Some Important Math You Missed in Undergraduate Physics

Michael Astwood

December 2018 Edition 0.3

Acknowledgements
I would like to thank William, Akshay, Oz, and Francois for showing their enthusiasm and providing feedback for this project (as well as doing some writing themselves).

1 Introduction

This is meant to be a document teaching supplementary useful mathematics for non mathematical physics students (and those mathematical physics who did not take extra courses).

There are a lot of topics which are glossed over or simply not covered at all in your average physics curriculum. The mathematics covered in this book will be covered briefly, with as many references to outside literature for further reading as possible. Due to the absolute bucketload of content contained in this book, it will be quite dense. **Be prepared to look things up in other books, look for exercises in other books, and look online for guidance. This book is simply meant to familiarize you with important concepts. It is not meant to be a solitary source of knowledge that you can learn everything from.**

There will be a few exercises at the end of each chapter which are necessary for understanding the material in this book. Do not skip any of them - they cover topics which are important but often used as homework questions in mathematics courses. If you would like to learn each of these topics properly you *need* to go study them from full textbooks, otherwise you will get nowhere.

I was inspired to do this because of the amount of discrepancy I saw between students who had a good foundation in “abstract” math and those who only knew what they needed to do calculations. Many people seem almost afraid of learning “abstract” mathematics, because either they believe it’s not useful, or they believe it is too hard to be worth the effort. I believe that if people are taught basic definitions and concepts from every applicable area of math, they will experience less math anxiety and feel more motivated to seek out topics in depth. **Having a growth mindset is important in every of area life, including learning mathematics. The earlier you familiarize yourself with the concepts of high level math, the easier it will become to learn that math in depth.**

1.1 How to Use This Book

This book is meant to be a brief (in many areas we will elaborate if there is no other choice) introduction to *many* concepts in mathematics that are useful in physics.

Each section will therefore contain a subsection called "further reading" dedicated to books, youtube channels, and other resources which are essential to learning said section to the depth required in a full class. At the very end of the book there is a section listing where you can find all of the exercises in each section. You can go through these and check them off if you wish.

The structure of each subsection will be that of a mathematics text book. That is to say, it will list Definition, Theorem, Proof, Example(s). In places where a proof takes too long we will defer you to the further reading section.

Some day we hope to turn this book into a youtube lecture series. Please stay tuned for these to come out.

1.2 Helpful Other Resources

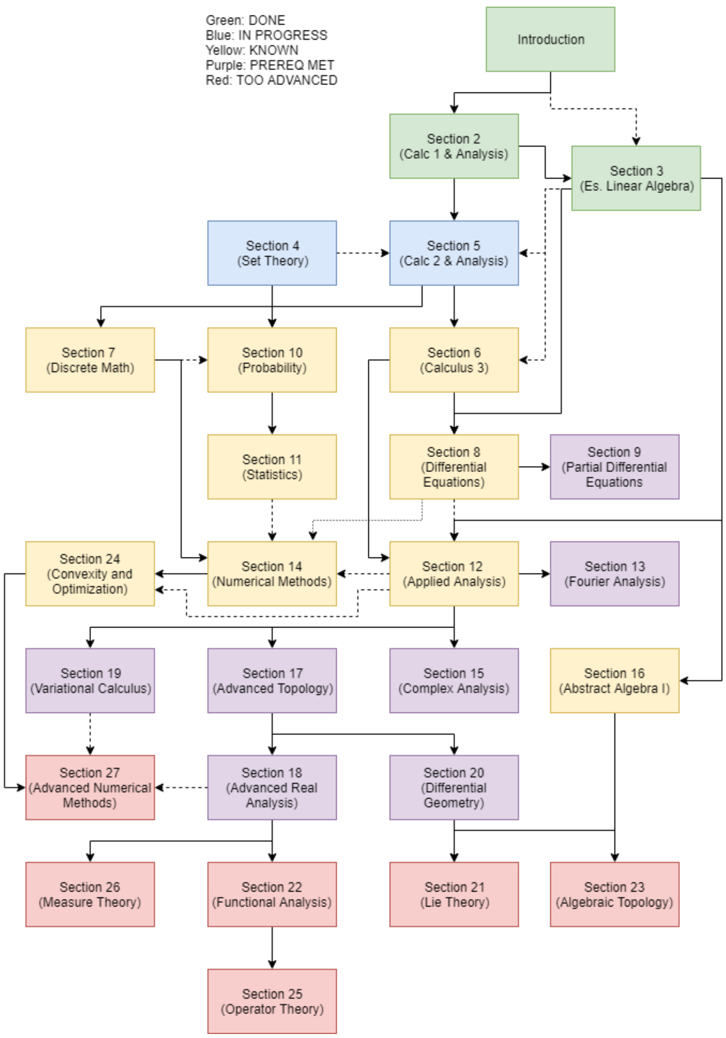
There is an amazing book out on the internet called *An Infinitely Large Napkin* by Evan Chen. It's a good place to start if you are looking for a brief overview of conceptual material but don't need the depth of knowledge required to solve problems. Hence you should not be using that book as primary study material. As Chen mentions on his website, "it is not the purpose of this book to train you to solve exercises or write proofs" [3]. On another note, *Napkin* isn't meant for physicists, it is meant for mathematicians. I recommend it to you if you have an interest in pure math and mathematical physics, but otherwise it may not be of much use.

The following is therefore a list of excellent textbooks, websites, and content creators who should help you along your physics journey. (not necessarily in order)

- *An Infinitely Large Napkin* - Chen
- *How to Prove It* - Velleman
- *Book of Proof*
- *Set Theory and Metric Spaces* - Kaplansky
- *Calculus* - Michael Spivak
- *Understanding Analysis* - Abbott
- *Real Analysis With Real Applications* - Donsig
- *Principles of Mathematical Analysis* - Rudin
- *Linear Algebra Done Right* - Axler
- *Linear Algebra Done Wrong* - Treil
- *Linear Algebra* - Friedberg et al.
- *Abstract Algebra* - Pinter
- *Algebra* - Artin
- *Abstract Algebra* - Dummit and Foote
- *Complex Analysis* - Ahlfors
- *Visual Complex Analysis* - Needham
- *Topology* - Munkres
- *Introduction to Smooth Manifolds* - Lee
- *A Comprehensive Introduction to Differential Geometry* - Spivak
- *Differential Geometry* - Kreyszig
- *Lie Algebras, Lie Groups, and Representations* - Hall
- *Functional Analysis* - Einsiedler
- *Functional Analysis* - Kolmogorov
- *Mathematical Physics* - Hassani
- *Mathematical Methods* - Boas

1.3 Flowchart of Sections

Here is a diagram showing our progress on the book and which chapters connect to which. Note how we plan to introduce chapters on subjects we do not yet know. This will take a while.



Note: Extra challenging sections are marked with a †. These are covered only briefly.

Contents

2 Calculus I and Analysis

This chapter concerns itself with the definitions, concepts, and theorems you miss in MATH127 and MATH137. As you might guess, this chapter deals with material from MATH147 and the basics of AMATH331. We focus on preparing you for AMATH331 and further applied analysis courses in order to keep the book physics-related. If you have finished MATH127 you have everything you need to understand this section (with some effort for a few of the parts marked with a †).

Mathematical analysis (analysis) is the branch of mathematics dealing with limits and related theories, such as differentiation, integration, measure, infinite series, and analytic functions. [6].

We note the difference between **Calculus** and analysis - Calculus is the set of tools and theorems specifically pertaining to differentiation and integration on real or complex functions. Analysis encompasses a much wider array of ideas, and does not only concern itself with differentiation and integration. In this chapter the form of analysis we will be talking about is **real analysis**: the analysis of real numbers and maps between real numbers.

Topology is the field of mathematics concerned with the properties of a set which are preserved under deformation.[9] A famous example is that if you squish a mug into the shape of a donut, the number of holes that the shape has is still one. "Holes" are one example of a property which is preserved by deformation. In order to do any analysis, we need to know about the topology of the set we are *analyzing*. Here are some resources you should check out for more help.

- *Calculus* - Spivak
- *Real Analysis with Real Applications* - Donsig
- *Understanding Analysis* - Abbott
- *Set Theory and Metric Spaces* - Kaplansky
- *Essence of Calculus* - 3Blue1Brown on youtube
- *Imaginary Numbers are Real* - Welch Labs on youtube
- *Principles of Mathematical Analysis* - Rudin (**Challenge Book**)

2.1 Sets

There are a few symbols which should be read as english words when you read them in a mathematical sentence. Here is a short guide:

$\forall$	"for all"
$\exists$	"there exists"
$:$	"such that"
$\iff$	"if and only if"
$\rightarrow$	"to" or "goes to"
$\in$	"in" or "is in"

Note the difference between "if and only if" and "if". If X is a true/false statement, and Y is true if X is true, then we say "Y if X".

But if Y is not true, we are not certain that Y is true, since "Y if X" does not mean "X if Y".

In the case that X if Y *and* Y if X, we say X if and only if Y.

Definition 2.1.1 (Set)**.** A set is a collection of objects. These can be numbers, vectors, functions, or other sets. We denote such a collection with squiggly brackets as such: $A = \{a_1, a_2, ...\}$

Example 2.1.1. $A = \{1, 2\}$ is a set of natural numbers.

$B = \{1.2123, \pi, 8\}$ is a set of real numbers.

$C = \{A, B\}$ is a set of sets.

$D = \{\text{hello, mydude}\}$ is a set of words. We have given the natural numbers and the real numbers properties, but we haven't really given words properties, so it's unlikely we will see sets like this one.

Figure 2.1.1: It is often helpful to imagine countable sets as a basket of fruit. You can see yourself picking fruit out of the basket, organizing them, performing operations on the fruit, and so on. I think this is a good metaphor for a set because every one of the fruits is unique even though some may be of the same species.

Definition 2.1.2 (Union)**.** The union of two sets is the collection of objects which are in both sets. This is denoted $A \cup B = \{x : x \in A \text{ or } x \in B\}$. Note that x in this definition can be in both A and B - in math, "or" implies "one or the other or both".

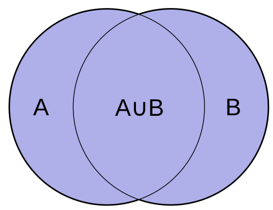


Figure 2.1.2: Diagram of the union of two sets. Note how the entire diagram is the same colour, indicating how the union encompasses the entirety of the elements in A and/or B

Definition 2.1.3 (Intersection)**.** The intersection of two sets is the collection of objects which are common to both sets. $A \cap B = \{x : x \in A \text{ and } x \in B\}$.

You will note that $A \cap B$ denotes A "and" B, while $A \cup B$ denotes A "or" B. This is quite helpful when thinking intuitively about these objects.

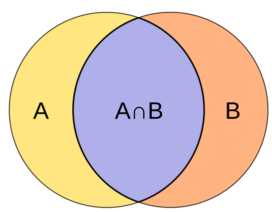


Figure 2.1.3: Diagram of the intersection between two sets. The intersection only encompasses elements which are in both A *and* B.

Definition 2.1.4. Intersections and unions can be chained together similarly to addition. We use the following notation:

$$\bigcap_{n}^{k} A_n = A_1 \cap ... \cap A_k$$

$$\bigcup_{n}^{k} A_n = A_1 \cup ... \cup A_k$$

Definition 2.1.5 (Subset)**.** A set A is called a subset of a set B if all of the elements of A are in B. If A could possibly equal B (but it is unknown whether equality holds), then we denote $A \subseteq B$. Otherwise A is called a proper subet of B: $A \subset B$.

Definition 2.1.6 (Complement)**.** If $A \subseteq U, A^c = \{x : x \in U, x \notin A\}$ denotes the "complement" of A in U. This is sometimes written as A'.

Definition 2.1.7 (Interval)**.** Let an **interval** be defined as so:
$[a, b] = \{x : a \leq x \leq b\}$
$(a, b] = \{x : a < x \leq b\}$
$[a, b) = \{x : a \leq x < b\}$
$(a, b) = \{x : a < x < b\}$

Example 2.1.2. $[a, b]^c = (-\infty, a) \cup (b, \infty)$.

Definition 2.1.8 (Difference)**.** If $A, B \subseteq U$, $B \setminus A = \{x : x \in B, x \notin A\}$ is the "difference" between A and B. This is different from the complement since A is not necessarily contained in B. This is sometimes written $B - A$.

Example 2.1.3. Let $A = [-100, 100], B = \{1\}$. Then $A - B = [-100, 1) \cup (1, 100]$. Note that $\forall x \in A - B, x \neq 1$ since if $x \in (1, 100]$, $x < 1$ and vice versa for $[-100, 1)$.

Theorem 2.1.1 (De Morgan's Laws)**.**

$$A^c \cup B^c = (A \cap B)^c$$

$$(A \cap B)^c = A^c \cup B^c$$

Proof. Let $x \in (A \cap B)^c$. Then $x \notin A \cap B$. This means that $x \notin A$ or $x \notin B$. We can rephrase that to say $x \in A^c$ or $x \in B^c$. So $x \in A^c \cup B^c$. Note the use of our nice grammatical rules defined earlier.

Conversely, let $x \in A^c \cup B^c$. then $x \in A^c$ or $x \in B^c$. Assume for contradiction that $x \notin (A \cap B)^c$. We can see that $x \notin A^c$ and $x \notin B^c$. So $x \notin (A \cup B)^c$. But this is a contradiction since $x \notin (A \cup B)^c \subseteq (A \cap B)^c$. So $X \in (A \cap B)^c$ and we are done. The proof of the other law is similar and left as an exercise to the reader.

Exercise 2.1.1. Prove that $(A \cap B)^c = A^c \cup B^c$

Definition 2.1.9. We denote the empty set as $\emptyset = \{\}$.

Exercise 2.1.2. Prove that if A is a set, then $\emptyset \subseteq A$

Definition 2.1.10 (Cardinality)**.** We denote the "cardinality" of a set A as $|A|$. If A has finitely many elements then $|A|$ is defined as the number of elements in A. The infinite case will be dealt with in the next section.

We will see in the sections about topology and analysis how finding the cardinality of arbitrary sets can be quite interesting. This is extremely important in quantum physics, where the size of certain calculations blows up to infinity. Depending on the "type" of infinity (again, see the next section) we need to perform corrections called "renormalization". It turns out that in statistical mechanics we use a lot of what is called "counting theory" - finding the cardinality of special sets in order to perform calculations.

Exercise 2.1.3. Prove the following:
(a) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
(b) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
(c) $(B \setminus A) \cap C = (B \cap C) \setminus A = B \cap (C \setminus A)$

Definition 2.1.11. The natural numbers are denoted $\mathbb{N} = \{1, 2, ...\}$
The integers are denoted similarly as $\mathbb{Z} = \{0, 1, -1, 2, -2, ...\}$.

Exercise 2.1.4. Consider a sequence of nonempty sets A_n. Let $\mathcal{A} = \{A_k : \forall n \in \mathbb{N}, A_{n-1} \subset A_n\}$.
Show that A (as defined below) is non-empty.

$$A = \bigcap_{n}^{\infty} A_n$$

Definition 2.1.12 (Cartesian Product)**.** The **cartesian product** of A and B is denoted as $A \times B = \{(a, b) : a \in A, b \in B\}$.

Example 2.1.4. $[a, b] \times [c, d]$ is a rectangle in $\mathbb{R}^2$.

Definition 2.1.13 (Power Set)**.** The **power set** of A is denoted $\mathcal{P}(A)$ and is the set of all possible subsets of A.

Example 2.1.5. The power set of $\{1, 2, 3\}$ is
$\{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$

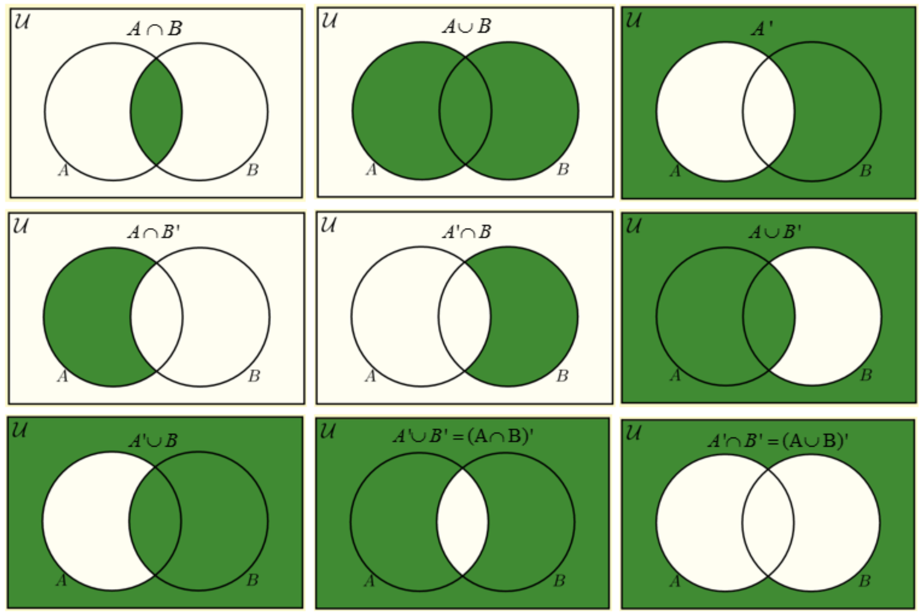


Figure 2.1.4: Venn diagrams are often used to demonstrate set theoretic concepts intuitively. Use these for help working through a problem.[7]

2.2 Relations, Classes, and Functions

Definition 2.2.1 (Relation)**.** A **binary relation** is a type of structure built from two mathematical objects which relates the two. We denote a relation $\diamond$ between a and b as follows:

If $a \sim b$ then $(a, b) \in A \sim B \subseteq A \times B$.

