Chapter 0

Preliminaries

This chapter contains some background material that you should have learned in high school but probably did not. Most likely your teachers skipped by these topics so that you could be trained to factor polynomials as fast as is humanly possible so that you would perform well on various mindless standardized tests.

Before reading further consider the following question: why does -1 times -1 equal 1? How could this be proved? And, you might ask, why don't you know this already?

1 Sets and Functions

1.1 Sets

A **set** is a collection of **elements**. The expression $p \in S$ means p is an element of the set S. A set may be defined in several ways: in ordinary English, *e.g.,* let A be the set of positive even integers; by listing its elements within braces, *e.g.,* let $A = \{2, 4, 6, 8, ...\}$; or by using "set builder" notation, *e.g.,* $A = \{n \in \mathbb{Z} \mid n > 0 \, \text{and} \, n \, \text{is even}\}$, read: A is the set of all integers n such that $n > 0$ and n is even ($\mathbb{Z}$ is the standard notation for the integers).

A set does not have an order. Thus $\{a, b\} = \{b, a\}$. An **ordered set** is a set together with an ordering. When we want to stress that a set has been endowed with an ordering we will use parenthesizes instead of braces: (a, b) is an ordered set and is not equal to (b, a).

The following notations are standard:

- $\phi = \{\}$, the empty set.
- $A \subset B$: read A is a subset of B, meaning, every element of A is an element of B.
- $A \cup B$: A union B, meaning, the set of all elements that are in A **or** in B.
- $A \cap B$: read A intersection B, meaning, the set of all elements that are in A **and** in B.
- $A - B$: read A minus B, meaning, the set of all elements of A that are not elements of B.
- $A \times B$: read A cross (product) B, meaning, the set of ordered pairs (a, b) where $a \in A$ and $b \in B$. Since there is a natural one-to-one correspondence between $(A \times B) \times C$ and $A \times (B \times C)$, $((a, b), c) \longleftrightarrow (a, (b, c))$, we shall ignore the distinction between them and use the notation $A \times B \times C$ for the set $\{(a, b, c) \mid a \in A, b \in B, \text{and } c \in C\}$. Other multiple cross products are defined similarly.
- $A^n = A \times \cdots \times A$, n times.

Some standard sets are:

- $\mathbb{Z}$: the integers (from the German *zummer*)
- $\mathbb{Q}$: the rationals (quotients)
- $\mathbb{R}$: the reals
- $\mathbb{C}$: the complex numbers

Remark. The sets $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ are normally given an ordering. Interestingly, $\mathbb{C}$ is not typically ordered.

Interval Notation:

$$\begin{array}{rclrcl}
[a,b] & = & \{x \in \mathbb{R} \mid a \le x \le b\} & [a,\infty) & = & \{x \in \mathbb{R} \mid a \le x\} \\
(a,b) & = & \{x \in \mathbb{R} \mid a < x < b\} & (a,\infty) & = & \{x \in \mathbb{R} \mid a < x\} \\
(a,b] & = & \{x \in \mathbb{R} \mid a < x \le b\} & (-\infty,b] & = & \{x \in \mathbb{R} \mid x \le b\} \\
[a,b) & = & \{x \in \mathbb{R} \mid a \le x < b\} & (-\infty,b) & = & \{x \in \mathbb{R} \mid x < b\}
\end{array}$$

Examples:

- $\{x \in \mathbb{R} \,|\, x \leq -\sqrt{7}\} \cup \{x \in \mathbb{R} \,|\, x \geq \sqrt{7}\}$ is the solution set for $x^2 - 7 \geq 0$.
- $\mathbb{R} - \{0\}$ is the natural domain of $1/x$.
- $\mathbb{R}^2$ is the plane. $\mathbb{R}^3$ is 3-dimensional space. $\mathbb{R}^4$ is 4-dimensional space. And so on.
- $\phi \subset A$, $\phi = A \cap \phi$, and $A = A \cup \phi$ are true statements for all sets A.
- $\{x \in \mathbb{R} \,|\, -2 \leq x < 5\} = [-2, 5) = [-2, 7] \cap (-10, 5)$.
- $\{2, 3\}^2 = \{(2, 2), (2, 3), (3, 2), (3, 3)\}$.
- $S = [0, 1] \times [0, 1]$ is the *unit square* in the plane $\mathbb{R}^2$ with corners (0,0), (1,0), (0,1), and (1,1).

1.2 Functions

Intuitively, a function f from a set A to a set B assigns to each element of A one element of B. Formally, f is a subset of $A \times B$ such that for every $a \in A$ there is one and only one $b \in B$ with $(a, b) \in f$. We normally write $f : A \to B$, and express $(a, b) \in f$ by $b = f(a)$.

A function $f : A \to B$ is **onto** if for every $b \in B$ there is at least one $a \in A$ such that $(a, b) \in f$, *i.e.,* such that $f(a) = b$. A function $f : A \to B$ is **one-to-one** if for every $b \in B$ there is at most one $a \in A$ with $f(a) = b$.

Let $f : A \to B$, $A' \subset A$, and $B' \subset B$. Then we define,

- $f(A') = \{b \in B \,|\, b = f(a) \text{ for at least one } a \in A'\}$ and is call the **image** of A' under f. We call $f(A)$ the **range** of f.
- $f^{-1}(b) = \{a \in A \,|\, b = f(a)\}$.
- $f^{-1}(B') = \{a \in A \,|\, a \in f^{-1}(b) \text{ for at least one } b \in B'\}$

If f is one-to-one and onto then $f^{-1}(b)$ always consists of a single element and we regard f^{-1} as a function from B to A. In this case we say f is **invertible**.

A **binary operation** is a function from the cross product of two sets to a third set. For example, the adding of two numbers is a binary operation from $\mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$. So is multiplication. For any binary operation $f : A \times B \to C$, if $a_1 = a_2 \in A$ and $b \in B$ then $f(a_1, b) = f(a_2, b)$. For multiplication this means for real numbers a, b, and c, if $a = b$ then $ac = bc$. Note that we have written $f(a, b)$ instead of $f((a, b))$ since this shorthand is customary.

Example 1. Let $S = \{\clubsuit, \diamondsuit, \heartsuit, \spadesuit, \Box, \bigcirc, \star\}$ and let $L = \{c, d, h, s, z\}$. Let $f : S \to L$ be defined as indicated by Figure 1. But what *is* f really? It is the set of arrows. But each arrow is a pictorial representative of an ordered pair. Thus $(\clubsuit, c) \in f$ but $(\diamondsuit, z) \notin f$. Or, equivalently, $f(\clubsuit) = c$ while $f(\diamondsuit) \neq z$. This function is not one-to-one since, for example, $f(\clubsuit) = f(\bigcirc)$. It is not onto since there is no $x \in S$ such that $f(x) = z$, that is, for every $x \in S$, $(x, z) \notin f$. Or, we could say z is not in the range of f.

