99C010I

Project Number: 51-BZS-2904

TREATISE OF FUNDAMENTAL TOPICS IN MATHEMATICS

An Interactive Qualifying Project Report

submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Bachelor of Science

by

Peter R. Minear

Peter R. Minear

Date: December 18, 1998

Approved:

Brigitte Servatius

Professor Brigitte Servatius, Major Advisor

Table of Contents

Abstract

WPI freshmen experience difficulties in their basic science courses because of poor background in mathematics. This IQP examines the reasons for this phenomenon and develops a treatise of fundamental topics in order to help students fill the gap between expected and actual knowledge.

Introduction

In recent years, much has been reported about the low achievement of American students in mathematics.

A study of 12 industrialized nations conducted in 1964 found that of "17-year-old students enrolled in a math-intensive college preparatory high school curriculum, that is, each country's best prepared students, the Americans ranked last" (see [5], p. 2). Even more disturbing is the magnitude of the gap in the mathematical competencies of American students, relative to the best-educated students from other countries. Four out of five Americans scored below the overall international average in math.

The Second International Mathematics Study, conducted from 1980-1982, found that the top five percent of America's best high school seniors "had only average scores in relation to international standards in algebra, functions and calculus, and slightly above average scores in geometry" (see [5], p. 3). The SIMS findings also show that the United States ranks among the lowest in student retention of math. The U.S. was classified as a low coverage and low retention system (see [18], pp. 134-140).

These educational studies indicate that the math curricula found in various countries are determined more by tradition than by fundamental educational principles. "American textbooks tend to develop ideas very slowly by progressing through a hierarchy of small, straightforward learning tasks. Texts from Asian countries and from the Soviet Union immerse students in much more demanding problem situations from the beginning (see [5], p. 27).

In 1995, the Educational Testing Service reported that although math achievement was on the rise, only "about five to ten percent of students are able to demonstrate satisfactory or in-depth performance on problem-solving tasks." Additionally, "many students perceive mathematics as mostly memorizing facts and not as a way to solve problems in real world situations" (see [4], p. 2). Other researchers note that students in math classrooms often learn rules for manipulating symbols, but fail to learn the meaning of the symbols and the principles which they represent. They plunge into problem solving without making sure they fully understand the problem (see [3], p. 72).

The Third International Mathematics and Science Study, in 1996, surveyed education in 50 countries and included students who are in their last year of secondary school. This comprehensive study included analysis of textbooks, curricula, and instructional practices, as well as questionnaires and student assessments (U.S. "The Third" 1). The Summary of Findings resulting from this study notes that math curricula standards in the U.S. "are unfocused and aimed at the lowest common denominator. In other words, they are a mile wide and an inch deep." Also cited is a "lack of common standards on what to teach and how to teach it" (see [16], p. 1). Even textbooks are criticized. "U.S. mathematics and science textbooks included far more topics than was typical internationally, but provided significantly less coverage than the international average for the five most emphasized topics in math and science" (2). Of the 21 nations whose students took part in the math assessment, American high school seniors came in 19th. American advanced students in math came in 15th among 16 nations (see [9]).

Schools have been adopting changes to cope with these dismal results, but the changes have been criticized. Interactive math has students working in groups to discover concepts with little direct instruction from a teacher (see [5], p. 66). Some new math textbooks contain poetry and essays about cultural differences in contrast to Japanese texts which are all about math (see [7], p. 182). In the first year after adopting these new math reforms, public school students in Palo Alto, California dropped from the 86th percentile nationally to the 58th, then went back up to the 77th percentile the next year when the schools moderated their approach (see [10]).

One of the problems in changing the way math is taught in American classrooms is that researchers cannot agree on standardized methods of defining and recording outcomes (see [11], pp. 87-88). This leaves school systems and teachers confused about which new teaching methods to implement.

Critics of math education in the United States have expressed their feelings to Congress. Senator Robert Byrd, in an address to the Senate, cited the results of the Third International Math and Science Study. He criticized a recent algebra textbook for its low mathematics content, and condemned the philosophy of teaching students that with computers, there is little need to do calculations and solve equations (see [2], p. 2). He expressed dismay at the current outcomes in math education, despite the billions of dollars the Federal government spends to improve education (1). In conclusion, he urged parents to "get heavily involved to reverse that trend now" (3).

Dr. Martha Schwartz, a representative of Mathematically Correct, an organization of concerned parents, mathematicians, scientists, and educators, recently expressed their concerns in a speech delivered to the House of Representatives. She stated that the U.S. is slipping further behind in math education, and that an increasing number of parents are growing dissatisfied with the math taught in their schools (see [14], p. 1). She blamed some of the problems on state governments, and referred to a Fordham Foundation report, which found that most states fail to delineate standards for math education (2).