Recall our fruit basket from section 2.1. You could imagine that for some fruit a and b, $a \sim b$ if and only if a is of the same species as b.

Example 2.2.1. Let $a, b \in \mathbb{Z}$. We say that a divides b if $ac = b$ for some $c \in \mathbb{Z}$. We denote this relation as $a|b$. $A|B = \{(2, 4), (3, 6), (4, 64), ...\} \subset A \times B$.

Definition 2.2.2 (Equivalence Relation)**.** A relation $\sim$ is called an equivalence relation if the following properties hold:

(a) $a \sim a$ (reflexive)

(b) $a \sim b \iff b \sim a$ (symmetric)

(c) If $a \sim b$ and $b \sim c$ then $a \sim c$ (transitive)

From here on out, assume $\sim$ denotes an equivalence relation.

Exercise 2.2.1. We define the relation "=" on the rational numbers as so. Prove that it is an equivalence relation.

$$\frac{a}{b} = \frac{c}{d} \iff ac = bd$$

$$\left(\frac{a}{b}, \frac{c}{d}\right) \in \mathbb{Q}[=]\mathbb{Q}$$

Definition 2.2.3 (Equivalence Class)**.** Since $\diamond$ has the transitive property, many things can be equivalent to each other. What happens if $a_1 \diamond a_2 \diamond ... \diamond a_n$?

We denote this $\{a_1, a_2, ..., a_n\} = [a_j]$ for some arbitrary j and call $[a_j]$ the **equivalence class** of a_j under $\diamond$.

Example 2.2.2. The equivalence class of $\frac{1}{2}$ under "=" in $\mathbb{Q}$ is $[\frac{1}{2}] = \{\frac{1}{2}, \frac{2}{4}, ...\}$ and so on as you can imagine.

It is commonplace to reduce elements of an equivalence class to a sufficiently nice element (such as reducing two quarters to one half).

Definition 2.2.4 (Map)**.** A **map** m is a relation m which relates a set A to another set B. We define the function by writing the following down:

$$m : A \to B$$

.

We call A the domain of m and B the co-domain of m.

The set associated with this relation is denoted $A \times m(A) = \{(a, m(a)), ...\}$.

Definition 2.2.5 (Image)**.** Consider $(a, m(a)) \in A \times m(A)$. We call the element $m(a) \in B$ associated with a the **image of a under** m. The set $m(A)$ is called the **image** or **range** of f and is usually written $\mathrm{im}(A)$. We denote the image of a under m b so that $m(a) = b$ defines our map in a more useful way.

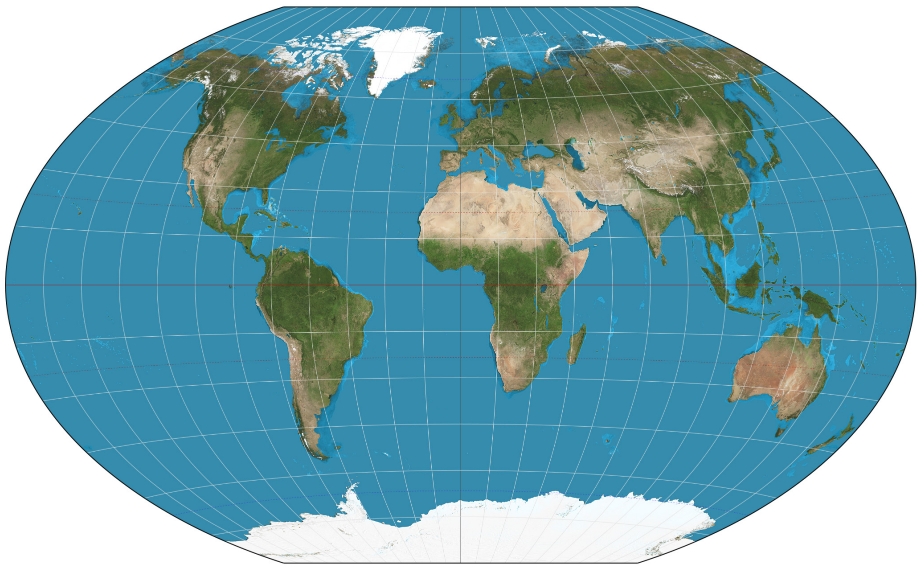


Figure 2.2.1: A world map is the image of a mathematical map called a projection which takes objects in higher dimensions and "projects" them down into a lower dimension. This is an example of the image of a Winkel-Triple projection.

2.2.1 Injections, Surjections, and Functions

Now we are going to throw a bunch of terminology at you. Be prepared to write these down, look up examples, and think about properties of these types of maps. They are extremely important for you to know.

Definition 2.2.6 (Injection)**.** If $m(a) = m(b) \Rightarrow a = b$ then we say that m is **injective** or **one to one**.

Definition 2.2.7 (Surjection)**.** If $\forall b \in B \exists a : m(a) = b$ then we say that m is **surjective** or **onto**.

Definition 2.2.8 (Bijection)**.** If m is both surjective and injective we say that m is **bijective**.

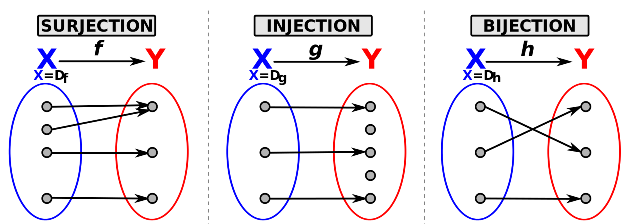


Figure 2.2.2: Diagram showing the difference between injection, surjection, and bijection.

Example 2.2.3. Recall the definition of a linear map from MATH114/136. It is clear that $m : \mathbb{R}^n \to \mathbb{R}^m$ is a surjection if and only if $\ker(m) = \{\mathbf{0}\}$.

Definition 2.2.9 (Fiber)**.** Consider $m : A \to B$. The **preimage** or **fiber** $m^{-1}(b)$ of $b \in B$ is defined as so:

$m^{-1}(b) = \{a \in B : m(a) = b\}$

Definition 2.2.10 (Inverse Map)**.** We call the map $m^{-1} : B \to A$ the **inverse map** of m.

Here are some useful properties of inverse maps:

- m is injective $\iff m^{-1}(m(x)) = x$
- m is surjective $\iff m(m^{-1}(x)) = x$
- m is bijective $\iff m$ invertible (in the usual sense)
- If m is invertible then $m^{-1}(b)$ has only one element.

Example 2.2.4. Let $m : \mathbb{R} \to \mathbb{R}$ be defined so that $m(x) = x^2$.

We can see that $m(\mathbb{R}) = [0, \infty)$.

But then the fiber $m^{-1}(x) = \{\sqrt{x}, -\sqrt{x}\}$.

This is fine, since we haven't required that the inverse map for m be a function. This motivates our next definition, since without defining "function" we might run into problems when trying to do work with maps that are not one-to-one.

Definition 2.2.11 (Function)**.** A **function** $f : A \to B$ is a map with the following property:

$f(a) = b$ for only one element b in B.

That is to say: there is only one output for every input.

Example 2.2.5. Let $f : \mathbb{R} \to \mathbb{R}$ be defined so that $f(x) = 3x$.

This is a function since if $x \in \mathbb{R}$, only one value of $y \in \mathbb{R}$ has the property that $f(x) = y$. There is only one output $3x$ for every input x.

Let $g : \mathbb{R} \to \mathbb{R}$ be defined so that $g(x) = x^2$.

This is also a function, since there is only one output x^2 for every input x.

On the other hand, for every y there are *two* x's so that $x^2 = y$, being $\sqrt{y}$ and $-\sqrt{y}$. Usually we simply take the positive branch: $g^{-1}(y) = +\sqrt{y}$.

2.3 Distance and Metrics

Definition 2.3.1 (Metric)**.** A **metric** $d : A \times A \to [0, \infty)$ is any function which obeys the following properties:

(a) $d(x, y) \geq 0$ (positive-definiteness)
(b) $d(x, y) = 0 \iff x = y$ (indiscernible identity)
(c) $d(x, y) = d(y, x)$ (symmetry)
(d) $d(x, z) \leq d(x, y) + d(y, z)$ (the triangle inequality)

It is common to call this function the "distance function" on A. This function is the defining feature of nearly every single set we work with. It determines size and distance - without it physics wouldn't exist.

Definition 2.3.2 (Metric Space)**.** A **Metric Space** (X, d) is a set X paired up with a metric $d : X^2 \to [0, \infty)$. We often just call the space "X" if it is obvious what the metric should be for that set.

Example 2.3.1. The function $d(x, y) = |x - y|$ defines a metric on $\mathbb{R}$.

Example 2.3.2. Similarly to example 2.9, $d(\mathbf{x}, \mathbf{y}) = |x_1 - y_1| + |x_2 - y_2|$ defines a metric on $\mathbb{R}^2$.

In general, $d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i|$ defines a metric on $\mathbb{R}^n$

This is called the "taxi-cab" metric (see the figure below). We call it this because roads are often arranged in grids, so the distance a taxi-cab needs to drive to get from x to y is based on the length of the sides of that grid, rather than the "straight-line" distance from x to y.

Example 2.3.3. The function $d : \mathbb{R}^2 \times \mathbb{R}^2 \to [0, \infty)$ defined by

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

Is a metric on $\mathbb{R}^2$. We will call this the *Euclidean* metric

You will be more familiar with this metric than the taxi-cab metric, because it is a more realistic picture of distance in physics.

Exercise 2.3.1. Show that the functions defined in Example 2.3.1, 2.3.2, and E2.3.3 are indeed metrics by proving the four properties we discussed in D2.25.

Theorem 2.3.1 (Triangle Inequality)**.** One of the required properties for a function to be a metric is the **Triangle Inequality**. We are restating it as a theorem because it is so important.

For a metric d, $d(x, z) \leq d(x, y) + d(y, z)$

Note how equality (and not inequality) sometimes holds for every x, y, z in a set:
$|x - z| \leq |x - y| + |y - z| = |x - z|$.

Some important corollaries of the triangle inequality are the following:

Theorem 2.3.2 (Reverse Triangle Inequality)**.** For $x, y \in \mathbb{R}$, define the size of x to be $|x| = |x - 0|$. Then the following inequality holds:

$$\big||x| - |y|\big| \leq |x - y| \text{ (reverse triangle inequality)}$$

The applications of such rigorous definitions of metrics and distance in physics will become apparent when you discuss analysis and general relativity, where your idea of distance depends on where you are in spacetime. In later reading you will have to discuss the "distance" between functions.

This is imperative for understanding functional analysis later on. Functional analysis is the analysis of infinite dimensional vector spaces, which turn out to include the variety of spaces you can construct out of sets of functions and operators.

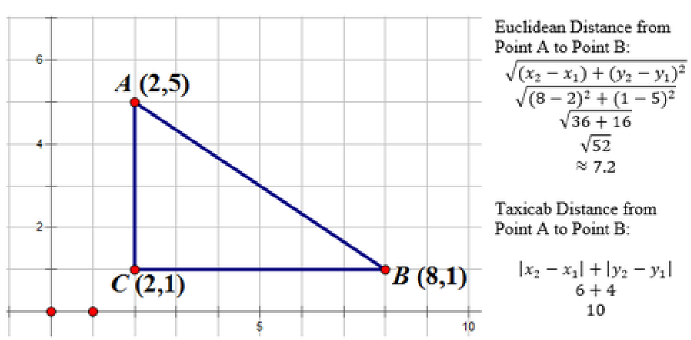


Figure 2.3.1: Diagram demonstrating the difference between the taxicab and Euclidean metrics.[4]

2.3.1 Supplementary Examples †

Example 2.3.4. Let $K = \mathcal{C}([a,b])$ be the set of continuous bounded functions on a subset $[a,b] \subseteq \mathbb{R}$.

Now let's define the following function: $d : K^2 \to \mathbb{R}^+$

$$d(f,g) = \int_a^b |f(x) - g(x)| dx$$

One can show that d defines a metric on K using the Cauchy-Schwarz inequality along with the properties of Riemann sums.

Example 2.3.5. Let $K = \mathcal{C}(\mathbb{R})$ be the set of continuous square-integrable functions on the reals.

Now let's define the following function: $d : K^2 \to \mathbb{R}^+$

$$d(f,g) = \left(\int_{-\infty}^{\infty} |f(x) - g(x)|^2 dx \right)^{\frac{1}{2}}$$

It turns out that d defines a metric on K. This metric is very useful when considering the convergence of Fourier Series.

Example 2.3.6 ($\mathcal{L}_\infty$ Distance)**.** Let $K = \mathcal{C}(\mathbb{R})$ be the set of continuous bounded functions on the reals.

Now let's define the following function: $d : K^2 \to \mathbb{R}^+$

$$d_p(f,g) = \left(\int_{-\infty}^{\infty} |f(x) - g(x)|^p dx \right)^{\frac{1}{p}}$$

$$d_\infty(f,g) = \lim_{p \to \infty} \left(\int_{-\infty}^{\infty} |f(x) - g(x)|^p dx \right)^{\frac{1}{p}}$$

$$d_\infty(f,g) = \sup_{x \in \mathbb{R}} (|f(x) - g(x)|)$$

It turns out that d_p also defines a metric on K. The metric d_∞ is useful when considering the convergence of Fourier Series.

Consider $\|\cdot\|_\infty : K \to \mathbb{R}^+$ so that $\|f\|_\infty = d_\infty(f,0), \|f - g\|_\infty = d_\infty(f,g)$. We call this the infinity-norm, maximum-norm, or sup-norm.

2.4 Open and Closed Sets

The topological properties of a space can be very important when considering domains of integration, regions of space-time, and the properties of interesting objects such as black holes. Here are some useful properties to consider:

Definition 2.4.1 (Neighborhood)**.** Consider a metric space (X, d). Let $r \in \mathbb{R}^+, p \in X$ and define the set $\mathcal{B}_r(p) = \{x \in X : d(x,p) < r\}$.
This set is called an **r-neighborhood** of p.

The interiors of circles in $\mathbb{R}^2$ and spheres in $\mathbb{R}^3$ are open balls.

Example 2.4.1. Any open interval $(a, b) \subseteq \mathbb{R}$ defines an open ball in $\mathbb{R}$:
$(a, b) = \mathcal{B}_r(c)$ for $r = |b - a|/2$ and $c = (b - a)/2$

Definition 2.4.2 (Open Set)**.** A subset $Y \subseteq X$ of a metric space (X, d) is called **open** if it obeys the following property: $\forall y \in Y, \exists \epsilon : \mathcal{B}_\epsilon(y) \subset Y$

That is to say, a set Y is open if every point in Y has a neighborhood which is contained *completely* in Y.

Example 2.4.2. Open intervals in R are open. We can see that for any $x \in (a, b)$, the neighborhood $\mathcal{B}_r(x) \subset (a, b)$ if $r < \min(|b - x|/2, |a - x|/2)$

Definition 2.4.3 (Cluster Point)**.** Consider $Y \subseteq (X, d)$. A **cluster point of** Y is a point which has the following property:
If y is a cluster point of Y, then for any choice of $r \in \mathbb{R}^+$, the set $\mathcal{B}_r(y) \cap Y$ has more than one element.

Theorem 2.4.1. If x is a cluster point in $A \subseteq B$, then x is a cluster point of B as well.

Example 2.4.3. Consider a sequence a_n in A (see section 2.5). We say that $a_n \to a$ if the following holds:
$\forall \epsilon > 0, \exists N : \forall n > N,\ d(a_n, a) < \epsilon$.
If we pick $A = \{a_n : n \in \mathbb{N}\}$, then we can say that a is a cluster point of A because every ϵ-neighborhood of a contains some $a_k \in A$. In this case, a is often called a "limit point" of A rather than a cluster point.

Definition 2.4.4 (Closed Set)**.** Consider a metric space X, d and a subset $Y \subseteq X$. We call Y **closed** if the following property holds:
If y is a cluster point of Y, then $y \in Y$.

Example 2.4.3 is very important in the next section. Usually in order to check whether a set is closed or not we will look at sequences in that set.

Theorem 2.4.2. The complement of an open set is closed. The complement of a closed set is open.

Exercise 2.4.1. Prove the previous theorem.

Some diagrams are left on the next page to help you with your intuition. Remember that drawing a diagram (usually) isn't a proof!

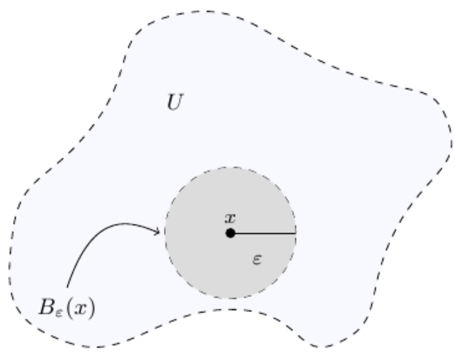


Figure 2.4.1: Diagram demonstrating the defining property of an open set[1]. You are always able to find a neighborhood of any point which does not contain points outside of U.