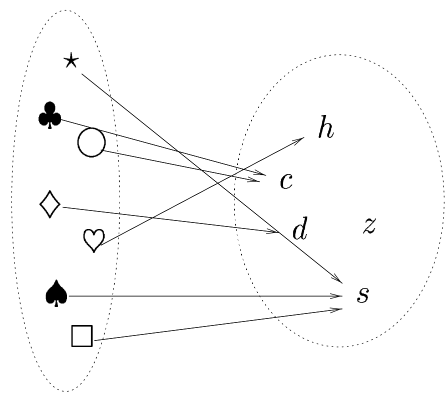

Figure 1: A function

If we order the elements of S and L then we can **graph** f. This is shown in Figure 2. Notice that the familiar *vertical line test* shows that f is not one-to-one. We can see that the graph of f is a subset of $S \times L$.

Additional Examples:

- Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x^2$. Then $f^{-1}(4) = \{-2, 2\}$, $f^{-1}([0, 1]) = [-1, 1]$, and $f^{-1}([1, 9]) = [-3, -1] \cup [1, 3]$.
- Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = \sin \pi x$. Then $f^{-1}(0) = \mathbb{Z}$, and

$$f^{-1}([0, 1]) = \cdots \cup [-4, -3] \cup [-2, -1] \cup [0, 1] \cup [2, 3] \cup \cdots$$

- Let $A = \{1, 2, 3, ...\}$. Then the set $\{(1, 2), (2, 3), (3, 4), ...\} \subset A \times A$, is the function $f : A \to A$ produced by adding a one: $f(n) = n + 1$. It is

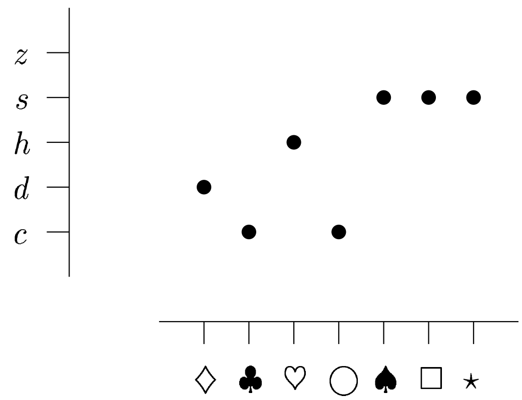

Figure 2: A graph of the function in Figure 1

one-to-one but not onto. But if we let $B = A - \{1\}$ and let $g : A \to B$ be addition by one, then g is onto.

- The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is the set $\{(x, y) \in \mathbb{R}^2 \mid y = x^2\}$.
- Let $A = \{2, 3\}$. Let $f = \{(2, 3), (3, 3)\}$, $g = \{(2, 3), (3, 2)\}$, and $h = \{(2, 2), (3, 2)\}$. Then f is a function which is not one-to-one or onto, g is a one-to-one onto function, while h is not a function. Check that $g^{-1}(f(3)) = 3$ and $f(g(f(x))) = g(f(g(x)))$ for all $x \in A$.

Problems:

1. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x, y) = \sqrt{x^2 + y^2}$. Draw a picture of $f^{-1}([4, 9])$.
2. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x, y) = \sin x \cos y$. Find $f^{-1}(1)$ and $f^{-1}(0)$.
3. Let $f : \mathbb{R} \to \mathbb{R}^2$ be defined by $f(x) = (x, x^2)$. Show that f is one-to-one but not onto.
4. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $f(x, y) = (3x + 2y, x - y)$. Show that f is one-to-one and onto. Find f^{-1}.
5. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $f(x, y) = (x + y, x + y)$. Show that f is neither one-to-one nor onto.

2 The Axioms of Arithmetic

Let's play a game. Let R be a nonempty set. We shall suppose that we have two binary operations on R. The first is called *addition*. Given a and b in R addition gives an element $a+b$ in R. The other is called *multiplication*. Given a and b in R multiplication gives $a \cdot b \in R$. We shall assume that these two operations obey the axioms listed below. The game is to prove facts about R based solely on these axioms.

Axioms: For all a, b and c in R the following hold.

a. $a+b=b+a$ (addition is commutative)

b. $a+(b+c)=(a+b)+c$ (addition is associative)

c. There is an element $z \in R$, independent of a, such that $z+a=a$ (an additive identity exists)

d. There is an $\bar{a} \in R$, which depends on a, such that $\bar{a}+a=z$ (additive inverses exist)

e. $a \cdot b = b \cdot a$ (multiplication is commutative)

f. $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (multiplication is associative)

g. There is an element $u \in R$, independent of a, such that $u \cdot a = a$ (a multiplicative identity exists)

h. If a is not an additive identity, there is an $\hat{a} \in R$, which depends on a, such that $\hat{a} \cdot a = u$ (multiplicative inverses exist)

i. $a \cdot (b+c) = a \cdot b + a \cdot c$ (multiplication distributes over addition)

Applications:

1. There is only one additive identity element in R. *Proof:* Suppose z_1 and z_2 are both additive identity elements. Then by (c) $z_1+z_2=z_2$ and $z_2+z_1=z_1$. But, by (a), $z_1+z_2=z_2+z_1$. Thus, $z_1=z_2$.

We are now justified in saying that the "zero element" is unique and shall denote it by 0.

2. There is only one multiplicative identity element in R. *Proof:* **Problem 1**. The unique "unity element" shall be denoted by 1.

3. Additive inverses are unique. *Proof 1:* Let $a \in R$. Suppose $\bar{a}_1$ and $\bar{a}_2$ are additive inverses of a. Then $\bar{a}_1 = 0 + \bar{a}_1 = (\bar{a}_2 + a) + \bar{a}_1 = \bar{a}_2 + (a + \bar{a}_1) = \bar{a}_2 + (\bar{a}_1 + a) = \bar{a}_2 + 0 = \bar{a}_2$. The reader should check that each step used exactly one of the axioms. *Proof 2:* $a + \bar{a}_1 = 0 \Rightarrow \bar{a}_2 + (a + \bar{a}_1) = \bar{a}_2 + 0 \Rightarrow (\bar{a}_2 + a) + \bar{a}_1 = \bar{a}_2 \Rightarrow 0 + \bar{a}_1 = \bar{a}_2 \Rightarrow \bar{a}_1 = \bar{a}_2$. Note that we have used a basic property of all binary functions in adding $\bar{a}_2$ to both sides of an equation and have freely used more than one axiom per step.

