon}_x \int [1, 4]_x \right\rangle.$$

To compute this integral, we'll need to write the ruler as a differential. The [differential law for base functions](#) helps!

$$\bar{\partial}(x^{3/2}) = \frac{3}{2} \cdot x^{1/2} \cdot \bar{\varepsilon}_x$$

Let's solve for our ruler.

$$\frac{3}{2} \cdot x^{1/2} \cdot \bar{\varepsilon}_x = \bar{\partial}(x^{3/2})$$

$$\sqrt[2]{x} \cdot \bar{\varepsilon}_x = \frac{2}{3} \cdot \bar{\partial}(x^{3/2})$$

$$\sqrt[2]{x} \cdot \bar{\varepsilon}_x = \bar{\partial}\left(\frac{2}{3} \cdot x^{3/2}\right)$$

We've found an anti-differential. Let's take the integral!

$$\begin{aligned} \left\langle \sqrt[2]{x} \cdot \bar{\varepsilon}_x \int [1, 4]_x \right\rangle &= \left\langle \bar{\partial}\left(\frac{2}{3} \cdot x^{3/2}\right) \int [1, 4]_x \right\rangle \\ &= \left\langle \frac{2}{3} \cdot x^{3/2} \int \bar{\partial}[1, 4]_x \right\rangle \\ &= \left\langle \frac{2}{3} \cdot x^{3/2} \int [4]_x \right\rangle - \left\langle \frac{2}{3} \cdot x^{3/2} \int [1]_x \right\rangle \\ &= \frac{2}{3} \cdot 8 - \frac{2}{3} \cdot 1 \\ &= \frac{14}{3} \end{aligned}$$

Example.

Let's calculate the integral.

$$\left\langle \log x \cdot \bar{\varepsilon}_x \int [1, \text{nat}]_x \right\rangle$$

To do so, we'll need to write the ruler

$$\log x \cdot \bar{\varepsilon}_x$$

as the differential of a function. This is easier said than done! Let's try taking the differential of the function $x \cdot \log x$.

$$\bar{\partial}(x \cdot \log x) = \log x \cdot \bar{\partial}x + x \cdot \frac{1}{x} \cdot \bar{\partial}x$$

$$\bar{\partial}(x \cdot \log x) = \log x \cdot \bar{\partial}x + \bar{\partial}x$$

Our ruler is a term in this equation. We'll solve for our ruler.

$$\log x \cdot \bar{\partial}x = \bar{\partial}(x \cdot \log x) - \bar{\partial}x$$

$$\log x \cdot \bar{\varepsilon}_x = \bar{\partial}(x \cdot \log x - x)$$

We've found an anti-differential! Let's use it to compute the integral.

$$\begin{aligned} \left\langle \log x \cdot \bar{\varepsilon}_x \int [1, \text{nat}]_x \right\rangle &= \left\langle \bar{\partial}(x \cdot \log x - x) \int [1, \text{nat}]_x \right\rangle \\ &= \left\langle x \cdot \log x - x \int \ddot{\partial}[1, \text{nat}]_x \right\rangle \\ &= \left\langle x \cdot \log x - x \int [\text{nat}]_x \right\rangle - \left\langle x \cdot \log x - x \int [1]_x \right\rangle \\ &= (\text{nat} \cdot 1 - \text{nat}) - (1 \cdot 0 - 1) \\ &= 1 \end{aligned}$$

13.5 Boxes

Boxes are built out of blips and intervals in a way that might remind you of [polyvectors](#).

Definition.

We build *boxes* as follows.

- Blips are boxes.
- Intervals are boxes.
- All other boxes are made by multiplying blips and intervals.

We require that each factor of a box have its own variable, x or y .

We enforce the following multiplication laws for boxes.

Laws.

A blip $b : \text{Blip}$ commutes with any box $q : \text{Box}$.

$$q \cdot b = b \cdot q$$

Intervals anti-commute

$$j \cdot i = -i \cdot j$$

where $i : \text{Interval}$ and $j : \text{Interval}$ are intervals.

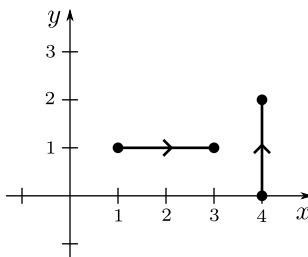
A blip is a degree zero box, and an interval is a degree one box. The degree of a box counts the number of interval factors and ignores blips.

Example.

Let's draw two examples of degree one boxes.

$$[1, 3]_x \cdot [1]_y : \text{Box}_{(x,y)}^1 \quad \text{and} \quad [4]_x \cdot [0, 2]_y : \text{Box}_{(x,y)}^1$$

The first takes x -values from one to three while fixing y at one. The second holds x at four and takes y -values from zero to two.



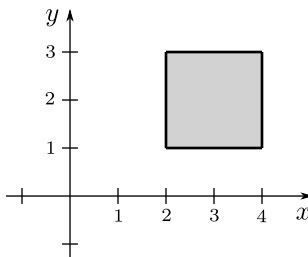
When we picture a degree two box, we give it a *clockwise orientation* or a *counterclockwise orientation*.

Example.

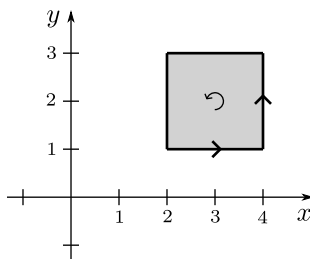
Consider the box

$$[2, 4]_x \cdot [1, 3]_y : \text{Box}_{(x,y)}^2.$$

This box contains all points with x -values between two and four, and y -values between one and three.



To understand orientation, we look at our box as giving ordered instructions: travel from $x = 2$ to $x = 4$, and then travel from $y = 1$ to $y = 3$. By following these instructions on the boundary, we determine the box's orientation.



We've found a counterclockwise orientation.

Negating a box reverses the box's orientation.

Example.

Let's consider the box

$$q : \text{Box}_{(x,y)}^2$$

$$q = [2, 4]_x \cdot [1, 3]_y$$

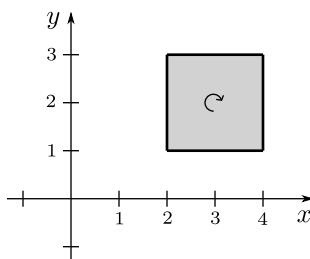
from the previous example, and its negative, $-q$. We may absorb the negative sign in any number of ways.

$$-q = [4, 2]_x \cdot [1, 3]_y$$

$$-q = [2, 4]_x \cdot [3, 1]_y$$

$$-q = [1, 3]_y \cdot [2, 4]_x$$

All of these ways of writing $-q$ give the same clockwise orientation.



13.6 Boundary

We've already defined the boundary operator, $\ddot{\partial}$, for blips and intervals.

$$\ddot{\partial}[x_0]_x = 0$$

$$\ddot{\partial}[x_1, x_2]_x = [x_2]_x - [x_1]_x$$

Let's extend the boundary operator to all boxes by using laws similar to the [differential laws for polyrulers](#).

Law.

The boundary law for addition states

$$\ddot{\partial}(q + r) = \ddot{\partial}q + \ddot{\partial}r$$

where $q : \text{Box}$ and $r : \text{Box}$ are boxes.

We must be careful with our multiplication laws!

Laws.

We have two boundary laws for multiplication

$$\ddot{\partial}(b \cdot q) = \ddot{\partial}b \cdot q + b \cdot \ddot{\partial}q$$

$$\ddot{\partial}(i \cdot q) = \ddot{\partial}i \cdot q - i \cdot \ddot{\partial}q$$

where

$$b : \text{Blip}, \quad i : \text{Interval}, \quad q : \text{Box}$$

are a blip, an interval, and a box.

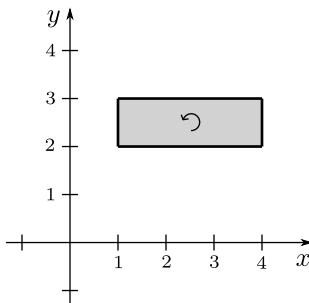
We can make sense of these laws with an example.

Example.

Let's take the boundary of the degree two box

$$[1, 4]_x \cdot [2, 3]_y : \text{Box}_{(x,y)}^2$$

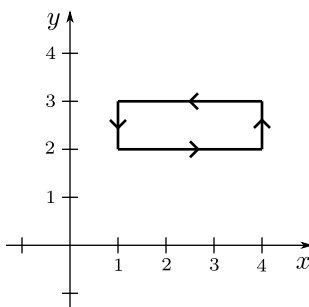
pictured as follows.



We calculate.

$$\begin{aligned} \ddot{\partial}([1, 4]_x \cdot [2, 3]_y) &= \ddot{\partial}[1, 4]_x \cdot [2, 3]_y - [1, 4]_x \cdot \ddot{\partial}[2, 3]_y \\ &= ([4]_x - [1]_x) \cdot [2, 3]_y - [1, 4]_x \cdot ([3]_y - [2]_y) \\ &= [4]_x \cdot [2, 3]_y - [1]_x \cdot [2, 3]_y - [1, 4]_x \cdot [3]_y + [1, 4]_x \cdot [2]_y \\ &= [4]_x \cdot [2, 3]_y + [1]_x \cdot [3, 2]_y + [4, 1]_x \cdot [3]_y + [1, 4]_x \cdot [2]_y \end{aligned}$$

The boundary is the sum of four boxes, each with degree one.



Notice how the boundary cycles, matching the box's orientation.

If we take the boundary of a boundary, we'll find many cancellations!

Theorem.

Suppose $q : \text{Box}_{(x,y)}^2$ is a box with degree two. The second boundary of q is zero.

$$\partial\partial q = 0$$

Why?

When we take the second boundary, we'll see each corner twice: once with a positive orientation and once with a negative orientation. These each cancel. If you wish to see the details, write q as

$$q = [x_1, x_2]_x \cdot [y_1, y_2]_y$$

and compute $\partial\partial q$.

13.7 Integrals of Boxes

Putting together polyrulers and boxes, we can measure area!

Definition.

Suppose we have a degree two polyruler and a degree two box

$$\bar{\partial}r : \text{Polyruler}_{(x,y)}^2 \quad q : \text{Box}_{(x,y)}^2$$

where $\bar{\partial}r$ is the differential of a ruler $r : \text{Ruler}_{(x,y)}$. The *integral* of q

$$\left\langle \bar{\partial}r \int q \right\rangle : \text{Number}$$

$$\left\langle \bar{\partial}r \int q \right\rangle = \left\langle r \int \ddot{\partial}q \right\rangle$$

is computed by trading the differential for a boundary.

The *area polyruler* for the (x, y) -coordinate plane is the polyruler

$$\bar{\varepsilon}_x \cdot \bar{\varepsilon}_y : \text{Polyruler}_{(x,y)}^2.$$

We get a simple formula when we calculate the area of a box.

Theorem.

The area integral of a box is given as follows

$$\left\langle \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y \int [x_1, x_2]_x \cdot [y_1, y_2]_y \right\rangle = (x_2 - x_1) \cdot (y_2 - y_1)$$

where

$$x_1 : \text{Number}, \quad x_2 : \text{Number}, \quad y_1 : \text{Number}, \quad y_2 : \text{Number}$$

are all constants.

Why?