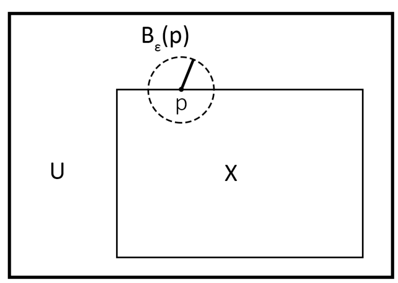


Figure 2.4.2: Diagram demonstrating the defining property of a closed set. You are allowed to place points on the boundary of the closed set X, meaning that you can always make a neighborhood around that point which contains points in $U - X$. Note that every point on the boundary of this set is a cluster point. What separates most closed sets from open sets is the fact that an open set usually does not contain the cluster points on its' boundary.

Definition 2.4.5 (Boundary)**.** The boundary of a set X is the set of all cluster points of X which are also cluster points of X^c. That is: $\partial X = \{x : \exists a_n \in X, a_n \to x \text{ and } \exists b_n \in X^c : b_n \to x\}$. These are points that can be approached from "both sides" - hence the name "boundary".

Definition 2.4.6 (Closure)**.** As you might imagine, the closure of a set X is $\overline{X} = X \cup \partial X$. This set is closed because it contains all the cluster points of X which are not in X.

Exercise 2.4.2. Prove that if X is closed then $\overline{X} = X$.

Exercise 2.4.3. Prove that the closure of X is the smallest closed set which contains X. That is: if $X \subseteq Y$ and Y is closed, then $\overline{X} \subseteq Y$ for any such choice of Y.

Definition 2.4.7 (Interior)**.** The interior of a set is all of the points in the set which are not on the boundary.

$$\text{int}(X) = X \setminus \partial X$$

2.5 Sequences

The notion of a sequence is one of the single most powerful concepts in mathematical analysis. Using sequences we can come up with some very interesting results about sets, and construct useful ways of analyzing functions and spaces.

Definition 2.5.1 (Sequence)**.** Consider a function $a : \mathbb{N} \to A$, which takes natural numbers and outputs elements of an arbitrary metric space (A, d).

Such a function is called a **sequence**. Standard notation has it so that the following are equivalent:

$a(k) = a_k$

$a(\mathbb{N}) = (a_n)_1^\infty = (a_n) \subset A$

Definition 2.5.2 (Bounded)**.** A sequence $(x_n) \subset X$ is **bounded** if there is some $M \in X$ so that $-M \leq x_n \leq M$.

Example 2.5.1. Consider the sequence $a_n = 1/n$. We can see that as n grows large, a_n gets very small, but never goes below zero. We can also see that $a_1 = 1$, $a_n < a_{n-1}$. Thus $0 < a_n \leq 1$. It is clear that a_n is bounded.

Exercise 2.5.1. Prove that the sequence $a_n = \frac{1}{n^2}$ is convergent using the epsilon-N definition of the limit. To do this, guess the value of the limit and then find N as a function of epsilon.

Example 2.5.2. The **Fibonacci Sequence** is defined as follows:

$a_0 = a_1 = 1$, $a_n = a_{n-1} + a_{n-2}$.

We can see that $(a_n) = (1, 1, 2, 3, 5, 8, 13, ...)$.

This sequence is unbounded, since $a_{n+1} > a_n$

Definition 2.5.3 (Convergent)**.** A sequence $a : \mathbb{N} \to A$ is **convergent** if the following property holds for some $a \in A$:

$\forall \epsilon > 0, \exists N : \forall n > N, d(a_n, a) < \epsilon$

In the case that a_n is convergent, we say that a_n converges to a and that $a = \lim_{n\to\infty} a_n$ is the **limit** of a_n.

Definition 2.5.4 (Increasing/Decreasing)**.** If a_n is defined so that $\forall n, a_{n+1} > a_n$, then we say that a_n is **strictly increasing**. If $a_{n+1} \geq a_n$ instead we say that it is **increasing**. The definitions for **decreasing** and **strictly decreasing** are similar.

Definition 2.5.5 (Monotone)**.** If a_n is increasing or decreasing, we say that a_n is **monotone**

Theorem 2.5.1 (Monotone Convergence Theorem)**.** If $a : \mathbb{N} \to X$ is bounded and monotone, then $\exists a \in X$ so that $a \to X$.

Exercise 2.5.2. Prove the monotone convergence theorem.

Theorem 2.5.2 (Nested Intervals Theorem)**.** Let $X_n = [a_n, b_n] \subset \mathbb{R}$, and $X_{n+1} \subseteq X_n$. Then this sequence converges to some non-empty $X = [a, b] \subset \mathbb{R}$.

Exercise 2.5.3. Prove the nested intervals theorem.

Definition 2.5.6 (Cauchy Sequence)**.** A sequence $(a_n) \in (A, d)$ is **Cauchy** if the following holds:
$\forall \epsilon > 0 \exists N : \forall n, m > N,\ d(a_n, a_m) < \epsilon$

Example 2.5.3 (Alternating Sequence)**.** The sequence (a_n) defined as shown below is not Cauchy because the space between terms is always greater than 2 - it does not tend to zero.

$$a_n = \begin{cases} -1 & n \text{ even} \\ 1 & n \text{ odd} \end{cases}$$

Theorem 2.5.3 (Bolzano-Weierstrass Theorem)**.** Every bounded sequence a_n has a convergent subsequence a_{b_k}.

Exercise 2.5.4 (Proof of Bolzano-Weierstrass Theorem)**.** The proof of Bolzano-Weierstrass is sufficiently convoluted that you probably won't get a lot out of reading it unless you do it yourself. For this purpose, the proof has been left as an exercise to the reader. Note that this is one of the most important theorems in analysis, and you should put some effort into doing it.

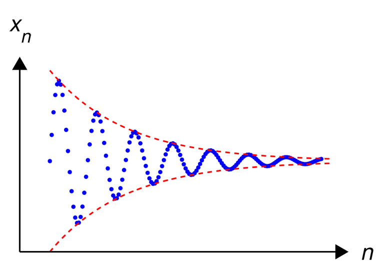


Figure 2.5.1: Diagram of how a Cauchy sequence converges. Each pair of terms gets closer and closer together as n increases.

2.6 Properties of Real Numbers

Up until now, we have taken for granted some distinct properties of the real numbers. It is important for a physics student to understand the difference between rational and irrational numbers as well as the deep implications of those differences.

Definition 2.6.1. A rational number is a number that can be written as a fraction: $q \in \mathbb{Q} \iff \exists a, b \in \mathbb{N} : q = [\frac{a}{b}]$.

Note that $\mathbb{Q}$ is actually the set of equivalence classes of fractions under $a/b \sim c/d \iff ad = bc$. Through a notational disguise we write $[a/b] = a/b$

Theorem 2.6.1. The number $2^{1/2}$ is irrational. There are no $a, b \in \mathbb{Z}$ so that $a/b = \sqrt{2}$.

Proof. Assume $\sqrt{2} = \frac{a}{b}$. Require that a and b share no prime factors without loss of generality.

Then $2 = \frac{a^2}{b^2}$.

Then $2b^2 = a^2$ implies that a^2 is even. So a is also even. Let $a = 2n$.

Then $2b^2 = 4n^2$. This implies that $b^2 = 2n^2$. But this means that b and a are both even. Since a and b share no factors, we have a contradiction. This means that a and b do not exist - $\sqrt{2}$ is irrational.

Now that you have seen that the rationals seem to be missing some numbers, we will construct the real numbers. There is a big list of properties (Tarski's Axioms) that a the real numbers need to have, so our construction needs to match those propertise.

Consider how every real number has a decimal expansion. This decimal expansion can be infinite, or it can end in zero after finitely many digits.

In any case, the sequence of decimal expansions is a sequence of rational numbers. If we take the limit of this sequence we get the real number back. What we are going to do next is define the real numbers as the equivalence class of sequences of rational numbers.

Two convergent rational sequences can be added or subtracted: $(x_n)+(y_n) = (x_n + y_n)$. Therefore we will define R as the set of all such sequences. Let $(x_n) \in R$ and $(y_n) \in R$.

We say that $(x_n) \sim (y_n)$ (x is equivalent to y) if $|x_n - y_n| \to 0$.

We can define inequalities as such:

$(x_n) \geq (y_n)$ if $x_n \sim y_n$ or if $\exists N : \forall n > N, x_n > y_n$.

Let $\mathbb{R}$ be the set of equivalence classes of $(x_n) \in R$ under the relation $\sim$: $(x_n) \sim (y_n) \iff |x_n - y_n| \to 0$.

Using the Bolzano-Weierstrass theorem, we can show that this construction satisfies Tarski's Axioms.

Definition 2.6.2 (Supremum)**.** The **least upper-bound** or **supremum** of an ordered set $X \subset Y$ is the number $x \in y$ so that:

$\forall x \in X, x \leq y$, and

$\forall b \in Y$, if b is an upper bound of X, $y \leq b$.

The least upper bound of a set is usually denoted $y = \sup(X)$

Definition 2.6.3 (Infimum)**.** The **greatest lower-bound** or **infimum** of an ordered set $X \subset Y$ is the number $x \in y$ so that:

$\forall x \in X, x \geq y$, and

$\forall b \in Y$, if b is a lower bound of X, $y \geq b$.

The infimum of a set is usually denoted $y = \inf(X)$

Example 2.6.1 (Least Upper-Bound Property)**.** Every (non-empty) subset of $\mathbb{R}$ which has an upper bound must have a supremum. In addition, every non-empty subset of $\mathbb{R}$ which has a lower bound must have an infimum.

This turns out to be one of the defining properties of $\mathbb{R}$ itself.

Example 2.6.2. Here's how to apply the idea to intervals:

$\sup([a, b]) = b. \inf([a, b]) = a$

$\sup((a, b)) = b. \inf((a, b)) = a$

Example 2.6.3. Consider the sequence $a_n = 1/n$. We know that $\forall n, a_n > 0$. We also know that $a_n \to 0$, so 0 is a cluster point of the set $\{a_n : n \in \mathbb{N}\}$. This implies that for any $\epsilon > 0$ there is some k so that $a_k < \epsilon$. So there is no way for any positive number to be a lower bound on a_n. This means that 0 is the greatest lower bound. $\inf_{n \in \mathbb{N}}(a_n) = 0$

2.7 Completeness and Compactness

2.7.1 Completeness

Definition 2.7.1 (Complete)**.** A metric space $(X, d) \subseteq (Y, d)$ is **complete** if the following properties hold:
(a) If $(x_n) \in X$ is Cauchy then $x_n \to x$
(b) If x is the limit of a Cauchy sequence, then $x \in X$.

Theorem 2.7.1 (Completeness Theorem)**.** The real numbers are a complete metric space.

Proof. Suppose $(a_n) \in \mathbb{R}$ is Cauchy.
Let $\epsilon/2 > 0$. Then $\exists N : \forall n, m > 0 |a_n - a_m| < \epsilon/2$.
Allow $x = \max\{|a_1|, ..., |a_{N-1}|, |a_N| + \epsilon/2\}$. Then $|a_n| < x \forall n \in \mathbb{N}$.
Since a_n is bounded, by Bolzano-Weierstrass there is some sequence b_n so that $a_{(b_n)}$ converges to $a \in \mathbb{R}$.
So $\exists M : \forall k > M$, $|a_{b_k} - a| < \epsilon/2$.
$\Rightarrow |a_{b_k} - a + a_n - a_n| < \epsilon/2$
Choose $n > b_k$. Then $|a_{b_k} - a| < \epsilon/2$ since a_{b_k} converges and $|a_{b_k} - a_n| < \epsilon/2$ since a_n is Cauchy.
$\Rightarrow |a - a_{b_k}| + |a_{b_k} - a_n| < \epsilon$
$\Rightarrow |a_n - a| < |a - a_{b_k}| + |a_{b_k} - a_n| < \epsilon$ by the Triangle Inequality
So $a_n \to a$.
Conversely, suppose $a_n \to a$. Then $\exists N : \forall n > N, |a_n - a| < \epsilon/2$. Pick $m > n$. Then $|a_n - a| + |a - a_m| < \epsilon/2 + \epsilon/2$.
Then $|a_n - a_m| < \epsilon$.

2.7.2 Compactness

Definition 2.7.2 (Cover)**.** Given a set X, a **cover** of X is a collection of sets $\mathcal{C}$ which have the following property:

$$X \subseteq \bigcup_{C_i \in \mathcal{C}} C_i$$

Definition 2.7.3 (Open Cover)**.** If all of the elements of a cover $\mathcal{C}$ are open, we call $\mathcal{C}$ an open cover.

Definition 2.7.4 (Compact)**.** A set K is called compact if for any open cover $\mathcal{C}$ of K, there is some **finite** subcover $\mathcal{A} \subseteq \mathcal{C}$ so that

$$K \subseteq \bigcup_{a \in \mathcal{A}} a$$

Theorem 2.7.2 (Sequentially Compact)**.** A subset of a metric space $A \subseteq (X, d)$ is compact if and only if for all sequences $(a_n) \in A$ there is a subsequence (a_{b_k}) which converges to some $a \in A$. We call this type of compactness "sequential compactness". The proof that this is equivalent to the original notion of compactness will be explored in later chapters. For now we take it for granted.

Theorem 2.7.3 (Bolzano Weierstrass II)**.** A subset $A \subseteq \mathbb{R}$ is compact if and only if it is closed and bounded.

Exercise 2.7.1. Prove theorem 2.7.3 using the Bolzano Weierstrass theorem.

Example 2.7.1. Consider the compact set $A = [-1, 1]$. The alternating sequence:

$$a_n = \begin{cases} -1 & n \text{ even} \\ 1 & n \text{ odd} \end{cases}$$

is completely contained in A, and has two trivial convergent subsequences: $a_{2n} = -1$ and $a_{2n+1} = 1$.

Definition 2.7.5 (Totally Bounded)**.** A compact set X is totally bounded if for any $\epsilon > 0$ there is some set of points $C = \{x_1, ..., x_n\} \subset X$ so that $\{\mathcal{B}_\epsilon(x_i) : x_i \in C\}$ is a finite open subcover of X.

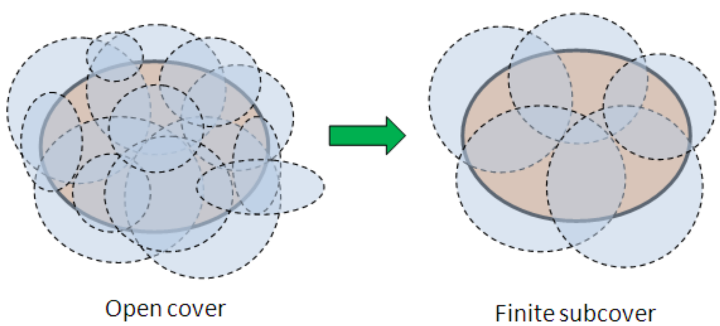


Figure 2.7.1: Demonstration of how compactness works. In a metric space it is always possible to come up with a finite subcover for X entirely composed of neighborhoods (ie. X is totally bounded)[10]

2.8 Continuity

The notion of continuity is essential in physics. A non-continuous equation of motion is nonphysical because it would imply that the object has been momentarily hit by an infinite force. You have already learned the basic notion of continuity, but maybe you haven't treated it with much care.

Definition 2.8.1 (Sequentially Continuous). Consider a function $f : A \to B$ and a sequence $(a_n) \in A$ which converges.

We say f is sequentially continuous if it has the following property:

$$\lim_{n\to\infty} f(a_n) = f(\lim_{n\to\infty} a_n)$$

This property is equivalent to the following:

$$\forall x \in A, \forall \epsilon > 0 \exists \delta > 0 : d_A(x, a) < \delta \Rightarrow d_B(f(x), f(a)) < \epsilon$$

This definition of continuity is very nice. It ensures that a function treats limit points in a nice way. On the other hand it does not guarantee that the function has no asymptotic discontinuities or holes. We will usually use the following stronger definition instead:

Definition 2.8.2 (Uniformly Continuous). We say that a function $f : A \to B$ is uniformly continuous if it has the following property:

$$\forall \epsilon > 0 \exists \delta > 0 : \forall x, y \in A, d_A(x, y) < \delta \Rightarrow d_B(f(x), f(y)) < \epsilon$$

This property is stronger than sequential continuity because your choice of δ is completely independent of the points x and y. Whereas in the previous definition your choice of δ is going to change depending on what point you are looking at. This property is therefore called "global continuity".

2.8.1 Extra Properties †

Theorem 2.8.1 (Inverse Preservation of Openness). $f : A \to B$ is sequentially continuous if and only if whenever $V \subseteq B$ is open, $f^{-1}(V)$ is also open.

Proof. Consider $V \subset B$ and f continuous. Since V is open, whenever $f(p) \in V$, $\mathcal{B}_\epsilon(f(p)) \subseteq V$. Since f is continuous there is some δ so that $f(\mathcal{B}_\delta(p)) \subseteq \mathcal{B}_\epsilon(f(p))$. Therefore $\mathcal{B}_\delta(p) \subseteq f^{-1}(\mathcal{B}_\epsilon(f(p))) \subseteq f^{-1}(V)$. So $f^{-1}(V)$ is open.