Yet, the public is partly to blame for the weakness in math education. Richard J. Shumway notes that traditionally the math topics taught in a school system are determined by the demands of society. He states that "There is continuing public pressure for a school mathematics program that guarantees minimal mathematical competence for survival" (393).

In light of these findings, I questioned whether students at WPI are experiencing problems due to weak high school math backgrounds. The technical nature of WPI's curriculum requires students to have a stronger mathematical background than the general population. I decided to continue my research on the WPI campus to determine if students here are affected by poor high school math preparation.

Selection of Topics

To determine if the problem of weak high school math backgrounds affects WPI students, I took a survey of the entire freshman class. Another reason for the survey was to assess what areas of math were least understood. The freshmen were asked to indicate their major/minor subjects, and to indicate how they well they felt they understood each of the following main topics: functions, complex numbers, logarithms, trigonometry, polar coordinates, geometry formulas, vectors, arithmetic & geometric series, and probability. Below is a summary of the weaknesses of the 181 freshmen who responded.

	Those Indicating Little or No Knowledge	
Topic	Number of Respondents	Percent of Respondents
Probability	72	40
Polar Coordinates	68	38
Vectors	49	27
Arithmetic & Geometric Series	48	27
Logarithms	35	19
Complex Numbers	30	17
Functions	7	4

Seven students mentioned they did not feel comfortable with any of the math topics.

Another method I used to find out more about typical math weaknesses of incoming WPI freshmen was to consult with several of WPI's MASH (Math and Science Help) leaders. I spoke with Carlos Calvo, Chris Cole, Melissa Curry, Patricia DeChristopher, Dennis Hubbard, and Vikki Tsefrikas who work with the calculus courses MA1020, MA1021, and MA1022. The most common failures reported were with sequences and series, probability, and logarithms, while geometry and trigonometry were also mentioned as weaknesses. Chris Cole noted that some of his students in MA1020 have trouble with elementary algebra.

The MASH leaders feel that many students can memorize and regurgitate almost any topic, but have trouble in applying the topics to solving word problems and other applications, because these require thinking analytically through the problem. A widespread difficulty in doing proofs was observed. Proofs require an understanding of why a theorem is true, as opposed to problem solving, which requires only plugging numbers into a theorem and solving for an answer.

I obtained the following information on the failure rates of students taking calculus courses from the Math Department Office.

Course	Percent of Students Failing	
MA1021 '96 - 97	79 / 409 =	19 %
MA1022 '96 - 97	150 / 622 =	24 %
MA1023 '96 - 97	116 / 719 =	16 %
MA1024 '96 - 97	125 / 661 =	19 %
MA1021 '97 - 98	120 / 465 =	26 %
MA1022 '97 - 98	156 / 607 =	26 %
MA1023 '97 - 98	125 / 586 =	21 %
MA1024 '97 - 98	61 / 418 =	15 %

I also spoke with several members of WPI's math faculty to find out their opinions regarding the common weaknesses of incoming WPI freshmen. Professors with whom I spoke include Brigitte Servatius, Peter Christopher, Joseph Fehribach, Joseph Petruccelli, Peter Schultz, Christopher Larsen, Roger Lui, and J. J. Malone. The faculty noted weaknesses in algebra, functions, advanced trigonometry formulas, logarithms, probability, and discrete math. Some professors observed that there are WPI students who don't know even the basics of high school mathematics.

The faculty shared the opinion that students are insufficiently prepared by their high school math education, pointing out that most students are not taught to think logically to analyze problems, but rather to memorize the necessary formulas. Memorization without understanding sometimes leads to confusion of simple concepts such as sine and cosine of 30 degrees. Students are really lost when they cannot find an appropriate formula to use. Retention of math learned by memorization is poor. This causes professors to take time away from the teaching of new material to review what should have been learned in high school.

Finally, I visited Doherty High School in Worcester to learn about a typical high school math curriculum. I spoke with Mr. Bertrand Bolduc, the head of the math department, regarding the math courses offered there and the math requirements to graduate. Each student must take a minimum of three years of math, in grades 9 - 12. The normal honors sequence is Algebra I, Geometry, Algebra II, Precalculus, Calculus (AP AB). Of the topics in my treatise, polar coordinates are not covered at all, because that topic is not needed until BC calculus. Functions, logarithms, trigonometry, geometry formulas, vectors, and arithmetic and geometric series are all covered, most of them in the precalculus course. There are three different levels in most of these courses, 'Honors', 'Level 1' and 'Level 2' and the thoroughness of coverage varies greatly between levels. A student in the honors classes learns each topic thoroughly, and a student in the lowest level (level 1) learns little or nothing about any of them. I reviewed tests from the Honors level and noticed that they consisted mainly of short problems.