We can write the area polyruler $\bar{\varepsilon}_x \cdot \bar{\varepsilon}_y$ as the differential of the ruler $r = x \cdot \bar{\varepsilon}_y$. Indeed,

$$\bar{\partial}r = \bar{\partial}(x \cdot \bar{\varepsilon}_y) = \bar{\partial}x \cdot \bar{\varepsilon}_y = \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y.$$

Let's use this to calculate the area A .

$$\begin{aligned} A &= \left\langle \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y \int [x_1, x_2]_x \cdot [y_1, y_2]_y \right\rangle \\ &= \left\langle \bar{\partial}(x \cdot \bar{\varepsilon}_y) \int [x_1, x_2]_x \cdot [y_1, y_2]_y \right\rangle \\ &= \left\langle x \cdot \bar{\varepsilon}_y \int \ddot{\partial}([x_1, x_2]_x \cdot [y_1, y_2]_y) \right\rangle \\ &= \left\langle x \cdot \bar{\varepsilon}_y \int \ddot{\partial}[x_1, x_2]_x \cdot [y_1, y_2]_y \right\rangle - \left\langle x \cdot \bar{\varepsilon}_y \int [x_1, x_2]_x \cdot \ddot{\partial}[y_1, y_2]_y \right\rangle \\ &= I_2 - I_1 \end{aligned}$$

So far our calculation has produced the difference of two integrals $I_2 - I_1$. Let's first work with I_1 .

$$\begin{aligned} I_1 &= \left\langle x \cdot \bar{\partial}y \int [x_1, x_2]_x \cdot \ddot{\partial}[y_1, y_2]_y \right\rangle \\ &= \left\langle x \cdot \bar{\partial}y \int [x_1, x_2]_x \cdot [y_2]_y \right\rangle - \left\langle x \cdot \bar{\partial}y \int [x_1, x_2]_x \cdot [y_1]_y \right\rangle \\ &= \left\langle x \cdot \bar{\partial}y_2 \int [x_1, x_2]_x \right\rangle - \left\langle x \cdot \bar{\partial}y_1 \int [x_1, x_2]_x \right\rangle \\ &= \left\langle 0 \int [x_1, x_2]_x \right\rangle - \left\langle 0 \int [x_1, x_2]_x \right\rangle \\ &= 0 \end{aligned}$$

We find that I_1 is zero because y_1 and y_2 are constants. Let's pick up our calculation of the

area with I_2 .

$$\begin{aligned}
A &= \left\langle x \cdot \bar{\partial} y \int \ddot{\partial}[x_1, x_2]_x \cdot [y_1, y_2]_y \right\rangle \\
&= \left\langle x \cdot \bar{\partial} y \int [x_2]_x \cdot [y_1, y_2]_y \right\rangle - \left\langle x \cdot \bar{\partial} y \int [x_1]_x \cdot [y_1, y_2]_y \right\rangle \\
&= \left\langle x_2 \cdot \bar{\partial} y \int [y_1, y_2]_y \right\rangle - \left\langle x_1 \cdot \bar{\partial} y \int [y_1, y_2]_y \right\rangle \\
&= x_2 \cdot \left\langle \bar{\partial} y \int [y_1, y_2]_y \right\rangle - x_1 \cdot \left\langle \bar{\partial} y \int [y_1, y_2]_y \right\rangle \\
&= x_2 \cdot \left\langle y \int \ddot{\partial}[y_1, y_2]_y \right\rangle - x_1 \cdot \left\langle y \int \ddot{\partial}[y_1, y_2]_y \right\rangle \\
&= x_2 \cdot (y_2 - y_1) - x_1 \cdot (y_2 - y_1) \\
&= (x_2 - x_1) \cdot (y_2 - y_1)
\end{aligned}$$

Phew! We found our answer.

Chapter 14

Paths

Imagine a crab scuttling around a beach, leaving a trail in the sand as it goes. There are several questions we might ask about the crab's motion.

- At any time, which direction is the crab moving in?
- How fast is the crab moving?
- How much does the crab change its height as it, say, climbs up a sandcastle?

We'll learn to answer these questions by studying motion along a path.

14.1 Paths

A *path* is defined as a point $(x_p, y_p) : \text{Point}_{(x,y)}$ whose location depends on time t . We'll write a path as

$$p : \text{Path}_{(x,y)} \quad \begin{cases} x_p \leftarrow t \\ y_p \leftarrow t \end{cases}$$

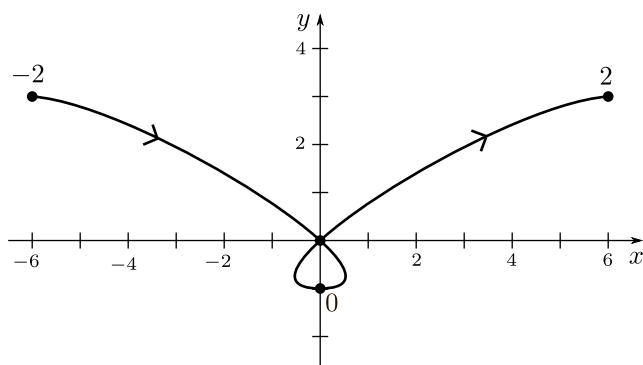
by describing how each coordinate x_p and y_p depends on time. We *plot a path* by drawing the point (x_p, y_p) for every allowed time t . By labeling this point with the time at several locations, we get some sense of how the path is traced. We'll refer to the point (x_p, y_p) as the path's *tracing point*.

Example.

Let's look at the path p for the interval i .

$$p : \text{Path}_{(x,y)} \quad i : \text{Interval}_t \quad \begin{cases} x_p = t^3 - t \\ y_p = t^2 - 1 \end{cases} \quad i = [-2, 2]_t$$

Our path starts at $(-6, 3)$ and ends at $(6, 3)$. The tracing point travels through the origin twice, once at time $t = -1$ and once at time $t = 1$.



t	(x_p, y_p)
-2	$(-6, 3)$
-1	$(0, 0)$
0	$(0, -1)$
1	$(0, 0)$
2	$(6, 3)$

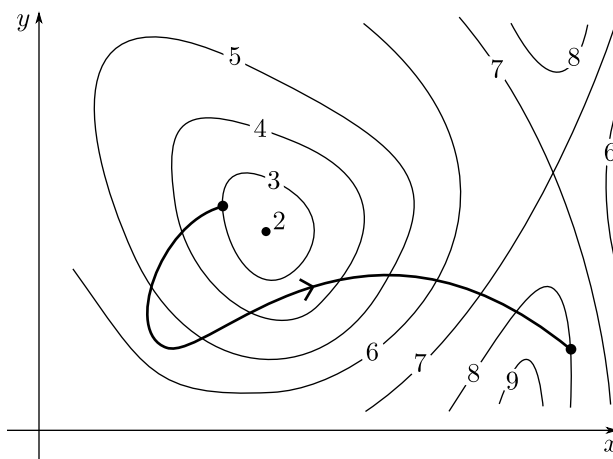
On the whole, the tracing point travels from left to right, but for a brief time around $t = 0$ it travels from right to left. By looking at the amount of ground covered between t -values, we can get a sense of how quickly the tracing point moves. Around time $t = 0$, do you expect the

tracing point is moving relatively slowly or relatively quickly? We'll address this question in the next section by taking a derivative.

As in the previous example, we'll typically define a path for a fixed interval of time. This means that the path has a starting point and an ending point. Let's see another example of a path, but this time sharing the (x, y) -coordinate plane with a [contour map](#).

Example.

Consider the following path on the contour map. What is the net change in height Δ_z experienced across this path?



The path starts at height $z = 3$ and ends at height $z = 8$. So the tracing point sees a net change of height of

$$\Delta_z = 8 - 3 = 5.$$

As always we subtract:

to – from.

14.2 Motion

In order to study motion, we'll need to take time derivatives of vectors.

Laws.

The derivative law for addition states

$${}^t\partial(v + w) = {}^t\partial v + {}^t\partial w$$

and the derivative law for multiplication states

$${}^t\partial(n \cdot v) = {}^t\partial n \cdot v + n \cdot {}^t\partial v$$

where

$$v : \text{Vector}, \quad w : \text{Vector}$$

are vectors and $n : \text{Number}$ is a number.

The standard vectors have zero derivative.

$${}^t\partial(\vec{\varepsilon}_x) = 0 \quad \text{and} \quad {}^t\partial(\vec{\varepsilon}_y) = 0$$

The standard vectors therefore pull out of the derivative

$${}^t\partial(n \cdot \vec{\varepsilon}_x) = {}^t\partial n \cdot \vec{\varepsilon}_x$$

$${}^t\partial(n \cdot \vec{\varepsilon}_y) = {}^t\partial n \cdot \vec{\varepsilon}_y$$

where $n : \text{Number}$ is any number.

Position, Velocity, and Speed

A path

$$p : \text{Path}_{(x,y)}$$

$$\begin{cases} x_p \leftarrow t \\ y_p \leftarrow t \end{cases}$$

has *position vector*, *velocity vector*, and *speed* as follows.

$$\text{pos } p : \text{Vector}_{(x,y)} \quad \text{vel } p : \text{Vector}_{(x,y)} \quad \text{speed } p : \text{Number}$$

$$\text{pos } p = x_p \cdot \vec{\varepsilon}_x + y_p \cdot \vec{\varepsilon}_y \quad \text{vel } p = {}^t\partial \text{pos } p \quad \text{speed } p = \text{mag vel } p$$

The velocity vector points in the direction of travel and so is sometimes called the *tangent vector*. But the velocity tells us more than just a direction: its magnitude tells us how fast our tracing point moves!

Example.

Let's revisit a path that we looked at in the previous section. Consider

$$p : \text{Path}_{(x,y)} \quad \begin{cases} x_p = t^3 - t \\ y_p = t^2 - 1 \end{cases}$$

for the time interval

$$i : \text{Interval}_t \\ i = [-2, 2]_t.$$

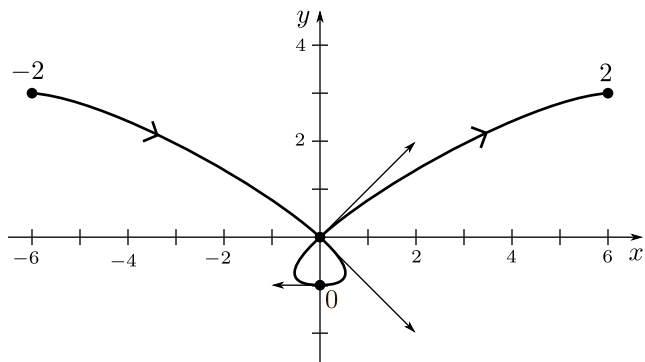
This path has position, velocity, and speed as follows.

$$\begin{aligned} \text{pos } p &= (t^3 - t) \cdot \vec{\epsilon}_x + (t^2 - 1) \cdot \vec{\epsilon}_y \\ \text{vel } p &= (3t^2 - 1) \cdot \vec{\epsilon}_x + 2t \cdot \vec{\epsilon}_y \\ \text{speed } p &= \sqrt[2]{9t^4 - 2t^2 + 1} \end{aligned}$$

Let's look at these for a few t -values.

t	$\text{pos } p$	$\text{vel } p$	$\text{speed } p$
-2	$-6 \cdot \vec{\epsilon}_x + 3 \cdot \vec{\epsilon}_y$	$11 \cdot \vec{\epsilon}_x - 4 \cdot \vec{\epsilon}_y$	$\sqrt[2]{137}$
-1	0	$2 \cdot \vec{\epsilon}_x - 2 \cdot \vec{\epsilon}_y$	$\sqrt[2]{8}$
0	$-\vec{\epsilon}_y$	$-\vec{\epsilon}_x$	1
1	0	$2 \cdot \vec{\epsilon}_x + 2 \cdot \vec{\epsilon}_y$	$\sqrt[2]{8}$
2	$6 \cdot \vec{\epsilon}_x + 3 \cdot \vec{\epsilon}_y$	$11 \cdot \vec{\epsilon}_x + 4 \cdot \vec{\epsilon}_y$	$\sqrt[2]{137}$

The velocity vector at time $t = 0$ points due West. The velocity at times, $t = -1$ and $t = 1$, points South-East and North-East, respectively.



Compared with the other times in our table, the tracing point is moving relatively slowly at time $t = 0$. That said, we can find times when the tracing point moves slower: you can check that the lowest speed

$$\text{speed } p = \sqrt[2]{\frac{8}{9}}$$

occurs at times $t = -1/3$ and $t = 1/3$. So oddly enough, $t = 0$ is a high point for speed!

14.3 Path Integrals

Let's see how we take an integral along a path.

Definition.

Suppose we have a ruler, a path, and a time interval

$$\bar{\partial}z : \text{Ruler}_{(x,y)} \quad p : \text{Path}_{(x,y)} \quad i : \text{Interval}_t$$

where $z \leftarrow (x, y)$ is a function. We define the *path integral*

$$\left\langle \bar{\partial}z \int^p i \right\rangle : \text{Number}$$

$$\left\langle \bar{\partial}z \int^p i \right\rangle = \left\langle z \int^p \ddot{\partial}i \right\rangle$$

by trading the differential for a boundary.

A path integral has many ingredients. An example should help us see how they all work together.

Example.