Conversely, consider the case where $f^{-1}(V)$ is open whenever $V \subseteq B$ is open. Let $p \in A$ so that $\mathcal{B}_\epsilon(f(p))$ is open. Then $f^{-1}(\mathcal{B}_\epsilon(f(p)))$ is open by hypothesis. Since $f^{-1}(\mathcal{B}_\epsilon(f(p)))$ is open then by the definition of "open" there is some $\delta > 0$ so that $\mathcal{B}_\delta(p) \subseteq f^{-1}(\mathcal{B}_\epsilon(f(p)))$. This is the same as saying that $f(\mathcal{B}_\delta(p)) \subseteq \mathcal{B}_\epsilon(f(p))$. This is the definition of continuous. $\square$

This proof is quite complicated at first glance. I advise you to study it closely and write it out step by step. Figure out why each sentence is true and think about it deeply. If you want, feel free to look up explanations in other books. There should be a good explanation in *Understanding Analysis* by Abbott, or maybe in *Calculus* by Spivak.

Theorem 2.8.2 (Inverse Preservation of Closedness)**.** $f : A \to B$ is sequentially continuous if and only if whenever $V \subseteq B$ is closed, $f^{-1}(V)$ is also closed.

Exercise 2.8.1. By studying the previous proof of 2.8.1, you should be able to come up with a proof of theorem 2.8.2. Good luck!

Theorem 2.8.3 (Preservation of Compactness)**.** If $f : A \to B$ is sequentially continuous and $C \subseteq A$ is compact then $f(C) \subset B$ is compact.

Exercise 2.8.2. Prove that continuous functions preserve compactness (ie; prove the previous theorem). This should be a lot easier than the proof of 2.8.1 and 2.8.2.

It may not be clear at the moment why the hell you need to know any of this. These theorems come back later when we talk about Fourier series, so keep them in mind. These theorems also capture some intuitive ideas about continuity, so try to think about them in a visual way if you can. Either way, it's good for your brain to study analysis. Learning to think like a mathematician is half of learning to think like a physicist (in the honour of Dr. Epp)

2.8.2 Intermediate Value Theorem

Theorem 2.8.4 (Intermediate Value Theorem)**.** Given a continuous function $f : [a, b] \to X$, for any $z \in [f(a), f(b)]$ there is some $c \in [a, b]$ so that $f(c) = z$.

Corollary 2.8.1 (Bolzano's Theorem)**.** Given a continuous function $f : [a, b] \to X$, if $f(a) \geq 0$ and $f(b) \leq 0$ or $f(b) \geq 0$ and $f(a) \leq 0$, there is some point $c \in [a, b]$ so that $f(c) = 0$.

Corollary 2.8.2. Given a continuous function $f : [a, b] \to X$, $f([a, b]) = [f(a), f(b)]$. That is, the image of an interval under a continuous function is an interval.

Theorem 2.8.5 (Composition Rules)**.** Here are some important rules to remember when doing proofs.

(a) The sum of two continuous functions is continuous.

(b) The product of two continuous functions is continuous. (c) The composition of two continuous functions is continuous.

Exercise 2.8.3. Prove the composition rules in theorem 2.8.5

2.9 Differentiability

The notion of a derivative is one you should know from Calculus 1. Knowing the precise definition of a derivative is something you probably know, but you might not know a few particular properties of the derivative. Recall the definition of the derivative:

Definition 2.9.1 (Derivative)**.** Given a continuous function $f : A \to B$, we define the derivative function $f' : A \to B$ as the following:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

In the case where only the one sided limit exists, we need to consider two alternate definitions:

Definition 2.9.2 (Forward Derivative)**.** Given a continuous function $f : A \to B$, in the case that only the forward limit exists:

$$f'(x) = \lim_{h \to 0^-} \frac{f(x+h) - f(x)}{h}$$

Definition 2.9.3 (Backward Derivative)**.** Given a continuous function $f : A \to B$, in the case that only the backwards limit exists:

$$f'(x) = \lim_{h \to 0^+} \frac{f(x+h) - f(x)}{h}$$

There is another way to calculate the derivative, which is very useful when estimating the derivative numerically.

Definition 2.9.4 (Secant Rule)**.** Given a continuous function $f : A \to B$, we can calculate the derivative using the secant rule:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x-h)}{2h}$$

It is also important to note the types of differentiability. There are a few classes of note:

Definition 2.9.5 (Differentiability Classes)**.** A function f is said to be in $\mathcal{C}^1(A)$ if the derivative f' is continuous on A.

Similarly, a function f is said to be in $\mathcal{C}^2(A)$ if the first and second derivatives f' and f'' are continuous on A.

To generalize, a function f is in $\mathcal{C}^k(A)$ if the first k derivatives are continuous on A.

Definition 2.9.6 (Smooth)**.** A function f is called **smooth** on $A \subseteq \mathbb{R}$ if it is in $\mathcal{C}^{\infty}(A)$. This means that it is continuously differentiable an indefinite amount of times.

Smooth functions are of immense importance to physics. Oftentimes we will be required to use properties of smooth functions (and their complex counterparts, analytic functions) in lengthy discussions about topology and space.

Example 2.9.1 (Weierstrass Function)**.** Pathological objects are objects which are the exact opposite of what you expect from a definition. You may expect that if a function is continuous everywhere, then it must be differentiable at least somewhere. There is a function $W(x)$ which is continuous everywhere but differentiable nowhere. We define this function as follows: Let $a \in (0,1)$, $b > 0$

$$W(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x)$$

$$ab > 1 + \frac{3}{2}\pi$$

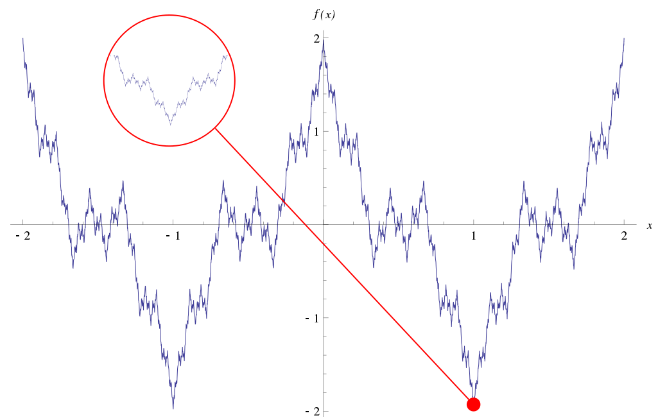


Figure 2.9.1: Plot of Weierstrass Function on [-2,2]

Definition 2.9.7 (Lipschitz Continuity)**.** A real function $f : A \to B$ is said to be Lipschitz if it obeys the following property:

$$\frac{|f(x) - f(y)|}{|x - y|} \leq K \text{ for some } K \in \mathbb{R}$$

K is called the **Lipschitz constant** for f.

Theorem 2.9.1 (Mean Value Theorem)**.** Let $f \in \mathcal{C}^1([a,b])$. Then $\exists c \in [a,b]$:

$$\frac{f(b) - f(a)}{b - a} = f'(c)$$

The expression $(f(b) - f(a))/(b - a)$ represents the slope of a secant line from point a to point b. You can see how this works intuitively in Figure 2.9.2.

Theorem 2.9.2. A function $f \in \mathcal{C}^1(\mathbb{R})$ is Lipschitz if and only if $|f'(x)| \leq K$ for some $K \in \mathbb{R}$. In general, K is the Lipschitz constant for f.

The proof of theorem 2.9.2 will be left as an exercise. The Lipschitz property is interesting because in general a physical trajectory needs to be Lipschitz $(K \leq c)$. While this is pointless in application, it is a nice way to think about it.

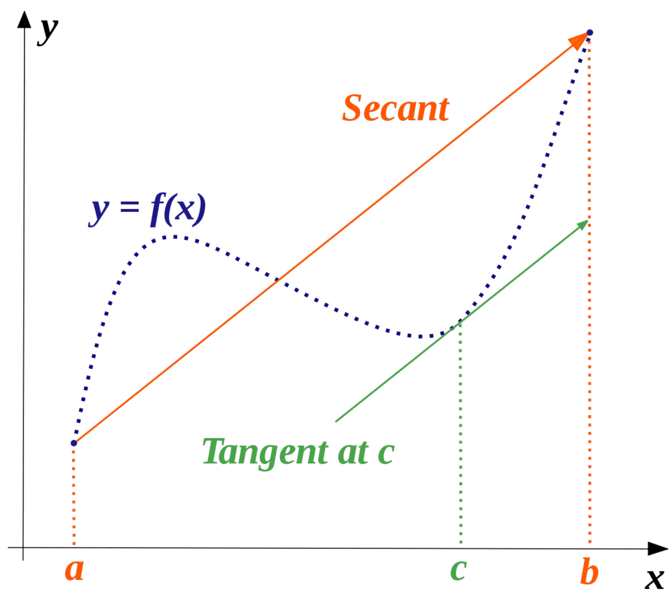


Figure 2.9.2: Diagram demonstrating the mean value theorem.

2.10 Complex Numbers

Complex numbers are often glossed over by physics students. You may never see them until Quantum Physics I if you are very unlucky, a fact which I find completely unacceptable. I think it is useful to have a small section about complex numbers, because they will show up in useful examples later on.

Definition 2.10.1. A **Complex Number** is an ordered pair of real numbers (a, b) which satisfies the following properties:
(1) $(a, b) + (c, d) = (a + c, b + d)$ (you can add them)
(2) $(a, b) \cdot (c, d) = (ac - bd, ad + bc)$ (you can multiply them in a special way)

Theorem 2.10.1. The complex numbers can be written in the form $z = a + bi$, where $i^2 = -1$. We call i the "imaginary unit", because there is no real number which squares to negative one.

Definition 2.10.2. We define the **conjugate** of a complex number z^*:
$z^* = (a, b)^* = (a, -b) = a - bi$

Definition 2.10.3. We define the "real part" and "imaginary part" of $z = a+bi$ as following: $\mathrm{Re}(z) = a$, the real part, and $\mathrm{Im}(z) = b$, the imaginary part.

Theorem 2.10.2. If $z = w$, $\mathrm{Re}(z) = \mathrm{Re}(w)$ and $\mathrm{Im}(z) = \mathrm{Im}(w)$. This may seem obvious but it is useful when solving equations.

Theorem 2.10.3. $z^* \cdot z = a^2 + b^2$

Definition 2.10.4. The **length/norm/modulus/size** of a complex number is the value $|z| = \sqrt{z^* z}$.

You can draw complex numbers on a plane. This plane is called the complex plane, and such drawings are called "Argand diagrams"

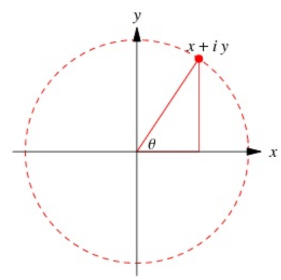


Figure 2.10.1: Argand Diagram of $x + iy$ [2]

2.10.1 Euler's Formula and de Moivre's Theorem

There is a connection between complex numbers and two dimensional vectors. It turns out that complex numbers are simply 2D vectors with a few extra properties. In this case it is useful to note the following properties:

Theorem 2.10.4. Given a complex number z, there is some angle θ so that the following identity holds (See Figure 2.10.1):

$$z = a + bi = |z|(\cos(\theta) + i\sin(\theta))$$

We call this the **Polar Form** or **Trigonometric Form** of z

Theorem 2.10.5. Using math and stuff, one can show that the following identity holds for all real numbers x:

$$e^{ix} = \cos(x) + i\sin(x)$$

This identity is called **Euler's Identity**

See *https://en.wikipedia.org/wiki/Euler%27s_formula#Proofs* for a proof. There are geometric proofs, multiple proofs using calculus (including the glorious Taylor Series expansion), and even crazier proofs out there.

Theorem 2.10.6. Combining Theorem 2.10.3 and 2.10.4 we arrive at the following: For all complex numbers z, there is some θ so that $z = |z|e^{i\theta}$. We call this the **Exponential Form** of z.

Theorem 2.10.7. Using Euler's Identity, one can show that the following holds:

$$(\cos(x) + i\sin(x))^n = \cos(nx) + i\sin(nx)$$

This is called **De Moivre's Theorem**.

Proof. $(\cos(x) + i\sin(x))^n = (e^{ix})^n = e^{i(nx)} = \cos(nx) + i\sin(nx)$

Using De Moivre's Theorem, we can prove all of the double-angle trigonometric identities. Using similar methods one can prove the angle-addition formulas.

Example 2.10.1. Here we prove the double angle identities.

$$\cos(2x) + i\sin(2x) = (\cos(x) + i\sin(2x))^2$$

$$(\cos(x) + i\sin(x))^2 = \cos^2(x) - \sin^2(x) + 2i\sin(x)\cos(x)$$

$$\cos^2(x) = 1 - \sin^2(x)$$

$$(\cos(x) + i\sin(x))^2 = 2\cos^2(x) - 1 + 2i\sin(x)\cos(x)$$

$$\mathrm{Re}(\cos(2x) + i\sin(2x)) = \mathrm{Re}(2\cos^2(x) - 1 + 2i\sin(x)\cos(x))$$

$$\mathrm{Im}(\cos(2x) + i\sin(2x)) = \mathrm{Im}(2\cos^2(x) - 1 + 2i\sin(x)\cos(x))$$

$$\cos(2x) = 2\cos^2(x) - 1$$

$$\sin(2x) = 2\sin(x)\cos(x)$$

2.11 Brief Intro to Conic Sections

For some reason Canada has refrained from teaching conic sections in high school. This is bad because many people come into university without knowing the relationship between a parabola, hyperbola, and ellipse.

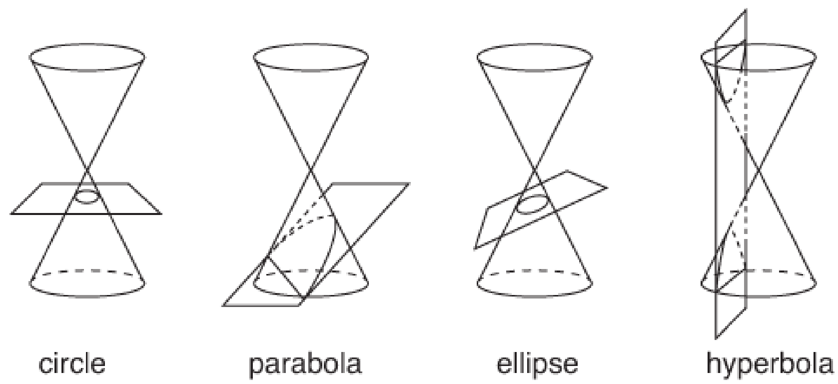


Figure 2.11.1: The Conic Sections [8]

Definition 2.11.1 (Cartesian Equations of Conics)**.** Note that in these equations, x and y correspond to their normal roles (independent and dependent variable). All other terms are constants.

1. (Circle) $(x-x_0)^2+(y-y_0)^2=R^2$
2. (Ellipse) $\frac{(x-x_0)^2}{a^2}+\frac{(y-y_0)^2}{b^2}=1$
3. (Parabola) $a(x-b)^2+y=c$
4. (Hyperbola) $x^2-y^2=k^2$

There are a few functions related to the hyperbola which are important for every physicist to know about. These show up in the solutions to an absolute ton of equations, and will be useful in special relativity as well. See the following page for diagrams.

Definition 2.11.2 (Hyperbolic Trig Functions)**.**

$$\cosh(x)=\frac{e^x+e^{-x}}{2}$$

$$\sinh(x)=\frac{e^x-e^{-x}}{2}$$

$$\tanh(x)=\frac{\sinh(x)}{\cosh(x)}=\frac{e^x-e^{-x}}{e^x+e^{-x}}$$

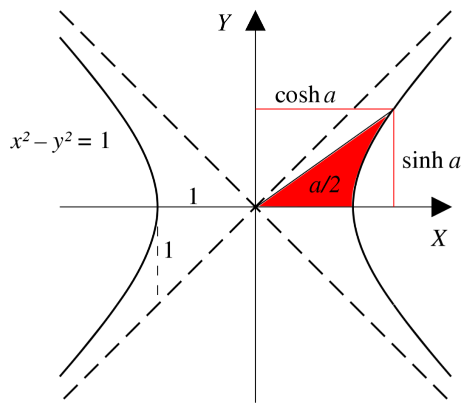


Figure 2.11.2: Diagram of Hyperbolic Trig Functions on the Unit Hyperbola

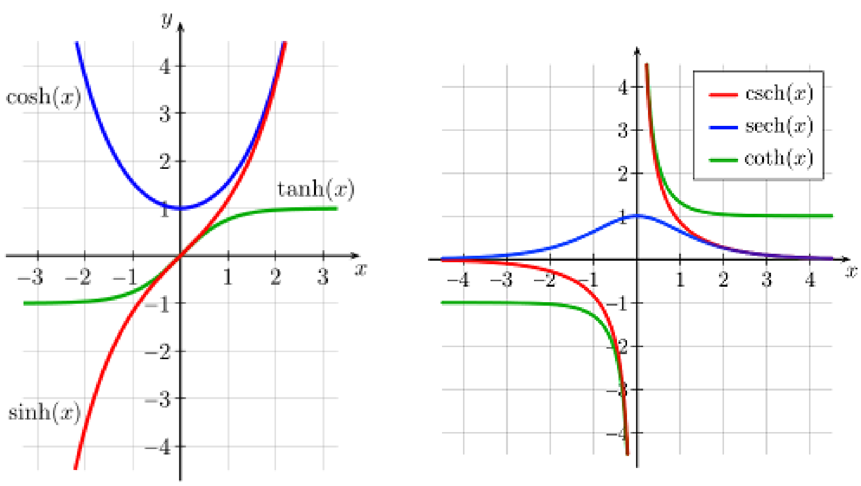


Figure 2.11.3: Plot of the hyperbolic trig functions

2.12 Brief Intro to Taylor Series

Taylor series are one of the single most powerful tools for a physicist. Our MATH137P professor in 2017 made sure to teach us how to use Taylor series as early as possible because they are just that useful. So what is a Taylor series? First we need to begin with a linear approximation.