4. Multiplicative inverses are unique. *Proof:* **Problem 2**.

5. Let $a \in R$. Then $a \cdot 0 = 0$. *Proof:* $a \cdot 0 = a \cdot 0 + 0 = a \cdot 0 + (a + \bar{a}) = (a \cdot 0 + a) + \bar{a} = (a \cdot 0 + a \cdot 1) + \bar{a} = a \cdot (0 + 1) + \bar{a} = a \cdot 1 + \bar{a} = a + \bar{a} = 0$. The reader should check each step to see which of the axioms are being applied.

6. Let $a \in R$. Then $\bar{\bar{a}} = a$. *Proof:* **Problem 3**. Hint: Start with $\bar{\bar{a}} = \bar{\bar{a}} + 0$.

7. Let $a \in R$. Then $\bar{a} = \bar{1} \cdot a$. *Proof:* Since additive inverses are unique we need only show that $a + \bar{1} \cdot a = 0$. $a + \bar{1} \cdot a = 1 \cdot a + \bar{1} \cdot a = a \cdot 1 + a \cdot \bar{1} = a \cdot (1 + \bar{1}) = a \cdot 0 = 0$. Notice the last step uses 5.

8. Let $a \in R - \{0\}$. Then $\hat{\hat{a}} = a$. *Proof:* **Problem 4**.

9. Let a and b be in R and suppose that $a \cdot b = 0$. Then either $a = 0$ or $b = 0$. *Proof:* **Problem 5**.

10. Let $a + c = b + c$. Then $a = b$. *Proof:* **Problem 6**.

11. **Problem 7**: Let $ac = bc$. Show that it need not follow that $a = b$.

If we let R be the real numbers $\mathbb{R}$ then the axioms apply to the normal addition and multiplication operations. It is customary to denote the additive inverse of a by $-a$ and its multiplicative inverse by a^{-1} or $1/a$, for $a \neq 0$.

Problem 8. Prove that $-1 \times -1 = 1$.

If we let R be rationals $\mathbb{Q}$, or the complex numbers $\mathbb{C}$, then the axioms still apply. This is clear for $\mathbb{Q}$. But for $\mathbb{C}$ it takes a bit of effort to show this. For the integers $\mathbb{Z}$ only axiom h fails to hold.

Problem 9. It is easy to check that $\mathbb{C}$ obeys axioms a through g and i. The only difficulty is axiom h. Let $a + ib \in \mathbb{C} - \{0\}$. Find $c + id \in \mathbb{C}$ such that $(a + ib)(c + id) = 1$, and thus establish axiom h.

Project 1. Let $\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\}$. Define addition and multiplication as follows. Let $a \oplus b$ be the remainder of $a + b$ divided by n and let $a \otimes b$ the remainder of $a \times b$ divided by n. For example, in $\mathbb{Z}_7$ we get $5 \oplus 5 = 3$, because $10 \div 7$ has remainder 3; and $4 \otimes 5 = 6$, because $20 \div 7$ has remainder 6. The set $\mathbb{Z}_n$ is called the *integers modulo* n and the operations are referred to as *modular arithmetic*.

(a) Show that $\mathbb{Z}_n$ satisfies axioms a through g and i.

(b) Show that $\mathbb{Z}_7$ satisfies axiom h but that $\mathbb{Z}_6$ does not.

(c) Study various $\mathbb{Z}_n$. Under what conditions does $\mathbb{Z}_n$ satisfy axiom h?

Appendix to Section 2: Why do proofs?

Many of you will find the concept of a proof difficult and frustrating. This is because (1) the concept is difficult, (2) the reason for doing proofs may not be clear to you, and (3) the public schools have watered down much of the mathematics curriculum.

Let's address (2) first.

Proofs are to mathematics what experiments are to science: the test of truth. But there is difference. Science is based on **inductive reasoning**, while mathematics is based on **deductive reasoning**. Scientists will repeat an experiment many times. The results may confirm a given hypothesis. But this does not prove the hypothesis since further testing may produce contrary evidence. Through many trails, careful measurements and statistical analysis scientists gradually form increasingly accurate models of the physical, biological, and more recently, the social and economic worlds.

Mathematics is a abstract science. It does not deal directly with the objects of the world. Mathematics deals with abstract structures: numbers, equations, sets, operations like multiplication and integration, and so on. Mathematics is useful because many of the models used in science are mathematical in nature. The logical structures within mathematics seem to mirror patterns in nature. No one fully understands why this is. But, insights gleamed from mathematical proofs develop thinking patterns that are useful in broader areas.

A second reason for doing mathematical proofs in frankly political. If university courses were taught on the basis that truth comes from authority then students would fail to incorporate democratic values. They would come to feel comfortable in authoritarian settings. There is a connection between the academic's push for critical thinking (asking why?, demanding proof!) and the citizen's demand for accountability from political and business leaders. The view of most employers is mixed. They value employees who can think of new ways of doing things, but are afraid of employees who challenge the boss'authority. However, having employees who speak out may well be good for the economy as a whole!

But if you are still not convinced you are left with a conundrum: if you are willing to believe in a formula or theorem without proof because your professor says so, then you have to accept that you should understand proofs because your professor says so. (Note: there is a fallacy in the previous sentence; can you find it?)

Let's work through one of the proofs from the last section in greater detail. This will begin to address point (1) above. We will redo Application 5.

Claim. Let R be a nonempty set that obeys the axioms listed in Section 2. Let $a \in R$, then $a \cdot 0 = 0$.

We will use the *two column format* for our proof:

STEP	REASON
Let $a \in R$	R is nonempty
$a \cdot 0 = 0 + a \cdot 0$	axiom c
$= (a + \bar{a}) + a \cdot 0$	axiom d
$= (\bar{a} + a) + a \cdot 0$	axiom a
$= \bar{a} + (a + a \cdot 0)$	axiom b
$= \bar{a} + (a \cdot 1 + a \cdot 0)$	axiom g
$= \bar{a} + a \cdot (1 + 0)$	axiom i
$= \bar{a} + a \cdot (0 + 1)$	axiom a
$= \bar{a} + a \cdot 1$	axiom c
$= \bar{a} + 1 \cdot a$	axiom e
$= \bar{a} + a$	axiom g
$= 0$	axiom d

It is not necessary to do each step separately. Here is a shorted version of the same proof:

STEP	REASON
Let $a \in R$	R is nonempty
$a \cdot 0 = 0 + a \cdot 0$	axiom c
$= (a + \bar{a}) + a \cdot 0$	axiom d
$= \bar{a} + (a + a \cdot 0)$	axioms a and b
$= \bar{a} + (a \cdot 1 + a \cdot 0)$	axiom g
$= \bar{a} + a \cdot (1 + 0)$	axiom i
$= \bar{a} + 1 \cdot a$	axioms a, c and e
$= 0$	axioms g and d

How many steps is it okay to combine? Your reader should be able to reconstruct a complete one-step-at-a-time proof from your proof. Indeed, mature writes rarely employ the two-column format and instead write in standard English. The main point is that each step has to be justified by an axiom, a hypothesis, a previously know result or a basic law of logic. If you are new to proofs it is best to stick with one-step-at-a-time proofs, at least for now.