Consider the path integral

$$\left\langle \bar{\partial}z \int^p i \right\rangle$$

where

$$z \leftarrow (x, y) \quad p : \text{Path}_{(x,y)} \quad i : \text{Interval}_t$$

$$z = xy \quad \begin{cases} x_p = 3t \\ y_p = t \end{cases} \quad i = [1, 4]_t.$$

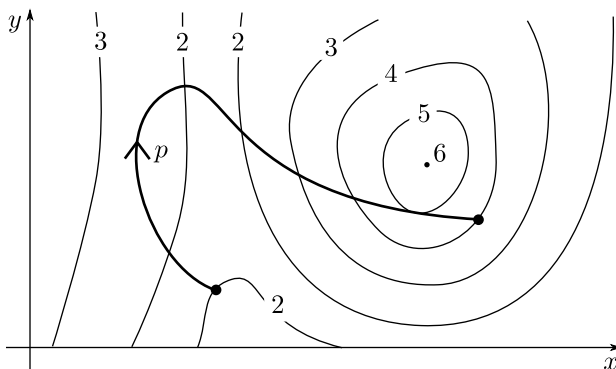
We calculate.

$$\begin{aligned}
 \left\langle \bar{\partial} z \int_p i \right\rangle &= \left\langle z \int_p \ddot{\partial}[1, 4]_t \right\rangle \\
 &= \left\langle z \int_p [4]_t \right\rangle - \left\langle z \int_p [1]_t \right\rangle \\
 &= \left\langle xy \int_p [4]_t \right\rangle - \left\langle xy \int_p [1]_t \right\rangle \\
 &= \left\langle 3t^2 \int_p [4]_t \right\rangle - \left\langle 3t^2 \int_p [1]_t \right\rangle \\
 &= 48 - 3 \\
 &= 45
 \end{aligned}$$

The path integral of a differential $\bar{\partial}z$ computes a change in height! Let's visualize a path integral using a contour map.

Example.

Consider the path p drawn on the contour map for a function $z \leftarrow (x, y)$.



Our path begins at height $z = 2$ and ends at height $z = 4$. The path integral with ruler $\bar{\partial}z$ computes the change in height.

$$\left\langle \bar{\partial} z \int_p i \right\rangle = \left\langle z \int_p \ddot{\partial} i \right\rangle = 4 - 2 = 2$$

14.4 Tame and Wild Rulers

Let's take a quick detour from integrals to look at tame and wild rulers.

Definition.

A ruler $r : \text{Ruler}_{(x,y)}$ is *tame* if its differential is identically zero.

$$\bar{\partial}r = 0$$

If the differential of a ruler is non-zero at any point, we'll say that the ruler is *wild*.

Every ruler is either tame or wild.

Example.

Let's see whether the ruler

$$\begin{aligned} r : \text{Ruler}_{(x,y)} \\ r = xy^2 \cdot \bar{\varepsilon}_x + x \cdot \bar{\varepsilon}_y \end{aligned}$$

is tame or wild.

$$\begin{aligned} \bar{\partial}r &= \bar{\partial}(xy^2 \cdot \bar{\varepsilon}_x + x \cdot \bar{\varepsilon}_y) \\ &= \bar{\partial}(xy^2 \cdot \bar{\varepsilon}_x) + \bar{\partial}(x \cdot \bar{\varepsilon}_y) \\ &= \bar{\partial}(xy^2) \cdot \bar{\varepsilon}_x + \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y \\ &= y^2 \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_x + 2xy \cdot \bar{\varepsilon}_y \cdot \bar{\varepsilon}_x + \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y \\ &= -2xy \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y + \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y \\ &= (1 - 2xy) \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y \end{aligned}$$

We see that the polyruler $\bar{\partial}r$ is not identically zero. Although there may be some points where $\bar{\partial}r$ is zero, say

$$\left(\frac{1}{2}, 1\right) : \text{Point}_{(x,y)},$$

it's easy to find points where the polyruler is non-zero. So r is wild.

The standard rulers $\bar{\varepsilon}_x$ and $\bar{\varepsilon}_y$ are examples of tame rulers. We get further examples of tame rulers from the differential.

Important.

For any function $z \leftarrow (x, y)$ the differential

$$\bar{\partial}z : \mathbf{Ruler}_{(x,y)}$$

is tame.

Why?

To check whether a ruler is tame, we take its differential.

$$\bar{\partial}\bar{\partial}z = 0$$

The [second ruler differential](#) is always zero.

If a ruler is tame, there's a good chance you'll be able to write the ruler as the differential of a function.

Example.

The ruler

$$r : \mathbf{Ruler}_{(x,y)}$$

$$r = y^2 \cdot \bar{\epsilon}_x + 2xy \cdot \bar{\epsilon}_y$$

is tame.

$$\begin{aligned} \bar{\partial}r &= \bar{\partial}(y^2 \cdot \bar{\epsilon}_x + 2xy \cdot \bar{\epsilon}_y) \\ &= \bar{\partial}(y^2) \cdot \bar{\epsilon}_x + \bar{\partial}(2xy) \cdot \bar{\epsilon}_y \\ &= 2y \cdot \bar{\epsilon}_y \cdot \bar{\epsilon}_x + 2 \cdot \bar{\partial}(xy) \cdot \bar{\epsilon}_y \\ &= 2y \cdot \bar{\epsilon}_y \cdot \bar{\epsilon}_x + 2y \cdot \bar{\epsilon}_x \cdot \bar{\epsilon}_y + 2x \cdot \bar{\epsilon}_y \cdot \bar{\epsilon}_y \\ &= -2y \cdot \bar{\epsilon}_x \cdot \bar{\epsilon}_y + 2y \cdot \bar{\epsilon}_x \cdot \bar{\epsilon}_y \\ &= 0 \end{aligned}$$

Can you find a function

$$z \leftarrow (x, y)$$

that has r as its differential?

If a ruler is wild, it cannot be written as the differential of any function.

Example.

As we saw in our first example, the ruler

$$r : \text{Ruler}_{(x,y)}$$

$$r = xy^2 \cdot \bar{\varepsilon}_x + x \cdot \bar{\varepsilon}_y$$

is wild. We'll never be able to find a function

$$z \leftarrow (x, y)$$

whose differential is r .

14.5 Path Integrals, Revisited

We've defined [path integrals](#) for rulers that are differentials. Unfortunately, this definition is useless for many rulers: a wild ruler can never be written as the differential of a function $z \leftarrow (x, y)$. To work with wild rulers, we'll need to broaden our definition of the path integral.

Definition.

Suppose we have a ruler, a path, and a time interval as follows.

$$r : \text{Ruler}_{(x,y)} \quad p : \text{Path}_{(x,y)} \quad i : \text{Interval}_t$$

We calculate the *path integral*

$$\left\langle r \int^p i \right\rangle : \text{Number}$$

by using the path p to rewrite the path integral as a time integral.

By moving path integrals from (x, y) -coordinates to time t we can take many more integrals!

Example.

Let's compute the path integral

$$\left\langle r \int^p i \right\rangle$$

with ruler, path, and interval given as follows.

$$\begin{aligned} r : \text{Ruler}_{(x,y)} & & p : \text{Path}_{(x,y)} & & i : \text{Interval}_t \\ r = y \cdot \bar{\epsilon}_x + (2y - x) \cdot \bar{\epsilon}_y & & \begin{cases} x_p = t^2 \\ y_p = t \end{cases} & & i = [0, 3]_t \end{aligned}$$

You can check that the ruler r is wild. So there is no hope of finding a function $z \leftarrow (x, y)$ that satisfies

$$\bar{\partial}z = y \cdot \bar{\epsilon}_x + (2y - x) \cdot \bar{\epsilon}_y.$$

Let's instead use our path p to rewrite r as a ruler that measures time.

$$\begin{aligned}
 r &= y \cdot \bar{\partial}x + (2y - x) \cdot \bar{\partial}y \\
 &= t \cdot \bar{\partial}(t^2) + (2t - t^2) \cdot \bar{\partial}t \\
 &= t \cdot 2t \cdot \bar{\partial}t + 2t \cdot \bar{\partial}t - t^2 \cdot \bar{\partial}t \\
 &= (2t^2 + 2t - t^2) \cdot \bar{\varepsilon}_t \\
 &= (t^2 + 2t) \cdot \bar{\varepsilon}_t
 \end{aligned}$$

To solve our integral, we should look for a function $z \leftarrow t$ that satisfies

$$\bar{\partial}z = (t^2 + 2t) \cdot \bar{\varepsilon}_t.$$

By anti-differentiating, we find

$$\begin{aligned}
 (t^2 + 2t) \cdot \bar{\varepsilon}_t &= (t^2 + 2t) \cdot \bar{\partial}t \\
 &= t^2 \cdot \bar{\partial}t + 2t \cdot \bar{\partial}t \\
 &= \frac{1}{3} \cdot 3t^2 \cdot \bar{\partial}t + 2t \cdot \bar{\partial}t \\
 &= \frac{1}{3} \cdot \bar{\partial}(t^3) + \bar{\partial}(t^2) \\
 &= \bar{\partial}\left(\frac{1}{3} \cdot t^3\right) + \bar{\partial}(t^2) \\
 &= \bar{\partial}\left(\frac{1}{3} \cdot t^3 + t^2\right).
 \end{aligned}$$

We can calculate our path integral as a time integral.

$$\begin{aligned}
 \left\langle r \int^p i \right\rangle &= \left\langle y \cdot \bar{\varepsilon}_x + (2y - x) \cdot \bar{\varepsilon}_y \int [0, 3]_t \right\rangle \\
 &= \left\langle (t^2 + 2t) \cdot \bar{\varepsilon}_t \int [0, 3]_t \right\rangle \\
 &= \left\langle \bar{\partial} \left(\frac{t^3}{3} + t^2 \right) \int [0, 3]_t \right\rangle \\
 &= \left\langle \frac{t^3}{3} + t^2 \int \ddot{\partial} [0, 3]_t \right\rangle \\
 &= \left\langle \frac{t^3}{3} + t^2 \int [3]_t \right\rangle - \left\langle \frac{t^3}{3} + t^2 \int [0]_t \right\rangle \\
 &= \left(\frac{27}{3} + 9 \right) - (0 + 0) \\
 &= 18
 \end{aligned}$$

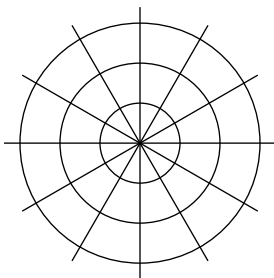
We've found our path integral!

Chapter 15

Angles

Upon completing the previous chapter on paths, we've accomplished the goals we set for ourselves! We learned to calculate the differential, and we've seen many geometric interpretations. We can analyze a graph's slope, bend, warp, gradient, etc. We can measure length and area using integrals. And we can understand the position, velocity, and speed of a point moving along a path. Sounds like it's about time to pack up and head home.

But wait! Before we call it quits, let's see what Calculus has to say about angles. So far, all of our tools have used standard (x, y) -coordinates. Metrics, vectors, rulers, and boxes, have all been written using x and y standards. But these tools are actually quite flexible. They can be used with non-standard coordinates, too.



In this last chapter we'll study angles in order to define polar coordinates. We'll investigate polar rulers and use them to measure the area of polar regions. The differential is a powerful idea that has applications in many contexts!

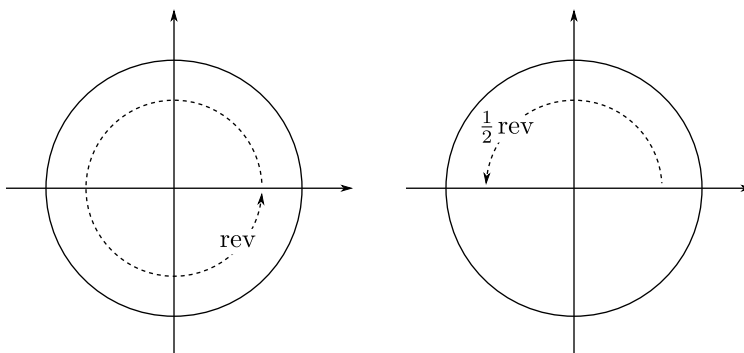
15.1 Angles

In a standard coordinate plane we'll draw angles starting from the positive side of the horizontal axis, and we'll move in the counterclockwise direction.