2.12.1 Linear Approximations

Definition 2.12.1 (Tangent at a point)**.** Consider a function $f \in \mathcal{C}^\infty$. Due to the property that the derivative of f at a point a is the slope of that graph, we can always find a line $y = m(x-a)+b$ which passes tangent to the function at that point. With a little bit of algebra we can find the exact definition:

$$f(x) \approx L_a(x) = f(a) + f'(a) * (x-a)$$

The "approximately" symbol in the previous expression means that for values of x very close to a the function is very close to $L_a(x)$. This is called the **linear approximation**.

Definition 2.12.2 (Quadratic Approximation)**.** Since we can approximate functions with lines, what about with quadratics? Here is how we can do this:

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a) \text{ when } x \approx a$$

Definition 2.12.3 (Taylor Series)**.** Well, if we can approximate functions with quadratics, why not just go right ahead and approximate functions with arbitrary polynomials? In the limit as the degree of the polynomial goes to infinity, your approximation becomes more and more perfect.

$$f(x) = f(a) + \sum_{n=1}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \text{ if } f \in \mathcal{C}^\infty$$

We will go into depth about how this and the quadratic approximation work in later chapters.

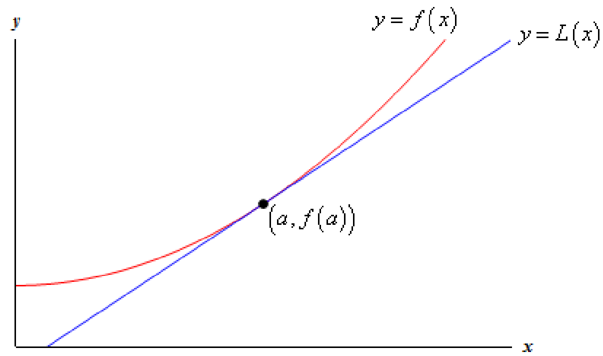


Figure 2.12.1: The Linear Approximation. You can see how the line is very close to the original function for values of x near a. [5]

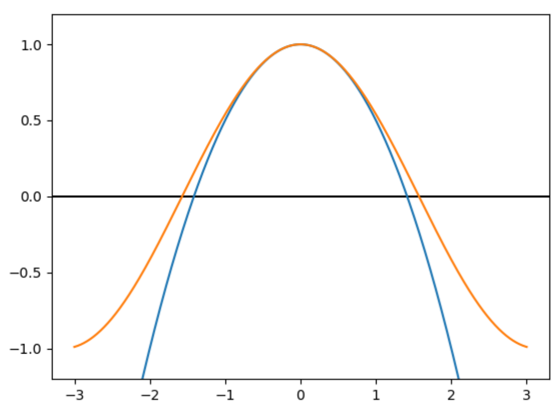


Figure 2.12.2: The Quadratic Approximation for $\cos(x)$.

2.13 Brief Intro to Multivariable Functions

2.13.1 Scalar Fields

Definition 2.13.1 (Scalar Field)**.** A scalar field in $\mathbb{R}^n$ is a function $f : \mathbb{R}^n \to \mathbb{R}$. We call it a scalar field because you can imagine a "field" where the height of the hills on the field determines the value of the function at that point.

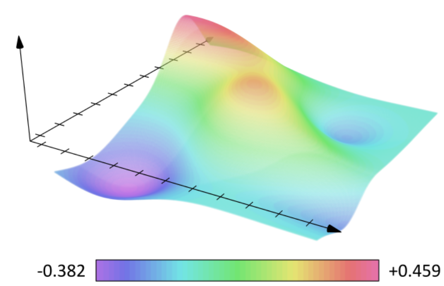


Figure 2.13.1: Plot of a scalar field $f : \mathbb{R}^2 \to \mathbb{R}$

The derivative of a scalar field is an interesting concept. You can imagine standing in the field and measuring the slope in any direction along the x-y plane.

Definition 2.13.2 (Partial Derivative)**.** The derivative of a scalar field $f(x_1, ..., x_n) = f(\mathbf{x})$ is dependent on which direction you point. It is nice to choose directions which are along the principal axes of the field. These are called the partial derivatives of f. To calculate these you simply treat all the variables but x_i as a constant.

$$\frac{\partial f}{\partial x_i} = \lim_{h\to 0} \frac{f(x_1, ..., x_i + h, ..., x_n) - f(x_1, ..., x_n)}{h} = \lim_{h\to 0} \frac{f(\mathbf{x} + h\mathbf{e}_i) - f(\mathbf{x})}{h}$$

(recall that the vector $\mathbf{e}_i$ is the i^{th} unit vector. $\mathbf{e}_1 = \hat{\imath}$, $\mathbf{e}_2 = \hat{\jmath}$, and $\mathbf{e}_3 = \hat{k}$)

Alternative notation includes: $D_{x_i} f(\mathbf{x}) = \partial_i f(\mathbf{x}) = f_{x_i}(\mathbf{x})$

Definition 2.13.3 (Gradient)**.** The gradient of a scalar field is the vector representing the direction of steepest ascent at a point. Imagine standing on the side of a cliff. If you point in the direction which takes you up the cliff the fastest, the projection of that vector onto the x-y plane is the gradient. Consider $f : \mathbb{R}^n \to \mathbb{R}$.

$$\text{grad}(f)(\mathbf{x}) = \vec{\nabla} f(\mathbf{x}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \mathbf{e}_i$$

Definition 2.13.4 (Directional Derivative)**.** Let's say you want to know the value of the derivative along a particular direction. You can imagine projecting the gradient along a unit vector $\hat{\mathbf{u}}$ in the specified direction. We write this in the following way:

$$D_{\hat{\mathbf{u}}} f(\mathbf{x}) = \left(\vec{\nabla} f(\mathbf{x}) \cdot \hat{\mathbf{u}} \right) \hat{\mathbf{u}}$$

One of the problems you might need to solve in the future is finding the direction along a field in which the height of the field does not change (finding a "curve of constant height" so to speak). To do this you would take the above expression, set it equal to zero, and solve the resulting system of equations.

Now we should concern ourselves with some concepts related to integration. You may be asked in PHYS122 to integrate a function in polar coordinates (such as an electric field around a wire). We should start off with multiple integration.

The principle behind integrating a scalar is the exact same as the principle behind taking the partial derivative. You treat all the other variables as constants and integrate variable-by-variable. Your "plus C" term becomes a function of all the variables you didn't integrate over (because you treated them as constants!)

$$\iint \frac{\partial^2 f}{\partial x \partial y} dx dy = \int \frac{\partial f}{\partial y} + C_1(y) dy$$
$$= f(x, y) + \int C_1(y) dy + C_2(x) = f(x, y) + C_2(x) + C_3(y)$$

You won't have to do any complicated integrals until calculus 3, but it helps to know the idea behind multiple integrals as soon as you can.

2.13.2 Vector Fields

The next step is knowing what a vector field is. Many things in physics are vector fields (Electric and magnetic fields, gravitational fields, velocity fields of fluids), so having a grasp on the concept behind them is essential as early as possible.

Definition 2.13.5 (Vector Field)**.** A vector field $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ is a function which spits out vectors at each point in space. We write these as so:

$$\mathbf{f}(\mathbf{x}) = \mathbf{y}$$

Example 2.13.1 (Gravitational Field)**.** Around any massive object there is a field which determines the force experienced by a mass towards that object. If we place our coordinate axes so that the object is at the origin, we can write this field as so:

$$\mathbf{F}_g(\mathbf{x}, \mathbf{y}, \mathbf{z}) = -\frac{GMm}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}(x, y, z) = -\frac{GMm}{r^3}\mathbf{r}$$

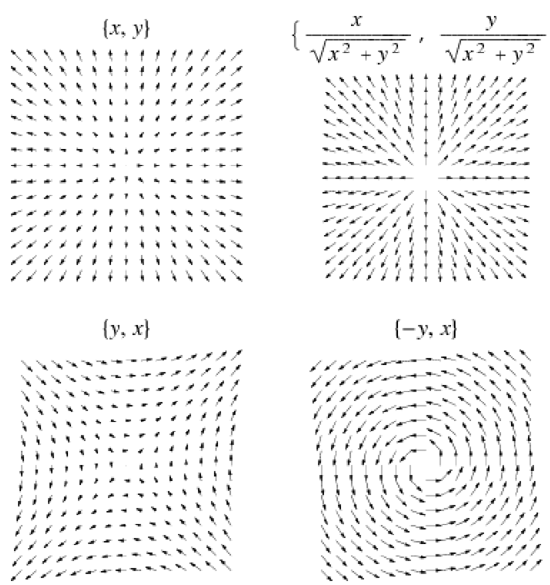


Figure 2.13.2: Some Vector Fields. [11]

2.14 Exercises for Section 2

Exercise 2.14.1. This exercise is good practice for *Proof by Contradiction*. We seek to prove that $a = b \iff \forall \epsilon, |a - b| < \epsilon$.

(a) Show that $\forall \epsilon > 0, a = b \Rightarrow |a - b| < \epsilon$

(b) Show that $\forall \epsilon > 0, |a - b| < \epsilon \Rightarrow a = b$

Exercise 2.14.2. This exercise is good practice for *Proof by Induction*. We see to show that the sequence defined by $x_n = \frac{3}{2}x_{n-1}^2, x_0 = 1$ is increasing.

(a) Show that $x_0 \leq x_1$

(b) Show that $x_{n-2} \leq x_{n-1} \Rightarrow x_{n-1} \leq x_n$

Exercise 2.14.3. Prove that if $a_n > 0$, $a_n \to a$, then $\sqrt{a_n} \to \sqrt{a}$ (thereby proving that $\sqrt{x}$ is continuous)

Exercise 2.14.4. Let $A_1 \geq A_2 \geq ... \geq A_k$. Prove the following

$$\max_{i \leq k}(A_i) = \lim_{n \to \infty}(A_1 + ... + A_k)^{1/n}$$

Exercise 2.14.5. Consider bounded subsets $A, B \subset \mathbb{R}$, $A + B = \{a + b : a \in A, b \in B\}$. Prove that $\sup(A + B) = \sup(A) + \sup(B)$

Exercise 2.14.6. Prove that for $0 < a < b$:

$$\sqrt{a} - \sqrt{b} \leq \frac{b - a}{2\sqrt{a}}, \ \sqrt{ab} < \frac{1}{2}(a + b)$$

Exercise 2.14.7 (Proof of Intermediate Value Theorem)**.** $\forall c, \exists z : f(z) = c$. Assume that $f(a) \leq f(b)$ without loss of generality.

We construct the set $H = \{x \in [a, b] : f(x) < c\}$. Let $\sup(H) = p$.

(a) Prove that if $f(p) < c$ and $a < c-\delta < c+\delta < b$, then there is a contradiction.

(b) Prove that if $f(p) > c$ and $a < c-\delta < c+\delta < b$ then there is a contradiction.

Exercise 2.14.8. Show that $\cosh(ix) = \cos(x)$ and that $\sinh(ix) = \sin(x)$.

Exercise 2.14.9. Show that $\cosh^2(x) - \sinh^2(x) = 1$

Exercise 2.14.10. Prove the following angle addition identity for tanh.

$$\tanh(x + y) = \frac{\tanh(x) + \tanh(y)}{1 + \tanh(x)\tanh(y)}$$

3 Essential Linear Algebra

3.1 Vector Spaces

The motivation for the study of linear algebra is the fact that often times in math and physics you will have some sort of objects which can be combined. When these combinations have some sort of structure to them then a lot of interesting results can be obtained. This will be the motivation for us to define what a vector space is and we will see in a moment that this "structure" that I refer to is called linearity. The study of algebraic structures like vector spaces and linearity is called Abstract Algebra, and is the foundation of much of modern mathematical and theoretical physics. It is extremely important to pay attention to the structure of the space you are working in, since symmetries, relations, and other properties of those spaces can have important physical consequences.

You have seen in your elementary linear algebra classes (114/136) that objects of the form $(a_1, ..., a_n) \in \mathbb{R}^n$ can be called "vectors", and that they can be added together and manipulated in very useful ways.

What we hope to show you in this chapter is that the study of a more general class of objects (what we really mean by "vector") will produce key insights into physical problems like quantum physics and relativity. If you have completed MATH235, feel free to skip a few subsections of this chapter. Note however that there are important examples that are used in physics but are glossed over in 235. Here are some things you should check out for more help.

- *Essence of Linear Algebra* - 3Blue1Brown on Youtube
- *Linear Algebra Done Right* - Axler
- *Linear Algebra Done Wrong* - Treil
- *Linear Algebra* - Friedberg et al.
- *Principles of Quantum Mechanics, Chapter 1.* - Shankar
- *Linear Algebra 1 & 2 Coursenotes* (MATH136/235) - Dan Wolczuck

Definition 3.1.1 (Field)**.** A field k is a set of objects which obey a few properties, most importantly that addition/substraction and multiplication/division are defined for all elements of the field. These operations need to follow a set of "field axioms": let $\alpha, \beta \in k$

(a) $\alpha + \beta \in k$ (closure under addition)
(b) $\exists 0 \in k : \alpha + 0 = \alpha$ (additive identity)
(c) $\forall \alpha, \exists(-\alpha) : \alpha + (-\alpha) = 0$ (additive inverses)
(d) $\alpha \bullet \beta \in k$ (closure under multiplication)
(e) $\exists 1 : \alpha \bullet 1 = \alpha$ (multiplicative identity)
(f) $\forall \alpha, \exists \alpha^{-1} : \alpha^{-1} \bullet \alpha = 1$ (multiplicative inverses)

Example 3.1.1 (Fields)**.** Some examples of fields are $\mathbb{R}, \mathbb{C}$ and $\mathbb{Q}$. $\mathbb{Z}$ is not a field because it does not have multiplicative inverses. Similarly, $\mathbb{N}$ is not a field because it has neither additive nor multiplicative inverses. A pathological example is the p-adic numbers (see wikipedia).

Finally, we can define a vector space:

Definition 3.1.2 (Vector Space)**.** A set V over a field k along with vector addition and scalar multiplication is called a **Vector Space** if the following properties hold:

Let $\mathbf{v}, \mathbf{w}, \mathbf{z} \in V$
Let $\alpha, \beta \in k$
(a) $\mathbf{v} + \mathbf{w} \in V$ (closure under vector addition)
(b) $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$ (commutativity of vector addition)
(c) $\mathbf{v} + (\mathbf{w} + \mathbf{z}) = (\mathbf{v} + \mathbf{w}) + \mathbf{z}$ (associativity of vector addition)
(d) $\exists \mathbf{0} \in V$ such that $\mathbf{0} + \mathbf{v} = \mathbf{v}$ (existence of an additive identity)
(e) $\exists(-\mathbf{v}) \in V$ such that $(-\mathbf{v}) + \mathbf{v} = \mathbf{0}$ (existence of an additive inverse)
(f) $\alpha \bullet \mathbf{v} \in V$ (closure under scalar multiplication)
(g) $(\alpha+\beta) \bullet \mathbf{v} = \alpha \bullet \mathbf{v} + \beta \bullet \mathbf{v}$ (scalar multiplication distributes scalar addition)
(h) $\alpha \bullet (\mathbf{v}+\mathbf{w}) = \alpha \bullet \mathbf{v} + \alpha \bullet \mathbf{w}$ (scalar multiplication distributes vector addition)
(i) $\alpha(\beta \bullet \mathbf{v}) = (\alpha\beta) \bullet \mathbf{v}$ (associativity of scalar addition)
(j) $\exists\ 1 \in k$ such that $1\mathbf{v} = \mathbf{v}$ (existence of a scalar multiplicative identity, $\mathbf{1}$)

Example 3.1.2. Let ℓ be the set of bounded sequences in $\mathbb{R}^n$. Define vector addition as follows:[a]

$$\{x_n\}_{n=0}^{\infty} + \{y_n\}_{n=0}^{\infty} = \{x_n + y_n\}_{n=0}^{\infty}$$

Define scalar multiplication (The field being $\mathbb{R}$) by:

$$c \bullet \{x_n\}_{n=0}^{\infty} = \{cx_n\}_{n=0}^{\infty}$$

Then this is a vector space. We can check that it satisfies the axioms (a-g):
a) The sum of two bounded sequences is also bounded by results from the previous section.
b) This follows from the fact that addition is commutative in the real numbers
c) This follows from associativity of the real numbers
d) The zero sequence is just $\{0, 0, ...\}$ and this is clearly bounded
e) If $\{x_n\}_{n=0}^{\infty}$ is bounded, then due to the provable linearity of limits, $\{-x_n\}_{n=0}^{\infty}$ is also bounded and $\{x_n\}_{n=0}^{\infty} + \{-x_n\}_{n=0}^{\infty} = \{0, 0, ...\}$
f) It's easy to show that $\{cx_n\}_{n=0}^{\infty}$ is bounded by cM if $\{x_n\}_{n=0}^{\infty}$ is bounded by M (simply factor M in/out of the inequality)
g-i) If you treat addition and scalar multiplication pointwise on the terms of each sequence, the proof follows from the properties of $\mathbb{R}$.
j) This is just $1 \in \mathbb{R}$.