The concept of an **axiom**, or rather of an **axiomatic framework**, may also be new to you. Mature branches of mathematics and logic are governed by a collections of basic assumptions or postulates called axioms. For the set of real numbers we have listed those that deal with arithmetic. There are others that deal with ordering properties, e.g., $a > b \Longrightarrow a + c > b + c$. And there are axioms that allow us to analyze notions convergence and continuity of real valued functions on the real line. The theory of sets is defined in terms of axioms, see for example *Axiomatic Theory of sets and Classes,* by Murray Eisenberg. This was forced on mathematicans because there appeared to be logical paradoxes early on in set theory. The point of axioms is to take mathematical thought and break it down in to its most basic units. These basic units can then be seen as building blocks which are combined in accordance with logic. This is what we mean by deductive reasoning. Because each step is simple, sometimes ridiculously so, the whole structure is sound. Unlike music, language, politics or religion, mathematical concepts are universal across all modern cultures. You might compare the axioms developed by mathematicians to the way chemists break down ordinary matter into atoms.

I first encountered axioms and sets in a middleclass public school in the 6th grade. Times have changed. The following project is an attempt to address point (3).

Project 2. Use the internet to look up the address of your high school. Write a letter to the principal outlining how the education you received there has helped or hindered your progress in college.

> [Some] think that truth is only what sounds nice. If truth should prove to be something statistical, dry, or factual, something difficult to find and requiring study, they do not recognize it as truth; it does not intoxicate them. — Bertolt Brecht [1898–1956], German playwrite.

3 The Principle of Mathematical Induction

The **Principle of Mathematical Induction (PMI)** is just the following observation. Let $P(n)$ be a statement for each positive integer n. If $P(1)$ is true and if $P(k) \implies P(k+1)$ for all positive integers k, then $P(n)$ is true for all positive integers n. In other words, if $P(1)$ and $P(k) \implies P(k+1)$ then $P(1) \implies P(2) \implies P(3) \implies P(4) \implies \cdots$.

We will give three examples of proofs that use the Principle of Mathematical Induction.

Example 1 (The power rule). We will take the the product rule for derivatives as given: $(fg)' = f'g + fg'$. Also assume that $x' = 1$ is given. We will prove that $(x^n)' = nx^{n-1}$ for all positive integers n. But first we will check the power rule for some small values of n.

$$(x^1)' = x' = 1 = 1x^0$$

$$(x^2)' = (xx)' = x'x + xx' = 1x + x1 = 2x = 2x^1$$

$$(x^3)' = (x^2x)' = (x^2)'x + x^2x' = 2xx + x^2 = 3x^2$$

$$(x^4)' = (x^3x)' = (x^3)'x + x^3x' = 3x^2x + x^3 = 4x^3$$

So far so good. But this could take a very long time. To get around the fact that we cannot possibly check a statement for every positive integer mathematicians invented the PMI. Let $n = k$ be a fixed but arbitrary positive integer and suppose that $(x^k)' = kx^{k-1}$ is true. Then let $n = k + 1$ and compute:

$$(x^{k+1})' = (x^kx)' = (x^k)'x + x^kx' = kx^{k-1}x + x^k = (k+1)x^k$$

Thus, the conditions for the PMI are in place. The power rule is true for $n = 1$ and if it is true for $n = k$ it is true for $n = k + 1$. Thus, by the PMI the power rule $(x^n)' = nx^{n-1}$ is true for all positive integers n.

Example 2. Prove that

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

for all positive integers n. *Proof:* The formula works for $n = 1$ since it reduces to $1 = 1$. Suppose that for an arbitrary fixed positive integer $n = k$ it is true that

$$\sum_{i=1}^{k} i = \frac{k(k+1)}{2}$$

Let $n = k+1$. Now

$$\begin{aligned}\sum_{i=1}^{n} i &= \sum_{i=1}^{k+1} i \\ &= \left(\sum_{i=1}^{k} i\right) + (k+1) \\ &= \frac{k(k+1)}{2} + (k+1) \\ &= \frac{k(k+1)+2(k+1)}{2} \\ &= \frac{(k+1)(k+2)}{2} \\ &= \frac{n(n+1)}{2}.\end{aligned}$$

Thus, by PMI the hypothesized summation formula holds true for all positive integers n.

Example 3. Let $f : \mathbb{R} \to \mathbb{R}$ be a function such that $f(ab) = f(a)f(b)$. Then $f(a^n) = (f(a))^n$. *Proof:* The equation is true for $n = 1$ since it reduces to $f(a) = f(a)$. Suppose that for an arbitrary fixed positive integer $n = k$ it is true that $f(a^k) = (f(a))^k$. Then if we let $n = k+1$ and $b = a^k$ we get $f(a^{k+1}) = f(aa^k) = f(a)f(a^k) = f(a)(f(a))^k = (f(a))^{k+1}$. Thus, by the PMI, $f(ab) = f(a)f(b)$ implies $f(a^n) = (f(a))^n$ for all positive integers n.

Problem 1. Let $(\frac{1}{x})' = -\frac{1}{x^2}$ and $(fg)' = f'g + fg'$ be given. Prove that $\left(\frac{1}{x^n}\right)' = -\frac{n}{x^{n+1}}$, for all positive integers n.

Problem 2. Prove that $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$, for all positive integers n.

Problem 3. Let $f : \mathbb{R} \to \mathbb{R}$ be a function such that $f(ab) = f(a) + f(b)$. Prove that $f(a^n) = nf(a)$ for all positive integers n.

Problem 4. Consider a list of n symbols where $n \geq 2$. We wish to switch the first symbol with last symbol leaving the others unchanged. But, we can only switch two adjacent symbols at a time; call this operation a *simple switch move*. Prove that the number of simple switch moves needed to switch the first and last symbol, leaving the others unchanged, is always an odd number. We will use this result later.

