# MAE 301 Foundations of Secondary School Mathematics

## Fall 2006 Syllabus

This course has several goals. First, to ensure that you have fully mastered the secondary school curriculum. Second, to develop an understanding of the interconnections among the different branches of secondary school mathematics. Third, to develop an understanding of the connections between the secondary school curriculum and your college level mathematics curriculum. Fourth, to enhance your problem solving abilities, and to develop an understanding of the processes involved in learning and understanding mathematics and in mathematical problem solving.

**Instructor:** Professor Bernard Maskit

**Office:** Math 5-112

**Phone**: 632-8257

**e-mail:** bernie@math.sunysb.edu

**Office Hours:** Mondays: 12:00 - 3:00 p.m., by appointment only

Tuesdays: 2:00 - 4:00 p.m.

Wednesdays: 2:00 - 3:00 p.m.

There is no textbook for this course; class notes can be found at http://www.math.sunysb.edu/~bernie/classnotes.pdf

Class worksheets, which will also contain homework assignments, will be distributed in class.

MAE 301 is part of a sequence of courses leading to certification as a teacher of mathematics in grades 7-12. In light of the professional nature of this course, neither excessive absence nor excessive tardiness will be tolerated.

**First Exam:** There will be an in-class exam on Wednesday, September 20. The questions on this exam are taken directly from the New York State A and B Mathematics Regents exams. The passing grade on this exam is 85; students who do not achieve a passing grade will have two opportunities to pass an equivalent exam. While we will attempt to follow the course outline below, to some extent, the amount of time spent on any one topic will depend on the exam scores on questions related to this topic.

**Other examinations:** In addition to the regents exam mentioned above, there will be a midterm exam and a final exam.

**Calculators:** New York State requires graphing calculators on its Regents exams; you will not be able to solve some of the problems on examinations in this course without a graphing calculator. However, overly fancy calculators, such as those that do symbolic manipulations, or have telephones, or have overly extensive memories, will not be permitted.

**Homework:** Homework will be assigned in most classes, usually as part of the classroom worksheet. In general, the homework will be due the following class. A selection of the assigned homework problems will be graded. At the end of the semester, the lowest two homework grades for each student will be dropped.

**Classwork:** Class assignments will vary. For some, you will be required to work in groups assigned by the instructor; for some, you will be permitted to work in groups of your own choosing; for some, you will be required to work alone.

**Grading:** The final grade has two components.

Component 1:

First Exam: 10%

Midterm Exam: 20%

Final Exam: 40%

Classwork and

Attendance: 15%

Homework: 15%

Component 2: If you do not achieve a grade of at least 85% on the First exam (or on a subsequent make-up exam), your final grade will be at most C-.

Note: If you have a physical, psychological, medical or learning disability that may impact on your ability to carry out assigned course work, you are urged to contact the staff in the Disabled Student Services office (DSS). They will review your concerns and determine, with you, what accommodations are necessary and appropriate. All information and documentation of disability is confidential. Note that we cannot make special arrangements for students with disabilities except for those determined by DSS.

Tentative Course Outline

1. Introduction, Natural numbers and their properties, sets, prime numbers, integers
2. Formal logic, Boolean algebra, sets and circuits, logic in natural language, the language of mathematics
3. Functions, Cartesian products, relations, equivalence relations
4. Fractions, rational numbers, infinite decimals,
5. Regents exam
6. real numbers, complex numbers, quarternions and vectors
7. Cardinality --- The transfinite world
8. Real functions, approximations, Mean Value theorem and Taylor's theorem, the use of calculators and computers
9. Continuation of measurements; use of trigonometry
10. Mathematical induction
11. Polynomials, factorization, the Euclidean algorithm, division, roots
12. Combinatorics: factorials and the binomial theorem
13. Basic probability: coin tosses, dice, Bernoulli trials
14. Basic statistics: sampling, plots, mean, median, mode, standard deviation
15. Regression and related concepts
16. The laws of exponents, exponential functions
17. Probability distribution functions; the normal curve
18. Review
19. Midterm exam (covers material up to and including exponential functions)
20. Logorithms
21. The basic figures of geometry, congruence, areas, perimeters and volumes
22. The Euclidean motions: translations, rotations and reflections
23. Circles, chords, diameters, tangents
24. Constructions with triangles and circles

25. The parallel postulate; similarities and dilations
26. Polar coordinates and the trigonometric functions
27. Tesselations of the plane
28. Review

# MAE 301/501, FALL 2006, LECTURE NOTES

BERNARD MASKIT

## 1. Introduction

1.1. **What is mathematics?** This is a seemingly interesting question. Unfortunately, the best known answer is that mathematics is what mathematicians do. And what is it that mathematicians do? Why, mathematics, of course.

While we can't really say what mathematics is, we can give an approximate answer to the question of what mathematics is all about. First of all, there are *mathematical objects*, such as numbers, sets, matrices, triangles and probabilities. Then there are *functions* or *processes*, such as adding two numbers, forming the complement of a set, finding the inverse of a non-singular matrix, constructing the medians of a triangle or writing a formula for the probability of a compound event. Finally, there are proofs or thoughts or solutions to problems; these are difficult to describe, and they more or less comprise what mathematicians do; one important example would be to view the set of all functions from one set to another as a new mathematical object.

Mathematicians really don't like to go around in circles, so we will not further pursue the question of what mathematics is, but rather start talking about mathematical objects, constructions and other processes, and thoughts. We are going to be primarily mathematical in this development; that is, we'll start with some undefined objects and processes, and then carefully, and logically, build up other objects, and processes or operations that work with them, even other kinds of objects and other kinds of processes. This development will, to some minor extent, mirror the historical development.

1.2. **Basic mathematical objects.** First of all, we need some kind of object, something to talk about. The usual objects with which mathematicians start are the natural numbers, $\mathbb{N}$, with which we count; that is, $1, 2, 3, 4, \ldots$. Notice that we start at 1, there is no 0, and there are no negative numbers. Associated with the natural numbers, we have two processes.

First, we can count in clumps; that is, we can add numbers. Addition satisfies the two rules:

*Commutativity of addition:* For all natural numbers $a$ and $b$, $a + b = b + a$;

and

*Associativity of addition:* For all natural numbers $a$,$b$ and $c$, $(a + b) + c = a + (b + c)$.

Next we observe that we can add in clumps, that is, multiply. Multiplication also satisfies two rules, which we call by the same names:

*Commutativity of multiplication:* For all natural numbers $a$ and $b$, $ab = ba$;

and

*Associativity of multiplication:* For all natural numbers $a, b, c$, $(ab)c = a(bc)$.

**Question:** Why are these rules concerning different operations called by the same names?

There is also a rule, the distributive rule, concerning the connection between these two operations: For all natural numbers, $a, b, c$, $a(b + c) = (ab) + (ac)$.

**Problem 1.1.** *What happens to this rule if you interchange addition and multiplication?*

1.3. **Inverse operations.** Is there an inverse operation (subtraction) to addition? When is it defined? That is, for which $a$ and $b$ can we solve the equation $a - x = b$?

**Problem 1.2.** *Is there a natural number $a$ so that the equation $a - x = b$ can always be solved? Never be solved?*

**Problem 1.3.** *Is there a natural number $b$ for which the equation $a - x = b$ can always be solved? Never be solved?*

We don't have the tools to prove it, but we know that subtraction is always unique; that is, if $a = b + x$ and $a = b + y$, then $x = y$.

We can likewise ask the question: Is there an inverse operation (division) to multiplication? That is, can we solve the equation $ax = b$ for $x$. When can we solve this equation?

**Problem 1.4.** *Is there a natural number $a$ so that the equation $ax = b$ always has a solution? Never has a solution?*

**Problem 1.5.** *Is there a natural number $b$ for which the equation $ax = b$ always has a solution? Never has a solution?*

We also don't have the tools to prove that division is always unique; that is, if $a = bx$ and $a = by$, then $x = y$.

1.4. **Order.** The natural numbers are also naturally and completely ordered. That is, for every pair of natural numbers, $a$ and $b$, either $a < b$, $b < a$, or $a = b$, and exactly one of these three possibilities holds.

The major rule of order is *transitivity*: If $a < b$ and $b < c$, then $a < c$. Also, for every $a \neq 1$, $1 < a$.

There are also relations between the arithmetic operations and order: For all numbers $a$ and $b$, $a < a + b$, and $b < a + b$. Also, if $b \neq 1$, then $a < ab$.

We observe that some numbers are multiples of 2, others are not; we need to be able to talk about all numbers that are (or are not) multiples of 2; i.e., even and odd numbers.

## 2. Sets

So far, we have one kind of mathematical objects: Natural numbers. In order to form more kinds of mathematical objects, we need to form collections of them, or *sets*. The objects within a set are called *elements.* We write $x \in A$ to mean that the element $x$ lies in the set $A$.

In order to avoid philosophical difficulties, we take the view that a set does not exist until it has been defined. It can be defined by listing its elements, or it can be defined as a subset of another set by some number of defining properties.

While we have started with the natural numbers, there are many other possible starting points, such as geometric objects, events in a probability space, etc. Wherever we start, we must have a *universal* set, which prescribes the universe within which we are working, and all sets under discussion are subsets of this universal set.

We can also talk about one set as being a subset of another; that is, $A$ is a subset of $B$ if every element of $A$ is also an element of $B$. In this case, we write $A \subset B$.

There is also an empty set, $\emptyset$, which has no elements, and is a subset of every set, including itself.

2.1. **The algebra of sets.** The algebra of sets has a rich structure, with three operations, union ($\cup$), intersection ($\cap$) and complement ($\sim$). Recall that $x \in A \cup B$ if either $x \in A$ or $x \in B$. (As always in mathematics, the conjunction "or" is taken in its weak sense; that is, "$x \in A$ or $x \in B$" includes the possibility that $x \in A$ *and* $x \in B$.) Also, $x \in A \cap B$ if $x \in A$ and $x \in B$. Finally, $x \in \sim A$ if $x \notin A$.

We will quickly review the laws governing these operations.

The operation of forming the union of sets is both commutative and associative:

$A \cup B = B \cup A$, and $(A \cup B) \cup C = A \cup (B \cup C)$.

The operation of forming the intersection of sets is both commutative and associative:

$A \cap B = B \cap A$, and $(A \cap B) \cap C = A \cap (B \cap C)$.

The operation of forming the complement of a set is an *involution:*

$\sim (\sim A) = A$.

There are two distributive laws relating union and intersection:

$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$, and $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Finally, there are two laws relating complements with unions and intersections:

$\sim (A \cup B) = (\sim A) \cap (\sim B)$, and $\sim (A \cap B) = (\sim A) \cup (\sim B)$.

**Problem 2.1.** *Show that $A \subset B$ if and only if $A \cup B = B$.*

**Problem 2.2.** *Show that $A \subset B$ if and only if $A \cap B = A$.*

**Problem 2.3.** *Show that $A \subset B$ if and only if $\sim B \subset \sim A$.*

2.2. **Venn diagrams.** It is often useful to use Venn diagrams to understand complicated combinations of these symbols.

2.3. **Sets of natural numbers.** We can now form several sets of natural numbers, such as the set of even numbers, the set of odd numbers, the set of number divisible by 7, the set of numbers leaving a remainder of 5 when divided by 8, etc. Then we can use the set operations to form new sets.

2.4. **Divisibility and prime numbers.** A number $a$ is *prime* if, whenever you write $a = bc$, (recall that we are only working with the natural numbers) then either $b = a$ or $c = a$, but not both (we do not want to include 1 as a prime number).

**Problem 2.4.** *List the first 20 prime numbers.*

We say that $a$ *divides* $b$, if there is a number $c$ so that $b = ac$. Note that, for all numbers $a$, $a$ divides $a$. Note also that 1 divides every number.

**Problem 2.5.** *Show that if $a$ divides $b$ and $a$ divides $c$, then $a$ divides $b + c$.*

**Problem 2.6.** *Is the converse true? That is, is it true that if $a$ divides $b+c$, then $a$ divides $b$ and $a$ divides $c$?*

**Problem 2.7.** *Show that if $a$ divides $b$, and $a$ does not divide $c$, then $a$ does not divide $b+c$.*

**Problem 2.8.** *Prove that for every number $a$, there is a prime number greater than $a$. (HINT: Since the natural numbers are the counting numbers, we know that there are only finitely many numbers less than any given number.)*

## 3. The Integers

Not being able to subtract is unsatisfying; we need a new kind of number to represent for example $3 - 5$. To this end, we form a new set of objects, negative integers: $-1, -2, \ldots$, and another new object: 0. The set of integers consists of the positive integers, the negative integers and zero. We know how to add and multiply these numbers; in fact, one can write down the rules in terms of addition and multiplication of natural numbers.