[a] Note the new notation: $(a_n) = \{a_n\}_{n=0}^{\infty}$

Later on we will find that ℓ is "simply" the (countable) infinite-dimensional vector space $\ell = \{x \in \lim_{n\to\infty} \mathbb{R}^n : \forall n, x_n < \infty\}$. Later in the chapter we will develop the ℓ^p inner product space. This leads to some nice properties in functional analysis and quantum physics.

Example 3.1.3 (Function Space)**.** The set $V = \mathcal{C}(K)$ composed of all continuous functions on $K \subset \mathbb{R}$ is a vector space.

Proof. Consider $f, g \in V$, we can prove by elementary analysis that $f + g$ is continuous. Take $k \in K$, then kf is also continuous. Since functions add and multiply point-wise, the distributivity/commutativity/associativity relations all follow from $\mathbb{R}$. Note that $0f = 0$ and $1f = f$, so that the remaining axioms are satisfied.

Definition 3.1.3 (Linear Map)**.** Consider some vector spaces V and W over a field k. A **Linear Map** is a map $L : V \to W$ which obeys the following properties:

Let $u, v \in V, a, b \in k$:

$L(au + bv) = aL(u) + bL(v)$

You may note that according to these rules, $L(0) = L(0v) = 0L(v) = 0$.

3.1.1 Subspaces of Vector Spaces

Definition 3.1.4 (Subspace)**.** Given a vector space V over a field k, a set $W \subseteq V$ is denoted a **subspace** of V if the following properties hold: Let $\mathbf{a}, \mathbf{b} \in V$ and $\alpha \in k$

(a) $\mathbf{a} + \mathbf{b} \in V$

(b) $\alpha \mathbf{a} \in V$

(c) $\mathbf{0} \in V$

Definition 3.1.5 (Span)**.** The **span** of a set of vectors A in a vector space V over a field k is the set of all linear combinations of such vectors:

$\text{Span}(A) = \{\sum(c_n v_n) : c_n \in k, v_n \in A\} \subseteq V$.

Exercise 3.1.1. Prove that the span of a set of vectors in V is a subspace.

Definition 3.1.6 (Linearly Independent)**.** A set of vectors $\{v_1, ..., v_n\}$ in a vector space V over k is linearly independent if there does not exist $\{a_1, ..., a_n\} \neq \{0, ..., 0\} \subset k$ so that:

$$\sum_{i=1}^{n} a_i v_i = 0$$

Recall the definition of a basis from your linear algebra classes:

Definition 3.1.7 (Basis)**.** A **basis** for a vector space V is a linearly independent set of vectors $\{v_1, ..., v_n\} = \mathcal{B}$ so that $\text{Span}(\mathcal{B}) = V$

Example 3.1.4 (Change of Basis)**.** A basis defines the way vectors are written. You represent elements of your vector space as linear combinations of the basis vectors. As an example:

$$\text{Let } \mathcal{A} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

$$\text{Let } \mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} \right\}$$

$$\begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} = (1) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + (1) \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} + (1.5) \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1.5 \end{bmatrix}_{\mathcal{B}}$$

For some bases $\mathcal{C}$ and $\mathcal{D}$ you can find a change of basis matrix for them by computing ${}_{\mathcal{A}}P_{\mathcal{D}}$ and ${}_{\mathcal{A}}P_{\mathcal{C}}$. These are easy to find, for example:

$${}_{\mathcal{A}}P_{\mathcal{D}} = \begin{bmatrix} d_{11} & d_{21} & d_{31} \\ d_{12} & d_{22} & d_{32} \\ d_{13} & d_{23} & d_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{d}_1 & \mathbf{d}_2 & \mathbf{d}_3 \end{bmatrix}$$

In the end, you usually compute change of basis matrices like this:

$${}_{\mathcal{A}}P_{\mathcal{D}}^{-1} \cdot {}_{\mathcal{A}}P_{\mathcal{C}} = {}_{\mathcal{D}}P_{\mathcal{A}} \cdot {}_{\mathcal{A}}P_{\mathcal{C}} = {}_{\mathcal{D}}P_{\mathcal{C}}$$

We will see in later chapters that if you are considering abstract vector spaces such as ℓ^p and L^p (see Example 3.1.1 and Example 2.3.6) you need to construct more convoluted (pun intended) methods of changing basis. This has applications in quantum physics 1, where you will need to change between the "position" and "momentum" bases.

I highly recommend doing lots of exercises for change of basis and finding eigenvectors/eigenvalues of matrices. These are the two most important linear algebra computations in PHYS234, so knowing them by heart will help.

For a visual interpretation of change of basis, I recommend 3Blue1Brown's video on the topic.

3.2 Matrices

You should already be familiar with matrices from your elementary linear algebra class. For a review of calculating the determinant of a matrix, see Appendix IV.

Recall the definition of an eigenvalue:

Definition 3.2.1. (Eigenvalue)

A scalar λ is called an eigenvalue of a matrix M if $Mv = \lambda v$ for some vector v. You can find these by calculating the roots of the characteristic polynomial $\chi(\lambda; M) = \det(M - \lambda\mathbb{I})$

Theorem 3.2.1. $\det(M) = \prod_{i=1}^{n} \lambda_i$ where λ_i are the eigenvalues of M

Definition 3.2.2 (Dot Product)**.**

$$\text{Let } \mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}, \mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$

We define the dot product between $\mathbf{v}$ and $\mathbf{w}$ as $\mathbf{v} \boldsymbol{\cdot} \mathbf{w} = \det\left(\mathbf{v}^T\mathbf{w}\right)$. Oftentimes we will just write $\mathbf{v}^T\mathbf{w}$ instead of the determinant as shorthand.

$$\det\left(\mathbf{v}^T\mathbf{w}\right) = \det\left[\begin{bmatrix} v_1 & \dots & v_n \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}\right] = \det\left[\sum_{i=1}^{n} v_i w_i\right] = \sum_{i=1}^{n} v_i w_i$$

The dot product is immediately essential in the definition of "work" in physics. Work is the amount of energy something has gained by accelerating along a certain path - you need to take the dot product of the force with the direction of motion in order to calculate this.

Theorem 3.2.2. Every matrix $M \in \mathcal{M}^{n\times m}(k)$ defines a linear map from k^m to k^n, so that

$L : k^m \to k^n$ has

$L(v) = Mv$. We call this the "action" of $\mathcal{M}^{n\times m}$ on k^m. You will see this term come up more and more, so check out our chapter on abstract algebra for a good overview.

Definition 3.2.3 (Standard Matrix)**.** The **standard matrix** of a linear map L is the matrix $[L]$ so that $\forall v, L(v) = [L]v$ (as long as both the map and the matrix are written with respect to the same basis)

Example 3.2.1. You can find the standard matrix of a given linear map by plugging in your basis vectors: $[L]_1 = L(v_1), ..., [L]_n = L(v_n)$.

Definition 3.2.4. The **Column Space** $\text{Col}(M)$ of a matrix $M = [M_1...M_n]$ is the space $\text{Span}(\{M_1, ..., M_n\})$.

Definition 3.2.5. The **Row Space** $\text{Row}(M)$ of a matrix $M = [M_1...M_n]$ is the space $\text{Span}(\{(M^T)_1, ..., (M^T)_n\}) = \text{Col}(M^T)$.

Definition 3.2.6. The **Null-Space** of a matrix $M = [M_1...M_n]$ is the space $\text{Nul}(M) = \{v \in A^n : Mv = 0\}$.

Definition 3.2.7. The **Left Null-Space** of a matrix $M = [M_1...M_n] \in \mathcal{M}^{n\times m}$ is the space $\text{Nul}(M^T) = \{w \in A^m : M^T w = 0\}$.

Theorem 3.2.3. Consider a matrix M and the associated linear map L so that $L(v) = Mv$. Then:

$\text{im}(L) = \text{Col}(M)$

$\ker(L) = \text{Nul}(M)$

Exercise 3.2.1. Prove that the sets defined in definitions 3.2.2 through 3.2.5 are indeed subspaces.

Definition 3.2.8 (Eigenspace)**.** Consider a matrix M with eigenvectors $\mathcal{E} = \{\mathbf{a}_1, ..., \mathbf{a}_n\}$. The space $E = \text{Span}(\mathcal{E})$ is called the **eigenspace** of M.

If M is invertible, performing a change of basis to the eigenspace *diagonalizes* M, since $[M]_{\mathcal{E}} = {}_{\mathcal{E}}P_{\mathcal{A}}([M]_{\mathcal{A}})_{\mathcal{A}}P_{\mathcal{E}}$, which if you recall from 114/136 is the diagonalization formula.

3.2.1 The Fundamental Theorem of Linear Algebra

Definition 3.2.9. The **direct sum** of two vector spaces A and B over a field k, with $A \cap B = \{\mathbf{0}\}$ is the set $A \oplus B = \{na + mb : a \in A, m \in B, n, m \in k\}$

Example 3.2.2. $\text{Span}(A) \oplus \text{Span}(B) = \text{Span}(A \cup B)$

Definition 3.2.10. The **orthogonal complement** of a subspace $A \subseteq V$ is the set $A^\perp = \{v \in V : \forall a \in A, a \cdot v = 0\}$

Exercise 3.2.2. Consider the vector space $\mathbb{R}^n$ and a subspace $A \subseteq \mathbb{R}^n$. Prove that $A \oplus A^\perp = \mathbb{R}^n$

Definition 3.2.11 (Dimension)**.** Given a linearly independent basis $\mathcal{B}$ for a vector space V, we say the "dimension" of V, denoted $\dim(V)$, is defined as the number of elements in $\mathcal{B}$. You can see that $\dim(\mathbb{R}^n) = n$.

For a thorough exploration of what it's like to live in a higher or lower dimensional world, many people suggest reading the book *Flatland: A Romance of Many Dimensions* by Edwin A Abbott.

Theorem 3.2.4 (The Fundamental Theorem of Linear Algebra)**.** Let $M \in \mathcal{M}^{n\times m}(k)$. Then:

$$k^n = \mathrm{Col}(M) \oplus \mathrm{Nul}(M^T) \text{ and } k^m = \mathrm{Row}(M) \oplus \mathrm{Nul}(M)$$

$$\mathrm{Col}(M) = \mathrm{Nul}(M^T)^{\perp} \text{ and } \mathrm{Row}(M) = \mathrm{Nul}(M)^{\perp}$$

Corollary 3.2.1 (Dimension Theorem)**.** Given $M \in M_n(k)$

$$\dim(\mathrm{Col}(M)) = \mathrm{rank}(M)$$

$$\dim(\mathrm{Nul}(M)) = \mathrm{nullity}(M)$$

Corollary 3.2.2 (Rank-Nullity Theorem)**.** Given $M : V \to W$

$$\mathrm{rank}(M) + \mathrm{nullity}(M) \leq \dim(V)$$

It is very efficient to compute the column space by first computing the left nullspace, which is useful when you want to figure out the output space of some linear operator.

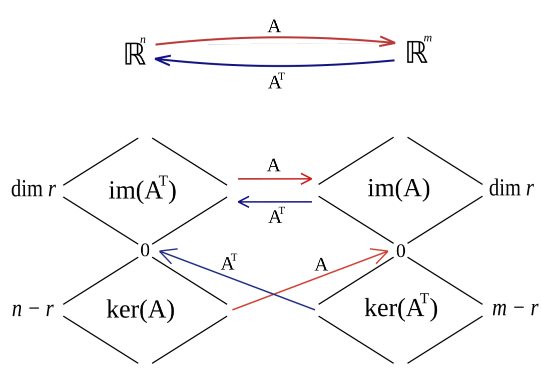


Figure 3.2.1: The Fundamental Theorem of Linear Algebra

3.3 Abstract Vector Spaces & Important Examples

Many linear algebra courses and books will avoid more abstract vector spaces out of fear of "intimidating" physics and engineering students. I find that this is a terrible idea, because there are tons of vector spaces that you need in physics (and even engineering) that aren't $\mathbb{R}^n$. Let's just start by abstracting the notion of the dot product, and then we will give you some examples of abstract vector spaces.

Definition 3.3.1 (Bilinear Product)**.** A bilinear product is a particular type of map between vector spaces that takes two inputs:

$L : V \times U \to W$

Bilinear means that L is linear in both inputs. Consider the cross product as a classic example of a bilinear product:

$(a+b)\times(c+d) = (a\times(c+d)) + (b\times(c+d)) = (a\times c + a\times d + b\times c + b\times d)$

$(sa \times zb) = sz(a\times b)$

Definition 3.3.2 (Inner Product)**.** Consider a vector space V over k. An **inner product** on V is a bilinear product defined as so:

$\langle\cdot,\cdot\rangle : V\times V \to k$

The inner product has a few properties it must satisfy:

(a) $\langle \mathbf{a},\mathbf{b}\rangle = \langle \mathbf{b},\mathbf{a}\rangle^*$

(b) $\langle s\mathbf{a},\mathbf{b}\rangle = s\langle \mathbf{a},\mathbf{b}\rangle$

(c) $\langle \mathbf{a}+\mathbf{c},\mathbf{c}\rangle = \langle \mathbf{a},\mathbf{b}\rangle + \langle \mathbf{c},\mathbf{d}\rangle$

(d) $\langle \mathbf{a},\mathbf{a}\rangle \geq 0$

(e) $\langle \mathbf{a},\mathbf{a}\rangle = 0 \iff \mathbf{a} = \mathbf{0}$

A space with an inner product is called an **Inner Product Space**.

Definition 3.3.3. (Norm) A **norm** on a vector space V over k is a function $\|\cdot\| : V \to k$ which has a few properties.

(a) $\|\mathbf{v}+\mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$ (triangle inequality)

(b) $\|v\| \geq 0$

(c) $\|\mathbf{v}\| = 0 \iff \mathbf{v} = \mathbf{0}$

(d) $\|\alpha\mathbf{v}\| = |\alpha|\,\|\mathbf{v}\|$

Example 3.3.1 (Standard Inner Product, 2-Norm)**.** The inner product $\langle \mathbf{a},\mathbf{b}\rangle = \det\left(\mathbf{a}^T\mathbf{b}\right)$ is called the standard inner product on k^n. The norm $\sqrt{\det(\mathbf{a}^T\mathbf{b})}$ is called the 2-norm, and is denoted $\|\cdot\|_2$ unless otherwise specified.

It is quite common to have $\|\mathbf{a}\| = \sqrt{\langle \mathbf{a},\mathbf{a}\rangle}$. This is the case in $\mathbb{R}^n$, $\mathbb{C}^n$, and the matrix space $\mathcal{M}_{m\times n}(k)$.

3.3.1 The Graham-Schmidt Procedure

Definition 3.3.4 (Orthogonal)**.** Two vectors $\mathbf{v}$ and $\mathbf{u}$ are orthogonal if $\langle \mathbf{v}, \mathbf{u} \rangle = 0$.

Theorem 3.3.1. If a set of vectors $A = \{\mathbf{v_1}, ..., \mathbf{v_n}\}$ is orthogonal that is if $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = K\delta_{ij},$ where $K \in k$ then it is linearly independent.

Definition 3.3.5 (Projection Operator)**.** The linear operator $P_{\mathbf{u}} : V \to V$ is defined as so:

$$P_{\mathbf{u}}\mathbf{v} = \text{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\mathbf{u}$$

Definition 3.3.6 (Orthonormal Basis)**.** If you have a (linearly independent/orthogonal) basis $\mathcal{B} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ for a normed vector space V, it is common practice to **normalize** each of the basis vectors:

$$\mathbf{v}_1 \mapsto \frac{\mathbf{v}_1}{\|\mathbf{v_1}\|}, \ldots, \mathbf{v}_n \mapsto \frac{\mathbf{v}_n}{\|\mathbf{v_n}\|}$$

Such a normalized set is called an **orthonormal basis**.