4 Vectors

On one level a vector is just a point; we can regard every point in $\mathbb{R}^2$ as a vector. When we do so we will write $\langle a, b\rangle$ instead of the usual (a, b). In most physical problems vectors are regarded as directed magnitudes; they have a size and a direction. To emphasize this vectors are drawn as arrows in the plane. Often the base of the arrow is at the origin and the head is at the point (a, b) for the vector $\langle a, b\rangle$. But in many applications the base of the arrow is placed elsewhere. For example, the velocity vector of a moving particle might be placed with its arrow based at the location of the particle.

The difference between points and vectors is that while points are purely geometric, the set of vectors in the plane is endowed with an algebraic structure. We define the following three binary operations:

- Vector addition: $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ is given by $\langle a_1, b_1\rangle + \langle a_2, b_2\rangle = \langle a_1 + a_2, b_1 + b_2\rangle$.
- Scalar multiplication: $\mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ is given by $r\langle a, b\rangle = \langle ra, rb\rangle$.
- The dot product: $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ is given by $\langle a_1, b_1\rangle \bullet \langle a_2, b_2\rangle = a_1b_1 + a_2b_2$.

Vectors and these vector operations can be defined for any $\mathbb{R}^n$. Let $\mathbf{v} = \langle v_1, v_2, \ldots, v_n\rangle$ and $\mathbf{u} = \langle u_1, u_2, \ldots, u_n\rangle$ be vectors in $\mathbb{R}^n$. Then:

- $\mathbf{v} + \mathbf{u} = \langle v_1 + u_1, v_2 + u_2, \ldots, v_n + u_n\rangle$.
- For $r \in \mathbb{R}$, $r\mathbf{v} = \langle rv_1, rv_2, \ldots, rv_n\rangle$.
- $\mathbf{v} \bullet \mathbf{u} = v_1u_1 + v_2u_2 + \cdots + v_nu_n = \displaystyle\sum_{i=1}^{n} v_iu_i$.

Theorem 4.1. *The vector operations have the following properties:*

(a) $\mathbf{u} \bullet \mathbf{u} \geq 0$, *and is 0 if and only if* $\mathbf{u} = \mathbf{0} \overset{def}{=} \langle 0, 0, ..., 0\rangle$.

(b) $\mathbf{v} \bullet \mathbf{u} = \mathbf{u} \bullet \mathbf{v}$.

(c) $\mathbf{v} \bullet (\mathbf{u} + \mathbf{w}) = \mathbf{v} \bullet \mathbf{u} + \mathbf{v} \bullet \mathbf{w}$.

(d) $r(\mathbf{u} \bullet \mathbf{v}) = (r\mathbf{u}) \bullet \mathbf{v}$.

(e) $r(\mathbf{u} + \mathbf{v}) = r\mathbf{v} + r\mathbf{u}$.

Problem 1. Your instructor will tell which of these two problems is right for your level:
(a) Prove Theorem 4.1 for vectors in $\mathbb{R}^3$
(b) Prove Theorem 4.1 for vectors in $\mathbb{R}^n$

Theorem 4.2. *Let* $\mathbf{v}$ *be a vector in* $\mathbb{R}^2$ *or* $\mathbb{R}^3$ *and define* $|\mathbf{v}|$ *to be* $\sqrt{\mathbf{v} \bullet \mathbf{v}}$. *Then* $|\mathbf{v}|$ *is the length of* $\mathbf{v}$. *If* $\mathbf{v}_1$ *and* $\mathbf{v}_2$ *are nonzero vectors in* $\mathbb{R}^2$ *or* $\mathbb{R}^3$ *and* θ *is the acute angle between them, then* $\mathbf{v}_1 \bullet \mathbf{v}_2 = |\mathbf{v}_1||\mathbf{v}_2| \cos\theta$. *In particular,*

$$\mathbf{v}_1 \perp \mathbf{v}_2 \iff \mathbf{v}_1 \bullet \mathbf{v}_2 = 0.$$

Problem 2. Your instructor will tell which of these two problems is right for your level:
(a) Prove Theorem 4.2 for vectors in $\mathbb{R}^2$
(b) Prove Theorem 4.2 for vectors in $\mathbb{R}^3$
Hint: Use the Law of Cosines.

Remark. For vectors in $\mathbb{R}^n$ the formulas in Theorem 4.2 are used to define length and angle in $\mathbb{R}^n$.

Problem 3. Let $\mathbf{u} = \langle 1, 2, 3 \rangle$, $\mathbf{v} = \langle 1, 1, 0 \rangle$, and $\mathbf{w} = \langle 2, -2, 0 \rangle$.
(a) Compute each of the following expressions or explain why the expression is nonsense: (i) $\mathbf{v} \bullet (\mathbf{u} + 3\mathbf{w})$, (ii) $3\mathbf{u} + \mathbf{v} \bullet \mathbf{w}$, (iii) $(\mathbf{v} \bullet \mathbf{u} - \mathbf{w} \bullet \mathbf{w})^2$, (iv) $|\mathbf{v}| + |\mathbf{u}|$, (v) $|\mathbf{w}|\mathbf{v}$, (vi) $|\mathbf{u}||\mathbf{w}| + \mathbf{v}$.
(b) Find the angles between $\mathbf{u}$ and $\mathbf{v}$, $\mathbf{u}$ and $\mathbf{w}$, and $\mathbf{v}$ and $\mathbf{w}$.

Problem 4. Let $\mathbf{u} = \langle 1, 2, 3, 4 \rangle$, and $\mathbf{v} = \langle 1, 1, 1, 1 \rangle$. Find the acute angle between them.

For vectors in $\mathbb{R}^2$ there is a useful graphical method of adding vectors. Suppose we want to add $\mathbf{v} = \langle 2, 1 \rangle$ and $\mathbf{u} = \langle 1, 2 \rangle$. Place an arrow with its tail at the origin and its head at $(2, 1)$ to represent $\mathbf{v}$. Next place an arrow with its tail at the origin and its head at $(1, 2)$ to represent $\mathbf{u}$. Now, move

$\mathbf{u}$ so that its base is at the head of $\mathbf{v}$. The head of $\mathbf{u}$ is now located at the point $(3,3)$. Draw an arrow with its tail at the origin and its head at $(3,3)$. This new vector $\langle 3,3\rangle$ is just $\mathbf{v}+\mathbf{u}$. See Figure 3.

Subtraction is similar. To continue the same example, let $\mathbf{w}=\langle 3,3\rangle$. To find $\mathbf{w}-\mathbf{u}$, start with by drawing the arrow for $\mathbf{w}$. Then draw $-\mathbf{u}$ by flipping $\mathbf{u}$ through the origin, so that $-\mathbf{u}$ has its head at $(-1,-2)$. Move $-\mathbf{u}$ so that its tail is at the head of $\mathbf{w}$. Now the head of $\mathbf{u}$ is at $(2,1)$ and we *see* that $\mathbf{w}-\mathbf{u}=\mathbf{v}$.