My research shows that the problem of American students having weaknesses in school math backgrounds is occurring among WPI students. After reviewing the feedback from students, tutors, and professors, the list of topics in my freshman survey appeared to be reasonable and comprehensive. Therefore, I have prepared my treatise on these topics. Since greater understanding rather than memorization is needed, I have included many explanations and proofs.

Algebra

Functions

A **function** from a set A to a set B is a mapping that corresponds each element in set A to exactly one element in set B. Set A is called the **domain**, or set of input values. Set B is called the **range**, or set of output values.

From this definition we can see some characteristics that are true of all functions. Each element in the domain must be mapped to exactly one element in the range, though the opposite is not true. Each element in the range may correspond with 1 or more than 1 element from the domain. A function for which every element in the range corresponds with exactly one element in the domain is a **one-to-one function**.

In algebra it is standard practice to represent a function by a set of ordered pairs (x,y) where the domain is the set of x values, and the range is the set of y values. The variable which represents the domain values is called the **independent variable**, and the variable which represents the range values is called the **dependent variable**.

Standard function notation is to use a lowercase letter to 'name' the function, and write it as f(x) = <expression> where x is the independent variable, and the <expression> is a formula for mapping each element in the domain. For example the equation $f(x) = x^2$ defines a function whose domain is the set of real numbers, and whose range consists of the squares of all the elements in the domain, which is the set of non-negative real numbers.

It is also possible to define functions by equations, for example the function above can be described by the equation $y = x^2$, where y is the dependent variable.

Domains of Functions

In most simple functions, the domain consists of all real numbers, however there are cases where some numbers are excluded. Consider a fraction with a polynomial function in the denominator, such as $f(x) = \frac{1}{x-5}$. Here trying to compute f(5) would require division by zero, therefore f(5) is undefined. All other values of x will produce a value f(x), therefore the domain of this function is all real numbers except 5.

Another example of a function which can have numbers excluded from its domain is a radical with an expression that can be negative, such as $f(x) = \sqrt{x}$. Since the square root operation is not defined for negative numbers, this function has a domain of all non-negative real numbers.

In each of the previous examples, the mapping formula was the same for all values in the domain. It is also possible to define a function different ways for different domain values, as long as the domain values do not overlap. For example, we can define a function g(x) which uses each of the previous definitions, $g(x) = \sqrt{x}$ for all $x \geq 0$, $\frac{1}{x-5}$ for all $x < 0$. This function has a domain of all real numbers.

Composition of Functions

Another way to define a function is by putting one function into another. For example, suppose we define $f(x) = x^2 + 5x - 6$, and $g(x) = x^2$. It is possible to **compose** these functions by defining a new function $h(x) = f(g(x))$. $f(g(x))$ simply means the expression for g(x) is substituted into the f(x) function.

Here $h(x) = f(g(x)) = (x^2)^2 + 5(x)^2 - 6$. It is possible to substitute f(x) into g(x) in the same way, obtaining $g(f(x)) = (x^2 + 5x - 6)^2$.

Inverse Functions

Some functions have the effect of 'undoing' each other, i.e. when they are composed with each other, they produce the original value of x. Such functions are called **inverses** of each other. Two functions f and g are inverses if and only if the domain of f is the same as the range of g (and vice-versa), and $f(g(x)) = g(f(x)) = x$. For example, the functions $f(x) = 3x - 7$ and $g(x) = \frac{1}{3}(x+7)$ are inverses of each other. The inverse of a function f(x) is denoted by $f^{-1}(x)$. Not all functions have an inverse, only one-to-one functions have an inverse.

Symmetry, Odd and Even Functions

Graphing a function often facilitates analyzing its features. Functions of one variable can be graphed in the xy plane, with the independent variable on the horizontal axis (called the x-axis) and the dependent variable on the vertical axis (called the y-axis).

Some functions have different forms of symmetry. **Graphically, a function is symmetric about a line if it produces a mirror image when folded over the line.** Two kinds of symmetry which are frequently used in algebra are symmetry with respect to the y-axis and symmetry with respect to the origin.

A function is symmetric with respect to the y-axis if $f(x) = f(-x)$ for all values of x in the domain. A function symmetric with respect to the y-axis is called an **even function**. Looking at the graph of an even function, it is easy to see that if it were folded over the y-axis, it would produce a mirror image.