**Problem 3.1.** *Write down the rules for addition of integers; that is, define the sum of a positive number and a negative number, and define the sum of two negative numbers.*

The rule for multiplication is more difficult, but only conceptually: Why should it be true that $(-1)(-1) = +1$? It is not easy, perhaps impossible, to give a real world explanation, but there is a straightforward mathematical explanation, which goes as follows: It is clear from the point of view of absolute value that the product must be either $+1$ or $-1$. It is also clear that, since $(+1)a = a$ for all $a$, $(+1)(-1) = -1$. So, if we want division to be unique, we must have that $(-1)(-1) = +1$.

It is a long and sometimes tedious job, once one has defined addition and multiplication of integers, to check that the rules of commutativity, associativity and distributivity still apply.

**Problem 3.2.** *Show that the commutative laws for both addition and multiplication hold for all integers.*

If $a > b$, then we already know what is $a - b$; that is, we can solve the equation $b + x = a$. The negative numbers have been defined so that we can solve the equation $b + x = a$, for all integers, $a$ and $b$. We also know that the solution is unique.

**Problem 3.3.** *It is not quite true that division of integers is unique. Find all cases where it is not unique.*

Now that we are working within the realm of integers, we can recall the Euclidean (division) algorithm:

**Theorem 3.1.** *Let $p$ and $q$ be positive integers, then there are non-negative integers, $s$ and $r < q$, so that $p = sq + r$.*

## 4. Logic

4.1. **Formal logic.** We will not do much with formal logic here, but we do need to understand something about it. The basic objects in formal logic are propositions, such as: "$1 + 1 = 2$", or "Ice is colder than water." Every proposition is either true or false[1]

**Problem 4.1.** *Write an English sentence that has the format of a proposition, but is neither true nor false.*

Propositions can be combined using the connectives 'and' and 'or'; note that the connective 'or' is the weak form of this word; that is; the proposition: "$a$ and $b$" is true if and only if both $a$ and $b$ are true; "$a$ or $b$" is true if either $a$ is true, or $b$ is true, or they are both true. (The strong form, where "$a$ or $b$" means that either $a$ is true or $b$ is true, but they are not both true, is never used in mathematics.)

The proposition "$a$ and $b$" is written as: $a \wedge b$, while the propostion "$a$ or $b$" is written as $a \vee b$.

There is also the unary operation of negation: The negation of $a$, written as $\sim a$, is true if and only if $a$ is false.

The rules for combining propositions along with negation are closely related to the rules for combining unions, intersections and complements of sets. That is:

- $a \vee b$ is true if and only if $b \vee a$ is true.
- $a \vee (b \vee c)$ is true if and only if $(a \vee b) \vee c$ is true.
- $a \wedge b$ is true if and only if $b \wedge a$ is true.
- $a \wedge (b \wedge c)$ is true if and only if $(a \wedge b) \wedge c$ is true.
- $a \wedge (b \vee c)$ is true if and only if $(a \wedge b) \vee (a \wedge c)$ is true.
- $a \vee (b \wedge c)$ is true if and only if $(a \vee b) \wedge (a \vee c)$ is true.
- $\sim (\sim a)$ is true if and only if $a$ is true.
- $\sim (a \vee b)$ is true if and only if $(\sim a) \wedge (\sim b)$ is true.
- $\sim (a \wedge b)$ is true if and only if $(\sim a) \vee (\sim b)$ is true.

[1]This raises an interesting problem in that one can write down English sentences that look like propositions, but which are neither true nor false.

The most important propositions for us are implications; these are statements of the form: "If $a$ then $b$", which is also written as: "$a$ implies $b$" or $a \Rightarrow b$.

The *converse* of $a \Rightarrow b$ is $b \Rightarrow a$.

**Problem 4.2.** *Give an example of a true proposition whose converse is false.*

The *contrapositive* of $a \Rightarrow b$ is $\sim b \Rightarrow \sim a$. Just as two sets are equal if they have exactly the same elements, so two propositions, $p$ and $q$ are *equivalent* if $p \Rightarrow q$ and $q \Rightarrow p$; that is, $q$ is true if and only if and $q$ is true; in this case, we write $p \iff q$.

**Problem 4.3.** *Show that $p \Rightarrow q$ and its contrapositive, $\sim q \Rightarrow \sim p$, are equivalent.*

4.2. **Truth Tables.** Just as one can use Venn diagrams to find out about complicated sets, so one can use truth tables to find out about complicated statements. For example, the *implication*: "If $a \Rightarrow b$", has the following truth table:

| | $b$ is True | $b$ is False |
|---|---|---|
| $a$ is True | True | False |
| $a$ is False | True | True |

You might notice that a false statement implies every statement!

**Problem 4.4.** *Show that the statements $a \Rightarrow b$ and its contrapositive, $\sim b \Rightarrow \sim a$ are equivalent.*

**Problem 4.5.** *Show that the statements, $a \Rightarrow b$ and $\sim a \vee b$ are equivalent.*

4.3. **Logic and sets.** Mathematicians often make statements about sets and elements of sets using quantifiers.

The universal quantifier, $\forall$, which is read as "for all', makes a statement about all elements of a set. For example, the statement, $\forall a \in \mathbb{N}$, $a \geq 1$, says "For all $a$ in the set of natural numbers, $a$ is greater than or equal to one", which is true.

Notice that the universal quantifier says nothing about existence.

The statement "$\forall a \in \mathbb{N}$ and $\forall b \in \mathbb{N}$, if $a < 1$, then $a > b$" is true. That is, this statement is actually an infinity of statements, one for each pair of natural numbers $a$ and $b$. For each such pair, the hypothesis that $a < 1$ is false, so the implication is true.

The existential quantifier, $\exists$, only says that something exists, it doesn't say anything more.

The negation of an existential statement is a universal statement. For example, the negation of "$\exists a \in \mathbb{N}$, where $a < 1$" is "$\forall a \in \mathbb{N}$, $a \geq 1$."

Similarly, the negation of a universal statement is an existential statement. The negation of the statement "All natural numbers are even", is the statement "There exists a natural number that is not even".

4.4. **Logic and language.** Mathematicians talk and write about mathematics in their own language. In the U.S., the language is American English, and mathematics is spoken and written in American English. If you look carefully at a mathematics textbook, you will see that it is written in complete English sentences! Even the long strings of formulae have appropriate punctuation.

We very rarely use formal logical symbols, such as $\vee$ or $\wedge$, but we do use complicated sentences that require work and concentration to parse. For example, the standard definition of what it means for a function $f(x)$ to be continuous at the point $a$ is:

> The function $f(x)$ is continuous at the point $a$ if, given any number $t > 0$, there is a number $s > 0$ so that if $|x - a| < s$, then $|f(x) - f(a)| < t$.

The above may look somewhat unfamiliar in that you usually see it written in the equivalent form:

> The function $f(x)$ is continuous at $x_0$ if, for every $\epsilon > 0$ there is a $\delta > 0$ so that $|f(x) - f(x_0)| < \epsilon$ whenever $|x - x_0| < \delta$.

You should be able to see that the two statements are logically equivalent, and you should be able to write down the negation of either one of them.

**Problem 4.6.** *Write down the negation of the above statement, in either form.*

## 5. Functions

Given two sets, $A$ and $B$, a *function* $f : A \to B$ assigns a unique element of $B$ for every element of $A$. For example, the function $m_2 : \mathbb{Z} \to \mathbb{Z}$, defined by $M_2(a) = 2a$, maps every number $a$ into the number $2a$. Another example is $p : \mathbb{Z} \to \mathbb{Z}$, where $p(1) = 1$, and, for $a > 1$, $p(a)$ is the greatest prime dividing $a$.

There are two important properties that a function may have. The function $f : A \to B$ is *one-to-one* if, for every $y \in B$, the equation $f(x) = y$ has at most one solution. The function $f : A \to B$ is *onto* if, for every $y \in B$, the equation $f(x) = y$ has at least one solution.

There are also two important sets associated with any function. The *domain* of $f : A \to B$ is the set $A$. In calculus, we often write down a function without explicitly identifying its domain, for example $\ln x$ is only defined for $x > 0$, so its domain is the set of positive real numbers. The *range* of $f$ is the set of elements of $y \in B$ for which the equation $f(x) = y$ has a solution.

**Problem 5.1.** *Is the function $m_2$ one-to-one? Is it onto? What is its range?*

**Problem 5.2.** *Is the function $p$ one-to-one? Is it onto? What is its range?*

## 6. Cartesian Products

The (Cartesian) *product* of two sets, $A$ and $B$ is the set of all *ordered* pairs $(x, y)$, where $x \in A$ and $y \in B$; we write the Cartesian product as $A \times B$. We emphasize the order for the special case that $A = B$; for example, in the Cartesian plane (named after its founder, René Descartes), the points $(2, 3)$ and $(3, 2)$ are quite different.

We can now regard addition and multiplication as functions. Addition is the function $a : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ defined by $a(x, y) = x + y$. Likewise multiplication is the function $m : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ defined by $m(x, y) = xy$.

**Problem 6.1.** *Is the function a one-to-one? Is it onto? What is its range?*

**Problem 6.2.** *Is the function $m$ one-to-one? Is it onto? What is its range?*

We observe that we can write the commutative law of addition in terms of the function $a$ defined above; that is, $a(x, y) = a(y, x)$. We can likewise write the associative law as $a(a(x, y), z) = a(x, a(y, z))$.

**Problem 6.3.** *Write the distributive law in terms of the functions $a$ and $m$.*

**Problem 6.4.** *Multiplication of matrices gives an example of an operation that is associative but not commutative. Natural examples of operations that are commutative but not associative are not easy to find. Construct such an example.*

6.1. **Composition of functions.** Now that we can define a fair number of functions, we can compose some of them. If $f : A \to B$, and $g : B \to C$, then the *composition* of $f$ and $g$ is the function $g \circ f : A \to C$ defined by $g \circ f(x) = g(f(x))$.

**Problem 6.5.** *Consider the functions $m_2$ and $p$ defined above.*

(1) *Find $p \circ m_2(60)$.*
(2) *Find $m_2 \circ p(60)$.*
(3) *Can you draw a conclusion from the above?*

6.2. **a set with 2 elements.** The set $\mathbb{Z}_2$ has 2 elements, denoted as 0 and 1, and has both addition and multiplication defined on it, where $0 + 0 = 1 + 1 = 0$ and $0 + 1 = 1 + 0 = 1$. Since there is some possibility for confusion, we denote multiplication with a dot. Then we have $0 \cdot 0 = 0 \cdot 1 = 1 \cdot 0 = 0$ and $1 \cdot 1 = 1$.

You can think of 0 as representing an arbitrary odd number, and 1 as representing an arbitrary even number; then this addition and multiplication is exactly correct.

Right now, we only need to have a set with two elements, which we will call by the same name, $\mathbb{Z}_2$. However, for the special purposes of this section only, we will label these two elements as $T$ (for True) and $F$ (for False).

6.3. **Relations.** A *relation* on a set $A$ is a function $r : A \times Z \to \mathbb{Z}_2$. That is, for every pair of elements, $a$ and $b$ from $A$, we have assigned to this pair, either the symbol 0 or the symbol 1.

The usual meaning of this is that the relation holds for the pair $(a, b)$ if $r(a, b) = T$, and the relation does not hold if $r(a, b) = F$.

6.3.1. *Example.* We look at the relation on the set of integers: $a < b$. That is, the value of the function $r_1$, for this relation, on the pair $(a, b)$ is 1 if and only if $a < b$. In particular, $r_1(3, 7) = T$ and $r_1(7, 3) = F$.

6.3.2. *Example.* $r_2(a, b) = 1$ if and only if $a + b$ is even.

6.3.3. *Example.* $r_3(a,b) = 1$ if and only if $a > 2b$ and $a < b$.

**Problem 6.6.** *Find all pairs of integers* $(a,b)$ *for which* $r_3(a,b) = T$.

The relation $r$, on the set $A$, is called *reflexive* if, for all $a \in A$, $r(a,a) = T$.
lThe relation $r$, on the set $A$, is called symmetric if, for all $a$ and $b$ in $A$, $r(a,b) = r(b,a)$.

**Problem 6.7.** *Write a short paragraph explaining the similarities and difference between symmetry and commutativity.*

The relation $r$ on the set $A$, is called *transitive* if, whenever $r(a,b) = T$ and $r(b,c) = T$, it follows that $r(a,c) = T$.