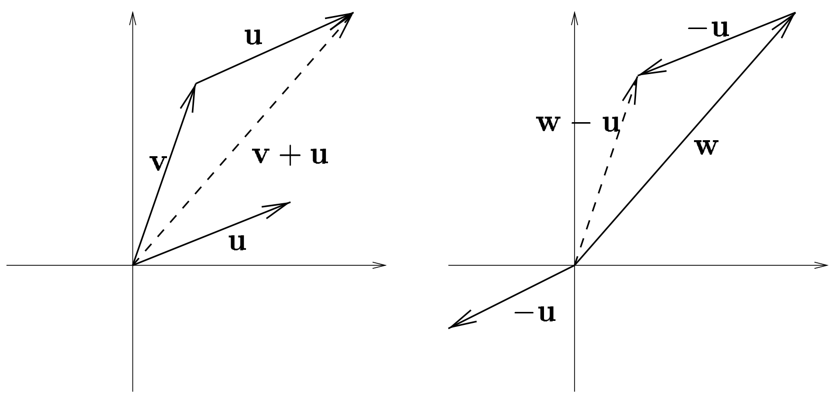

Figure 3: Vector addition

The scalar multiples of a vector are even easier to visualize. Figure 4 shows $3\mathbf{v}$, $-2\mathbf{v}$, and $\frac{1}{2}\mathbf{v}$.

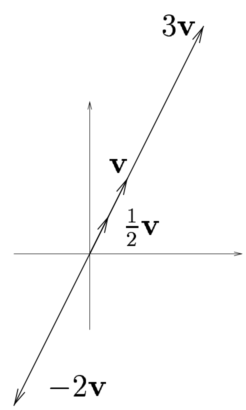

Figure 4: Scalar multiplies

Optional Problems

The problems below will not be needed for what follows. Your instructor may or may not want you to skip them.

Problem 5. (a) Prove that for vectors in any $\mathbb{R}^n$ the following holds:

$$|\mathbf{u}+\mathbf{v}|^2 + |\mathbf{u}-\mathbf{v}|^2 = 2|\mathbf{u}|^2 + 2|\mathbf{v}|^2.$$

Hint: Use only the properties of vectors in Theorem 4.1.
(b) This formula is called the *Parallelogram Law*. Draw some pictures for vectors in $\mathbb{R}^2$ and explain why the formula has this name.

Problem 6. The *Cauchy-Schwartz inequality* says $|\mathbf{u}\bullet\mathbf{v}| \leq |\mathbf{u}||\mathbf{v}|$. Prove this as follows. Observe that for all $r \in \mathbb{R}$, $0 \leq |r\mathbf{u}-\mathbf{v}|^2$. Expand the square and use the substitution $r = (\mathbf{u}\bullet\mathbf{v})/|\mathbf{u}|$. (The result is obvious when $\mathbf{u} = \mathbf{0}$.)

Problem 7. The *Triangle inequality* says $|\mathbf{u}+\mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}|$.
a) Prove this. You will need the Cauchy-Schwartz inequality.
b) Draw some vectors in the plane and explain why this inequality has its name.

Let I be a given interval in the real line. A function $f : \mathbb{R} \to \mathbb{R}$ is *square integrable with respect to* I if the definite integral over I of $f^2(x)$ exists and is finite: $\int_I f^2(x)\,dx < \infty$. Define $f \bullet g$ to be $\int_I f(x)g(x)\,dx$. Define $|f| = \sqrt{f \bullet f}$.

Problem 8. For any given interval I prove that Theorem 4.1 holds for square integrable functions. For Theorem 4.1(a) you may restort to an intuitive agrument since a formal proof involves ideas you probably did not encounter in calculus.

Problem 9. Restate the Parallelogram Law, the Cauchy-Schwartz and Triangle inequalities, as theorems about integration. Since their proofs only used Theorem 4.1 they remain valid.

5 Lines and Planes

The material here probably goes a bit beyond the typical high school level.

5.1 Lines in the Plane

Every line of points L in $\mathbb{R}^2$ can be expressed as the solution set for an equation of the form $Ax + By = C$. The equation is not unique for if we multiply both sides by any nonzero number the solution set is unchanged. Any line L can also be expressed by a pair of parametric equations of the form:

$$\begin{aligned} x(t) &= at + b \\ y(t) &= ct + d \end{aligned}$$

These can be rewritten in vector form: $\langle x, y\rangle = \langle a, c\rangle\, t + \langle b, d\rangle$. The vectors $\langle a, c\rangle$ and $\langle b, d\rangle$ have a nice geometric/physical interpretation. Regard t as time. Let $\mathbf{p}(t) = \langle x(t), y(t)\rangle$ and call it the *position vector*. One can image a particle moving along L in accordance with the given parametric equations. Then $\mathbf{p}(0) = \langle b, d\rangle$ is the *initial position*. Notice,

$$\frac{d\mathbf{p}}{dt} = \langle x'(t), y'(t)\rangle = \langle a, c\rangle$$

Thus, we call $\mathbf{v} = \langle a, c\rangle$ the *velocity vector*. It is parallel to L. It is customary to place its base point on L. See Figure 5(left side).

We now give a geometric interpretation for the "ABC" form of an equation of a line. First, suppose $C = 0$; this just means the line L goes through the origin. Let $\mathbf{n} = \langle A, B\rangle$, and again set $\mathbf{p} = \langle x, y\rangle$. Then we have $\mathbf{n} \bullet \mathbf{p} = 0$. That is the vectors $\mathbf{n}$ and $\mathbf{p}$ are at right angles to each other. Thus, the line L for $Ax+By = 0$ is the set of all points (x, y) such that $\langle x, y\rangle$ is perpendicular to $\langle A, B\rangle$.

Now we consider the general case: $Ax + By = C$. Pick some particular point on the line and call it (x_0, y_0). Then $C = Ax_0 + By_0$. Therefore, for any point (x, y) on L we have $Ax + By = Ax_0 + By_0$. We can rewrite this as

$$\begin{aligned} Ax - Ax_0 + By - By_0 &= 0 \\ A(x - x_0) + B(y - y_0) &= 0 \\ \langle A, B\rangle \bullet \langle x - x_0, y - y_0\rangle &= 0 \\ \mathbf{n} \bullet (\langle x, y\rangle - \langle x_0, y_0\rangle) &= 0 \\ \mathbf{n} \bullet (\mathbf{p} - \mathbf{p}_0) &= 0 \end{aligned}$$

In the last line we have let $\mathbf{p}_0 = \langle x_0, y_0\rangle$. The vector $\mathbf{p} - \mathbf{p}_0$ can be thought of as lying in L with its tail at (x_0, y_0) and its head at (x, y).