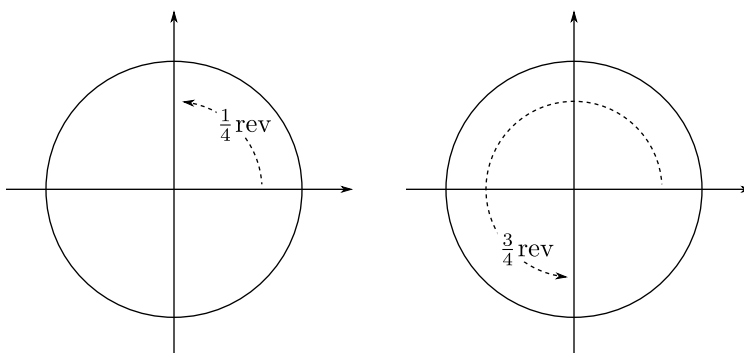
Definition.

We'll use rev to denote one complete revolution.

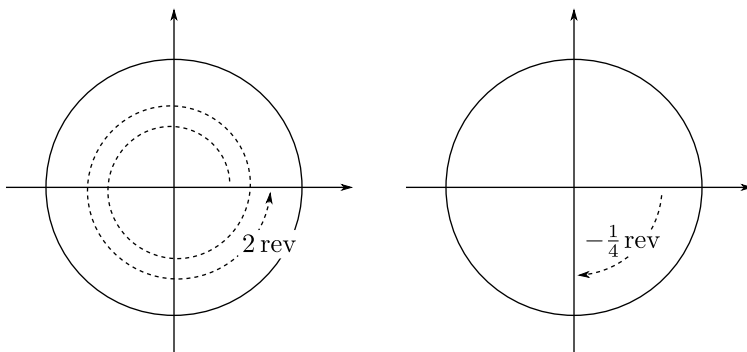
We'll record angles as fractions of a revolution. Here is a complete revolution, and here is half of a revolution.



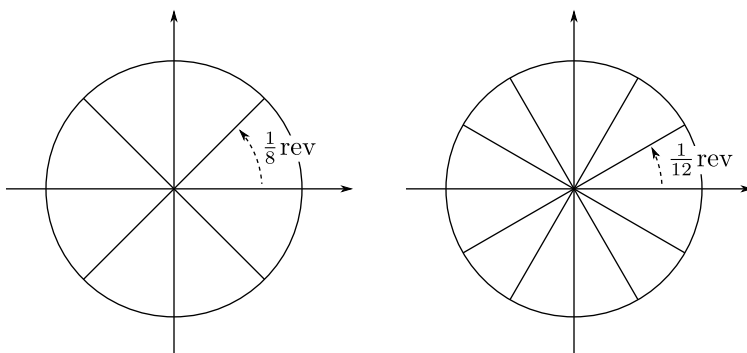
And here are the other possible quarters of a revolution. Notice that we can record a right angle as $\frac{1}{4} \text{ rev}$.



We may consider angles larger than rev by going beyond one revolution. Negative angles move in the clockwise direction.



We'll be interested in eighths of a revolution and twelfths of a revolution. We get these fractions by further subdividing each quarter evenly into two pieces for eighths, or into three pieces for twelfths.



Degrees

By describing angles as fractions of a revolution, we have managed to delay making an age old decision: degrees or radians. Let's see how our approach to angles works with degrees. In the next section and beyond, we'll switch to using radians.

Definition.

A full revolution is three hundred sixty degrees.

$$\text{rev} = 360^\circ$$

Let's write down a few common angles in degrees.

$$\begin{array}{lll} \frac{1}{2} \text{ rev} = 180^\circ & \frac{1}{4} \text{ rev} = 90^\circ & \frac{1}{6} \text{ rev} = 60^\circ \\ \frac{1}{8} \text{ rev} = 45^\circ & \frac{1}{12} \text{ rev} = 30^\circ & \frac{1}{24} \text{ rev} = 15^\circ \end{array}$$

Conveniently, each of these angles is an integer when written in degrees.

Example.

Let's write the angle $\frac{7}{8}$ rev in degrees.

$$\frac{7}{8} \text{ rev} = \frac{7}{8} \cdot 360^\circ = 7 \cdot 45^\circ = 315^\circ$$

And now let's write the angle 150° as a fraction of a revolution.

$$150^\circ = 150^\circ \cdot \frac{\text{rev}}{360^\circ} = \frac{5}{12} \text{ rev}$$

15.2 Unit Circle

When we work with degrees, we define a revolution as

rev : Number

$$\text{rev} = 360^\circ.$$

For differentials, another definition of rev works much better.

Definition.

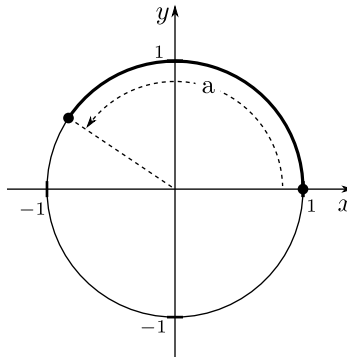
When writing angles in *radians* we define rev to be equal to the circumference of the unit circle.

rev : Number

$$\text{rev} = 6.282 \dots$$

We cannot write rev exactly in decimal notation: rev is an irrational number.

On the unit circle an angle written in radians is equal to the distance traveled along the circle.



Cosine and Sine

We define the cosine and sine functions by looking at coordinates on the unit circle.

Definition.

On the unit circle

$$x = \cos a$$

$$y = \sin a$$

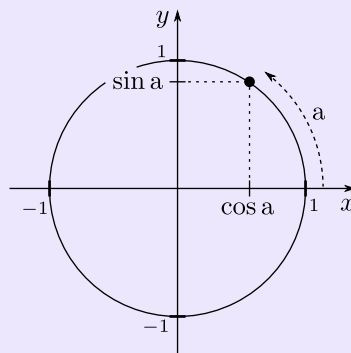
where a : Number is any angle.

Take note that the cosine and sine functions

$$x \xleftarrow{\cos} a$$

$$y \xleftarrow{\sin} a$$

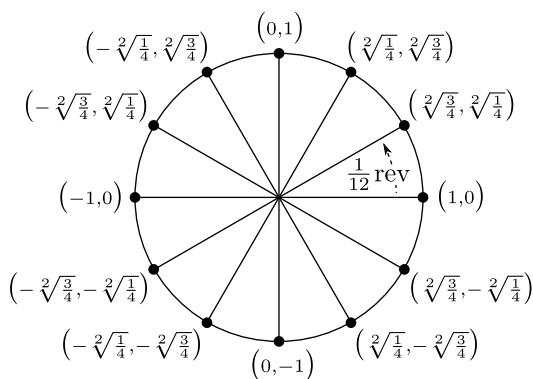
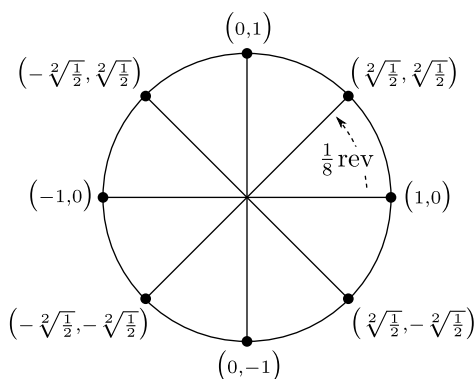
use x and y as output variables.



Since the cosine and sine functions describe a point on the unit circle, they satisfy the famous equation

$$(\cos a)^2 + (\sin a)^2 = 1.$$

Let's summarize some familiar cosine and sine values by splitting the unit circle into eighths, like a large pizza, and twelfths, like a clock face.



These values can all be found using geometry and algebra. I'd like to encourage the student to

get out pen and paper, and find geometric proofs for the cosine and sine values:

$$\begin{aligned}\cos\left(\frac{1}{8}\text{ rev}\right) &= \sqrt[2]{\frac{1}{2}} & \sin\left(\frac{1}{8}\text{ rev}\right) &= \sqrt[2]{\frac{1}{2}} \\ \cos\left(\frac{1}{12}\text{ rev}\right) &= \sqrt[2]{\frac{3}{4}} & \sin\left(\frac{1}{12}\text{ rev}\right) &= \sqrt[2]{\frac{1}{4}}.\end{aligned}$$

Of course the square root of one-quarter is one-half, but we'll write $\sqrt[2]{1/4}$ in solidarity with the other cosine and sine values.

15.3 Differential Laws

Let's take a look at differential laws for cosine and sine.

Laws.

The differential law for cosine states

$$\partial(\cos a) = -\sin a \cdot \partial a$$

and the differential law for sine states

$$\partial(\sin a) = \cos a \cdot \partial a$$

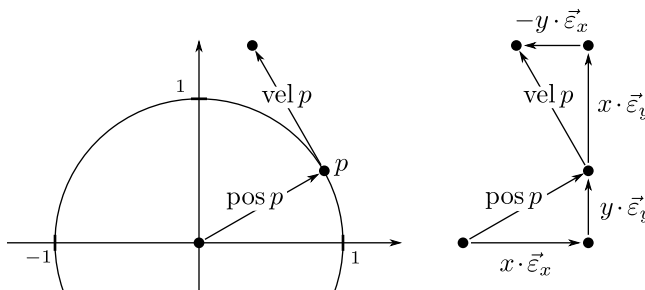
where $a : \text{Number}$ is an angle.

Why?

Consider the following path that travels counterclockwise around the unit circle.

$$p : \text{Path}_{(x,y)} \begin{cases} x = \cos t \\ y = \sin t \end{cases}$$

Since we write our angles in radians, our path's tracing point moves with a constant speed of one. Let's draw the position vector and the velocity vector in the plane. The velocity vector must be tangent to the circle, have magnitude one, and be pointing in the counterclockwise direction.



A little geometry allows us to relate the parts of the velocity vector to the parts of the position

vector.

$$\begin{aligned} \text{pos } p &: \text{Vector}_{(x,y)} & \text{vel } p &: \text{Vector}_{(x,y)} \\ \text{pos } p &= x \cdot \vec{\varepsilon}_x + y \cdot \vec{\varepsilon}_y & \text{and} & \quad \text{vel } p = -y \cdot \vec{\varepsilon}_x + x \cdot \vec{\varepsilon}_y \end{aligned}$$

The velocity vector is the t -derivative of the position vector, and so

$$\frac{t}{\partial} x = -y \quad \text{and} \quad \frac{t}{\partial} y = x.$$

The differential laws follow.

Let's use our differential laws for sine and cosine to calculate something!

Example.

Let's find the derivative for the function

$$\begin{aligned} z &\leftarrow a \\ z &= \frac{\sin a}{\cos a}. \end{aligned}$$

We compute.

$$\begin{aligned} \frac{a}{\partial} \left(\frac{\sin a}{\cos a} \right) &= \frac{1}{\cos a} \cdot \frac{a}{\partial} (\sin a) - \frac{\sin a}{(\cos a)^2} \cdot \frac{a}{\partial} (\cos a) \\ &= \frac{1}{\cos a} \cdot \cos a \cdot \frac{a}{\partial} a - \frac{\sin a}{(\cos a)^2} \cdot (-\sin a) \cdot \frac{a}{\partial} a \\ &= 1 + \frac{(\sin a)^2}{(\cos a)^2} \\ &= \frac{1}{(\cos a)^2} \end{aligned}$$

We've just calculated the derivative of another famous function.

$$\tan a = \frac{\sin a}{\cos a}$$

Although they have similar names, the *trigonometric tangent function*, $\tan a$, should not be confused with the [tangent metric](#).

15.4 Polar Coordinates

On the unit circle we saw that the cosine and sine functions tell us how the x and y -coordinates depend on an angular coordinate a . We can move off the unit circle by introducing a radial coordinate r .

Definition.

In the (x, y) -coordinate plane we may locate any point using a *radius* $r : \text{Number}$ and an *angle* $a : \text{Number}$.

$$x = r \cdot \cos a$$

$$y = r \cdot \sin a$$

We'll call the variables r and a *polar coordinates* to distinguish them from standard coordinates, x and y .

To keep things simple we'll only allow positive r -values, $r \geq 0$, when working with polar coordinates.

Theorem.

Suppose $r : \text{Number}$ is the radial polar coordinate. Then

$$x^2 + y^2 = r^2.$$

Why?