**Problem 6.8.** *Which of the relations* $r_1$, $r_2$, $r_3$ *are reflexive? Which are symmetric? which are transitive?*

A relation on a set that satisfies all three of these properties; that is, it is reflexive, symmetric and transitive, is called an *equivalence* relation.

Examples of equivalence relations: Equality of numbers, equality of matrices, equivalence of matrices, equality of functions, congruence of triangles, similarity of triangles, congruence of circles.

**Problem 6.9.** *Show that the three qualities, reflexivity, symmetry and transitivity, are independent. That is, for each one of these three, construct a relation that satisfies the other 2, but not this one.*

The crucial point about equivalence relations is that an equivalence relation on a set divides the set into disjoint equivalence classes. That is, an equivalence class, which is necessarily not empty, consists of some element of the original set, together with all elements related to it.

**Problem 6.10.** *Let* $r$ *be an equivalence relation on a set* $S$, *and let* $a$ *and* $b$ *be elements of* $S$. *Let* $A$ *be the set of all elements of* $S$ *related to* $a$, *and let* $B$ *be the set of all elements of* $S$ *related to* $b$. *Show that either* $A = B$, *or* $A \cap B = \emptyset$.

## 7. Fractions and Rational Numbers

We are now in a position to construct the rational numbers; we need them because we want to be able to divide numbers, and the equation $5x = 3$ has no solution in the set of integers.

First we need the fractions; these are usually thought of as the set of ordered pairs of integers, where the second element is not zero. However, we can just as easily consider the fractions to be $\mathbb{F} = \mathbb{Z} \times \mathbb{N}$. This is the set of ordered pairs, where the first is an integer, and the second is a natural number.

Now we can form equivalence classes of fractions. Two fractions, $(a,b)$ and $(c,d)$ are equivalent if $ad = bc$.

**Problem 7.1.** *Show that equivalence of fractions is indeed an equivalence relation.*

Define addition of fractions by $(a,b)+(c,d)=(ad+bc,bd)$.

**Problem 7.2.** *Show that if $(a,b)$ and $(c,d)$ are equivalent fractions, and $(e,f)$ and $(g,h)$ are equivalent fractions, then $(a,b)+(e,f)$ is equivalent to $(c,d)+(g,h)$.*

Define multiplication of fractions by $(a,b)\cdot(c,d)=(ac,bd)$.

**Problem 7.3.** *Show that if $(a,b)$ and $(c,d)$ are equivalent fractions, and $(e,f)$ and $(g,h)$ are equivalent fractions, then $(a,b)\cdot(e,f)$ is equivalent to $(c,d)\cdot(g,h)$.*

The rational numbers $\mathbb{Q}$ ($Q$ for quotients) are the set of equivalence classes of fractions. The two problems above show that addition and multiplication of rational numbers is well defined.

Notice that we have $\mathbb{Z}$ as a subset of $\mathbb{Q}$, where the integer $a$ corresponds to the equivalence class of $(a,1)$. This includes the rational number 0, defined as the equivalence class of $(0,1)$, and the rational number 1, defined as the equivalence class of $(1,1)$.

Observe that, for every fraction $(a,b)$, we have its negative, $-(a,b)=(-a,b)$, so that $(a,b)+(-(a,b))=0$.

**Problem 7.4.** *Show that if $(a,b)$ is equivalent to $(c,d)$, then $-(a,b)$ is equivalent to $-(c,d)$, showing that the negative of a rational number is well defined.*

We also have, for every fraction $(a,b)$, where $a\neq 0$, its multiplicative inverse $(a,b)^{-1}=(b,a)$, so that $(a,b)\cdot(a,b)^{-1}=1$.

**Problem 7.5.** *Show that if $(a,b)$ is equivalent to $(c,d)$, then $(a,b)^{-1}$ is equivalent to $(c,d)^{-1}$.*

**Problem 7.6.** *Can you explain in non-mathematical terms why $(-1)(-1)=+1$, why the product of the fractions $\frac{a}{b}$ and $\frac{c}{d}$ is $\frac{ac}{bd}$, and why the quotient of the fractions: $\frac{a}{b}$ over $\frac{c}{d}$ is $\frac{ad}{bc}$?*

The rational numbers form a field, that is, addition, subtraction, multiplication and division are always possible, except of course that we cannot divide by zero. However, there are still more numbers that are not in $\mathbb{Q}$. For example, there is no rational number $a$, which when squared is equal to 2, or 3, or 5 or 6, or any other number that is not a perfect square.

## 8. Infinite Decimals and Real Numbers

Even though there is no rational number whose square is 2, we can find a sequence of rational numbers, $\{a_n\}$ so that $a_n^2\to 2$. Now that we have calculators, we no longer learn the algorithm for extracting square roots in school, but we can nevertheless construct such a sequence. We will illustrate one possible such procedure.

We first observe that $1^2=1<2$, and that $2^2=4>2$, so our first number is 1.

Next we observe that $(1.4)^2<2$, while $(1.5)^2>2$, so our next number is 1.4.

Next we observe that $(1.41)^2<2$, while $(1.42)^2>2$, so our next number is 1.41.

And so on.

8.1. **Base 2.** For this section only, we write all numbers in base 2; that is, we express every number as a finite sum $\sum_{j=-n}^{m} a_j 2^j$, where each $a_j$ is either 0 or 1. Then the number that in decimal notation is "2", is 10 in this notation.

**Problem 8.1.** *Write the number 37 in base 2.*

**Problem 8.2.**
*Write the number 2/3 in base 2. (Hint: The multiplication and division algorithms work in any base, including base 2.)*

**Problem 8.3.** *Find the first 4 "decimal" places of $\sqrt{10}$ in base 2. (The number inside the square root sign is not the number of fingers on both hands, but is the number of people in a couple.)*

One of the main advantages of the real numbers over the rational numbers is the existence of a real number whose square is 2; more generally, we can use Newton's method for approximating roots of polynomials, provided the polynomial has a real root. Polynomials of odd degree necessarily have real roots (they are continuous, tend to $-\infty$ as $x \to -\infty$, and tend to $\infty$ as $x \to \infty$). However, polynomials of even degree, such as $x^2 + 1$, do not necessarily have any real roots.

**Problem 8.4.** *Find a number $B > 0$, so that the polynomial function $x^4 + x^2 - x + 1 \geq B$ for all $x$.*

## 9. The complex numbers

A *complex number* has the form, $z = x + iy$, where $x$ and $y$ are real numbers, and $i$ has the special property that $i^2 = -1$. The number $x$ is called the *real part* of $z$, and the number $y$ is called the *imaginary part*. One can add, subtract and multiply complex numbers by using the commutative, associative and distributive laws of addition and multiplication of real numbers, and by assuming that $i$ commutes with all real numbers. Then the additive identity is $0 = 0 + i0$, and the muliplicative identity identity is $1 = 1 + i0$. We observe that both addition and multiplication of complex numbers are commutative and associative, and that the usual distributive law holds.

Division is somewhat more interesting. We need the *complex conjugate* $\bar{z} = x - iy$, and we observe that $z\bar{z} = |z|^2 = x^2 + y^2$ is real and is equal to 0 only for $z = 0$. Then, if $z \neq 0$, we can define $1/z = \frac{\bar{z}}{z\bar{z}}$ so that $z(1/z) = 1$.

We note that, in terms of addition and subtraction, the set of complex numbers is not different from the set of 2-dimensional real vectors; however, the multiplication is different.

One of the crucial facts about the complex numbers is the "Fundamental Theorem of Algebra", which states that every polynomial of positive degree with complex coefficients has at least one complex root; in particular, every such polynomial with integer coefficients has at least one real root.

We will return to the complex numbers when we discuss polar coordinates.

## 10. Infinity and beyond

We first need a mathematical formulation for the process of counting. Let $S$ be a non-empty set. We say that a function $f : S \to \mathbb{N}$ is a counting function if $f$ is one-to-one, and it satisfies the following two properties.

(1) There is an $x \in S$ so that $f(x) = 1$, and
(2) For every $x \in S$, and for every integer $m < f(x)$, there is a $y \in S$ so that $f(y) = m$.

The set $S$ is countable if either it is empty or there is such a counting function $f : S \to \mathbb{N}$. In the case that the counting function is such that there is a number $N \in \mathbb{N}$ so that $f(x) \leq N$ for all $x \in S$, then $S$ is finite. We will see below that there are uncountable sets.

Observe that if $S$ is countable and not finite, then the counting function $f : S \to \mathbb{Z}$ is both one-to-one and onto; that is, it is a one-to-one equivalence. This leads to the following definitions. Two sets $S$ and $T$ have the *same cardinality* if there is a one-to-one equivalence $f : S \to T$.

The above defines the concept of two sets having the same cardinality, but does not necessarily name this cardinality. A finite set with $N$ elements in it has *cardinality* $N$. Then we need a name for the set of all natural numbers; the *cardinality* of $\mathbb{N}$ is $\aleph_0$.

Note that it is possible to show that two sets have the same cardinality without our knowing the cardinality of either set.

**Problem 10.1.** *If $S$ is countable and $T$ is a subset of $S$, then $T$ is countable.*

**Problem 10.2.** *Show that the union of a countable set and a finite set is countable.*

**Problem 10.3.** *If $S$ and $T$ are both countable, then $S \cup T$ is also countable.*

**Theorem 10.1.** *The sets $\mathbb{Z}$ and $\mathbb{N}$ have the same cardinality.*

*Proof.* The proof is basically a picture, which is somewhat difficult to describe in words. Think of the elements of $\mathbb{Z}$ as laid out on the number line, and start counting at 0, and alternately counting positive and negative integers: $0, 1, -1, 2, -2, 3, -3, \ldots$. □

**Theorem 10.2.** *$\mathbb{Q}$ and $\mathbb{N}$ have the same cardinality.*

*Proof.* The proof consists of several steps. First we show that the set of positive fractions with non-zero denominator is countable. This is again a picture, but now the picture is of an infinite matrix with $1/1, 1/2, 1/3, 1/4, \ldots$ in the top row, then $2/1, 2/2, 2/3, 2/4, \ldots$ in the second row, etc. We count the elements of this infinite matrix by starting at the top left, going right one step, then diagonally down and to the left until we meet the left hand edge, then down one step, then diagonally up and to the right until we meet the top edge, etc.

The next step is to look at each positive rational number as a fraction in lowest terms; this yields the positive rationals as a subset of the positive fractions; hence the positive rationals are countable.

There is an obvious one-to-one mapping of the negative rationals onto the positive rationals, hence the negative rationals are countable. It then follows that the set of all rational numbers is countable. □

**Problem 10.4.** **An algebraic number is a root of a polynomial with integral (integer) coefficients. Write such a polynomial as $a_n x^n + \ldots + a_0$. Show that the set of algebraic numbers is countable.*

**Theorem 10.3.** $\mathbb{R}$ *is not countable.*

*Proof.* Suppose we could count the real numbers. Then we could surely count the real numbers lying between 0 and 1. Hence we can assume that we have a real number $x_j$, $0 \le x_j \le 1$ assigned to each integer, $j$. We can write $x_j$ as an infinite decimal, perhaps in several different ways; if we have a choice between the infinite decimal ending in all zeroes or ending in all nines, we choose it to end in all zeroes. Having made this choice, we can write $x_j = 0.a_{j1}a_{j2} \cdots a_{jj} \cdots$.

We next construct a new number $b$ that is necessarily not equal to $a_j$ for every $j$. We write this new number as $b = 0.b_1 b_2 \cdots b_j \cdots$, where we still have to choose $b_1$, $b_2$, etc. We choose $b_1$ so that it is not equal to 0, $a_{11}$ or 9. With this choice, we are guaranteeing that $b \ne a_1$, and that our final number $b$ will not have more than one decimal expansion. We next choose $b_2$ so that it is not equal to 0 or 9, and not equal to $a_{22}$; this guarantees that $b \ne a_2$. We continue in this manner, choosing $b_j$ so that it is different from $a_{jj}$, and also different from 0 and 9. Since the number $b$ that we have constructed lies between 0 and 1, and is not equal to any of the number $a_1, \ldots$, we have reached a contradiction. □

There is an easy way to construct new sets from old. If $S$ is any non-empty set, then the power set of $S$ is the set of all subsets of $S$. One can generalize the above proof to show the following.

**Problem 10.5.** *A non-empty set $S$ and its power set do not have the same cardinality.*

**Problem 10.6.** *Show that if $S$ is finite and has $n$ elements, then the power set of $S$ has $2^n$ elements.*

## 11. Measurements

While we don't usually think of measurement as a mathematics topic, it is in both the NCTM (National Council of Teachers of Mathematics) and New York State syllabi. Historically, mathematics has always been used to compute things that cannot be directly measured; one famous example is Eratosthenes' computation of the circumference of the Earth. Other uses include the use of trigonometry to compute heights of trees; the use of proportions to compute the speed of a car; and numerical solutions to differential equations to predict the weather. In practice, all of these require measurements, and this brings up the question of how closely one can or should measure something.