Thus, L is the unique line perpendicular to the vector $\mathbf{n} = \langle A, B \rangle$ that passes through (x_0, y_0). See Figure 5(right side). The vector $\mathbf{n}$ is called a *normal vector* for the line L. Given a vector to use as normal vector and a point we can easily find an equation for the corresponding line.

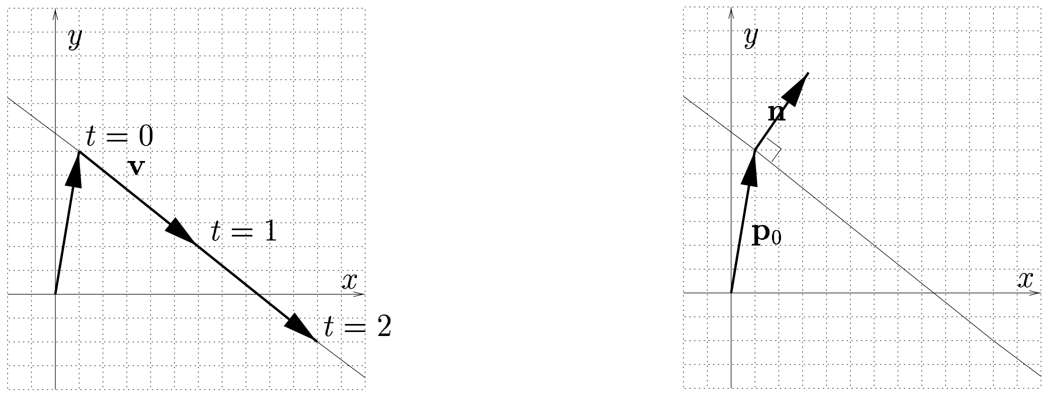

Figure 5: Left: A parametric line. Right: Normal vector to a line.

5.2 Lines and Planes in 3-space

The three dimensional set $\mathbb{R}^3$ is the set of all triples (x, y, z) where x, y, and z are real numbers. Such a triple is called the xyz-coordinates of a point. These are also called *rectilinear coordinates*. The set $\{(x, 0, 0) \mid x \in \mathbb{R}\}$ is the x-axis. The y and z axes are defined similarly. They are clearly lines. The set $\{(x, y, 0) \mid x \in \mathbb{R},\ y \in \mathbb{R}\}$ is the xy-plane. The yz and xz planes are defined similarly. Visualizing structures in three dimensions takes practice.

Any line L in $\mathbb{R}^3$ can expressed parametrically in the form:

$$\begin{aligned} x(t) &= at + b \\ y(t) &= ct + d \\ z(t) &= et + f \end{aligned}$$

or, in vector form, $\langle x, y, z \rangle = \langle a, c, e \rangle\, t + \langle b, d, f \rangle$. As with lines in $\mathbb{R}^2$ it is useful to think of $\langle a, c, e \rangle$ as a velocity vector and $\langle b, d, f \rangle$ as the position at $t = 0$.

However, there is no way to express a line in $\mathbb{R}^3$ as a single equation in three variables. In fact, we will show that "typically" solution sets of

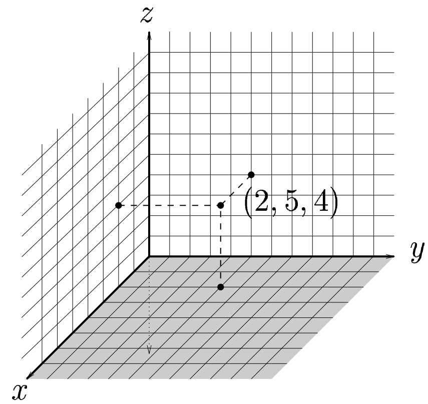

Figure 6: Three dimensional space: $\mathbb{R}^3$

equations of the form $Ax + By + Cz = D$ are planes and that every plane in $\mathbb{R}^3$ is the solution set of some equation in this form. Note: If $A = B = C = D = 0$, the solution set is all of $\mathbb{R}^3$; if $A = B = C = 0$ but $D \neq 0$ the solution set is empty.

Example 1. Convince yourself of the following:

- If $A = B = D = 0$ and $C \neq 0$ then $Ax + By + Cz = D$ is the xy-plane.
- If $A = C = D = 0$ and $B \neq 0$ then $Ax + By + Cz = D$ is the xz-plane.
- If $B = C = D = 0$ and $A \neq 0$ then $Ax + By + Cz = D$ is the yz-plane.

Let's consider the case where $D = 0$. Let $\mathbf{n} = \langle A, B, C\rangle$ and $\mathbf{p} = \langle x, y, z\rangle$. Then the equation $Ax + By + Cz = 0$ becomes $\mathbf{n} \bullet \mathbf{p} = 0$. Thus, the solution set is the plane P, passing through the origin of $\mathbb{R}^3$ whose points, when regarded as vectors, are perpendicular to $\mathbf{n}$.

We return to the general case: $Ax+By+Cz = D$. Let $\mathbf{p}_0 = \langle x_0, y_0, z_0\rangle$ be some fixed point that satisfies the given equation. We leave it to the reader to show that

$$\mathbf{n} \bullet (\mathbf{p} - \mathbf{p}_0) = 0.$$

Thus, the solution set of $Ax + By + Cz = D$ is the unique plane passing through $\mathbf{p}_0$ and perpendicular to $\mathbf{n} = \langle A, B, C\rangle$.

Example 2. Consider $x + y + z = 1$. The points $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$ satisfy the equation. We can connect them with line segments and visualize the triangle thus formed. This triangle sits in the plane. If we place the tail of $\mathbf{n} = \langle 1, 1, 1\rangle$ at any point of the triangle it is easy to see that it is perpendicular to the plane.

Example 3. Let P_1 be the plane given by $2x + 3y - z = 5$ and let P_2 be the plane given by $x + y + z = 1$. Find parametric equations for the line $L = P_1 \cap P_2$.

Solution.

$$\left.\begin{array}{rcl} 2x + 3y - z &=& 4 \\ x + y + z &=& 1 \end{array}\right\} \implies y - 3z = 2.$$

Let $z = t$. Then $y = 3t + 2$. Also, $x = 1 - y - z = -4t - 1$. We rewrite these as $\langle x, y, z\rangle = \langle -4, 3, 1\rangle\, t + \langle -1, 2, 0\rangle$. □

Example 4. Find an equation for the plane passing through the three points $(1, 1, 1)$, $(1, 2, 3)$, and $(2, -1, 0)$.