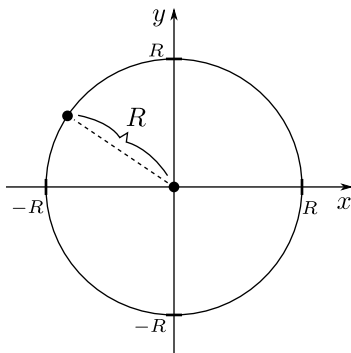
This is easy to check.

$$x^2 + y^2 = (r \cdot \cos a)^2 + (r \cdot \sin a)^2 = r^2 \cdot ((\cos a)^2 + (\sin a)^2) = r^2$$

Suppose $R : \text{Number}$ is a constant and $R > 0$. Fixing the radial coordinate

$$\{r = R\}$$

describes the circle with radius R .

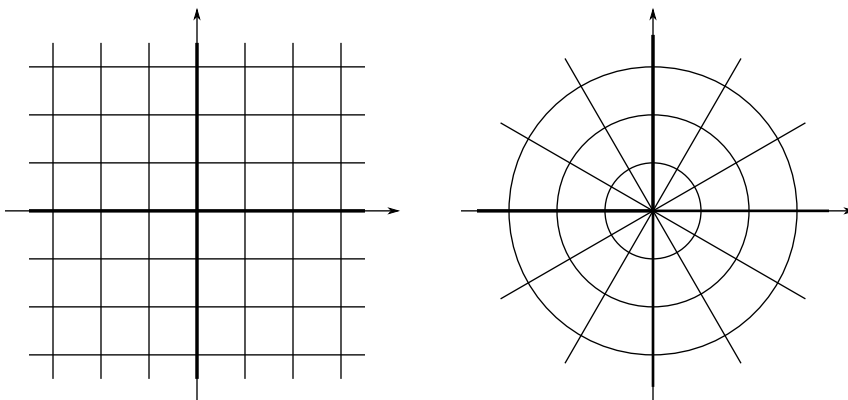


If we instead set r equal to zero

$$\{r = 0\}$$

we find the origin, only.

Our graph paper looks quite different when we switch from standard coordinates to polar coordinates. Each positive radius, $r > 0$, gives a circle, and each angle α gives a ray that starts at the origin.



Example.

Consider the point p : Point $_{(x,y)}$ with polar coordinates

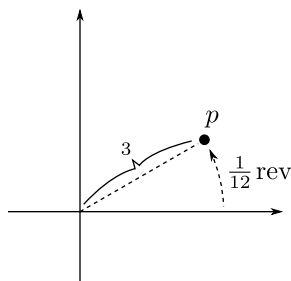
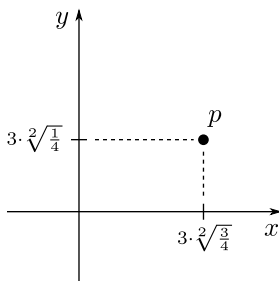
$$r = 3 \quad \text{and} \quad \alpha = \frac{1}{12} \text{ rev} .$$

Let's find the standard coordinates for this point.

$$x = 3 \cdot \cos\left(\frac{1}{12} \text{ rev}\right) = 3 \cdot \sqrt[2]{\frac{3}{4}}$$

$$y = 3 \cdot \sin\left(\frac{1}{12} \text{ rev}\right) = 3 \cdot \sqrt[2]{\frac{1}{4}}$$

We can locate our point in the plane using standard graph paper or polar graph paper.



15.5 Polar Differentials

The goal of this section is to learn about the polar differentials $\bar{\partial}r$ and $\bar{\partial}a$ by relating them to the standard differentials $\bar{\partial}x$ and $\bar{\partial}y$.

Formulas.

We can write the polar differentials as

$$\bar{\partial}r = \frac{x}{r} \cdot \bar{\partial}x + \frac{y}{r} \cdot \bar{\partial}y \quad \text{and} \quad \bar{\partial}a = \frac{-y}{r^2} \cdot \bar{\partial}x + \frac{x}{r^2} \cdot \bar{\partial}y.$$

Why?

We relate the standard coordinates to polar coordinates using

$$x = r \cdot \cos a$$

$$y = r \cdot \sin a.$$

Let's take differentials of x and of y .

$$\bar{\partial}x = \cos a \cdot \bar{\partial}r - r \cdot \sin a \cdot \bar{\partial}a$$

$$\bar{\partial}y = \sin a \cdot \bar{\partial}r + r \cdot \cos a \cdot \bar{\partial}a$$

Our plan is to solve for $\bar{\partial}r$ and $\bar{\partial}a$. We multiply each equation by r

$$r \cdot \bar{\partial}x = r \cdot \cos a \cdot \bar{\partial}r - r^2 \cdot \sin a \cdot \bar{\partial}a$$

$$r \cdot \bar{\partial}y = r \cdot \sin a \cdot \bar{\partial}r + r^2 \cdot \cos a \cdot \bar{\partial}a$$

so that we can make substitutions.

$$r \cdot \bar{\partial}x = x \cdot \bar{\partial}r - ry \cdot \bar{\partial}a$$

$$r \cdot \bar{\partial}y = y \cdot \bar{\partial}r + rx \cdot \bar{\partial}a$$

We can now use these two equations to either eliminate the $\bar{\partial}a$ terms, or to eliminate the $\bar{\partial}r$ terms.

$$rx \cdot \bar{\partial}x + ry \cdot \bar{\partial}y = (x^2 + y^2) \cdot \bar{\partial}r$$

$$-ry \cdot \bar{\partial}x + rx \cdot \bar{\partial}y = r \cdot (y^2 + x^2) \cdot \bar{\partial}a$$

Simplifying using $x^2 + y^2 = r^2$,

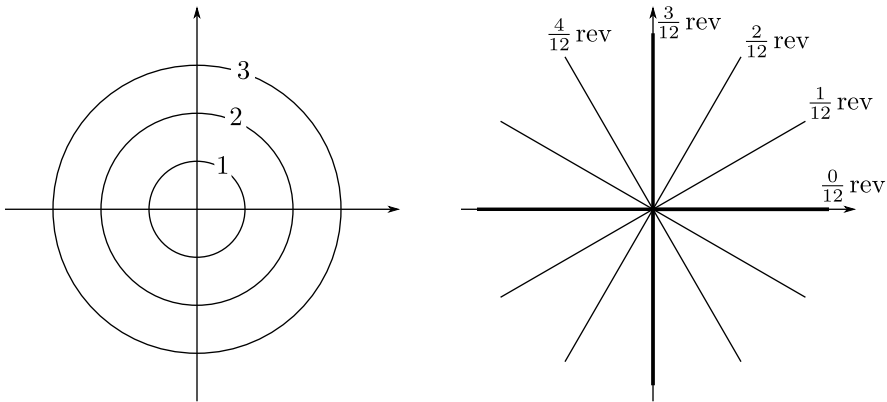
$$\begin{aligned} rx \cdot \bar{\partial}x + ry \cdot \bar{\partial}y &= r^2 \cdot \bar{\partial}r \\ -ry \cdot \bar{\partial}x + rx \cdot \bar{\partial}y &= r^3 \cdot \bar{\partial}a \end{aligned}$$

we can solve for the polar differentials!

If we calculate the magnitudes of our polar differentials, we find,

$$\text{mag } \bar{\partial}r = 1 \quad \text{and} \quad \text{mag } \bar{\partial}a = \frac{1}{r}.$$

The radial ruler $\bar{\partial}r$ is a **unit ruler**, but the angular ruler $\bar{\partial}a$ is not! This makes some sense if we look at our polar graph paper: the markings of $\bar{\partial}a$ are close together near the origin, and become further apart as we move away from the origin.



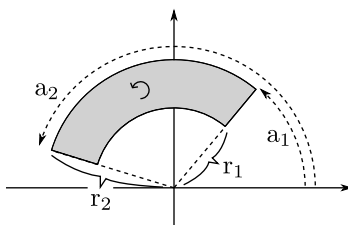
We can find a unit angular ruler by scaling. In particular,

$$\bar{\partial}r = \frac{x}{r} \cdot \bar{\partial}x + \frac{y}{r} \cdot \bar{\partial}y \quad \text{and} \quad r \cdot \bar{\partial}a = \frac{-y}{r} \cdot \bar{\partial}x + \frac{x}{r} \cdot \bar{\partial}y$$

are unit rulers.

15.6 Polar Area

A *polar box* $[r_1, r_2]_r \cdot [a_1, a_2]_a$ describes the region between two fixed radii and two fixed angles.



We'd like to use Calculus to find the area enclosed by a polar box.

Formula.

In polar coordinates, the [area polyruler](#) can be written as

$$r \cdot \bar{\partial} r \cdot \bar{\partial} a.$$

Why?

We compute.

$$\begin{aligned} r \cdot \bar{\partial} r \cdot \bar{\partial} a &= \bar{\partial} r \cdot (r \cdot \bar{\partial} a) \\ &= \left(\frac{x}{r} \cdot \bar{\epsilon}_x + \frac{y}{r} \cdot \bar{\epsilon}_y \right) \cdot \left(\frac{-y}{r} \cdot \bar{\epsilon}_x + \frac{x}{r} \cdot \bar{\epsilon}_y \right) \\ &= \frac{x^2}{r^2} \cdot \bar{\epsilon}_x \cdot \bar{\epsilon}_y - \frac{y^2}{r^2} \cdot \bar{\epsilon}_y \cdot \bar{\epsilon}_x \\ &= \frac{x^2}{r^2} \cdot \bar{\epsilon}_x \cdot \bar{\epsilon}_y + \frac{y^2}{r^2} \cdot \bar{\epsilon}_x \cdot \bar{\epsilon}_y \\ &= \frac{x^2 + y^2}{r^2} \cdot \bar{\epsilon}_x \cdot \bar{\epsilon}_y \\ &= \bar{\epsilon}_x \cdot \bar{\epsilon}_y \end{aligned}$$

Sure enough, we've found the area polyruler.

This formula will let us compute the area of a polar box!

Theorem.

Consider the polar box

$$[r_1, r_2]_r \cdot [a_1, a_2]_a$$

where

$$r_1 : \text{Number}, \quad r_2 : \text{Number}, \quad a_1 : \text{Number}, \quad a_2 : \text{Number}$$

are constants. The polar box has area

$$\frac{r_2^2 - r_1^2}{2} \cdot (a_2 - a_1).$$

Why?

We can measure the area by integrating the area polyruler.

$$\begin{aligned} A &= \left\langle r \cdot \bar{\partial} r \cdot \bar{\partial} a \int^{(r,a)} [r_1, r_2]_r \cdot [a_1, a_2]_a \right\rangle \\ &= \left\langle \bar{\partial} \left(\frac{r^2}{2} \cdot \bar{\partial} a \right) \int^{(r,a)} [r_1, r_2]_r \cdot [a_1, a_2]_a \right\rangle \\ &= \left\langle \frac{r^2}{2} \cdot \bar{\partial} a \int^{(r,a)} \bar{\partial} ([r_1, r_2]_r \cdot [a_1, a_2]_a) \right\rangle \\ &= \left\langle \frac{r^2}{2} \cdot \bar{\partial} a \int^{(r,a)} \bar{\partial} [r_1, r_2]_r \cdot [a_1, a_2]_a \right\rangle - \left\langle \frac{r^2}{2} \cdot \bar{\partial} a \int^{(r,a)} [r_1, r_2]_r \cdot \bar{\partial} [a_1, a_2]_a \right\rangle \end{aligned}$$

The integral on the right will be zero because $\bar{\partial} a_2 = 0$ and $\bar{\partial} a_1 = 0$. Let's continue our

calculation with the integral on the left.

$$\begin{aligned}
A &= \left\langle \frac{r^2}{2} \cdot \bar{\partial} a \int^{(r,a)} \ddot{\partial} [r_1, r_2]_r \cdot [a_1, a_2]_a \right\rangle \\
&= \left\langle \frac{r^2}{2} \cdot \bar{\partial} a \int^{(r,a)} [r_2]_r \cdot [a_1, a_2]_a \right\rangle - \left\langle \frac{r^2}{2} \cdot \bar{\partial} a \int^{(r,a)} [r_1]_r \cdot [a_1, a_2]_a \right\rangle \\
&= \left\langle \frac{r_2^2}{2} \cdot \bar{\partial} a \int^a [a_1, a_2]_a \right\rangle - \left\langle \frac{r_1^2}{2} \cdot \bar{\partial} a \int^a [a_1, a_2]_a \right\rangle \\
&= \frac{r_2^2}{2} \cdot \left\langle \bar{\partial} a \int^a [a_1, a_2]_a \right\rangle - \frac{r_1^2}{2} \cdot \left\langle \bar{\partial} a \int^a [a_1, a_2]_a \right\rangle \\
&= \frac{r_2^2 - r_1^2}{2} \cdot \left\langle \bar{\partial} a \int^a [a_1, a_2]_a \right\rangle \\
&= \frac{r_2^2 - r_1^2}{2} \cdot \left\langle a \int^a \ddot{\partial} [a_1, a_2]_a \right\rangle \\
&= \frac{r_2^2 - r_1^2}{2} \cdot (a_2 - a_1)
\end{aligned}$$

Our theorem gives us a formula for the area enclosed within a circle.