If we are measuring the length of a table, and we are measuring it to the nearest foot, then we can reasonably expect that we will get the same answer no matter how many times we measure it. However, if we are measuring it to the nearest inch, then we might sometimes get one answer, sometimes another; we might even see three or four different answers if we make sufficiently many measurements. Going further, if we measure the table to the nearest eighth of an inch (the usual markings on a carpenter's rule), then we will likely get about 8 distinct answers if we measure it 10 times. This leaves us with some question as to whether there is such a thing as *the exact length of the table*. We don't need to, but it is convenient for us to postulate that there is a real number that we postulate to be the *correct* measurement for any given object, such as the length of a table, even if in reality we cannot ever know what it is.

By making several measurements, we can be reasonably confident that the correct measurement lies somewhere between our largest and smallest measurements. Then, we can take an average of all our measurements to arrive at our *measured* value. By going through this procedure, we arrive at both a *measured value* and an estimate of the difference between the measured value and the *correct* value. That is, if the maximum measurement is $M_{\max}$, the minimum measurement is $M_{\min}$, the correct measurement (which we do not know) is $M_c$, and the average or mean of all our measurements is $M$, then we have that $|M - M_c| \le \max(M - M_{\min}, M_{\max} - M)$.

We remark that, for example, if this difference, $|M - M_c|$ is less than .05 inches, then we say that we have measured our table to within .1 inches. Similarly, if $|M - M_c| < 1/2$ inch, then we say that we have measured our table to within one inch.

The first kind of question that arises in this context is given by the following problem.

**Problem 11.1.** *You want to measure the height of a tree. You know that the transit (angle measuring device) measures angles to the nearest degree. You also know (to make the problem easier) that you have placed the transit exactly 25 feet from the base of the tree. If the transit shows that the angle of elevation to the top of the tree is 58 degrees. How high is the tree? How good is your approximation to the height of the tree?*

Solving this is relatively easy. You know that the height of the tree is given by $h = 25\tan(58^\circ)$ feet. Your calculator will show this to be $40.00836\ldots$. It should be clear that 40.008 feet is already too exact. Since the transit measures angles to within the nearest degree, we know that in fact, the height of the tree lies somewhere between $h_{\min} = 25\tan(57.5^\circ) = 39.242\ldots$ feet and $h_{\max} = 25\tan(58.5^\circ) = 40.796\ldots$ feet. So, a reasonably correct answer is that the tree is 40 feet high, and our approximation is correct to within .8 feet.

The following is a somewhat more difficult problem, where we seem to have a mismatch between the information given and the information we need. Since we are not given the measurement of the angle, we cannot so easily compute the error; however, we can approximate the error by using the mean value theorem of the differential calculus.

**Problem 11.2.** *You want to measure the height of a tree. You know that the transit (angle measuring device) measures angles to the nearest degree. You also know that the tree is at least 10 feet high and at most 200 feet high, and we assume, to make the problem easier, that you have placed the transit exactly 25 feet from the base of the tree. How good is your approximation to the height of the tree? Did you use all the information you were given to solve this problem.*

**Problem 11.3.** *You are on one side of a street that is 24 feet wide. (To make the problem easier, we assume this is a precise measurement.) You want to know the distance, to the nearest foot, from you to a person on the other side of the street and some distance along the street. You cannot measure this distance directly, as you would prefer not to be hit by a passing car, bus or truck, but you can measure the distance from where you stand to a point directly opposite the other person. How closely do you have to measure this distance?*

## 12. Mathematical Induction

The principle of mathematical induction concerns a sequence of mathematical propositions; call them $P_1, P_2, \ldots$. The principle states the following:

**Theorem 12.1.** *Suppose we know the following:*

(1) *$P_1$ is true; and*
(2) *if $P_{n-1}$ is true, then $P_n$ is true.*

*Then $P_n$ is true for all $n = 1, 2, \ldots$.*

It is important to note that this principle works equally well for a sequence of propositions starting with index 0, or any other integer and continuing onward; we illustrate this with the following example.

**Theorem 12.2.** *Let $S$ be a set containing $n$ elements. Then the number of distinct subsets of $S$ is $2^n$.*

*Proof.* We start with $n = 0$; that is, $S = \emptyset$. Then $S$ contains exactly one subset; namely $S$ itself, and $2^0 = 1$.

Now assume we know that every set containing $n - 1 \geq 0$ elements, has $2^{n-1}$ distinct subsets. Let $S$ be a set containing $n$ elements. Let $T$ be some subset of $S$ containing $n - 1$ elements, and let $x$ be the elements of $S$ that is not in $T$.

Every subset of $S$ that does not contain $x$ is also a subset of $T$, and two such subsets of $S$ are distinct if and only if they are distinct as subsets of $T$. Hence the number of such subsets is, by the induction hypothesis, $2^{n-1}$.

For every subset $R$ of $S$ that does contain $x$, $R \cap T$ is a subset of $T$, and again two such subsets are distinct as subsets of $S$ if and only if they are distinct as subsets of $T$. Hence, by the induction hypothesis, the number of such subsets is $2^{n-1}$.

Since every subset of $S$ either does or does not contain $x$, but not both, we have that the number of distinct subsets of $S$ is equal to $2^{n-1} + 2^{n-1} = 2 \cdot 2^{n-1} = 2^n$. □

### 12.1. **The Fibonacci Numbers.**

The Fibonacci sequence $\{a_n\}$ is defined inductively by:
$a_0 = 0$;
$a_1 = 1$;
for $n > 1$, $a_n = a_{n-1} + a_{n-2}$.

Thus the first few terms of the Fibonacci sequence are: 0, 1, 1, 2, 3, 5, 8, 13, 21, 34 ….

The Fibonacci sequence is intimately connected with the "Golden Ratio". The ancient Greeks thought that a rectangle with side lengths $a$ and $b$ had the most pleasing appearance if these sides were in the ratio $a : b = b : a + b$. Writing the ratio of the sides as $r = b/a$, we obtain the following quadratic equation for $r$: $r^2 - r - 1 = 0$. This quadratic has two solutions, $\alpha = \frac{1+\sqrt{5}}{2}$ and $\beta = \frac{1-\sqrt{5}}{2}$.

The connection between the Fibonacci sequence and the golden ratio is given by the following fact, which can be proven by induction.

**Problem 12.1.** *Show that* $a_n = \alpha^n + \beta^n$.

We need a seemingly stronger form of mathematical induction for these problems; this other form, which is actually logically equivalent to the usual form of mathematical induction is as follows. Suppose we are given a sequence of propositions $P_n$, satisfying
$P_1$ is true, and
if $P_j$ is true for all $j < n$, then $P_n$ is true.
Then $P_n$ is true for all $n$.

**Problem 12.2.** *Show that* $a_n < (4/7)^n$.

## 13. Polynomials

A polynomial (in one variable) is an expression of the form: $P = a_n x^n + \cdots + a_1 x + a_0$. In this expression, $x$ is the *variable*, the $a_i$ are the *coefficients*, and the integer $n$ is the *degree*. Note that the degree $n$ is an integer $\geq 0$.

Since we want to be able to add and multiply polynomials, we require that the coefficients lie in some number field, such as the integers, real numbers, etc. Our only requirements on this number field is that addition and multiplication satisfy the commutative, associative and distributive laws, and that this set of numbers include both the additive identity (0) and the multiplicative identity (1).

A polynomial is in some sense an algebraic object, depending on the coefficients; for example, the set of polynomials with real coefficients of degree at most 4 can be regarded as a vector space of dimension 5 over the field of real numbers.

We can also regard a polynomial $P$ as a function from the number field the coefficients lie in to itself; in this case, we will often write $P(x)$, rather than $P$. A number $x_0$ is a *root* of $P = a_n x^n + \cdots + a_0$ if, as a function, $P(x_0) = 0$.

The polynomial $P$ of degree $n$ is called *monic* if $a_n = 1$. Note that if we are dealing with polynomials over an actual field, where division by any number other than zero is always possible, then for all questions involving roots, we can assume that $P$ is monic.

We can divide one polynomial by another, using a process analogous to long division. Using this process (algorithm), we arrive at the statement of the *Euclidean algorithm* for polynomials, which is analogous to the Euclidean algorithm for numbers.

**Theorem 13.1.** *Given the polynomials $P$ and $Q$, there exist polynomials $S$ and $R$, where the degree of $R$ is less than the degree of $Q$, so that $P = SQ + R$.*

As in division of numbers, the polynomial $R$ is called the *remainder*. If in the above, $R = 0$, then we say, as with numbers, that $Q$ divides $P$.

**Theorem 13.2.** *The number $r$ is a root of the monic polynomial $P$ if and only if $x - r$ divides $P$.*

*Proof.* Use the Euclidean algorithm to write $P = S(x - r) + R$, and consider these as functions, so that $P(x) = S(x)(x - r) + R(x)$.

If $R = 0$, then $P(r) = S(r)(r - r) = 0$, so $r$ is a root of $P$.

If $r$ is a root of $P$, then $0 = P(r) = S(r)(r - r) + R(r) = 0 + R(r) = R(r)$. However, since the degree of $R$ is less than the degree of $x - r$, which is 1, $R$ has degree 0; that is, $R$ is a constant. It then follows that $R = 0$; i.e., the polynomial, $x - r$, divides $P$. □

## 14. Basic combinatorics

The basic problem in combinatorics is to count the number of ways something can happen. A simple example is the following, from the NY State Math A Regents exam.

**Problem 14.1.** *Lee is the leader of the team of 6 people and walks in front; the other members of the team are lined up behind her. How many different ways are there for them to line up?*

To solve this, we see that there are 5 possibilities for the person next in line behind Lee. For each of these 5 possibilites, there are 4 possibilites for the next person; then for each of these, there are 3 possibilites for the next person; then, for each of these, there are 2 possibilities for the next person in line; finally there is only one possible person left to be last. So the total number of possibilities is $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$.

In general, the function $n!$, called *$n$ factorial*, is defined inductively by the following;
$0! = 1$; and
$n! = n(n - 1)!$.

The statement that $0! = 1$ seems strange, as it asserts that the number of ways to order 0 objects is 1. However, it will turn out that with this definition, we will not need to consider separately certain special cases.

In general, $n!$ gives us the number of ways to line up $n$ objects; that is, it is the number of permutations of the numbers, $1, \dots, n$. However, we sometimes need to count objects without regard to order. For example, we could start with the $n$ vertices of a regular polygon, and ask the question: How many distinct lines are there passing through exactly two of these vertices. (We note that since these are the vertices of a regular polygon, no three of these vertices are collinear.) We can rephrase this question as: How many ways are there of choosing 2 objects out of $n$ objects, where the order of the 2 objects is immaterial.

To answer this question, we start by picking two vertices, a first and then a second. There are $n$ choices for the first vertex, and, for each of these, there are $n-1$ choices for the second vertex, so there are a total of $n(n - 1)$ ways of picking a first vertex followed by a second one. Of course, each of our lines will have been counted twice, so the correct answer for the number of lines is $n(n - 1)/2$.

Now ask the same question about triangles. How many distinct triangles are there whose vertices lie among the $n$ vertices of the regular polygon. (We are assuming $n > 3$.) (Here we regard two triangles with the vertices labeled in different orders as being the same triangle.) As above, there are $n$ choices for the first vertex, there are $n-1$ choices for the second vertex, and $n-2$ choices for the third vertex. Since there are 3! permutations of the vertices of any one triangle, each triangle will appear $3! = 6$ times in this count. Hence the number of such triangles is $n(n-1)(n-2)/6$.

In general, the number of distinct ways of choosing $m$ objects out of $n \geq m$ objects is written as

$$\binom{n}{m} ={}_n C_m = \frac{n!}{m!(n-m)!},$$

and is pronounced, “binomial n, m”, or “n choose m”.

14.1. **The Binomial Theorem.** We write the binomial theorem in its general form, as it concerns polynomials in two variables. One can set the variable $y = 1$, and so deal only with polynomials of one variable.