It is often helpful to take a set of (not necessarily orthogonal) vectors and transform them into an orthonormal set. To do this one must apply the Graham-Schmidt procedure:

Definition 3.3.7 (Graham Schmidt Procedure)**.** Take $A = \{\mathbf{v_1}, ..., \mathbf{v_n}\}$. We want to turn it into an orthonormal set $E = \{\mathbf{e}_1, ..., \mathbf{e}_n\}$. Apply the following procedure:

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1 \\
\mathbf{u}_2 &= \mathbf{v}_2 - P_{\mathbf{u_1}}\mathbf{v}_2 \\
\mathbf{u}_3 &= \mathbf{v}_3 - P_{\mathbf{u_1}}\mathbf{v}_3 - P_{\mathbf{u_2}}\mathbf{v}_3 \\
&\vdots \\
\mathbf{u}_n &= \mathbf{v}_3 - \sum_{i=1}^{n} P_{\mathbf{u_i}}\mathbf{v}_3 \\
\mathbf{e}_i &= \frac{\mathbf{u}_i}{\|\mathbf{u}_i\|}
\end{aligned}$$

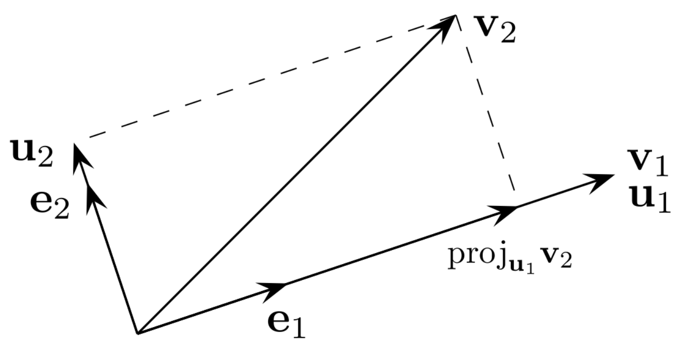


Figure 3.3.1: How the Graham-Schmidt procedure works in two dimensions

Example 3.3.2 (Extending to a Basis)**.** We are often tasked with finding an orthonormal basis for a vector space while only knowing one or two of the basis vectors. We can employ the Graham Schmidt procedure to find vectors which are orthogonal to $\mathbf{x}_1$ and $\mathbf{x}_2$ and thereby extend our initial set of vectors to a basis. You are given the vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ as defined below:

$$\mathbf{x}_1 = \frac{1}{\sqrt{6}}\begin{bmatrix}2\\1\\1\end{bmatrix},\ \mathbf{x}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\-1\\1\end{bmatrix}$$

Our next step is to take a guess at a possible third vector. Since our vectors share the $\hat{k}$ component, it is reasonable to begin with the following guess:

$$\mathbf{v_3} = \begin{bmatrix}0\\0\\1\end{bmatrix}$$

We then employ the Graham-Schmidt procedure:

$$\mathbf{u}_3 = \begin{bmatrix}0\\0\\1\end{bmatrix} - \frac{1}{6}\begin{bmatrix}2\\1\\1\end{bmatrix} - \frac{1}{2}\begin{bmatrix}0\\-1\\1\end{bmatrix}$$

$$\mathbf{u}_3 = \frac{1}{3}\begin{bmatrix}-1\\1\\1\end{bmatrix} \Rightarrow \mathbf{x}_3 = \frac{1}{\sqrt{3}}\begin{bmatrix}-1\\1\\1\end{bmatrix}$$

This same technique is applied to many vector spaces and can be done iteratively in order to find bases for infinite dimensional vector spaces as you will see in the next section.

3.3.2 Polynomial Spaces †

The simplest example of a vector space is the set of polynomials with coefficients in k. We denote this $k[x]$.

Definition 3.3.8 (Polynomial Space)**.** Let k be a field/ring (such as $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, or $\mathbb{C}$). We define the polynomial space $k[x]^+$ to be the set of polynomials with coefficients in k:

$$k[x] = \left\{ \sum_{n=0}^{\infty} a_n x^n : a_n \in k \right\}$$

The notation A^+ means that you can only add and subtract polynomials - remember, multiplication is not allowed in a vector space!

Exercise 3.3.1. Prove that $\mathbb{R}[x]^+$ is a vector space.

Definition 3.3.9 ($\mathbb{P}_n(k)$)**.** The subspace $\mathbb{P}_n(k)$ is the set

$$\mathbb{P}_n(k) = \left\{ \sum_{i=0}^{n} a_i x^i : a_n \in k \right\}$$

Note that $k[x]^+$ is infinite dimensional - you need an infinitely large basis to describe the entirety of the vector space.

Before we can define a basis, we need to define an inner product. There are a few different options for us - let's start with one that is very useful in quantum physics:

Example 3.3.3 (Hermite Polynomials)**.** Consider the inner product defined as follows:

$$\langle f(x), g(x) \rangle = \int_{-\infty}^{\infty} f(x) g(x) e^{-x^2} dx$$

One set of polynomials $H_n(x)$ which are orthonormal with respect to this inner product are the Hermite Polynomials. They are given by

$$H_n(x) = \left(2x - \frac{d}{dx} \right)^n \cdot 1$$

Where $H_0 = 1, H_1 = 2x, H_2 = 4x^2 - 2, ...$

You will see in quantum physics 1 and 2 that these polynomials are the solutions to the Schrodinger equation for a harmonic oscillator:

$$\left(\frac{\hbar^2}{2m} \frac{\partial}{\partial z} + \frac{1}{2} m \omega z^2 \right) \Psi(z, t) = E \Psi(z, t)$$

Example 3.3.4 (Laguerre Polynomials)**.** Another option for our inner product is the following one:

$$\langle f, g \rangle = \int_0^\infty e^{-x} f(x) g(x) dx$$

The set of polynomials which are orthogonal with respect to this inner product are called the Laguerre Polynomials.

$$L_n(x) = \frac{1}{n!} \left(\frac{d}{dx} - 1 \right)^n x^n$$

Which yields: $L_0 = 1, L_1 = 1 - x, L_2 = \frac{1}{2}(x^2 - 4x + 2), ...$

These polynomials have multitudes of uses in physics, such as in the solution to the Schrodinger equation for an electron in a hydrogen atom:

$$\Psi(r, \theta, \phi) = R_{n\ell}(r) Y_{m\ell}(\theta, \phi)$$

$$\frac{-\hbar^2}{2\mu} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) - \frac{\ell(1-\ell)R}{r^2} \right] + VR = ER$$

3.3.3 Function Spaces †

As you saw in the section on polynomial spaces, an inner product is often defined in terms of an integral. We will now direct our attention to some special spaces which are important in quantum physics.

Definition 3.3.10 (Hilbert Space)**.** A Hilbert Space is an abstract vector space V which has an inner product. The vector space can be infinite dimensional (as is the case with function spaces) or finite dimensional. In physics we mostly encounter infinite dimensional Hilbert spaces.

Example 3.3.5 (L^p Space)**.** We define the L^p space as the space of functions which are bounded under a certain norm called the p-norm:

$$f \in L^p \iff \|f\|_p = \left(\int_{-\infty}^{\infty} |f(x)|^p dx \right)^{\frac{1}{p}} < \infty$$

The only L^p space which is also an inner product space is L^2, where the inner product is defined as follows:

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)^* g(x) dx$$

Note that $g(x)$ is complex conjugated in the inner product - this allows us to define L^2 on complex valued functions.

You will often see this inner product written as $\langle f|g \rangle = \langle f, g \rangle$. This is called bra-ket notation and will be explored more later.

3.3.4 Brief overview of Operator Spaces and Algebras †

Example 3.3.6 (Operator Space)**.** Let's look at a normed vector space $(V, \|\cdot\|_V)$ and a vector space $(W, \|\cdot\|_W)$ over k. Consider the set of linear operators $L : V \to W$ with the following property:

$$\exists M \in k : \forall \mathbf{v} \in V, \|L\mathbf{v}\|_W \leq M \|\mathbf{v}\|_V$$

These operators are called bounded and this set is called the space of bounded linear operators.

See the Wikipedia page on "Bounded Operators" for more information.

Exercise 3.3.2. Show that the set of bounded operators follows the vector space axioms, and hence that it is a vector space over k.

Definition 3.3.11 (Algebra)**.** An **algebra** is a vector space with a bilinear product.

Example 3.3.7. Any inner product space is an algebra with respect to the inner product.

Example 3.3.8. $\mathbb{R}^n$ and $\mathbb{C}^n$ are algebras with respect to the cross product.

Example 3.3.9. The space of matrices $\mathcal{M}_{n\times n}$ is an algebra with respect to matrix multiplication.

Example 3.3.10 (Operator Algebra)**.** To continue Example 3.3.5, we define an operator algebra. Certain sets of linear operators which are continuous/bounded can be studied as an algebra with respect to composition:

Consider the bounded linear operators $L : V \to W$, $M : W \to L$. The compositions $L \circ M$ and $M \circ L$ are bounded linear operators, but are defined in different operator spaces ($V \to V$ and $W \to W$). We can see that composition is a bilinear form, and that it has interesting relationships with operator spaces. Studying this bilinear form is extremely important in high level quantum physics and functional analysis. We will see more of these in later chapters.

Now that we have seen a lot of different vector spaces, it would be nice to learn about some applications.

3.4 Complex Matrices and Vectors

It is imperative to discuss complex matrices. Consider the matrix space $\mathcal{M}_{n\times n}(\mathbb{C})$. In this space, every matrix has n eigenvalues (up to multiplicity) due to the fundamental theorem of algebra). If one or more of these eigenvalues is zero then the matrix is non-invertible. Consider $M \in \mathcal{M}_{n\times n}(\mathbb{C})$.

$$M = \begin{bmatrix} m_{11} & \dots & m_{n1} \\ \vdots & \ddots & \vdots \\ m_{1n} & \dots & m_{nn} \end{bmatrix}$$

Definition 3.4.1 (Complex Transpose)**.** The Hermitian adjoint (often called the complex transpose) of a matrix is defined as the following:

$$M^{\dagger} = (M^*)^T = \begin{bmatrix} m_{11}^* & \dots & m_{1n}^* \\ \vdots & \ddots & \vdots \\ m_{n1}^* & \dots & m_{nn}^* \end{bmatrix}$$

In general the Hermitian adjoint of a linear operator is a concept that can be generalized even to spaces in which the concept of a matrix does not make sense. In an infinite dimensional vector space such as a function space we can't make every linear operator into a matrix. As such, we have the following definition:

Definition 3.4.2 (Hermitian Adjoint)**.** The Hermitian adjoint of an abstract linear operator $A : V \to W$ is the unique operator $A^{\dagger} : W \to V$ satisfying the following identity for $\mathbf{v} \in V$ and $\mathbf{w} \in W$

$$\langle (A\mathbf{v}), \mathbf{w}\rangle = \langle \mathbf{v}, (A^{\dagger}\mathbf{w})\rangle$$

Exercise 3.4.1. Show that definition 3.4.1 and 3.4.2 are equivalent in a finite-dimensional vector space over $\mathbb{C}$.

Definition 3.4.3 (Unitary)**.** A linear operator M is unitary if $M^{-1} = M^{\dagger}$.

Theorem 3.4.1. The columns of a unitary matrix are orthonormal. Consequently, unitary matrices correspond to rotations.

Intuitively, if you imagine the unitary matrix M taking the unit vectors and moving them to the positions represented by M's columns, you will see that because the length of the vectors does not change, and the angle between them doesn't change, that the vectors must have only rotated around by some angle.

Definition 3.4.4 (Hermitian)**.** A linear operator M is Hermitian if $M = M^{\dagger}$.

Hermitian matrices are some of the most important matrices you will encounter in physics. One of the postulates of quantum mechanics is that every observable quantity/quantum number corresponds to the eigenvalue of a Hermitian operator (called an "observable"). To demonstrate why this is important we introduce the Schrodinger Equation in the Heisenberg form:

$$\hat{H}\Psi = E\Psi, \text{ where } \hat{H} = \hat{H}^\dagger$$

The operator $\hat{H}$ is called the Hamiltonian, and the eigenvalues E_i correspond to all of the measurable energies that the particle represented by the vector Ψ can have.

In the case where there are only a finite number of states we can always find a diagonal matrix so that the Schrodinger equation reduces to a simple linear system of equations.

We follow this up with a nice computational trick:

Theorem 3.4.2. Consider an eigenvalue λ of M. Then λ^* is also an eigenvalue of M.

Something else you encounter often is Schur decomposition. There are many different ways to factor matrices, and the Schur decomposition is probably the most interesting one. We will explore something similar later on called the Jordan Normal form of a matrix.

Definition 3.4.5 (Schur Decomposition)**.** Given a matrix M, one may write $M = QUQ^{-1}$, where Q is unitary and U is an upper triangular matrix. The process of finding Q and U is **unitary triangularization** or **Schur decomposition**.

This decomposition leads us to our following definition, since we would like to know what conditions are required for the Schur matrix to be diagonal.

Definition 3.4.6 (Normal Matrix)**.** A linear operator M is normal if $M^\dagger M = MM^\dagger$

Theorem 3.4.3 (Spectral Theorem)**.** Every normal operator N is unitarily diagonalizable, that is to say:

$$N = U^\dagger DU, \text{ where } U \text{ is some unitary operator}$$

Theorem 3.4.4. Unitary, symmetric, and Hermitian matrices are normal.

Due to the importance of the Spectral Theorem, we will often call the unitary diagonalization of a matrix the **spectral decomposition**. In future chapters we will discuss the importance of this decomposition to the field of *Spectral Theory*, particularly in the context of functional analysis.

An important skill for you to have is diagonalizing complex matrices. This is one of the primary problem solving tools in PHYS234. You should do as many practice problems for these as you can.

Example 3.4.1 (Complex Diagonalization)**.** We will unitarily diagonalize the following Hermitian matrix:

$$F = \begin{bmatrix} 3 & 2+i \\ 2-i & 7 \end{bmatrix}$$

Step one is to solve the characteristic polynomial:

$$\det \begin{bmatrix} 3-\lambda & 2+i \\ 2-i & 7-\lambda \end{bmatrix} = \lambda^2 - 10\lambda + 16 = (\lambda - 2)(\lambda - 8)$$

Then we find the eigenvectors:

$$F - \lambda_1 \mathbb{I} = \begin{bmatrix} 1 & 2+i \\ 2-i & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 2+i \\ (2+i)(2-i) & 5(2+i) \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2+i \\ (2+i)(2-i) & 5(2+i) \end{bmatrix} \sim \begin{bmatrix} 1 & 2+i \\ 5 & 5(2+i) \end{bmatrix} \sim \begin{bmatrix} 1 & 2+i \\ 1 & 2+i \end{bmatrix} \sim \begin{bmatrix} 1 & 2+i \\ 0 & 0 \end{bmatrix}$$

$$\text{So } (\mathbf{v}_1, \lambda_1) = \left(\begin{bmatrix} 2+i \\ -1 \end{bmatrix}, 2 \right)$$

$$F - \lambda_2 \mathbb{I} = \begin{bmatrix} -5 & 2+i \\ 2-i & -1 \end{bmatrix} \sim \begin{bmatrix} -5 & 2+i \\ 5 & -1(2+i) \end{bmatrix}$$

$$\begin{bmatrix} -5 & 2+i \\ 5 & -1(2+i) \end{bmatrix} \sim \begin{bmatrix} -5 & 2+i \\ 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -(2+i)/5 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1/(2-i) \\ 0 & 0 \end{bmatrix}$$

$$\text{So } (\mathbf{v}_2, \lambda_2) = \left(\begin{bmatrix} -1 \\ -2+i \end{bmatrix}, 2 \right)$$

Recall that eigenvectors are orthogonal, so constructing a unitary diagonalizing matrix only requires normalizing the eigenvectors.

$$\frac{\mathbf{v}_i}{\|\mathbf{v}_i\|} = \frac{1}{\sqrt{6}} \mathbf{v}_i \text{ (since both eigenvectors have the same length)}$$

$$F = UDU^\dagger \Rightarrow \begin{bmatrix} 3 & 2+i \\ 2-i & 7 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 2+i & -1 \\ -1 & -2+i \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} 2-i & -1 \\ -1 & -2-i \end{bmatrix}$$

Example 3.4.2 (Diagonalization in Calculus †)**.** Here's a wacky example courtesy of FlammableMaths on youtube. By the time you are reading this you are probably somewhat familiar with integration. Let's take a look at how we can integrate an interesting function (twice).

$$I = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\Big(-\Big\langle \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}\Big\rangle\Big) dxdy$$

Let's make some notational shortcuts for a moment to spare us pain when writing.

$$I = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp(-\langle \mathbf{v}, A\mathbf{v}\rangle) dxdy$$

Recall that the inner product produces a *scalar*, so we are actually integrating something familiar to us. This integral looks fairly complicated: Let's expand it into terms we understand more.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\Big(-\Big\langle \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} 2x-y \\ -y+2x \end{bmatrix}\Big\rangle\Big) dxdy$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\big(-(2x^2 - xy - y^2 + 2xy)\big) dxdy$$

Unfortunately, the terms containing xy are extremely difficult (if not impossible) to integrate. How can we simplify this integral?

Well, we should note that $A = A^T$. Because A is symmetric it is *normal*, and so by the Spectral Theorem we can unitarily (orthogonally in this case because it is a real matrix) diagonalize it. Let's look at an interesting property of symmetric matrices:

$$N^T = N \Rightarrow \langle \mathbf{v}, N\mathbf{v}\rangle = \mathbf{v}^T N\mathbf{v} = (N^T\mathbf{v})^T\mathbf{v} = (N\mathbf{v})^T\mathbf{v} = \langle N\mathbf{v}, \mathbf{v}\rangle$$

Because orthogonal matrices are (skew) symmetric, we can take a moment to look at the diagonalized version of A.

$$A = NDN^T \Rightarrow \langle \mathbf{v}, A\mathbf{v}\rangle = \langle \mathbf{v}, N^TDN\mathbf{v}\rangle = \langle N\mathbf{v}, DN\mathbf{v}\rangle$$

Aha! This might be useful. Let's write $\mathbf{u} = N\mathbf{v}$ (a classic u-substitution), and check whether we need to change our integral much to make it in terms of $\mathbf{u}$. Recall that when we make a u-substitution we need to convert $dx \mapsto \frac{du}{dx}dx$. Now remember that a determinant is the amount by which a linear map "stretches" areas on a grid. Because the only thing changing about the infinitesimal area $dxdy$ is the size, we can just use the determinant: $dxdy = |N|du_1du_2$. But remember: the absolute value of the determinant of a "unit"ary matrix is 1 - meaning that nothing changes at all! Applying these discoveries gives us:

$$I = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp(-\langle \mathbf{u}, D\mathbf{u}\rangle) du_1du_2$$

Now we start the arduous process of finding D.