A function is symmetric with respect to the origin if $f(-x) = -f(x)$ for all values of x in the domain. A function symmetric with respect to the origin is called an **odd function**. The graph of an odd function will produce a mirror image when folded over both the x-axis and y-axis. All polynomial functions containing only terms of odd degree are odd functions, and all polynomial functions containing only terms of even degree are even functions.

Arithmetic Operations on Functions

It is possible to apply the basic operations of arithmetic, namely addition, subtraction, multiplication, and division to functions. The domain of sum, difference, product, and quotient functions of functions f and g is the intersection of the domain of f and the domain of g. (In the case of quotient functions, any value which produces a zero in the divisor function must be eliminated.) The range is dependent on the functions themselves.

Synthetic Division

Division of a polynomial of degree greater than 1 by a polynomial of degree 1 can be done by a method called **Synthetic Division**. When dividing a polynomial $f(x) = a_nx^n + a_{n-1}x^{n-1} + \ldots + a_2x^2 + a_1x + a_0$ by a polynomial $x - k$, the result will be a polynomial

$g(x) = b_{n-1}x^{n-1} + \ldots + b_2x^2 + b_1x + b_0 + \frac{r}{x-k}$, where the quotient g(x) is a polynomial of degree one less than f(x), and r is the constant remainder.

To use Synthetic Division, write the coefficients of the dividend polynomial in decreasing order and put in a zero for the coefficient of any missing term. b_{n-1}, the leading coefficient of g(x), is the same as a_n, the leading coefficient of f(x). All other coefficients of the quotient polynomial are computed using the formula $b_i = a_{i+1} + k(b_{i+1})$ for $i = n-2, n-3, \ldots, 1, 0$. $r = a_0 + k(b_0)$.

Example: Divide the polynomial $2x^3 - 29x + 33$ by $x - 3$.

Here $k = 3$, and we must put in a zero for the missing x^2 term.

3 |	2	0	-29	33	dividend coefficients
		6	18	-33	
	2	6	-11	0 (remainder = 0)	quotient coefficients

Remainder Theorem

The Remainder Theorem states that if a polynomial f(x) is divided by (x – k) the remainder is equal to f(k). If the remainder is zero, then by applying the **Remainder Theorem** we can claim that x – k is a factor of f(x), and k is a root of f(x).

Finding Roots of Functions

A **root** of a function (also called a **zero**) is a member of the domain, which when applied to the function, produces a value of zero. There are many ways to determine the roots of a polynomial function. Every polynomial of degree n ($n \geq 1$) has at most n real roots, exactly n real or complex roots. The easiest way, when feasible, is to factor the polynomial and set each factor equal to zero, and solve.

Polynomials of degree two can always be solved using the **quadratic formula**, $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$, where a, b, c are the coefficients of x^2, x, and the constant term. However there is no simple formula for finding roots of polynomials of degree greater than two. Graphically, roots occur at all points where the function touches the x-axis, but this does provide an exact value. There are a few useful theorems to help us determine these roots.

Rational Root Test

The **Rational Root Test** states how to find rational roots in a polynomial. Given a polynomial $f(x) = a_nx^n + a_{n-1}x^{n-1} + \ldots + a_2x^2 + a_1x + a_0$ which has all integer coefficients, all rational roots of f must be of the form $\frac{p}{q}$, where p is a factor of the constant term a_0, and q is a factor of the leading coefficient a_n.

Intermediate Value Theorem

The **Intermediate Value Theorem** states that if f is a polynomial function such that $a < b$ and $f(a) \neq f(b)$ then every value between f(a) and f(b) occurs in the interval [a,b]. This theorem can be used to approximate a root if values of a and b can be found such that f(a) and f(b) have opposite sign. The closer a and b are, the better the approximation can be.

Descartes' Rule of Signs

Let f(x) be a polynomial of the form $a_nx^n + a_{n-1}x^{n-1} + \ldots + a_2x^2 + a_1x + a_0$, where all coefficients are real and $a_0 \neq 0$. Descartes' Rule of Signs states that

1) The number of positive real roots of f is either equal to the number of changes in sign of f(x) or is less than that number by an even integer,

2) The number of negative real roots of f is either equal to the number of changes in sign of f(-x) or is less than that number by an even integer.

Complex Numbers

There is no real number solution to the equation $x^2 = -1$. Mathematicians have invented a number system called the **imaginary numbers** where the number **i** is defined as the square root of -1, and the set of imaginary numbers consists of all numbers of the form bi, where b is any real number. All of basic arithmetic operations, such as addition, subtraction etc. can be performed on imaginary numbers just as on real numbers.

Complex numbers are created when a real number is added to an imaginary number. All complex numbers can be expressed in the form a + bi, where a and b are real numbers, and i is the imaginary unit. Complex numbers whose real component, a, is zero are called **pure imaginary numbers**.