**Theorem 14.1.** *For every non-negative integer $n$,*

$$(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^j y^{n-j}.$$

*Proof.* The proof is by induction; the step for $n = 0$ is trivial. We assume that we know the result for $n-1$. Then

$$\begin{aligned}
(x+y)^n = (x+y)(x+y)^{n-1} &= (x+y)\sum_{j=0}^{n-1}\binom{n-1}{j}x^j y^{n-1-j} \\
&= \sum_{j=0}^{n-1}\binom{n-1}{j}(x^{j+1}y^{n-1-j} + x^j y^{n-j}) \\
&= \sum_{j=0}^{n-1}\binom{n-1}{j}x^{j+1}y^{n-1-j} + \sum_{j=0}^{n-1}\binom{n-1}{j}x^j y^{n-j} \\
&= \sum_{k=1}^{n}\binom{n-1}{k-1}x^j y^{n-k} + \sum_{k=0}^{n-1}\binom{n-1}{k}x^k y^{n-k}.
\end{aligned}$$

For the last line, we used the substitution $k = j+1$ in the first sum, and $k = j$ in the second.

We complete the proof by observing that $\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}$, which we leave as an exercise. □

**Problem 14.2.** *For $n > k \geq 1$,*

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}.$$

**Problem 14.3.** *If the set $S$ contains 7 elements, how many distinct subsets of $S$ are there containing 3 elements?*

**Problem 14.4.** *Robin, the captain of a team of 7, leads the team in a parade. Robin is followed by 3 team members walking together; they are followed by another 3 team members walking together. How many possible such arrangements are there (ignoring the order of each set walking 3 abreast). The answer should be a number, and your explanation should not use any formulas.*

**Problem 14.5.** *Similar problem to the above. Here the team has 11 members. The captain leads the parade; the captain is followed by 3 team members walking abreast; they are followed by 2 team members walking abreast; they are followed by 3 team members walking abreast; finally 2 team members walking together bring up the rear.*

## 15. Finite Probability Spaces

One of the major applications of basic combinatorics is to compute probabilities. A *finite probability space* is a finite set $A$, where each element of $a \in A$ has assigned to it a *probability*, $p(a)$, which is a non-negative number. These number $p(a)$ are required to satisfy $\sum_{a\in A} p(a) = 1$.

The simplest example is the tossing of a true coin, where there are two possible outcomes: Heads (H) or Tails (T), and each of these two elements, $H$ and $T$, has probability 1/2.

We can then go on to toss the coin twice, in which case there are 4 possible outcomes: $HH$, $HT$, $TH$ and $TT$; each of these has probability 1/4. We can likewise toss the coin $n$ times, in which case there are $2^n$ possible outcomes. Here our probability space is a set of sequences of $n$ letters, where each letter is either $H$ or $T$, and each such sequence has probability $2^{-n}$.

We illustrate the use of combinatorics to compute probabilities with some examples, all based on the following scenario.

15.1. **Examples.** A coin is tossed 7 times, so our probability space consists of $2^7 = 128$ sequences of 7 letters each.

(1) Assuming that the coin is true, what is the probability that the first 3 coin tosses land on heads?

    Exactly half of the sequences have their first letter equal to $H$. Of these 64 sequences, exactly half have their second letter equal to $H$, and of these 32 sequences, exactly half, 16, have their first letter equal to $H$. So the answer is $16/128 = 1/8$.

    One could also look at this problem as follows: The probabality of the first toss landing on heads is 1/2. Independent of what happens on the first toss, the probability of the second toss landing on heads is 1/2. Also, independent of what happens on the first two tosses, the probability of the third toss landing on heads is 1/2. Therefore, the probability of all three landing on heads is $(1/2)^3 = 1/8$.

(2) Assuming that the coin is true, what is the probability that the first three tosses are heads and the last four tosses are tails.

Here there is exactly one sequence out of 128 that matches, namely, $HHHTTTT$, so the probability is $1/128$.

We could also say that the 7 tosses are independent, and we are requiring a particular outcome for each toss, so the probability is $1/2^7 = 1/128$.

(3) Assuming that the coin is badly unbalanced, so that the probability of its landing on heads is .6, and the probability of its landing on tails is .4, what is the probability that the first three tosses are heads.

Here the probability assigned to a sequence of $H$'s and $T$'s depends on the number of $H$'s in the sequence. If there are $h$ $H$'s, and $(7-h)$ $T$'s, then the probability assigned to this sequence is $(.6)^h(.4)^{7-h}$.

In this case, the probability for the first toss to land on heads is .6; independent of what happens on the first toss, the probability that the second toss lands on heads is .6, etc. Hence the probability here is $(.6)^3 = .216$.

**Problem 15.1.** *Use the binomial formula to show that the sum of the probabilities over all possible sequences is equal to 1.*

(4) Assuming that the coin is badly unbalanced, so that the probability of its landing on heads is .6, and the probability of its landing on tails is .4, what is the probability that the first three tosses are heads and the last 4 are tails.

In this case, there is only possible sequence containing 3 $H$'s and 4 $T$'s, so the requisite probability is $(.6)^3(.4)^4 = .0055\ldots$.

(5) Assuming that the coin is badly unbalanced, so that the probability of its landing on heads is .6, and the probability of its landing on tails is .4, what is the probability that there are exactly 3 heads out of the 7 tosses.

We need to look at the sequences where there are exactly three $H$'s and four $T$'s. The probability for each such sequence is $(.6)^3(.4)^4$. Since there are $\binom{7}{3} = 35$ such sequences, the answer is $35(.6)^3(.4)^4 = .1935\ldots$.

(6) Assuming that the coin is badly unbalanced, so that the probability of its landing on heads is .6, and the probability of its landing on tails is .4, what is the probability that there are at most 3 heads out of the 7 tosses.

Exactly as above, we can compute the probability that there are exactly 0 heads, exactly 1 head, exactly 2 heads and exactly 3 heads. Adding these together, we obtain the probability of there being at most 3 heads.

**Problem 15.2.** *Assuming that the coin is badly unbalanced, so that the probability of its landing on heads is .6, and the probability of its landing on tails is .4, what is the probability that there are at most 3 heads out of the 7 tosses.*

**Problem 15.3.** *Assuming that the coin is badly unbalanced, so that the probability of its landing on heads is .6, and the probability of its landing on tails is .4, what is the probability that there are at least 3 heads out of the 7 tosses.*

## 16. Basic Statistics — Data Analysis

We start with a set of data, perhaps scores on an exam, perhaps measurements of the length of a table, or perhaps something quite different. As a particular example, we give the following measurements of the length and width of a table. These are paired by the fact that each of the 11 rows of measurements was made by one student.

| Length | Width |
|---|---|
| 59.5 | 29.8 |
| 60.0 | 29.5 |
| 60.0 | 30.0 |
| 59.6 | 29.6 |
| 59.6 | 29.7 |
| 59.5 | 30.0 |
| 60.0 | 29.6 |
| 60.4 | 30.5 |
| 60.0 | 30.8 |
| 59.7 | 29.5 |
| 58.9 | 30.4 |

The first set of data, the first column, consists of 11 numbers, which, in order to be more general, we denote as $x_1, \ldots, x_{11}$; we sometimes refer to this vector in $\mathbb{R}^{11}$ as simply x. We likewise have the second column, or vector of 11 numbers as $y = \{y_1, \ldots, y_{11}\}$.

In order to have some consistency, we have reported all the measurements to the nearest tenth of an inch.

The *mean* or average of a set of data $x$ is, in general, given by the formula:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

The mean is also often denoted by the Greek letter $\mu$.

**Problem 16.1.** *Compute the means of the $x$ and $y$ data given in the above table.*

Note that most graphing calculators include a statistical package. These compute all of the statistics called for here, and many more. Our problem is to understand what these statistics mean, so that we can understand the answers we get from the calculator, when we call for these statistics.

Note that we should report the mean to the same accuracy as we have reported the original measurements. Our computations of the mean might very well yield more decimal places, but, since our original measurements are only accurate to the nearest tenth of an inch, it would be misleading to report the mean with seemingly greater accuracy.

One can view the mean as being a *balance point* for the first moment. That is, again in the general case, if we take the sum of the distances of the data from the mean, we obtain

$$\sum_{i=1}^{n}(x_i - \bar{x}) = \sum_{i=1}^{n} x_i - n\bar{x} = n\bar{x} - n\bar{x} = 0.$$

That is, the sum of the distances from the mean on the negative side exactly balances the sum of the distances from the mean on the positive side.

The next question we need to address is: How much variation is there in this data; that is, are all the data points relatively close to the mean, or are they more scattered. A measure of this is the *standard deviation*of a set of data, defined by

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

There are, unfortunately, two distinct standard deviations, this is the *population standard deviation*, which is more accessible, in terms of understanding; the other is the *sample standard deviation*, defined by:

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

From here on, we will use the population standard deviation, and make no further mention of the sample standard deviation.

One can understand the standard deviation in terms of its square, the *variance*, which is given by

$$\sigma^2 = \frac{1}{n}\sum(x_i - \bar{x})^2.$$

The variance can be viewed as the average second moment about the mean. The first moment of one data point is the difference between the data point and the mean of the data; the second moment for one data point is the square of this difference.

**Problem 16.2.** *Compute the variance and the standard deviation for both of the data sets given at the beginning of this section.*

While the mean of a set of data gives some information, it is only one number. It is often useful to appropriately chunk the data and then draw a picture of it. We first find the *median* of a set of data. We start by sorting the data so that the smallest number comes first, then the next smallest and so on. Call the sorted data $\tilde{x}_1, \dots, \tilde{x}_n$, so that $\tilde{x}_1 \le \tilde{x}_2 \le \cdots \le \tilde{x}_n$. After we have sorted the data in this fashion, the median is the half-way point. That is, if $n$ is odd, the median is $x_{(n+1)/2}$, while if $n$ is even, the median is the mean of $x_{n/2}$ and $x_{(n+2)/2}$.

**Problem 16.3.** *Find the medians of the sets of data given at the beginning of this section.*

Another sometimes useful statistal measure is the *mode* of a set of data. It could happen that we have two equal data points; that is, for some $i$ and $j$, $x_i = x_j$. It could happen that there are also other equal data points; for example, in the set of lengths listed above, the number 60.0 occurs four times, and this is more than the occurrence of any other value. In this case, we say that 60.0 is the mode. If there is no single value that occurs more often than all the other values that do occur, then we say that there is no mode. For example, there is no mode in the case that all the data points are distinct.

**Problem 16.4.** *Does the set of width measurements have a mode; if so what is it?*

In addition to the median, which we can think of as the halfway point, we also want to pick out the quarter way points, the 25-th and 75-th percentile points. At the 25-th percentile point, one-quarter of the data points lie below, and three-quarters of the data points lie above; this is exactly reversed at the 75-th percentile point.

**Problem 16.5.** *Find the 25-th and 75-th percentile points for the length and width data given at the beginning of this section.*

If we have a lot of data, it is easier to understand if we group the data, so that nearby data points are grouped together. For example, we could group the length data by half inch intervals; this would give us the following *Frequence Distribution Table* for our length data.

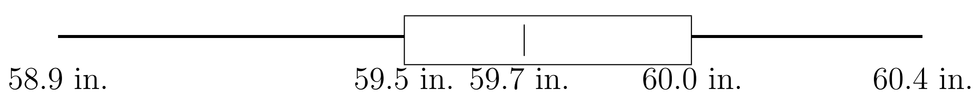
| Interval | Number |
|---|---|
| 58.5 to 58.95 inches | 1 |
| 59.0 to 59.45 inches | 0 |
| 59.5 to 59.95 inches | 5 |
| 60.0 to 60.45 inches | 5 |

For other purposes, we might want to know how many data points there are below certain points; this would be given by a *Cumulative Frequency Distribution Table*. For the purpose of illustration, we show the Cumulative Frequency Distribution Table for the Length data given at the beginning of this section, in intervals of .2.

| Data points less than | Number |
|---|---|
| 58.8 inches | 0 |
| 59.0 inches | 1 |
| 59.2 inches | 1 |
| 59.4 inches | 1 |
| 59.6 inches | 3 |
| 59.8 inches | 6 |
| 60.0 inches | 6 |
| 60.2 inches | 10 |
| 60.4 inches | 10 |
| 60.6 inches | 11 |

Our final example of how to visual one set of data is the *Box and Whiskers* plot. This has a horizontal line from the minimum value of the data to the 25-th percentile point. Then there is a box from the 25-th to the 75-th percentile point, with a vertical bar at the median separating it into two chambers, and then there is a horizontal line out to the maximum value.

We illustrate this with the Length data, where the minimum is at 58.9 inches; the 25-th percentile is at 59.5 inches; the median is at 59.7 inches ; the 75-th percentile is at 60.0 inches; and the maximum is at 60.4 inches.




16.1. **Regression and Correlations.** Suppose we have two sets of paired data, such as our length and width data, which are paired by the fact that each pair of data, length and width, was measured by the same person.