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$\det(A - \lambda \mathbb{I}) = \det \begin{bmatrix} 2-\lambda & -1 \\ -1 & 2-\lambda \end{bmatrix} = \lambda^2 - 4\lambda + 3 = (\lambda - 1)(\lambda - 3)$$

$$\begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} \sim \begin{bmatrix} -1 & -1 \\ 0 & 0 \end{bmatrix} \Rightarrow \mathbf{k}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \mapsto \frac{1}{\sqrt{2}}\mathbf{k}_1$$

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \Rightarrow \mathbf{k}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mapsto \frac{1}{\sqrt{2}}\mathbf{k}_2$$

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}, D = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{u} = U\mathbf{v} \text{ and } UDU^\dagger = A$$

$$I = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(-\langle \mathbf{u}, D\mathbf{u} \rangle) du_1 du_2$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\Big(- \Big\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \Big\rangle \Big) du_1 du_2$$

Now that we have the integral in terms where you can apply the exponential rule $\exp(a+b) = \exp(a)\exp(b)$, it is possible to split the integral in two (for a rigorous reason why this is allowed, see Fubini's theorem).

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\big(-(3u_1^2 + u_2^2)\big) du_1 du_2 = \int_{-\infty}^{\infty} \exp\big(-3u_1^2\big) du_1 \int_{-\infty}^{\infty} \exp\big(-u_2^2\big) du_2$$

As these are Gaussian integrals, you do not have to do any more work. Just look up or memorize the answer because integrating these yourself is a pain. You will learn how to do it yourself in calculus 3 (or in calculus 2 if you want to teach yourself - it isn't conceptually difficult, just tedious).

$$I = \int_{-\infty}^{\infty} \exp\big(-3u_1^2\big) du_1 \int_{-\infty}^{\infty} \exp\big(-u_2^2\big) du_2 = \sqrt{\frac{\pi}{3}} \sqrt{\frac{\pi}{1}} = \frac{\pi}{\sqrt{3}}$$

Performing integrals like this one has use in statistics, particularly in the case of a three dimensional Gaussian distribution.

3.5 Misc. Topics

This subsection concerns itself with the various extra things you should learn a bit about before second year. In particular, having familiarity with the ins and outs (pun intended) of Dirac notation is going to help greatly if you want to read ahead (or take PHYS234 over the summer). So let's begin with Dirac notation:

3.5.1 Dirac Notation and Outer Products

Dirac notation (also called Bra-Ket notation) is the way we take all of our strange notions of vector spaces and merge them into one.

A vector (in the regular sense) is written as $|\psi\rangle$. Whatever space this vector is in is not of great concern as long as the vector transforms in the way we want it to (with respect to change of base matrices and so on).

Similarly, a "row vector" (a dual vector) is written as $\langle\psi|$. We therefore denote the inner product $\langle\psi, \phi\rangle = \langle\psi|\phi\rangle$. Let's look at some examples:

Example 3.5.1 ($\mathbb{C}^3$)**.**

$$|\phi\rangle = \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}, |\psi\rangle = \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix}$$

$$\langle\phi|\psi\rangle = |\phi\rangle^\dagger |\psi\rangle = |\phi\rangle \cdot |\psi\rangle = a_1 a_2 + b_1 b_2 + c_1 c_2$$

Example 3.5.2 (L^2 Space)**.**

$$|\xi\rangle = \xi(x), |\zeta\rangle = \zeta(x)$$

$$\langle\xi|\zeta\rangle = \int_{-\infty}^{\infty} \xi(x)^* \zeta(x) dx$$

For a thorough overview of doing linear algebra with Dirac notation, check out *Principles of Quantum Mechanics* by R. Shankar. The whole first chapter is dedicated to getting you intricately familiar with Dirac notation while giving you extra practice in linear algebra as a good refresher. Another amazing book in the context of quantum physics is *Quantum Mechanics: A Paradigms Approach* by D. McIntyre. This book is great and also begins with Dirac notation. It is often used as the book for PHYS234. Let's look at one last definition then we will be done with linear algebra for now.

Definition 3.5.1 (Outer Product on a Finite Vector Space)**.** Consider a vector space $V = k^n$. The "inner product" $\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^T \mathbf{w}$ has a simple interpretation. The **outer product (dyadic product)** is defined similarly:

$\mathbf{v} \otimes_D \mathbf{w} = \mathbf{v}\mathbf{w}^T = |v\rangle\langle w|$

If you calculate this explicitly, you get an $n \times n$ matrix. It turns out that this is a special case of something called a "tensor product", which we will explore in later chapters.

Example 3.5.3 (Density Matrix)**.** The density matrix of a quantum system is the matrix which represents the state of a quantum system with multiple (usually non-interacting) objects.

Let's define our system so that the outcome of a single measurement is one of the vectors $|+\rangle$ or $|-\rangle$ (up or down). Let's imagine that there are two particles for now. This means that there is some probability p_i of measuring $|i\rangle$, where $|1\rangle = (|+\rangle, |+\rangle)$, $|2\rangle = (|+\rangle, |-\rangle)$ and $|3\rangle = (|-\rangle, |-\rangle)$.

We define the matrix ρ as the density matrix:

$$\rho = \sum_{i=1}^{3} p_i \, |i\rangle\langle i|$$

Recall that measurements in quantum mechanics are the eigenvalues of an operator A called an observable. Imagine that you operate on the system ρ with A, and want to know the average value $\langle A \rangle$ taken by your measurements. The following is known from quantum mechanics:

$$\langle A \rangle = \sum_{i=1}^{3} p_i \, \langle i|A|i\rangle$$

From this, one can derive the following alternative expression, which happens to be more versatile in general.

$$\langle A \rangle = \mathrm{Tr}\{\rho A\}$$

The above defines the general expression for an expectation value. The density matrix is nice because we can compute it once and be able to resolve the expectation values of any known observables. The density matrix hence summarizes all of the expectation values for an entire system.

3.6 Exercises for Section 3

Exercise 3.6.1. Diagonalize each of the following matrices over $\mathbb{C}$, or explain why (using theorems making sure to check that *all* of the hypothesis are satisfied) the matrix can not be diagonalized.

(a)

$$\begin{bmatrix} 2 & 1 \\ 4 & 5 \end{bmatrix}$$

(b)

$$\begin{bmatrix} 1 & -i \\ i & -1 \end{bmatrix}$$

(c)

$$\begin{bmatrix} 1 & -i & 2 \\ -1 & 0 & 2i \\ 1 & i & 1-i \end{bmatrix}$$

(d)

$$\begin{bmatrix} 0 & -i & i \\ i & 0 & i \\ -i & -i & 0 \end{bmatrix}$$

**You may use a calculator to factor the characteristic polynomial for part c*

Exercise 3.6.2. Let $\langle \cdot, \cdot \rangle$ be the standard inner product on $\mathbb{R}^n$ (the dot product) calculate $\langle \mathbf{u}, \mathbf{v} \rangle$ where,

(a)

$$\mathbf{u} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -1 \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \mathbf{v} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 1 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

(b)

$$\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 1 \\ 2 \end{bmatrix}$$

Exercise 3.6.3. Calculate the dyadic outer product $\mathbf{u} \otimes_D \mathbf{v}$ of the following vectors, where $\mathbf{u} \otimes_D \mathbf{v} = \mathbf{u}\mathbf{v}^T$.

(a)

$$\mathbf{u} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -1 \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \mathbf{v} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 1 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

(b)

$$\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 1 \\ 2 \end{bmatrix}$$

4 Appendix I: Proofs

Doing mathematical proofs is an essential skill for any mathematician. However, it is not absolutely essential for a physicist to be completely rigorous and structured with their arguments. On the other hand, learning to do proofs will help you make clearer arguments and even in physics it is helpful to show that your methodology actually works in the first place. If you want to go into Applied Mathematics and/or Mathematical Physics, doing proofs is going to be your job. You better learn to do them. There isn't really a good way to write a chapter about "how to do proofs". In my experience the only way to really learn them well is to do examples. Here is some helpful supplementary material:

- *MATH135 Coursenotes*[1]
- *How to Prove It* - Velleman
- *Book of Proof* - Hammack
- *The Art of Problem Solving* - Rusczyk and Lehoczky
- *Set Theory and Metric Spaces* - Kaplansky
- *Calculus* - Spivak

[1] https://cs.uwaterloo.ca/ cbruni/Math135Resources/index.php

5 Appendix II: Useful Courses at UW

5.0.1 Mathematics

- MATH135: If you haven't already considered taking this, you should. The course is pretty useful in all higher level mathematics courses, and will help you solve problems quickly and neatly.
- MATH235: This course isn't the most useful all the time, but everything you learn has some application somewhere in physics. It is definitely a good course
- MATH239: Combinatorics and graph theory are useful if you want to do computational physics, in which case you will be required to put a lot of thought into data structures and optimizing code. Applications may also include some theoretical physics.

5.0.2 Applied Mathematics

- AMATH242 (Computational Methods): If you want to do computational physics or just want a better course than PHYS236, this is a great choice
- AMATH271 (Theoretical Mechanics): Incredible alternative to PHYS263. This course is challenging but well designed.
- AMATH331 (Applied Real Analysis): The course itself does not have a lot of applications in it, but if you look at the textbook (Donsig) you will see that following the content of the course itself is chapters and chapters of material which is incredibly useful to physics. Such applications include Fourier Analysis, Wavelets, and Optimization
- AMATH342 (Computational DEs): Very useful for computational physics
- AMATH363 (Continuum Mechanics): This is just another physics course. Essential for geophysics and fluid dynamics, and helpful for relativity.
- AMATH391 (Wavelets): Good for doing computational and engineering physics. Will be essential if you do optics or signal processing.
- AMATH442 (Computational PDEs): Sequel to AMATH342.
- AMATH456 (Variational Calculus): Very useful course in modern classical mechanics, engineering physics, and various other applications.
- AMATH463 (Fluid Mechanics): Another physics course. Applications are obvious.

5.0.3 Pure Mathematics

- PMATH333 (Real Analysis 0): Great if you want to do theoretical and mathematical physics
- PMATH336 (Applied Group Theory): Useful introduction to groups. Not as in depth as 347 and may not be as well structured.
- PMATH347 (Groups and Rings): Amazing course. Provides deep insight into the mathematical structure of physics. Provides enough background to start looking into the algebraic structures in physics such as modules, operator algebras, etc.
- PMATH351 (Real Analysis 1): Essential for theoretical and mathematical physics. Very difficult.
- PMATH352 (Complex Analysis): Interesting and beautiful subject. The course doesn't teach enough computation (take 332 or teach yourself) but should be useful on its own.
- PMATH365 (Differential Geometry): Not perfect on its own, but a good enough intro to Diff Geo that GR may be slightly easier. If you take this you should follow up with 465.
- PMATH445 (Representation Theory): Useful intro to representations. Only applicable if you follow this up with a thorough study of Lie theory.
- PMATH446 (Commutative Algebra): Useful in Mathematical Physics. Very interesting topics in pure algebra. Provides some insight into the mathematical viewpoint of tensors.
- PMATH450 (Real Analysis 2): Useful in that it leads well into Functional Analysis
- PMATH451 (Measure Theory): Could be useful in quantum theory. Mostly just hardcore as hell
- PMATH453 (Functional Analysis): Essential topic in the mathematical study of quantum physics. One of the primary tools in mathematical physics
- PMATH463 (Algebraic Geometry): Esoteric and wacky. Take this for fun.
- PMATH465 (Manifolds): Very useful in GR and QFTs. Take this for a prelude to high level Lie theory and relativity.
- PMATH467 (Algebraic Topology): Wacky topic, but our only proper course on topology. Probably useful someday.

5.0.4 Combinatorics and Optimization

- CO250 (Optimization): Good for engineering, computational, and applied physics applications.
- CO342 (Graph Theory): In case LQG ever gets off the ground
- CO367 (Nonlinear Optimization): Useful in computational physics (particularly QIP)
- CO456 (Game Theory): Maybe useful in QIP applications
- CO485 (Public Key Crypto): Useful in QIP
- CO487 (Applied Crypto): Useful in QIP

5.0.5 Statistics

- STAT230 (Probability): If you aren't in math phys, you should take this as an elective. It is fun and useful.
- STAT231 (Statistics): Also useful. There is a reason it's required for math phys (even though we always put it off)
- STAT330 (Mathematical Statistics): If you really like statistical mechanics and/or other statistical physics concepts, this is a good idea.
- STAT332 (Experimental Design): Probably very useful for experimental physics
- STAT340 (Simulation): This is an awesome idea if you want to do computational physics.
- STAT341 (Data Analysis): Also a good idea for comp phys and experimental physics.

6 Appendix III: Miscellaneous

6.0.1 The Greek Alphabet

letter, Letter	Name	English Pronounciation (Greek Pronounciation)
α, A	"alpha"	"al fah"
β, B	"beta"	"bey tah"
γ, Γ	"gamma"	"gah mah"
δ, Δ	"delta"	"del tah"
ε/ϵ, E	"epsilon"	"epp see lawn"
ζ, Z	"zeta"	"zee tah"
η, H	"eta"	"ee tah"
θ, Θ	"theta"	"thee tah"
ι, I	"iota"	"eye oh tah" ("yota")
κ, K	"kappa"	"ka pah"
λ, Λ	"lambda"	"lam da"
μ, M	"mu"	"mew" ("mee")
ν, N	"nu"	"new" ("nee")
ξ, Ξ	"xi"/"ksi"	"zeye" ("ksee")
o, O	"omicron"	"oh meecron"
π, Π	"pi"	"pie" ("pee")
ρ, P	"rho"	"row"
σ/ς, Σ	"sigma"	"sig ma" ("seeg ma")
τ, T	"tau"	"t-ow" ("taff")
υ, Υ	"upsilon"	"up see lon" ("eep see lon")
ϕ, Φ	"phi"	"fee"
χ, X	"chi"	"k-eye"
ψ, Ψ	"psi"	"psy"
ω, Ω	"omega"	"oh meh gah"

7 Appendix IV: Planned Content

- Set Theory Notes [Section 4]
- Calculus II [Section 5]
- Calculus III [Section 6]
- Discrete Mathematics [Section 7]
- Differential Equations [Section 8]
- Partial Differential Equations [Section 9]
- Probability [Section 10]
- Statistics [Section 11]
- Applied Real Analysis [Section 12]
- Topics in Fourier Analysis [Section 13]
- Numerical Methods [Section 14]
- Complex Analysis [Section 15] - Left this late because some complex methods will be covered in earlier sections
- Abstract Algebra [Section 16]
- Advanced Topology [Section 17]
- Advanced Real Analysis [Section 18]
- Variational Calculus [Section 19] - Left this late because variational methods will be discussed earlier
- Differential Geometry [Section 20]
- Lie Groups and Lie Algebras [Section 21]
- Functional Analysis [Section 22]
- Algebraic Topology [Section 23] - left this late because it isn't as useful
- Convexity and Optimization [Section 24] - same as 24
- Operator Theory [Section 25]
- Measure Theory [Section 26]
- Advanced Numerical Methods [Section 27]

List of Exercises

8 References

[1] *About Topology*. URL: http://www.uncg.edu/mat/comgrouptop/abouttop.html.

[2] *Argand Diagram*. URL: http://mathworld.wolfram.com/ArgandDiagram.html.

[3] Evan Chen. *An Infinitely Large Napkin*. DRAFT. URL: http://web.evanchen.cc/napkin.html.

[4] *Distance in Taxi-cab*. URL: https://mat292project.weebly.com/distance-in-taxi-cab.html.

[5] *Linear Approximations*. URL: http://tutorial.math.lamar.edu/Classes/CalcI/LinearApproximations.aspx.

[6] *Mathematical analysis*. Nov. 2018. URL: https://en.wikipedia.org/wiki/Mathematical_analysis.

[7] *Shading Venn Diagrams*. URL: https://www.onlinemathlearning.com/shading-venn-diagrams.html.

[8] *The Four Conic Sections*. URL: https://www.cliffsnotes.com/study-guides/algebra/algebra-ii/conic-sections/the-four-conic-sections.

[9] *Topology*. Dec. 2018. URL: https://en.wikipedia.org/wiki/Topology.

[10] *Topology: Sequentially Compact Spaces and Compact Spaces*. Feb. 2013. URL: https://mathstrek.blog/2013/02/24/topology-sequentially-compact-spaces-and-compact-spaces/.

[11] *Vector Field*. URL: http://mathworld.wolfram.com/VectorField.html.