Formula.

The area inside a circle with radius R is

$$\frac{R^2}{2} \cdot \text{rev}.$$

Why?

We can describe the region as the polar box $[0, R]_r \cdot [0, \text{rev}]_a$. We just apply our theorem!

Appendices

Appendix A

Are Variables Dummies?

There are two schools of thought when using variables in mathematics.

Dummy Variables.

Input variables are used merely as placeholders in function definitions. Any name for a variable is as good as any other. We prefer to use function names over variables when possible. Output variables are largely unnecessary.

Meaningful Variables.

A variable should be given meaning when possible. Choosing good variable names helps simplify notation. We should only reuse a variable name when that variable refers to the same value: it is a mistake to have both $x = 2$ and $x = 5$ at the same time in the same context.

Both schools of thought on this matter are perfectly valid! That said, the author tends to prefer using meaningful variables. Calculus is already difficult enough, and so we should be on the lookout for ways to be more easily understood. Let's take a look at how we compose functions using meaningful variables, and how this works using dummy variables.

Example.

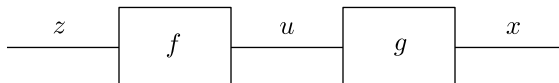
Consider the functions

$$\begin{array}{ll} z \stackrel{f}{\leftarrow} u & u \stackrel{g}{\leftarrow} x \\ z = 3u & \text{and} \quad u = x^2 + 1. \end{array}$$

Let see how z depends on x by making a substitution.

$$z = 3 \cdot (x^2 + 1)$$

We'll call u an *intermediate variable* because it is an input variable for f and an output variable for g .



In the previous example our intermediate variable, u , is a meaningful variable: reusing the variable name allows us to infer what substitution to make. Let's redo the example using dummy variables to see how that style works.

Example.

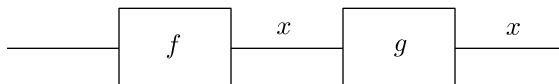
Consider the functions

$$f\ x = 3x \quad \text{and} \quad g\ x = x^2 + 1$$

where we regard both uses of x as dummy variables. Let's find $f\ g\ x$ by substituting.

$$f\ g\ x = f\ (x^2 + 1) = 3 \cdot (x^2 + 1)$$

This works, but it would be very easy to mix something up. Let's see why by looking at the schematic for this composite.



We've used the same variable name x for both the input and the output of g . In situations like this, we'll prefer to choose distinct variable names to help reduce confusion.

Appendix B

The Natural Exponential

Let's look at exponential functions in greater detail so that we can explain the differential law for the natural exponential.

Theorem.

A function $z \leftarrow^f x$ is exponential if it satisfies the rules:

1. $f\ 0 = 1$.
2. $f(m + n) = f\ m \cdot f\ n$ for any numbers $m : \text{Number}$ and $n : \text{Number}$.
3. The function has positive output, $z > 0$, only.

Why?

Let's start by defining the constant b as the output corresponding to the input $x = 1$.

$$b = f\ 1$$

We'd like to show that our function takes the form

$$z = b^x.$$

Let's check a few x -values.

$$f\ 0 = 1 = b^0$$

$$f\ 1 = b = b^1$$

$$f\ 2 = f(1 + 1) = (f\ 1) \cdot (f\ 1) = b^2$$

$$f\ 3 = f(2 + 1) = (f\ 2) \cdot (f\ 1) = (f\ 1) \cdot (f\ 1) \cdot (f\ 1) = b^3$$

Sure enough, using the assumed rules, we're finding

$$f\ x = b^x.$$

You may wish to check that the formula holds for other inputs, say $x = -1$, or $x = 1/2$, or $x = 1/3$.

We can use the previous theorem to give a new formula for the natural exponential.

Formula.

We may write the natural exponential function as the sum of infinitely-many terms

$$\exp x = 1 + \frac{1}{1!} \cdot x + \frac{1}{2!} \cdot x^2 + \frac{1}{3!} \cdot x^3 + \cdots$$

where $j!$ is j *factorial*.

$$j! = 1 \cdot 2 \cdot 3 \cdots j$$

Why?

Let's define a new function f which is equal to the right hand side of the above equation.

$$z \stackrel{f}{\leftarrow} x$$

$$z = 1 + \frac{1}{1!} \cdot x + \frac{1}{2!} \cdot x^2 + \frac{1}{3!} \cdot x^3 + \cdots$$

We wish to show that f is equal to the natural exponential. To do this, we'll apply our theorem. Let's check the three assumptions.

1. We compute.

$$\begin{aligned} f(0) &= 1 + \frac{1}{1!} \cdot 0 + \frac{1}{2!} \cdot 0^2 + \frac{1}{3!} \cdot 0^3 + \dots \\ &= 1 \end{aligned}$$

2. Let's check that f takes a sum of inputs to a product of outputs.

$$f(m+n) = f(m) \cdot f(n)$$

We'll begin by looking at the left hand side.

$$\begin{aligned} f(m+n) &= 1 + \frac{(m+n)}{1} + \frac{(m+n)^2}{2} + \frac{(m+n)^3}{6} + \dots \\ &= 1 + m + n + \frac{m^2}{2} + mn + \frac{n^2}{2} + \frac{m^3}{6} + \frac{m^2n}{2} + \frac{mn^2}{2} + \frac{n^3}{6} + \dots \end{aligned}$$

Now let's look at the right hand side.

$$f(m) \cdot f(n) = \left(1 + \frac{m}{1} + \frac{m^2}{2} + \frac{m^3}{6} + \dots\right) \cdot \left(1 + \frac{n}{1} + \frac{n^2}{2} + \frac{n^3}{6} + \dots\right)$$

When we multiply these out, we find the terms:

$$\begin{array}{ccccccc} 1 \cdot 1 & 1 \cdot \frac{n}{1} & 1 \cdot \frac{n^2}{2} & 1 \cdot \frac{n^3}{6} & \dots & & \\ \frac{m}{1} \cdot 1 & \frac{m}{1} \cdot \frac{n}{1} & \frac{m}{1} \cdot \frac{n^2}{2} & \dots & & & \\ \frac{m^2}{2} \cdot 1 & \frac{m^2}{2} \cdot \frac{n}{1} & \dots & & & & \\ \frac{m^3}{6} \cdot 1 & \dots & & & & & \\ \vdots & & & & & & \end{array}$$

Sure enough, these match the terms for the left hand side!

3. Let's check that our function f has positive output for any input. If the input is positive, $x > 0$, then the output is positive because it is the sum of positive terms.

$$f x = 1 + \frac{x}{1} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

If the input is negative, it is not at all obvious from the definition that the output will be positive. Instead, notice that the calculation

$$f x \cdot f(-x) = f(x - x) = f 0 = 1$$

shows that $f x$ and $f(-x)$ are reciprocals. Since our output is positive for $x > 0$, it must also be positive for $x < 0$.

We've shown that our function f meets all three of the assumptions for the theorem, so we may conclude that f is exponential. We find the base by evaluating our function at one.

$$\begin{aligned} f 1 &= 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots \\ &= \text{nat} \end{aligned}$$

We've found the natural base!

We can now calculate the differential law for the natural exponential.

Law.

The differential law for the natural exponential function states

$$\partial(\exp u) = \exp u \cdot \partial u$$

where u : Number is a number.

Why?

Let's calculate the u -derivative.

$$\begin{aligned}\partial^u(\exp u) &= \partial^u \left(1 + \frac{1}{1!} \cdot u + \frac{1}{2!} \cdot u^2 + \frac{1}{3!} \cdot u^3 + \frac{1}{4!} \cdot u^4 + \cdots \right) \\&= \partial^u(1) + \partial^u \left(\frac{1}{1!} \cdot u \right) + \partial^u \left(\frac{1}{2!} \cdot u^2 \right) + \partial^u \left(\frac{1}{3!} \cdot u^3 \right) + \partial^u \left(\frac{1}{4!} \cdot u^4 \right) + \cdots \\&= \frac{1}{1!} \cdot \partial^u(u) + \frac{1}{2!} \cdot \partial^u(u^2) + \frac{1}{3!} \cdot \partial^u(u^3) + \frac{1}{4!} \cdot \partial^u(u^4) + \cdots \\&= 1 + \frac{1}{2!} \cdot 2u + \frac{1}{3!} \cdot 3u^2 + \frac{1}{4!} \cdot 4u^3 + \cdots \\&= 1 + \frac{1}{1!} \cdot u + \frac{1}{2!} \cdot u^2 + \frac{1}{3!} \cdot u^3 + \cdots \\&= \exp u\end{aligned}$$

The differential law follows.

Appendix C

Multi-Components

A *multi-component* is a component that allows any number of input wires. In this appendix we'll look at two examples of multi-components: sum and product. Also we'll take a look at duplication. Although duplication is not a multi-component, it is something quite similar: duplication will allow any number of output wires.

Sums

As we well know, addition takes two inputs and produces an output:

$$z = x + y.$$

By nesting additions, we can add more than two numbers together.

$$(5 + 1) + 6 = 12$$

We know from experience that it makes no difference what order or groupings we use to add the numbers.

$$5 + (1 + 6) = 12$$

$$(6 + 1) + 5 = 12$$

$$1 + (5 + 6) = 12$$

We'll usually leave out parentheses since they are unnecessary.

$$5 + 1 + 6 = 12$$

Let's introduce another notation for the *sum* of a list of numbers.

$$\text{sum}(5, 1, 6) = 12$$

Sums makes sense for lists of any length.

$$\begin{aligned}\text{sum}() &= 0 \\ \text{sum}(a) &= a \\ \text{sum}(a, b) &= a + b \\ \text{sum}(a, b, c) &= a + b + c \\ &\vdots\end{aligned}$$

The following sum law assures us that we've defined a sensible multi-component!

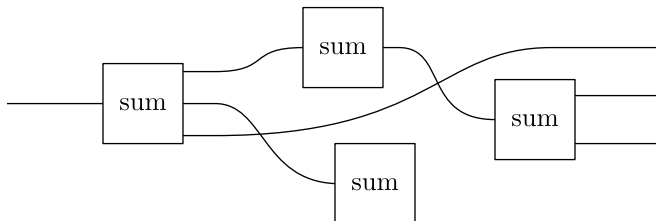
Law.

Any composite of sums is, itself, a sum.

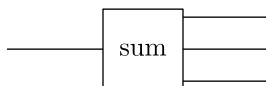
Let's see how this works by looking at a schematic.

Example.

Consider the following composite of sums.



Without looking at this schematic in any detail, we know that it must be equivalent to the sum with three input wires.



Sure enough, using a little algebra we can confirm:

$$(y + z) + 0 + x = x + y + z.$$

Products

Just as we add a list of numbers to get the sum, we multiply a list of numbers to get the *product*.

$$\begin{aligned}\text{product}() &= 1 \\ \text{product}(a) &= a \\ \text{product}(a, b) &= a \cdot b \\ \text{product}(a, b, c) &= a \cdot b \cdot c \\ &\vdots\end{aligned}$$

These definitions ensure that the following product law holds.

Law.

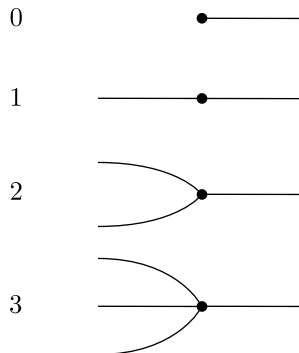
Any composite of products is, itself, a product.

In short, we may regroup and reorder a product of numbers without fear of changing the result.

$$y \cdot (1 \cdot (x \cdot w)) \cdot 1 \cdot z = w \cdot x \cdot y \cdot z$$

Duplication

We've seen how a [splitter](#) will take an input and produce two copies as output. More generally, we may *duplicate* an input on any number of output wires.