In general, we have two paired data sets, $x = x_1, \ldots, x_n$, and $y = y_1, \ldots, y_n$, where we have some reason to associate $x_i$ with $y_i$. We can look at these paired data in terms of a *scatter plot*, which is the $x, y$ plane with the points $(x_i, y_i)$, $i = 1, \ldots, n$ plotted on it.

The *regression line* is the straight line that best approximates $y$ as a function of $x$; there is also another regression line that best approximates $x$ as a function of $y$. This is an approximation in terms of functions; that is, we write the regression line as $y = mx + b$, and we want to compare this function with the values $y_i(x_i)$, $i = 1, \ldots, n$. Our measure of the difference is that of least squares; that is, we want to choose the slope $m$ and the intercept $b$ so as to minimize the expression:

$$E = \sum_{i=1}^{n} (y_i - (mx_i + b))^2. \tag{1}$$

In the above, the $y_i$ and $x_i$ are known, we treat this as a function of the two variables, $m$ and $b$, and minimize.

We know from the calculus of several variables that the minimum will occur at a point where the partial derivatives with respect to these two variables vanish. Before we start this computation, we perform a translation of our axes. We introduce new variables, $\tilde{x} = x - \bar{x}$, and $\tilde{y} = y - \bar{y}$, so that, for these new variables, the mean of $\tilde{x}$ is equal to the mean of $\tilde{y}$ is equal to zero. We notice that this transformation does not change the slope $m$, but it does change the intercept, $b$. In fact, we will see below that our regression line for these new variables goes through the origin.

$$
\begin{aligned}
\frac{\partial E}{\partial m} &= \sum_{i=1}^{n} -2\tilde{x}_i(\tilde{y}_i - (m\tilde{x}_i + b)) \\
&= -2\sum_{i=1}^{n} \tilde{x}_i\tilde{y}_i + 2m\sum_{i=1}^{n} \tilde{x}_i^2 - 2b\sum_{i=1}^{n} \tilde{x}_i \\
&= -2\sum_{i=1}^{n} \tilde{x}_i\tilde{y}_i + 2m\sum_{i=1}^{n} \tilde{x}_i^2,
\end{aligned} \tag{2}
$$

where we have used the fact that $\sum \tilde{x}_i = 0$. We also have

$$
\begin{aligned}
\frac{\partial E}{\partial b} &= \sum_{i=1}^{n} -2(\tilde{y}_i - m\tilde{x}_i + b) \\
&= -2nb.
\end{aligned} \tag{3}
$$

where we have used that $\sum \tilde{x}_i = \sum \tilde{y}_i = 0$.

Setting these both equal to zero, we see that $b = 0$, and that

$$
m = \frac{\sum_{i=1}^{n} \tilde{x}_i\tilde{y}_i}{\sum_{i=1}^{n} \tilde{x}_i^2}, \tag{4}
$$

This gives us our regression line $\tilde{y} = m\tilde{x}$, in terms of the variables $\tilde{x}$ and $\tilde{y}$. Translating back to our original variables, $x$ and $y$, we obtain that the equation of the least squares regression line is $y = mx + b$, where

$$
m = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}. \tag{5}
$$

and $b = \bar{y} - m\bar{x}$.

16.2. **Standardized Data.** In our example, we had two paired measurements of length, but that is unusual; it is more usual to have for example, a score on a math exam and a score on a history exam. In these cases, we will want to adjust not only for different means, but also for different variability. To this end, we introduce *standardized* scores. The standardized scores $\hat{x}_1, \ldots, \hat{x}_n$ corresponding to the scores $x_1, \ldots, x_n$, are given by $\hat{x}_i = (x_i - \bar{x})/\sigma$.

**Problem 16.6.** *Show that standardized scores have mean of 0, and standard deviation of 1.*

We now look again at the formula for the mean square difference between the regression line and the points $(\hat{x}_i, \hat{y}_i)$ given by standard scores. In this case, we call the slope of the regression line $r = \sum \hat{x}_i \hat{y}_i$, and the intercept is, as above, 0. Then the error $E$ made by approximating our data by the regression line is given by:

$$
\begin{aligned}
E &= \sum (\hat{y}_i - r\hat{x}_i)^2 \\
&= \sum \hat{y}_i^2 - 2r \sum \hat{y}_i \hat{x}_i + rm^2 \sum \hat{x}_i^2 \\
&= 1 - r^2
\end{aligned} \tag{6}
$$

The number $r$ is called the *correlation coefficient*. Notice that it is the slope of the regression line for predicting from $x$ to $y$, and also of the regression line for predicting from $y$ to $x$.

**Problem 16.7.** *Consider $\hat{x} = \hat{x}_1, \ldots, \hat{x}_n$ and $\hat{y} = \hat{y}_1, \ldots, \hat{y}_n$ as vectors in $n$-space. Show that the correlation coefficient $r$ is also the cosine of the angle between these two vectors.*

## 17. Exponential Functions

Our exploration of the exponential functions begins with a sequence of problems, starting with the basic algebraic property of converting addition into multiplication. For mathematical purposes, we also need some additional properties.

### 17.1. **Defining properties for the exponential functions.**

(1) for all $x$ and $y$, $f(x + y) = f(x)f(y)$;
(2) $f(x)$ is defined and continuous for all $x$;
(3) $f(x)$ is not constant; that is, there are two numbers, $x_1$ and $x_2$ so that $f(x_1) \neq f(x_2)$; and
(4) $f(x)$ is differentiable at zero.

The main point of the additional hypotheses concerning continuity and differentiability is to make the connection between the known properties of $a^r$, where $a$ and $r$ are both rational, and the transcendental function $e^x$. We also note that, from the mathematical point of view, we do not as yet know that any such function exists.

### 17.2. **Exponential properties.**

**Problem 17.1.** *Show that $f(0) = 1$.*

**Problem 17.2.** *Show that, for all $x$, $f(x)f(-x) = 1$.*

**Problem 17.3.** *Show that, for all $x$, $f(x) \neq 0$.*

**Problem 17.4.** *Show that, for all $x$, $f(x) > 0$.*

**Problem 17.5.** *Show that, for all $x$, $f(-x) = 1/f(x)$.*

**Problem 17.6.** *Show that, for all $x$, $(f(x))^2 = f(2x)$.*

**Problem 17.7.** *Use mathematical induction to show that, for all $x$, and for all natural numbers $n$, $(f(x))^n = f(nx)$.*

**Problem 17.8.** *Show that, for all $x$, and for all natural numbers $n$, $(f(x))^{1/n} = f(x/n)$.*

**Problem 17.9.** *Show that, for all $x$, and for all rational numbers $r$, $(f(x))^r = f(rx)$.*

**Problem 17.10.** *Use the definition of the derivative to show that, since $f$ is differentiable at 0, it is differentiable at every $x$. Further, $f'(x) = f'(0)f(x)$.*

**Problem 17.11.** *Show that, since $f(x)$ is not constant, $f'(0) \neq 0$. (HINT: It is a consequence of the mean value theorem that a function whose derivative is everywhere equal to zero is constant.)*

**Problem 17.12.** *Show that $f$ has derivatives of all orders at every point.*

**Problem 17.13.** *Show that if $f'(0) > 0$, then $f$ is increasing at every $x$; if $f'(0) < 0$, then $f$ is decreasing at every $x$.*

Let $a = f(1)$. We know that $f(0) = 1$, and $f(x)$ is either increasing or decreasing. It follows that either $0 < f(1) < 1$, or $f(1) > 1$.

**Problem 17.14.** *Show that if $f'(0) > 0$, then $a > 1$. Likewise, if $f'(0) < 0$, then $a < 1$.*

**Problem 17.15.** *Show that, for every rational number $r$, $f(r) = a^r$.*

We now define the number $e$ by specifying that $f'(0) = 1$, and setting $e = f(1)$. We then define, for every $x$, the function $e^x$ to be the solution to the first order differential equation, $f'(x) = f(x)$, and $f(0) = 1$.

In the above, we use the existence of solutions of the initial value problem for the differential equation $y' = y$.

We still need to observe that this function $e^x$ satisfies our properties. The first property follows from the uniqueness of solutions of the initial value problem as follows. Consider the function $g_1(x) = e^{x+a}$, for some real number $a$. Taking derivatives, we see that $g_1'(x) = g_1(x)$, and $g_1(0) = e^a$. Next look at the function $g_2(x) = e^a e^x$. Taking derivatives, we obtain that $g_2'(x) = g_2(x)$, and $g_2(0) = e^a$. We conclude that $g_1(x) = g_2(x)$; i.e., $e^{x+a} = e^x e^a$, for every $x$ and for every $a$.

The other property that we need for this function is that $e^{kx} = (e^x)^k$ for every real number $k$. Since this function $e^x$ satisfies the properties listed in 17.1, we can use what we already know about these functions to conclude that $e^{rx} = (e^x)^r$ for every rational number $r$; in particular, $e^{-x} = 1/e^x$. Since we have not as yet defined the quantity $a^x$, when $x$ is irrational, we can define $(e^x)^a = (e^a)^x = e^{ax}$, for all $a$ and for all $x$.

We still need to know that the function $e^x$ maps the real line onto the positive real numbers. To this end, we consider a new function $h(x) = e^x - x$. We observe that $h(0) = 1$, and that $h'(x) = e^x - 1$. We know that $e^0 = 1$, and that $e^x$ is increasing for all $x$. We conclude that $h'(x) > 0$ for all $x > 0$; that is, $h(x)$ is an increasing function of $x$ for all positive $x$. Since

$h(0) = 1$, and $h(x) = e^x - x$ is increasing, we conclude that $h(x) \geq x$ for all $x > 0$. It follows that, as $x \to \infty$, $e^x \to \infty$. Since $e^{-x} = 1/e^x$, it follows that, as $x \to -\infty$, $e^x \to 0$.

Putting together the information above, we have shown that the function $e^x$ maps the real line in a one-to-one manner onto the positive real numbers.

**Problem 17.16.** *Since $e^x$ has derivatives of all orders, we can write down its Taylor series at the origin; call it*

$$\sum_{n=0}^{\infty} a_n x^n.$$

*We know that $e^0 = a_0 = 1$. Differentiate term by term to show that*

$$a_n = \frac{a_{n-1}}{n}$$

*. Then use mathematical induction to conclude that $a_n = 1/n!$ for all $n$.*

**Problem 17.17.** *Show that the Taylor series*

$$\sum_{n=0}^{\infty} \frac{1}{n!} x^n$$

*converges for all real numbers $x$; conclude that*

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n.$$

We next return to the general exponential function. Again, we set $a = f(1)$, and we know that $a > 1$ if $f'(0) > 0$, and $a < 1$ if $f'(0) < 0$. We set $k = e^a$, and we solve the initial value problem, $y' = ky$, with $y(0) = 1$. We know that this problem has a unique solution, and we observe that $y = e^{kx} = (e^k)^x = a^x$ is that solution. Also, as above, we note that, for any fixed real number $b$, $a^{x+b}$ satisfies the same differential equation, $y' = ky$, with a different intial value; here, $y(0) = a^b$. We conclude that $a^{x+b} = a^b a^x$, which is our defining property for exponential functions.

We still need to know, as above, that $a^x$ maps the real line in a one-to-one manner onto the positive real numbers. Since we know that the derivative of this function, $e^a a^x$ is either everywhere positive or everywhere negative, we know that $a^x$ is a one-to-one function. Since $a^{-x} = 1/a^x$, it suffices to consider only the case that $a > 1$, in which case, $a^x$ is increasing.

**Problem 17.18.** *Let $a > 1$. Set $h(x) = a^x - e^a x$. Show that $h'(x) > 0$ for all $x > 0$. Conclude that $a^x \to \infty$ as $x \to \infty$.*

## 18. Continuous probability distrubutions — The Normal Curve

We've already seen finite probability spaces, each containing a finite number of elements, where each element is given a non-negative real number, called its probability, so that the sum of the probabilities is 1.

A Probability Distribution Function (PDF), is a continuous function $f(x)$, defined for some real interval, $a < x < b$, so that for all such $x$, $f(x) \geq 0$ and

$$\int_a^b f(x)\,dx = 1.$$

If $f(x)$ is the PDF for the event $E$ (whatever that might be), then the probability that the event $E$ occurs between the values $c$ and $d$, is given by

$$\int_c^d f(x)\,dx = 1.$$

An example of this is the wait-time distribution. This starts with the following problem. You make a phone call and are put on hold. How long do you have to wait until someone answers the phone? The wait-time PDF is defined for all positive $x$, and is given by the function $ke^{-kx}$, where $k$ is a parameter that depends on various aspects of the situation (how many phone calls on average do they receive; how long on average does it take to answer a phone call; how many people do they have on average answering these phone calls.)