The **complex conjugate** of a complex number is the complex number which has the same real component, and the opposite imaginary component. In other words, the complex conjugate of (a + bi) is (a – bi).

Arithmetic Operations on Complex Numbers

Addition and subtraction of complex numbers can be done by separately adding or subtracting the real and imaginary coefficients. For example, the sum of (1 + 2i) + (3 + 4i), is simply (1 + 3) + (2i + 4i) = 4 + 6i.

Multiplication of complex numbers is done by first expressing each complex number as a binomial, with real term and imaginary term, and multiplying the binomials. Using the same complex numbers as in the previous example, (1 + 2i)(3 + 4i) = 3 + 4i + 6i + $8i^2$. Since i was defined to be $\sqrt{-1}$, i^2 is –1. So the expression simplifies to 3 + 10i - 8 = -5 + 10i.

To perform **division** of complex numbers, it is first necessary to convert the denominator into a real number. This is done by multiplying both numerator and denominator by the complex conjugate of the denominator. The product of the conjugates produces a real number, which is divided into each component of the numerator. To find quotient $\frac{1+2i}{3+4i}$, first multiply $(\frac{1+2i}{3+4i})(\frac{3-4i}{3-4i}) = \frac{3-4i+6i-8i^2}{9-12i+12i-16i^2}$ $= \frac{11+2i}{25} = \frac{11}{25} + \frac{2}{25}i$.

Powers of i

One final operation that is worth knowing is how to compute powers of i. Since i was defined to be $\sqrt{-1}$, i^2 is –1. $i^3 = (i^2)(i) = -i$. $i^4 = (-i)(i) = -i^2 = 1$. Since $i^4 = 1$, we need not compute any further powers of i. i^k, where k > 4, is equal to $(i^{k-4})(i^4) = i^{k-4}$. So, to compute any power of i, take the remainder when the exponent is divided by 4. If the remainder is 0, the answer is 1, if the remainder is 1, the answer is i if the remainder is 2, the answer is -1 if the remainder is 3, the answer is –i.

Complex and Irrational Roots of Polynomials

If all of the coefficients of a polynomial are real, complex roots will occur in conjugate pairs such as $5 + 2i$ and $5 - 2i$.

If all of the coefficients of a polynomial are rational, irrational roots will occur in conjugate pairs such as $1 + \sqrt{3}$ and $1 - \sqrt{3}$.

Logarithms

Given an exponential equation $y = a^x$, we say that $\log_a y = x$. This is read 'The base a logarithm of y equals x.' **A logarithm is an exponent**. In the previous equation, x is a logarithm, it is the power to which a must be raised to obtain y. An example using numbers is $\log_5 125 = 3$, because 3 is the power to which 5 must be raised to obtain 125.

Every exponential equation can be rewritten as a log equation, and vice-versa.

Basic properties of logarithms, along with the corresponding exponential equation:

a) $\text{Log}_x(1) = 0 \quad \Leftrightarrow \quad x^0 = 1$

b) $\text{Log}_x(x) = 1 \quad \Leftrightarrow \quad x^1 = x$

c) $\text{Log}_x(x^y) = y \quad \Leftrightarrow \quad x^y = x^y$

These identities are obvious.

a) Follows because any non-zero number raised to the 0 power equals 1.

b) Is true because any number raised to the 1st power equals itself.

c) Any number raised to another power equals the same number raised to the same power!

Logarithm Bases and the Number e

Logarithms can be defined in any positive number base, except base 1. When there is no base specified, base 10 is assumed.

A special logarithm that is often used in mathematics and has many applications is called the **natural logarithm**. The natural logarithm involves taking the base e log, and is denoted by ln. i.e. $\ln(x) = \log_e(x)$.

The irrational number e, rounded to a few digits equals 2.71828 . . . There are many applications of natural logarithms; they are frequently used in calculus.

Domain and Range of Exponential and Logarithmic Functions

The domain of any exponential function, such as $f(x) = e^x$, is all real numbers, while the range is all positive numbers. The domain of any logarithmic function, such as $f(x) = \ln(x)$, is all positive numbers, while the range is all real numbers. The logarithm of a quantity between zero and one will be negative.

Other Properties of Logarithms:

Product Rule

1) $\log_a(uv) = \log_a(u) + \log_a(v)$

To prove this, suppose we have two logarithm equations, $x = \log_a(u)$, and $y = \log_a(v)$. Multiplying the corresponding exponential forms of the these equations, $u = a^x$, and $u = a^y$, gives us the equation $uv = a^x a^y$. Taking the logarithm to base a of both sides of the equation gives $\log_a(uv) = \log_a(a^{x+y}) = x + y$. Substituting $x = \log_a(u)$, and $y = \log_a(v)$ completes the proof.