One-fold duplication just reproduces its input as output: this is equivalent to a plain wire. Zero-fold duplication produces no output. We'll call 0-fold duplication a *forgetter*. Similar to sums and products, we have a duplication law!

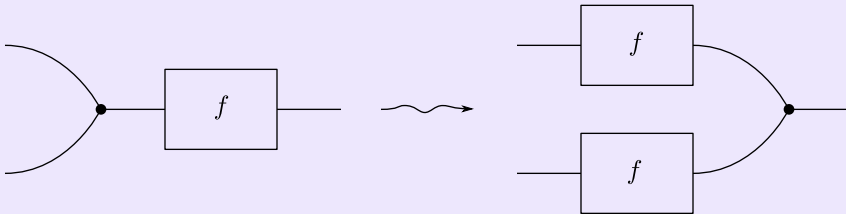
Law.

Any composite of duplications is, itself, a duplication.

While we're investigating splitters in some detail, let's take a look at one more law.

Law.

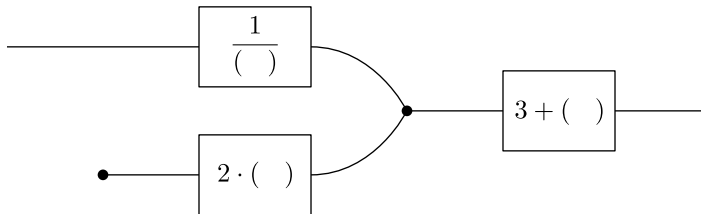
Pushing a splitter to the right past a component makes copies of that component.



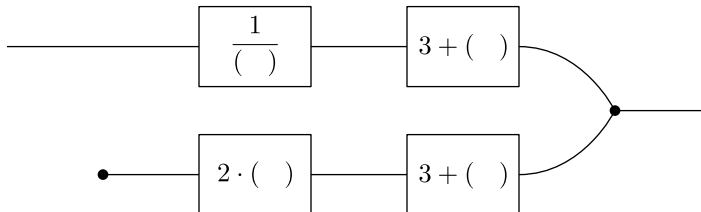
We can ensure that splitters always live at the far right of our schematics by applying this law repeatedly.

Example.

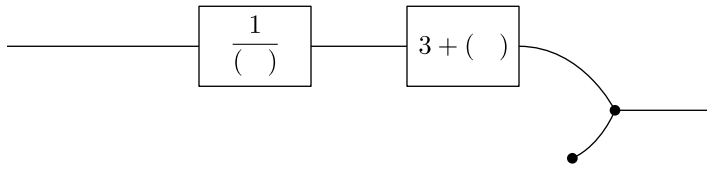
Let's clean up the following schematic.



We push the splitter right, duplicating the three-adder.



Now we push the forgetter right, removing components as we go.



The adjacent forgetter and splitter are equivalent to a plain wire.



Appendix D

Independence

Recall that the partials $\overset{x}{\partial}$ and $\overset{y}{\partial}$ assume that the input variables x and y are independent.

$$\overset{x}{\partial}y = 0 \quad \text{and} \quad \overset{y}{\partial}x = 0$$

The independence laws allow us to take shortcuts when computing partials: we can treat certain expressions as constant. I hope to convince you that taking shortcuts is not always a good thing!

Example.

Consider the function described by

$$z \leftarrow (x, y)$$

$$z = x^2y.$$

The partials for this function are

$$\overset{x}{\partial}z = 2xy$$

$$\overset{y}{\partial}z = x^2.$$

If we now assert dependence,

$$y \leftarrow x$$

$$y = x^2$$

then a substitution let's us rewrite the x -partial as

$$\overset{x}{\partial}z = 2xy = 2x^3.$$

But this is not correct! A direct calculation gives the actual x -partial.

$$z = x^2y = x^4$$

$$\frac{\partial}{\partial x} z = \frac{\partial}{\partial x} (x^4) = 4x^3$$

Where did our example go wrong? Just by using partials, we implicitly assumed that x and y were independent. We then broke our assumption when we defined $y = x^2$. The differential does not assume independence, so it won't run into this same problem.

Example.

Let's redo the above example

$$z \leftarrow (x, y)$$

$$z = x^2y$$

but this time using the differential.

$$\begin{aligned} \partial z &= \partial(x^2y) \\ &= y \cdot \partial(x^2) + x^2 \cdot \partial y \\ &= 2xy \cdot \partial x + x^2 \cdot \partial y \end{aligned}$$

Again, let's assert dependence.

$$y \leftarrow x$$

$$y = x^2$$

We continue our computation by substituting x^2 for y .

$$\begin{aligned} \partial z &= 2xy \cdot \partial x + x^2 \cdot \partial y \\ &= 2x \cdot x^2 \cdot \partial x + x^2 \cdot \partial(x^2) \\ &= 2x^3 \cdot \partial x + x^2 \cdot 2x \cdot \partial x \\ &= 2x^3 \cdot \partial x + 2x^3 \cdot \partial x \\ &= 4x^3 \cdot \partial x \end{aligned}$$

Sure enough, this agrees with the direct calculation.

$$z = x^4$$

$$\partial z = \partial(x^4) = 4x^3 \cdot \partial x$$

Appendix E

Applying a Metric

Suppose

$$m : \text{Metric}_{(x,y)}$$

$$m = a \cdot \hat{\varepsilon}_x + b \cdot \hat{\varepsilon}_y$$

is a metric, and

$$r : \text{Number} \quad \text{and} \quad s : \text{Number}$$

are numbers. By *applying* the metric m we produce a number

$$m[r; s] : \text{Number}$$

$$m[r; s] = a \cdot r + b \cdot s.$$

Example.

Consider the metric

$$m : \text{Metric}_{(x,y)}$$

$$m = 2 \cdot \hat{\varepsilon}_x + 3 \cdot \hat{\varepsilon}_y.$$

Let's apply our metric for a few different choices of values.

$$m[2; 0] = 2 \cdot 2 + 3 \cdot 0 = 4$$

$$m[0; 2] = 2 \cdot 0 + 3 \cdot 2 = 6$$

$$m[1; 1] = 2 \cdot 1 + 3 \cdot 1 = 5$$

In each case we just substitute the value on the left for $\hat{\varepsilon}_x$ and the value on the right for $\hat{\varepsilon}_y$.

The **tangent metric** calculates changes for a function's **tangent plane**.

Theorem.

Suppose

$$\Delta_x : \text{Number} \quad \text{and} \quad \Delta_y : \text{Number}$$

are numbers that represent changes in input. Then

$$\text{tangent } z [\Delta_x; \Delta_y]$$

is the change in height for the tangent plane.

Why?

When we substitute Δ_x and Δ_y into the tangent metric, we find the equation for the tangent plane.

$$\begin{aligned} \text{tangent } z [\Delta_x; \Delta_y] &= \overset{x}{\text{slope } z} \cdot \Delta_x + \overset{y}{\text{slope } z} \cdot \Delta_y \\ &= \overset{x}{\text{slope } z} \cdot (x_T - x) + \overset{y}{\text{slope } z} \cdot (y_T - y) \\ &= z_T - z \end{aligned}$$

Let's see how this works!

Example.

Consider the function

$$\begin{aligned} z &\leftarrow (x, y) \\ z &= x^2y - xy^2. \end{aligned}$$

Let's find the tangent metric and localize it at the point $p = (3, 1)$.

$$\begin{aligned} \text{tangent } z &= \hat{\partial} z \\ &= (2xy - y^2) \cdot \hat{\varepsilon}_x + (x^2 - 2xy) \cdot \hat{\varepsilon}_y \end{aligned}$$

Localizing at p , we find

$$\text{tangent } z = 5 \cdot \hat{e}_x - 3 \cdot \hat{e}_y.$$

By applying the tangent metric, we can find changes in height due to changes in input.

$$\text{tangent } z [2; 0] = 10$$

$$\text{tangent } z [0; 2] = -6$$

$$\text{tangent } z [1; 1] = 2$$

If we do not travel far from the center of the tangent plane,

$$\Delta_x \approx 0 \quad \text{and} \quad \Delta_y \approx 0,$$

then a change of height for the surface is approximately equal to the change of height for the tangent plane

$$\Delta_z \approx \text{tangent } z [\Delta_x; \Delta_y].$$

Appendix F

Ruler Explanations

Here we'll provide explanations for a few theorems that deal with the geometry of rulers. Unfortunately, these explanations are fairly complicated: working through each will not necessarily lead to a greater understanding.

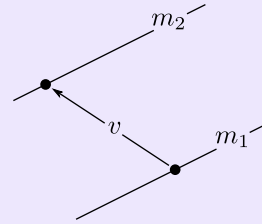
Theorem.

Consider a ruler and a vector.

$$r : \text{Ruler}_{(x,y)} \quad v : \text{Vector}_{(x,y)}$$

If the vector v travels from a marking m_1 to a marking m_2 , then the measurement is equal to the difference of the markings.

$$\langle r \mid v \rangle = m_2 - m_1$$



Why?

Suppose the ruler $r = a \cdot \bar{\epsilon}_x + b \cdot \bar{\epsilon}_y$ is centered at the point $p = (x_p, y_p)$, and suppose the vector $v : \text{Vector}_{(x,y)}$ travels from $p_1 = (x_1, y_1)$ to $p_2 = (x_2, y_2)$. On the one hand, we can start by rewriting the difference of markings.

$$\begin{aligned} m_2 - m_1 &= (a \cdot (x_2 - x_p) + b \cdot (y_2 - y_p)) - (a \cdot (x_1 - x_p) + b \cdot (y_1 - y_p)) \\ &= a \cdot (x_2 - x_1) + b \cdot (y_2 - y_1) \end{aligned}$$

On the other, we can rewrite the measurement.

$$\begin{aligned} \langle r \mid v \rangle &= \langle r \mid \text{pos } p_2 - \text{pos } p_1 \rangle \\ &= \langle a \cdot \bar{\epsilon}_x + b \cdot \bar{\epsilon}_y \mid (x_2 - x_1) \cdot \vec{\epsilon}_x + (y_2 - y_1) \cdot \vec{\epsilon}_y \rangle \\ &= a \cdot (x_2 - x_1) + b \cdot (y_2 - y_1) \end{aligned}$$

We've found the same result!

Theorem.

Suppose $r : \text{Ruler}_{(x,y)}$ is a ruler.

- The markings of r are perpendicular to the transpose vector, $\text{tran } r$.
- The transpose, $\text{tran } r$, points in the direction of increasing markings.

Why?

Let's write our ruler r and its transpose, $v = \text{tran } r$, in their standard forms

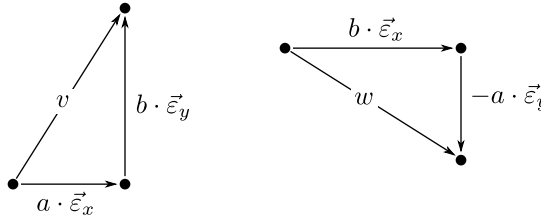
$$r = a \cdot \bar{\epsilon}_x + b \cdot \bar{\epsilon}_y,$$

$$v = a \cdot \vec{\epsilon}_x + b \cdot \vec{\epsilon}_y.$$

Looking to geometry, we see that the vector

$$w = b \cdot \vec{\epsilon}_x - a \cdot \vec{\epsilon}_y$$

is perpendicular to v .



We'll have verified our first statement if we can show that w is parallel to the markings of r . In other words, it suffices to check that measuring w with r gives zero.

$$\begin{aligned} \langle r | w \rangle &= \langle a \cdot \bar{\epsilon}_x + b \cdot \bar{\epsilon}_y | b \cdot \vec{\epsilon}_x - a \cdot \vec{\epsilon}_y \rangle \\ &= ab \cdot \langle \bar{\epsilon}_x | \vec{\epsilon}_x \rangle - a^2 \cdot \langle \bar{\epsilon}_x | \vec{\epsilon}_y \rangle + b^2 \cdot \langle \bar{\epsilon}_y | \vec{\epsilon}_x \rangle - ba \cdot \langle \bar{\epsilon}_y | \vec{\epsilon}_y \rangle \\ &= ab - ba \\ &= 0. \end{aligned}$$

To explain the second statement, we can use r to measure its transpose vector, v .

$$\begin{aligned} \langle r | v \rangle &= \langle a \cdot \bar{\epsilon}_x + b \cdot \bar{\epsilon}_y | a \cdot \vec{\epsilon}_x + b \cdot \vec{\epsilon}_y \rangle \\ &= a^2 \cdot \langle \bar{\epsilon}_x | \vec{\epsilon}_x \rangle + ab \cdot \langle \bar{\epsilon}_x | \vec{\epsilon}_y \rangle + ba \cdot \langle \bar{\epsilon}_y | \vec{\epsilon}_x \rangle + b^2 \cdot \langle \bar{\epsilon}_y | \vec{\epsilon}_y \rangle \\ &= a^2 + b^2 \end{aligned}$$

This measurement is positive and so the transpose must point from smaller to larger markings.