Notice that this distribution is only defined for positive $x$. The meaningn of the PDF is that the probability that you will have to wait between 1 and 2 minutes after placing your phone call, is given by

$$P = k\int_1^2 e^{-kx}\,dx$$

Since there is no chance of them answering the phone before you've called, there is no need to try to define this function for negative $x$.

We can define the mean of the PDF in analogy to our definition of the mean of a set of data. The mean is the point $\mu$ so that the total first moment about $\mu$ is equal to 0; that is:

$$k\int_0^\infty (x-\mu)e^{-kx}\,dx = 0.$$

Solving for $\mu$, we obtain

$$\mu = \frac{k\int_0^\infty xe^{-kx}\,dx}{k\int_0^\infty e^{-kx}\,dx} = \frac{1}{k},$$

where we have used the fact that

$$k\int_0^\infty e^{-kx}\,dx = 0,$$

and we have used integration by parts, to evaluate

$$\int_0^\infty kxe^{-kx}\,dx.$$

We can also define the median of a PDF as being that number $M$, so that, in this case,

$$\int_0^M ke^{-kx}\,dx = \frac{1}{2}.$$

Integrating and solving for $M$, we obtain that the median for this distribution is at $M = \frac{\ln 2}{k}$.

Finally, we can also find the standard deviation of this PDF; this is the square root of the integral of the second moment about the mean; that is

$$\sigma^2 = k\int_0^\infty (x-\mu)^2 e^{-kx}\,dx$$

Integrating by parts, we obtain that $\sigma^2 = \frac{1}{k^2}$, so $\sigma = \frac{1}{k}$..

18.1. **The normal curve.** The normal (PDF), usually called the normal distribution or normal curve is defined, in its standard form, by

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}.$$

It is clear that this function is defined and positive. The proof that its integral is equal to 1 goes as follows.

$$\begin{aligned}
(\int_{-\infty}^\infty e^{-\frac{x^2}{2}}\,dx)^2 &= \\
&= (\int_{-\infty}^\infty e^{-\frac{x^2}{2}}\,dx)(\int_{-\infty}^\infty e^{-\frac{y^2}{2}}\,dy) \\
&= \int_{-\infty}^\infty\int_{-\infty}^\infty e^{-\frac{x^2}{2}} e^{-\frac{y^2}{2}}\,dxdy \\
&= \int_{-\infty}^\infty\int_{-\infty}^\infty e^{-\frac{x^2+y^2}{2}}\,dxdy \\
&= \int_0^{2\pi}\int_0^\infty e^{-\frac{r^2}{2}}\,rdrd\theta \\
&= 2\pi.
\end{aligned}$$

Since $f(x)$ is even, which means that $f(-x) = f(x)$, or, equivalently, that the graph of $f(x)$ is symmetric about the $y$-axis, it follows at once that both the mean and the median are at zero.

Using integration by parts (just once!), we compute the standard deviation as follows.

$$\begin{aligned}
\sigma^2 &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty x^2 e^{-\frac{x^2}{2}}\,dx \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-\frac{x^2}{2}}\,dx \\
&= 1.
\end{aligned}$$

The (unstandardized) normal distribution function is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma}}. \tag{7}$$

**Problem 18.1.** *Use the change of variables $y = \frac{x-\mu}{\sigma}$ to show that*

$$\int_{-\infty}^{\infty} f(x)\,dx = 0,$$

*where $f(x)$ is defined by equation 7.*

**Problem 18.2.** *Use the same change of variables as above to show that the Normal PDF, given by equation 7 has mean $\mu$ and standard deviation $\sigma$.*

18.2. **Cumulative Distribution Functions.** We usually want to know the probability that a distribution, such as the normal distribution, lies between certain values, and these can be difficult to compute; for example, the normal distribution function cannot be directly integrated using elementary functions. The Cumulative Distribution Function (CDF) is the integral of the PDF, that is, its value at the point $x$ is the probability that the associated event is less than $x$. (In the world of continuous probabilities, the probability of an event falling at a particular point is 0; so there is no difference between the probability of the event being less that $x$, or of it being less than or equal to $x$.)

For the wait-time distribution, where the PDF is given by $ke^{-kx}$, the corresponding CDF is defined by

$$g(x) = \int_0^x ke^{-kt}\,dt = 1 - e^{-kx}.$$

We cannot write down the CDF for the normal distribution. but there are tables of its values available. For example, New York State distributes with its Math B regents exam a chart that shows the standardized normal distribution, with labels on the $x$-axis, and corresponding vertical lines, at $\pm.5, \pm 1, \pm 1.5$ etc., and the integral of the standardized normal distribution between each of these points is shown on the chart as a percentage.

## 19. Logorithms

We saw in Section 17 that the exponential function, $a^x$, where $a > 0$ and $a \neq 1$, is a one-to-one differentiable function from the entire real line onto the positive real numbers. This exponential function is monotone increasing if $a > 1$, and is monotone decreasing if $a < 1$. The main properties that we need now are that $a^{x+y} = a^x a^y$, and that the derivative of the function $a^x$ is $ka^x$, where $e^k = a$.

In this section, we will explore the inverse function, $\log_a(x)$. The meaning of inverse function in this case is that $a^{\log_a x} = x$ for all $x > 0$, and $\log_a(a^x) = x$ for all real numbers $x$.

For the special case that $a = e$ ($e^x$ is the exponential function whose derivative at 0 is equal to 1.), we write the inverse function as $\ln x$; that is, $\ln(e^x) = x$ and $e^{\ln x} = x$.

Once we have these functions defined we can write the derivative of $a^x$ as $\ln a a^x$.

**Problem 19.1.** *Show that* $\log_a(1) = 0$.

**Problem 19.2.** *Show that, for all* $x > 0$ *and* $y > 0$, $\log_a(xy) = \log_a(x) + \log_a(y)$.

**Problem 19.3.** *Show that, for all* $x > 0$, $\log_a(\frac{1}{x}) = -\log_a(x)$.

**Problem 19.4.** *Show that, for all* $x > 0$ *and for all* $y$, $\log_a(x^y) = y \log_a(x)$.

**Problem 19.5.** *Show that* $\frac{d}{dx} \log_a(x) = \frac{1}{(\ln a)x}$.

**Problem 19.6.** *Show that* $\ln x < x$ *for all* $x > 0$.

**Problem 19.7.** *Find* $\log_2(4)$.

**Problem 19.8.** *Find* $2^{\log_2(4)}$.

**Problem 19.9.** *Find* $4^{\log_2(4)}$.

**Problem 19.10.** *Solve for* $x$*:* $\log_x(16) = 2$.

**Problem 19.11.** *Show that, for all* $x$, $\frac{\log_2(x)}{\log_4(x)} = 2$.

**Problem 19.12.** *Solve for* $x$*:* $2^{2\log_2(6+x)+2} = 16$.

**Problem 19.13.** *Show that for every* $a > 0$ *and for every* $b > 0$, *there is a constant* $C$ *so that*

$$\frac{\log_a(x)}{\log_b(x)} = C.$$

**Problem 19.14.** *Show that for all* $a > 0$ *and for all* $b > 0$, $(\log_a(b))(\log_b(a)) = 1$.

## 20. Geometry

In this section, our goal is to present the proofs of several basic geometric facts, such as the Pythagorean theorem; the formula for the area of a triangle; the law of sines; the law of cosines; and the addition formulae for the sine and cosine. We will also discuss some less well-known facts concerning existence of triangles.

We assume throughout that the usual axioms of (Euclidean) geometry are known. In particular, we assume known that there is a unique parallel to a line through a point not on the line, and that this is equivalent to the usual statement concerning alternate interior angles. We also assume that the laws of congruence of triangles are known, and that the basic facts about circles are known.

20.1. **Existence and congruence of triangles.** We specify a triangle by its vertices; we write $\triangle ABC$. Following high school notation, we write $m\angle A$ to denote the measure, in degrees, of the angle at $A$. In this section, we regard the measure of an angle as positive; that is, we do not distinguish the angle from $a$ to $b$, from the angle from $b$ to $a$. The side opposite $A$ is denoted by $a$. We do not follow the high school notation for the length or measure of side $a$; we use the notation $|a| = |BC|$ to denote the length of side $a$.

Two triangles $\triangle ABC$ and $\triangle A'B'C'$ are congruent if corresponding angles have equal measure and corresponding sides have equal length; in this case, we write $\triangle ABC \cong \triangle A'B'C'$. Note that while it is always true that $\triangle ABC \cong \triangle ABC$, in general, $\triangle ABC \not\cong \triangle BCA$ (in fact, $\triangle ABC \cong \triangle BCA$ if and only if this triangle is equilateral.

The laws of congruence, $SAS$, $ASA$, $SSS$ and $AAS$, are well known. It is also well known that $SSA$ is not a law of congruence. (That is, it is known that we can have two non-congruent triangles, $\triangle ABC$ and $\triangle A'B'C'$, where $|a| = |a'|$, $|b| = |b'|$ and $m\angle A = m\angle A'$.)

**Problem 20.1.** *If one is given three positive numbers, $a$, $B$ and $c$, find necessary and sufficient conditions for there to be a triangle having sides of lengths $a$ and $b$, and included angle of measure $B$.*

**Problem 20.2.** *If one is given three positive numbers, $A$, $b$ and $C$, find necessary and sufficient conditions for there to be a triangle having angles of measure $A$ and $B$, and have the side between them of length $b$.*

**Problem 20.3.** *If one is given three positive numbers, $a$, $b$ and $c$, find necessary and sufficient conditions for there to be a triangle having sides of lengths $a$, $b$ and $c$.*

**Problem 20.4.** *If one is given three positive numbers, $a$, $B$ and $c$, find necessary and sufficient conditions for there to be a triangle having sides of lengths $a$ and $b$, and included angle of measure $B$.*

**Problem 20.5.** *If one is given three positive numbers, $A$, $B$ and $a$, find necessary and sufficient conditions for there to be a triangle having adjacent angles of measure $A$ and $B$, where one of the sides other than the one between these two angles has length $a$.*

**Problem 20.6.** *If one is given three positive numbers, $a$, $b$ and $A$, under what conditions on these three numbers is it true that there is no triangle having two adjacent sides of lengths $a$ and $b$, where the angle opposite the side of length $a$ has measure $A$? Under what conditions is it true that all such triangles are congruent? Under what conditions is it true that there are exactly two non-congruent such triangles? Can there be more than two?*

One can ask similar questions about quadrilaterals. Just as $SAS$ is a congruence rule for triangles, so one could have $SASAS$ as a congruence rule for quadrilaterals. (In fact, this is a congruence rule; it is clear that if we are given 3 sides and the two included angles of a quadrilateral, then the endpoints of the sides are all determined, so the fourth side, and remaining angles, are determined by this information. However, if one is given 3 lengths and 2 angles, even if the sum of the angles is less than 360°, there may not be a quadrilateral having 3 sides of these lengths, with the included angles having these measures.)

To start with, it is not clear how much information we need to specify in order to obtain a congruence rule. It is immediate that all 8 pieces of information suffices, since that is essentially the definition of congruence. Since the sum of the angles of a quadrilateral is equal to 360°, if we know 3 of the angles, then we know all 4. That leaves only one possibility for 7 pieces of information; that is, $ASASASA$. We can cut each of the quadrilaterals into two corresponding triangles, and then it is easy to show that the corresponding triangles are congruent.

For 6 pieces of information, again, if 3 of these are angles, then we can compute the fourth angle, putting us back into a situation that we have already resolved. This leaves the possibilities that we are given the lengths of 4 sides and the measures of 2 angles, or that we are given the lengths of 2 sides and the measures of all 4 angles. In the first case, regardless of whether the 2 angles are adjacent or opposite, we can construct diagonals cutting both quadrilaterals into two triangles, where corresponding triangles are congruent, so the quadrilaterals are congruent. In the second case, if the two sides are opposite, then there is no congruence rule, as can be seen by looking at two rectangles, where one has sides of length 1 and 2, and the other has sides of length 1 and 3. If the 2 sides are adjacent, then, as above, the quadrilaterals are congruent.