Quotient Rule

2) $\log_a(\frac{u}{v}) = \log_a(u) - \log_a(v)$

To prove this, use the logarithm equations, $x = \log_a(u)$, and $y = \log_a(v)$.
Dividing the corresponding exponential forms of the these equations, $u = a^x$, and $u = a^y$, gives us the equation $\frac{u}{v} = a^x a^y = a^{x-y}$. Taking the logarithm to base a of both sides of

the equation gives $\log_a(\frac{u}{v}) = \log_a(a^{x-y}) = x - y$. Substituting $x = \log_a(u)$, and $y = \log_a(v)$ completes the proof.

Power Rule

3) $\log_a(u^n) = n \log_a(u)$.

This can be proven by performing n applications of property #1, e.g. $u^n = u*u*u \ldots$ hence $\log_a(u^n) = \log_a(u*u*u \ldots) = \log_a(u) + \log_a(u) + \log_a(u) + \ldots = n\log_a(u)$.

Change of Base Rules

4) $\log_a(x) = \dfrac{\log_b x}{\log_b a}$

To prove this, let $y = \log_a(x)$, which implies $a^y = x$. Taking the logarithm of both sides to base b, we get $\log_b(a^y) = \log_b(x)$. By applying property #3 to the left side of this equation, and dividing both sides by $\log_b(a)$, we complete the proof.

5) $\log_a(x) = \dfrac{1}{\log_x(a)}$

To prove #5, simply apply property #4, substituting x for b, thus obtaining

$$\log_a(x) = \frac{\log_b x}{\log_b a} = \frac{\log_x x}{\log_x a} = \frac{1}{\log_x a}$$

Solving Logarithmic and Exponential Equations

Most logarithmic and exponential equations can be solved using these few basic properties.

To solve a logarithmic equation of the form $\log_a x = b$ (where x is the unknown variable) it is helpful to exponentiate both sides, i.e. raise both sides to the power of a, obtaining $a^{\log a(x)} = a^b$.

To solve an exponential equation of the form $a^x = b$, it is helpful to take a logarithm of both sides, i.e. take the logarithm to base a of both sides, obtaining $\log_a(a^x) = x = \log_a(b)$.

Trigonometry

Trigonometry is a branch of mathematics which deals with the study of sides and angles of triangles.

Degrees and Radians

Angles can be measured in degrees or radians, but most of the study of trigonometry deals with radian measure. A **degree** is an angle equal to $\frac{1}{360}$ of a revolution around a circle. A **radian** is an arc length equal to one radius of a circle. Since the circumference of a circle = $2\pi r$, there are 2π radians in a circle. Since there are 360 degrees in a circle, it is possible to convert an angle from radians to degrees by multiplying by $\frac{360}{2\pi}$, and from degrees to radians by dividing by this factor. The radian measure of a central angle or a circle equals the length of the intercepted arc divided by the length of the radius. It follows that a straight angle has measure π radians, and a right angle has measure $\frac{\pi}{2}$ radians.

Unit Circle and Quadrants

The **unit circle** is the circle given by the equation $x^2 + y^2 = 1$. We can imagine the real number line wrapped around the unit circle starting with 0 corresponding to the point (1,0). Positive numbers are wrapped in a counter-clockwise direction, and negative numbers are wrapped in a clock-wise direction. Every real number θ will then correspond to exactly one point on the circle.

The unit circle can be divided into four **quadrants**. The 1st quadrant contains angles from 0 to $\frac{\pi}{2}$ radians, the 2nd quadrant contains angles from $\frac{\pi}{2}$ to π radians, the 3rd quadrant contains angles from π to $\frac{3\pi}{2}$ radians, and the 4th quadrant contains angles from $\frac{3\pi}{2}$ to 2π radians. For angles larger than 2π radians, or less than 0 radians, add or subtract multiples of 2π until it fits one of the above categories.

Quadrants from x and y Coordinates

Another way to determine the quadrant is from the x and y coordinates. A point which has a positive x coordinate and a positive y coordinate lies in the 1st quadrant. A point which has a negative x coordinate and a positive y coordinate lies in the 2nd quadrant. A point which has a negative x coordinate and a negative y coordinate lies in the 3rd quadrant. A point which has a positive x coordinate and a negative y coordinate lies in the 4th quadrant. Note that angles which are exact multiples of $\frac{\pi}{2}$ (points that have either the x or y coordinate equal to zero) are not considered part of any quadrant.