Solution. We have three conditions and these give us three equations in four unknowns.

$$\left.\begin{array}{rcrcrcl} A &+& B &+& C &=& D \\ A &+& 2B &+& 3C &=& D \\ 2A &-& B && &=& D \end{array}\right\} \implies$$

$$\left.\begin{array}{rcrcrcr} A &+& B &+& C &=& D \\ && B &+& 2C &=& 0 \\ && -3B &-& 2C &=& -D \end{array}\right\} \implies$$

$$\left.\begin{array}{rcrcrcr} A &+& B &+& C &=& D \\ && B &+& 2C &=& 0 \\ && && C &=& -D/4 \end{array}\right\} \implies$$

$$\left.\begin{array}{rcrcrcr} A &+& B && &=& 5D/4 \\ && B && &=& D/2 \\ && && C &=& -D/4 \end{array}\right\} \implies \begin{array}{rcr} A &=& 3D/4 \\ B &=& D/2 \\ C &=& -D/4 \end{array}$$

Any nonzero value of D will do. Let $D = 4$. Then $3x + 2y - z = 4$ is an equation for our plane. □

Problem 1. Consider the three points $(1,1,1)$, $(2,0,2)$, and $(4,-1,4)$. Show that they do not determine a unique plane because they lie on the same line. Find an equation for this line; write it in vector form.

Problem 2. Let P be the plane given by $x+2y-3z=1$. Let L_{xy} be the intersection of P with the xy-plane. Define L_{xz} and L_{yz} similarly. Find equations for these three lines in "ABC" form.

Problem 3. Graph, separately, each of the planes determined by these three equations: $2x+2y-3z=1$, $x+2y+4z=-1$, and $3x-2y-2z=7$.

Problem 4. Find the point of intersection of the three planes determined by these three equations: $2x+2y-3z=1$, $x+2y+4z=-1$, and $3x-2y-2z=7$.

Problem 5. Show that the two planes determined by $2x+2y-3z=1$ and $4x+4y-6z=0$ do not intersect and are thus parallel.

Problem 6. Let P be the plane given by $2x+3y-2z=1$. Let L be the line given by $\langle x,y,z\rangle = \langle 1,1,1\rangle\, t + \langle 1,0,1\rangle$. Find the point where they meet.

Problem 7. Show that these four points lie in the same plane: $(1,1,-1)$, $(-1,0,0)$, $(-1,1,-\frac{1}{2})$, and $(1,-1,0)$. Find an equation for this plane.

5.3 Parametric Equation for a Plane

There is another form for equations of planes in $\mathbb{R}^3$ that is the analog of the parametric form for equations of a line. The difference is we will need two parameters, r and s, instead of one. Of course, the time metaphor is no longer useful.

Let P be a plane given by $Ax+By+Cz=D$. Assume that $C \neq 0$. Then we can solve for z and get $z = D/C - A/Cx - B/Cy$. (If $C=0$ solve for x or y instead.) Think of z as the height above the xy-plane. Now let $x=r$ and $y=s$, and think of r and s as free parameters. We can now write

$$\begin{aligned} \langle x,y,z\rangle &= \langle r,s,D/C - A/Cr - B/Cs\rangle \\ &= \langle 0,0,D/C\rangle + r\,\langle 1,0,-A/C\rangle + s\,\langle 0,1,-B/C\rangle \end{aligned}$$

This equation is far from unique. We can start with any point $(x_0,y_0,z_0) \in P$, regard it as a vector $\mathbf{p}_0 = \langle x_0,y_0,z_0\rangle$ and add multiplies of $\langle 1,0,-A/C\rangle$ and

$\langle 0, 1, -B/C \rangle$ to it and stay in the plane. Furthermore, if we let $\mathbf{v}_1$ and $\mathbf{v}_2$ be nonzero multiples of $\langle 1, 0, -A/C \rangle$ and $\langle 0, 1, -B/C \rangle$, respectively then

$$\mathbf{p} = \mathbf{p}_0 + r\mathbf{v}_1 + s\mathbf{v}_2$$

gives the same plane P. Indeed, we could use any pair of vectors in P with tails at $\mathbf{p}_0$ as long as they point in different directions.

We will use this formulation to place a coordinate system on P. Take a point (x_0, y_0, z_0) on P and call it the origin of P. Then any point on P can be gotten to by adding multiples of $\mathbf{v}_1$ and $\mathbf{v}_2$ to $\mathbf{p}_0$. Thus, for any point on P we can think of it as having coordinates (r, s). See Figure 7.

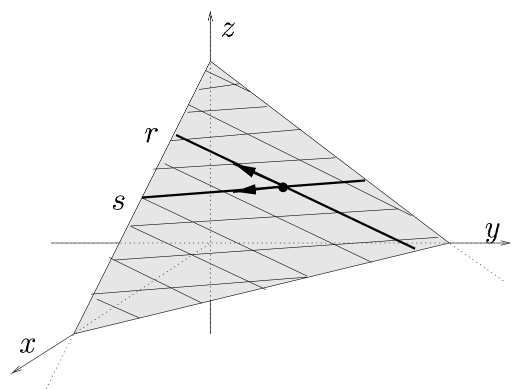

Figure 7: Coordinates for a plane: The dark lines are the r and s-axes

Example 5. Define a plane P by

$$\langle x, y, z \rangle = \langle 1, 2, 3 \rangle + r \langle 1, 1, 0 \rangle + s \langle 0, 1, 1 \rangle$$

Show that the point (0,2,4) is on P and find its rs-coordinates.

Solution. We have three equations and two unknowns.

$$\left.\begin{array}{l} 0 = 1 + 1r + 0s \\ 2 = 2 + 1r + 1s \\ 4 = 3 + 0r + 1s \end{array}\right\} \implies \begin{array}{c} r = -1 \\ s = 1 \end{array}$$

Thus, $(0, 2, 4) \in P$ and it has rs-coordinates $(-1, 1)$ relative to the given parametric equation. $\square$

Problem 8. Using the same plane P in Example 5, find the rs-coordinates of $(3, 3, 2)$.

Problem 9. Show that the point $(1, 2, -1)$ is not on the plane P of Example 5.

Problem 10 (Hard). The equation $2r + 3s = 1$ determines a line L in the plane P of Example 5, using rs-coordinates. Find a parametric equation for L in xyz-coordinates.