**Problem 20.7.** *Suppose we are given two quadrilaterals, $ABCD$ and $A'B'C'D'$, where $m\angle A = m\angle A'$, $m\angle B = m\angle B'$, $m\angle C = m\angle C'$, $m\angle D = m\angle D'$, sides $|AB| = |A'B'|$, and $|BC| = |B'C'|$. Show that quadrilateral $ABCD$ is congruent to quadrilateral $A'B'C'D'$.*

**Problem 20.8.** *Suppose one is given that two quadrilaterals have equal lengths of 3 sides and equal measures of 2 angles. How many distinct possible such configurations are there? For each of these configurations, state whether or not the two quadrilaterals are necessarily congruent, and give reasons for your answers.*

20.2. **Area and the Pythagorean theorem.** A rectangle is a quadrilateral with all four angles having measure equal to 90°. If $ABDC$ is a rectangle, then it follows from alternate interior angles that sides $AB$ and $CD$ are parallel, and also that sides $AD$ and $BC$ are parallel. We draw the diagonal $AC$, and observe, again using alternate interior angles, that $m\angle BAC = m\angle ACD$, and that $m\angle CAD = m\angle BCA$. It then follows that $\triangle ABC \cong \triangle CDA$, from which it follows that the opposite sides of the rectangle have equal length.

The starting point for knowledge of *area* is that if $ABCD$ is a rectangle with side lengths $a$ and $b$, then the area of $ABCD$ is $ab$. Since congruent figures have equal area (this is one of the basic assumptions of plane geometry), the area of the right triangle, $\triangle ABC$ is equal to $\frac{1}{2}|a||c|$.

**Problem 20.9.** *Consider $\triangle ABC$. Draw the perpendicular $h$ from $A$ to $a$. (This is the altitude.) Let $D$ be the point where it hits the line $a$. Show that the area of $\triangle ABC$ is equal to $\frac{|a||h|}{2}$ by computing the areas of $\triangle ABH$ and $\triangle AHC$.*

Let $a$ and $b$ be the legs of a right triangle and let $c$ be its hypoteneuse. Draw the square with side length $|a|+|b|$ in the coordinate plane so that its vertices are at $(0,0)$, $(|a|+|b|,0)$, $(|a|+|b|,|a|+|b|)$, and $(0,|a|+|b|)$. Construct a quadrilateral by connecting the points $(|a|,0)$, $(|a|+|b|,|a|)$, $(|b|,|a|+|b|)$, and $(0,|b|)$.

**Problem 20.10.** *Show that this quadrilateral is a square of side length $|c|$, and that the complement of this square in the bigger square consists of four congruent right triangles of side lengths $|a|$, $|b|$ and $|c|$. Conclude that the Pythogorean theorem is true; i.e., $|a|^2+|b|^2 = |c|^2$.*

20.3. **SIMILAR TRIANGLES AND THE DEFINITIONS OF THE TRIGONOMETRIC FUNCTIONS.** One of the basic facts in plane geometry is that similar triangles have proportional corresponding sides. The proof of this starts with a construction involving areas of rectangles.

Let $\triangle ABC$ be a right triangle, where $\angle A$ is a right angle. Let $D$ be some point on side $b$. Draw the line through $D$ parallel to side $c$, and let $E$ be the point where this line intersects $a$. Construct the line through $B$ parallel to side $b$, and let $F$ be the point where it intersects the line determined by $D$ and $E$. Draw the line through $C$ parallel to $c$, and let $G$ be the point where it intersects the line determined by $B$ and $F$. Finally, draw the line through $E$ parallel to $b$. Let $H$ be the point where this line intersects the line determined by $A$ and $B$, and let $I$ be the point where this line intersects the line determined by $C$ and $G$.

**Problem 20.11.** *Show that $ABGD$, $AHED$, $ABFD$, $DEIC$, and $DFBC$ are all rectangles.*

**Problem 20.12.** *Show that $\triangle ABC \cong \triangle GCB$.*

**Problem 20.13.** *Show that $\triangle EHB \cong \triangle EFB$.*

**Problem 20.14.** *Show that $\triangle CED \cong \triangle CIE$.*

**Problem 20.15.** *Conclude from the above that rectangles $ABFD$ and $HBGI$ have equal areas.*

**Problem 20.16.** *Conclude from the above that*

$$\frac{AB}{AC} = \frac{BH}{HE};$$

*that is, the similar triangles, $\triangle BAC$ and $\triangle BHE$, have the ratio of their respective legs in the same proportion.*

**Problem 20.17.** *Use the Pythagorean theorem to conclude that the same proportion applies to the hypoteneuse; that is*

$$\frac{AB}{BC} = \frac{HB}{BE}.$$

**Problem 20.18.**
*Conclude that, for any acute angle $\alpha$, there is a right triangle with angle $\alpha$, and so $\sin\alpha$ equals opposite/hypoteneuse and $\cos\alpha$ equals adjacent/hypoteneuse are well defined. Then $\tan\alpha = \frac{\sin\alpha}{\cos\alpha}$, $\cot\alpha = \frac{\cos\alpha}{\sin\alpha}$, $\sec\alpha = \frac{1}{\cos\alpha}$ and $\csc\alpha = \frac{1}{\sin\alpha}$ are also all well defined.*

20.4. **THE LAW OF SINES.** Let $\triangle ABC$ be such that angles $A$ and $C$ are acute. Draw the altitude from $B$ to side $b$; this is the line passing through $B$ and perpendicular to side $b$; call the altitude $h$, and let $D$ be the point where $h$ ends on side $b$. Since angles $A$ and $C$ are acute, $D$ lies between $A$ and $C$.

**Problem 20.19.** *Compute $\sin A$ and $\sin C$, and use these to show that $\frac{\sin A}{a} = \frac{\sin C}{c}$.*

Since we have not as yet defined the sine of an obtuse angle, we do not yet know the full law of sines for a triangle with an obtuse angle.

**Problem 20.20.** *Show that, if $\triangle ABC$ is a right triangle, then the law of sines holds; that is*

$$\frac{\sin A}{|a|} = \frac{\sin B}{|b|} = frac{\sin C}{|c|}.$$

20.5. **THE LAW OF COSINES.** Draw the same picture as above. Note that $D$ splits side $b$ into two pieces, call them $b_1$ and $b_2$, where $|b_1|^2 + |h|^2 = |c|^2$. (Then of course $|b_2|^2 + |h|^2 = |a|^2$.)

**Problem 20.21.** *Use the definition of the cosine, together with the Pythogorean theorem to prove the law of cosines in this case; that is, show that*

$$a^2 = b^2 + c^2 - 2bc\cos A$$

As above, we have not as yet defined the cosine of an obtuse angle. Notice that for a right angle, the law of cosines reduces to the Pythagorean theorem.

20.6. **THE ADDITION FORMULAE FOR SINE AND COSINE.** Draw the same picture as for the law of sines; assume also that angle $B$ is acute. Now let $\angle ABD = \alpha$, and let $\angle CBD = \beta$.

**Problem 20.22.** *Compute the sine and cosine of angles $A$ and $C$.*

**Problem 20.23.**
*Use the law of sines to express $\sin(\alpha+\beta)$ in terms of other (known) quantities, and show that $\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$.*

**Problem 20.24.**
*Use the law of cosines to express $\cos(\alpha+\beta)$ in terms of other (known) quantities, and show that $\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$.*

## THE DOUBLE ANGLE AND HALF ANGLE FORMULAE

**Problem 20.25.** *Set $\alpha = \beta$ in the addition formula for the sine to arrive at the double angle formula for the sine: $\sin 2\alpha = 2\sin\alpha\cos\alpha$.*

**Problem 20.26.** *Set $\alpha = \beta$ in the addition formula for the cosine to arrive at the double angle formula for the cosine: $\cos 2\alpha = \cos^2\alpha - \sin^2\alpha$.*

**Problem 20.27.** *Use the double angle cosine formula for a half-angle to arrive at the half angle formulae for sine and cosine:*

$$\sin\frac{\alpha}{2} = \pm\sqrt{\frac{1-\cos\alpha}{2}},$$

*and*

$$\cos\frac{\alpha}{2} = \pm\sqrt{\frac{1+\cos\alpha}{2}}.$$

20.7. **The Trigonometric functions.** As with the other transcendental (not rational) functions, the connection between the trigonometric functions as defined for angles of a right triangle and the periodic functions defined for all real numbers $x$ is given by calculus. We first need to compute two limits.

Let $O$ be the origin, let $A$ be the point $(0,1)$, let $B = (x,y)$ be some point in the first quadrant on the unit circle; let $\theta = \angle AOB$, and let $C$ be the point where the line $(1,t)$ intersects the line determined by $O$ and $B$. Then the line segment $\overline{A}C$ has length equal to $\tan\theta$, $\sin\theta = y$, and $\cos\theta = x$. We observe that the triangle with vertices at $O$, $(x,0)$ and $B$ is contained in the circular wedge $OAB$, which in turn is contained in $\triangle OAC$. We conclude that

$$\frac{\sin\theta}{2} < \frac{\theta}{2} < \frac{\tan\theta}{2}. \tag{8}$$

Dividing by $\sin\theta$, we obtain

$$1 < \frac{\sin\theta}{\theta} < \frac{\tan\theta}{\sin\theta} = \frac{1}{\cos\theta}. \tag{9}$$

Since $\lim_{\theta\to 0}\cos\theta = 1$, we see at once that

$$\lim_{\theta\to 0}\frac{\theta}{\sin\theta} = 1. \tag{10}$$

We will also need the following limit.

$$\lim_{\theta\to 0}\frac{1-\cos\theta}{\theta} = \lim_{\theta\to 0}\tfrac{(1+\cos\theta)(1-\cos\theta)}{\theta(1+\cos\theta)} \tag{11}$$

$$= \lim_{\theta\to 0}\tfrac{1-\cos^2\theta}{\theta(1+\cos\theta)} \tag{12}$$

$$= \lim_{\theta\to 0}\tfrac{\sin\theta}{\theta}\tfrac{\sin\theta}{1}\tfrac{1}{1+\cos\theta} \tag{13}$$

$$= 1\cdot 0\cdot \tfrac{1}{2} \tag{14}$$

$$= 0. \tag{15}$$

**Problem 20.28.** *Use the above two limits, together with the addition formula for the sine, in the definition of the derivative to show that* $\frac{d}{dx}\sin x = \cos x$.

**Problem 20.29.** *Use the above two limits, together with the addition formula for the cosine, in the definition of definition of the derivative to show that* $\frac{d}{dx}\cos x = -\sin x$.

**Problem 20.30.** *Show that* $\frac{d}{dx}\tan x = \sec^2 x$, $\frac{d}{dx}\cot x = -\csc^2 x$, $\frac{d}{dx}\sec x = \sec x\tan x$, *and* $\frac{d}{dx}\csc x = -\csc x\cot x$.

**Problem 20.31.** *Show that* $\sin x$ *is the unique solution to the initial value problem,* $y''+y = 0$, $y(0) = 0$, *and* $y'(0) = 1$, *while* $\cos x$ *is the unique solution to the initial value problem* $y''+y=0$, $y(0)=1$, $y'(0)=0$.

**Problem 20.32.** *Using the above definitions of sine and cosine as solutions to the above initial value problems, show that* $\frac{d}{dx}(\sin^2 x+\cos^2 x) = 0$*, and conclude that* $\sin^2 x+\cos^2 x = 1$*.*

**Problem 20.33.** *Let* $\triangle ABC$ *be a right triangle, where* $B$ *is the right angle.*

**Problem 20.34.** *Use the Pythagorean theorem to show that* $\sin^2\theta + \cos^2\theta = 1$*.*

20.8. **Approximation of $\pi$.** Consider a circle of unit diameter, so that its circumference has length $\pi$. We can approximate $\pi$ by the following procedure, which we define inductively.

For the first step, construct an inscribed square, and a circumscribed square. Let $p_1$ be the perimeter of the inscribed square, and let $P_1$ be the perimeter of the circumscribed square. Then $p_1 < \pi < P_1$.

Assume we have constructed an inscribed regular polygon with $2^{n+1}$ sides, $n > 1$, call its perimeter $p_n$, and we have constructed a circumscribed regular polygon with $2^{n+1}$ sides, call its perimeter $P_n$. Draw the perpendicular bisector of each of the sides of each of these polygons. Each perpendicular bisector passes through the center of the circle; we can use the point where it hits the circle as the vertex of a regular incsribed or circumscribed polygon with $2^{n+2}$ sides. Consider a triangle formed by one of the sides of the old polygon, and the two corresponding sides of the new one. By induction, we know the length of the side of the old polygon, and we know the sine and cosine of the corresponding central angle. We can use the half angle formulas to find the sine of the angles of the new triangle, and so compute the length of the new side.

It is important to observe that we need only the sine and cosine of the central angle at the first step, where the angle is $\pi/2$. We can then inductively find the values of the trigonometric functions we need by using the half-angle formulae.

**Problem 20.35.**
*Carry out the first two steps of the above procedure. That is, find the perimeters of the inscribed and circumscribed square and octogon.*