Trigonometric Functions

There are six trigonometric functions defined for angles on the unit circle and on triangles. They are sine, cosine, tangent, cotangent, secant, and cosecant. Below is a list of formulas to compute each trigonometric function for an angle θ. The first formula listed after each function is for an angle on the unit circle ($r = \sqrt{x^2 + y^2} = 1$), the second formula is for an acute angle in a right triangle.

$$\mathrm{Sin}(\theta) = \frac{y}{r} = \frac{\text{opposite leg}}{\text{hypotenuse}} \qquad \mathrm{Cos}(\theta) = \frac{x}{r} = \frac{\text{adjacent leg}}{\text{hypotenuse}}$$

$$\mathrm{Tan}(\theta) = \frac{y}{x} = \frac{\text{opposite leg}}{\text{adjacent leg}} \qquad \mathrm{Cot}(\theta) = \frac{x}{y} = \frac{\text{adjacent leg}}{\text{opposite leg}}$$

$$\mathrm{Sec}(\theta) = \frac{r}{x} = \frac{\text{hypotenuse}}{\text{adjacent leg}} \qquad \mathrm{Csc}(\theta) = \frac{r}{y} = \frac{\text{hypotenuse}}{\text{opposite leg}}$$

Worth noting above is that cotangent, secant, and cosecant are **reciprocals** of tangent, cosine, and sine, respectively.

Domain and Range of Trigonometric Functions

The domain of each of the six trigonometric functions is all real numbers. Thinking of the definitions on the unit circle, we can see that sine and cosine have a range of -1 to +1 inclusive or [-1, 1]. Tangent and cotangent have a range of all real numbers or $(-\infty, \infty)$. Secant and cosecant have a range of $(-\infty, -1] \cup [1, \infty)$, where $\cup$ **is the symbol for union of sets**. The trigonometric functions are periodic and are not one-to-one.

Trigonometric Functions of Special Angles

The values of the trigonometric functions of most angles cannot be calculated without a calculator or computer. However some 'special' angles have trigonometric functions which are simple. There are two kinds of right triangles that are familiar from geometry, the 90°-45°-45°, and the 90°-60°-30°.

In the **90°-45°-45° triangle**, both legs have equal length x. From the Pythagorean theorem, we derive that the hypotenuse has length equal to $\sqrt{2}$ times each leg, or $x\sqrt{2}$. Using the formulas for the three basic trig functions, we conclude the following:

$$\text{Sin}(45°) = \frac{\text{opposite leg}}{\text{hypotenuse}} = \frac{\sqrt{2}}{2} \qquad \text{Cos}(45°) = \frac{\text{adjacent leg}}{\text{hypotenuse}} = \frac{\sqrt{2}}{2}$$

$$\text{Tan}(45°) = \frac{\text{opposite leg}}{\text{adjacent leg}} = 1$$

The **90°-60°-30° triangle** comes from slicing an equilateral triangle in half along one of the altitudes. Therefore the shorter leg has length equal to half of the hypotenuse. Using the Pythagorean Theorem, we can determine that the three sides are in the ratio $1 : \sqrt{3} : 2$. From the ratio of the sides, we can easily figure out the values of the trigonometric functions:

$$\text{Sin}(30°) = \frac{\text{opposite leg}}{\text{hypotenuse}} = \frac{1}{2} \qquad \text{Cos}(30°) = \frac{\text{adjacent leg}}{\text{hypotenuse}} = \frac{\sqrt{3}}{2}$$

$$\text{Tan}(30°) = \frac{\text{opposite leg}}{\text{adjacent leg}} = \frac{\sqrt{3}}{3} \qquad \text{Sin}(60°) = \frac{\text{opposite leg}}{\text{hypotenuse}} = \frac{\sqrt{3}}{2}$$

$$\text{Cos}(60°) = \frac{\text{adjacent leg}}{\text{hypotenuse}} = \frac{1}{2} \qquad \text{Tan}(60°) = \frac{\text{opposite leg}}{\text{adjacent leg}} = \sqrt{3}$$

These are the only angles between 0 and 90 degrees that have simple trigonometric functions. It is possible, however, to use half-angle formulas and sum and difference formulas to compute trigonometric functions of some other angles. These formulas will be explained later.

Reference Angles

When dealing with angles greater than 90 degrees, it is impossible to define trig functions in terms of the parts of a right triangle. One way to define them is to use the unit circle definition, another way is to determine their values at the corresponding **reference angles**. A reference angle θ' of an angle θ is the acute angle formed by the terminal side of θ and the nearest horizontal axis.