Theorem.

A ruler's magnitude is equal to the measurement the ruler gives for the radius of the unit circle.

Why?

Let's consider a ruler

$$r : \text{Ruler}_{(x,y)}$$

and its transpose vector

$$v : \text{Vector}_{(x,y)}$$

$$v = \text{tran } r.$$

The transpose is perpendicular to the markings of r . Dividing v by its magnitude

$$v' = \frac{v}{\text{mag } v}$$

we find a unit vector pointing in the same direction as v . To measure the unit circle with our ruler r , we can measure the unit vector v' .

$$\langle r \mid v' \rangle = \left\langle r \mid \frac{v}{\text{mag } v} \right\rangle = \frac{\langle r \mid v \rangle}{\text{mag } v} = \frac{\langle r \mid \text{tran } r \rangle}{\text{mag } \text{tran } r} = \frac{(\text{mag } r)^2}{\text{mag } r} = \text{mag } r$$

We've found the magnitude of r .

Appendix G

Rulers in Space

Let's sketch how our theory of rulers extends to three-dimensional space. In (x, y, z) -coordinate space, we have three standard rulers.

$$\bar{\epsilon}_x : \text{Ruler}_{(x,y,z)} \quad \bar{\epsilon}_y : \text{Ruler}_{(x,y,z)} \quad \bar{\epsilon}_z : \text{Ruler}_{(x,y,z)}$$

We have an anti-commutative multiplication

$$s \cdot r = -r \cdot s$$

where $r : \text{Ruler}_{(x,y,z)}$ and $s : \text{Ruler}_{(x,y,z)}$ are any two rulers. We can write polyrulers in standard forms.

degree	polyruler standard form
0	a
1	$a \cdot \bar{\epsilon}_x + b \cdot \bar{\epsilon}_y + c \cdot \bar{\epsilon}_z$
2	$a \cdot \bar{\epsilon}_y \cdot \bar{\epsilon}_z + b \cdot \bar{\epsilon}_z \cdot \bar{\epsilon}_x + c \cdot \bar{\epsilon}_x \cdot \bar{\epsilon}_y$
3	$a \cdot \bar{\epsilon}_x \cdot \bar{\epsilon}_y \cdot \bar{\epsilon}_z$

We produce polyrulers with degree three as the product of the rulers. The polyruler

$$\bar{\epsilon}_x \cdot \bar{\epsilon}_y \cdot \bar{\epsilon}_z$$

is the volume polyruler for (x, y, z) -coordinate space.

Formulas.

The following formulas describe how to measure a polyvector with a polyruler.

$$\begin{aligned}\langle a_r \mid a_v \rangle &= a_r a_v \\ \langle a_r \cdot \bar{\varepsilon}_x + b_r \cdot \bar{\varepsilon}_y + c_r \cdot \bar{\varepsilon}_z \mid a_v \cdot \vec{\varepsilon}_x + b_v \cdot \vec{\varepsilon}_y + c_v \cdot \vec{\varepsilon}_z \rangle &= a_r a_v + b_r b_v + c_r c_v \\ \langle a_r \cdot \bar{\varepsilon}_y \cdot \bar{\varepsilon}_z + b_r \cdot \bar{\varepsilon}_z \cdot \bar{\varepsilon}_x + c_r \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y \mid a_v \cdot \vec{\varepsilon}_y \cdot \vec{\varepsilon}_z + b_v \cdot \vec{\varepsilon}_z \cdot \vec{\varepsilon}_x + c_v \cdot \vec{\varepsilon}_x \cdot \vec{\varepsilon}_y \rangle &= a_r a_v + b_r b_v + c_r c_v \\ \langle a_r \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y \cdot \bar{\varepsilon}_z \mid a_v \cdot \vec{\varepsilon}_x \cdot \vec{\varepsilon}_y \cdot \vec{\varepsilon}_z \rangle &= a_r a_v\end{aligned}$$

Grad, Curl, Div, and All That

The ruler differential $\bar{\partial}$ satisfies all [abstract differential laws](#), and we interpret:

$$\bar{\partial}x = \bar{\varepsilon}_x, \quad \bar{\partial}y = \bar{\varepsilon}_y, \quad \bar{\partial}z = \bar{\varepsilon}_z.$$

We may calculate the differential of polyrulers using the familiar laws

$$\begin{aligned}\bar{\partial}(p \cdot \bar{\varepsilon}_x) &= \bar{\partial}p \cdot \bar{\varepsilon}_x \\ \bar{\partial}(p \cdot \bar{\varepsilon}_y) &= \bar{\partial}p \cdot \bar{\varepsilon}_y \\ \bar{\partial}(p \cdot \bar{\varepsilon}_z) &= \bar{\partial}p \cdot \bar{\varepsilon}_z\end{aligned}$$

where $p : \text{Polyruler}_{(x,y,z)}$ is any polyruler. Or, if we prefer partials, we can find the differential using the following formulas.

Formulas.

A 0-polyruler a has differential

$$\overset{x}{\partial}a \cdot \bar{\varepsilon}_x + \overset{y}{\partial}a \cdot \bar{\varepsilon}_y + \overset{z}{\partial}a \cdot \bar{\varepsilon}_z.$$

A 1-polyruler $a \cdot \bar{\varepsilon}_x + b \cdot \bar{\varepsilon}_y + c \cdot \bar{\varepsilon}_z$ has differential

$$\left(\overset{y}{\partial}c - \overset{z}{\partial}b\right) \cdot \bar{\varepsilon}_y \cdot \bar{\varepsilon}_z + \left(\overset{z}{\partial}a - \overset{x}{\partial}c\right) \cdot \bar{\varepsilon}_z \cdot \bar{\varepsilon}_x + \left(\overset{x}{\partial}b - \overset{y}{\partial}a\right) \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y.$$

A 2-polyruler $a \cdot \bar{\varepsilon}_y \cdot \bar{\varepsilon}_z + b \cdot \bar{\varepsilon}_z \cdot \bar{\varepsilon}_x + c \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y$ has differential

$$\left(\overset{x}{\partial}a + \overset{y}{\partial}b + \overset{z}{\partial}c\right) \cdot \bar{\varepsilon}_x \cdot \bar{\varepsilon}_y \cdot \bar{\varepsilon}_z.$$

These formulas let us demonstrate an important fact: the second differential is identically zero

$$\bar{\partial}\bar{\partial}p = 0$$

for a polyruler $p : \text{Polyruler}_{(x,y,z)}$ of any degree!

Summaries

Summary of Abstract Differential Laws

The following differential rules hold for the differential in all of its forms: the derivative, partials, the metric differential, and the ruler differential. We assume that

u : Number, v : Number, c : Number

are numbers, c is constant, and any necessary restrictions are present.

function	differential law
constant	$\partial(c) = 0$
constant-adder	$\partial(u + c) = \partial u$
constant-multiplier	$\partial(c \cdot u) = c \cdot \partial u$
addition	$\partial(u + v) = \partial u + \partial v$
negation	$\partial(-v) = -\partial v$
subtraction	$\partial(u - v) = \partial u - \partial v$
multiplication	$\partial(u \cdot v) = v \cdot \partial u + u \cdot \partial v$
reciprocal	$\partial\left(\frac{1}{v}\right) = \frac{-1}{v^2} \cdot \partial v$
division	$\partial\left(\frac{u}{v}\right) = \frac{1}{v} \cdot \partial u - \frac{u}{v^2} \cdot \partial v$
squaring	$\partial(u^2) = 2u \cdot \partial u$
cubing	$\partial(u^3) = 3u^2 \cdot \partial u$

square root	$\partial(\sqrt[2]{u}) = \frac{1}{2 \cdot \sqrt[2]{u}} \cdot \partial u$
cube root	$\partial(\sqrt[3]{u}) = \frac{1}{3 \cdot (\sqrt[3]{u})^2} \cdot \partial u$
base	$\partial(v^c) = c \cdot v^{c-1} \cdot \partial v$
exponential	$\partial(c^u) = c^u \cdot \log c \cdot \partial u$
power	$\partial(v^u) = v^u \cdot \log v \cdot \partial u + u \cdot v^{u-1} \cdot \partial v$
natural exponential	$\partial(\exp u) = \exp u \cdot \partial u$
natural logarithm	$\partial(\log u) = \frac{1}{u} \cdot \partial u$
cosine	$\partial(\cos u) = -\sin u \cdot \partial u$
sine	$\partial(\sin u) = \cos u \cdot \partial u$

Summary of Vectors, Rulers, Metrics

	vector	ruler	metric
standard form	$a \cdot \vec{\epsilon}_x + b \cdot \vec{\epsilon}_y$	$a \cdot \bar{\epsilon}_x + b \cdot \bar{\epsilon}_y$	$a \cdot \hat{\epsilon}_x + b \cdot \hat{\epsilon}_y$
standards	$\vec{\epsilon}_x$ and $\vec{\epsilon}_y$	$\bar{\epsilon}_x$ and $\bar{\epsilon}_y$	$\hat{\epsilon}_x$ and $\hat{\epsilon}_y$
magnitude	length	unit circle measurement	
scaling	grows	shrinks	
operation	measurement	measurement	application
multiplication	anti-commutes	anti-commutes	commutes
poly-		length, area	slope, bend, warp
differential		$\bar{\partial}z$	$\hat{\partial}z$

Summary of Notation

name	notation	traditional notation
function	$z \overset{f}{\leftarrow} (x, y)$	$f : \mathbb{R}^2 \rightarrow \mathbb{R}$
dependence	$z \overset{\cdot}{\leftarrow} (x, y)$	$(x, y) \mapsto z$
differential	∂z	dz
derivative	$\overset{x}{\partial} z$	$\frac{dz}{dx}$
slope	$\overset{x}{\text{slope}} z$	
bend	$\overset{x}{\text{bend}} z$	
natural base	nat	e
natural exponential	exp	
natural logarithm	log	ln
partials	$\overset{x}{\partial} z$ and $\overset{y}{\partial} z$	$\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$
slopes	$\overset{x}{\text{slope}} z$ and $\overset{y}{\text{slope}} z$	
bends	$\overset{x}{\text{bend}} z$ and $\overset{y}{\text{bend}} z$	
warp	warp z	
standard metrics	$\hat{\varepsilon}_x$ and $\hat{\varepsilon}_y$	
metric differential	$\hat{\partial} z$	
tangent metric	tangent z	

Hessian polymetric	$\text{hess } z$	
saddle discriminant	$\text{disc } z$	
standard vectors	$\vec{\varepsilon}_x$ and $\vec{\varepsilon}_y$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$
standard rulers	$\bar{\varepsilon}_x$ and $\bar{\varepsilon}_y$	$\begin{bmatrix} 1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \end{bmatrix}$
measurement	$\langle r \mid v \rangle$	$\begin{bmatrix} x_r & y_r \end{bmatrix} \begin{bmatrix} x_v \\ y_v \end{bmatrix}$
magnitude	$\text{mag } v$	$\ v\ $
transpose	$\text{tran } r$	
ruler differential	$\bar{\partial} z$	
directional slope	$\text{direc } z^v$	
gradient	$\text{grad } z$	$\ \nabla z\ $
blip	$[x_0]_x$	
interval	$[x_1, x_2]_x$	$[x_1, x_2]$
boundary	$\ddot{\partial} b$	
integral	$\left\langle a \cdot \bar{\partial} x \int [x_1, x_2]_x \right\rangle$	$\int_{x_1}^{x_2} a \, dx$
position vector	$\text{pos } p$	
velocity vector	$\text{vel } z$	
speed	$\text{speed } z$	
one revolution	rev	2π