All trig functions of θ are the same as those of the reference angle θ', with the possible exception of their sign. To see why this is true, consider the unit circle. Any angle in the 2nd or 3rd quadrant will have its reference angle formed using the negative horizontal axis, while an angle in the 1st or 4th quadrant has its reference angle formed using the positive horizontal axis. Suppose we construct an acute angle α in each of the 4 quadrants, by starting on the positive horizontal axis and going up for the 1st quadrant, down for the 4th quadrant, and starting on the negative horizontal axis going up for the 2nd quadrant, and down for the 3rd. It is obvious by properties of symmetry, that the x and y values of each of these angles all have the same magnitude, just different signs.

Signs of the Trig Functions

Using the formulas for finding trig functions in terms of the unit circle, we can determine which functions are positive in which quadrants. In the 1st quadrant, both x and y are positive, so the ratios $\frac{y}{r}$, $\frac{x}{r}$, and $\frac{y}{x}$ (that is, sine, cosine, and tangent respectively) will be positive. (Remember r always equals 1 in the unit circle.) In the 2nd quadrant, x is negative and y is positive, so $\frac{y}{r}$ is positive, $\frac{x}{r}$ is negative, and $\frac{y}{x}$ is negative. In the 3rd quadrant, x and y are both negative, hence $\frac{y}{r}$ and $\frac{x}{r}$ will both be negative, and $\frac{y}{x}$ is positive. In the 4th quadrant, x is positive and y is negative, so $\frac{y}{r}$ is negative, $\frac{x}{r}$ is positive, and $\frac{y}{x}$ is negative.

To summarize, sine (and its reciprocal cosecant) are positive in the 1st and 2nd quadrants, negative in the 3rd and 4th. Cosine (and its reciprocal secant) are positive in the 1st and 4th quadrants, negative in the 2nd and 3rd. Tangent (and its reciprocal cotangent) are positive in the 1st and 3rd quadrants, negative in the 2nd and 4th.

Identities

There are many different kinds of algebraic identities involving trigonometric functions, which are useful in applications. Following is a list of them, along with simple proofs.

Pythagorean Identities

1) $\sin^2\theta + \cos^2\theta = 1$
2) $\tan^2\theta + 1 = \sec^2\theta$
3) $\cot^2\theta + 1 = \csc^2\theta$

Number 1 is easy to see by looking at the unit circle. Construct a right triangle inside the unit circle. A radius of the circle will become the hypotenuse. One leg will be constructed by drawing a vertical line from the point to the horizontal axis, and the other leg is constructed by drawing a line from the origin along the horizontal axis until it meets the other leg. Here the angle θ is the central angle of the unit circle, formed by the hypotenuse and the horizontal leg. If we substitute the trig formulas for the unit circle into this identity, we get $(\frac{y}{r})^2 + (\frac{x}{r})^2 = 1$. If we multiply each term in this equation by r^2, we get $y^2 + x^2 = r^2$, which is exactly the Pythagorean theorem, i.e. the sum of the squares of the legs of a right triangle equals the square of the hypotenuse.

Identities 2 and 3 follow easily from identity 1. Dividing each term in 1 by $\cos^2\theta$ gives identity 2, and dividing each term in 1 by $\sin^2\theta$ gives identity 3.

CoFunction Identities

1) $\sin(\frac{\pi}{2} - \theta) = \cos\theta$

2) $\cos(\frac{\pi}{2} - \theta) = \sin\theta$

3) $\tan(\frac{\pi}{2} - \theta) = \cot\theta$

4) $\cot(\frac{\pi}{2} - \theta) = \tan\theta$

5) $\sec(\frac{\pi}{2} - \theta) = \csc\theta$

6) $\csc(\frac{\pi}{2} - \theta) = \sec\theta$

Identities 1 and 2 follow from the definition of sine and cosine of an acute angle of a right triangle. The adjacent leg to one acute angle is opposite leg to the other acute angle. Identity 3 comes from dividing 1 by 2, and 4 comes from dividing 2 by 1. Identities 5 and 6 come from taking reciprocals of 2 and 1.

Negative angle identities

1) $\sin(-\theta) = -\sin\theta$

2) $\cos(-\theta) = \cos\theta$

3) $\tan(-\theta) = -\tan\theta$

4) $\cot(-\theta) = -\cot\theta$

5) $\sec(-\theta) = \sec\theta$

6) $\csc(-\theta) = \csc\theta$

Here numbers 1 and 2 can be seen from the unit circle. Starting from the point (1,0) if we move along the circle a distance θ in both the positive and negative directions, the resulting positions will have the same x coordinate, but opposite y coordinates. Just as in the previous section, we obtain 3 and 4 by dividing 1 by 2, and 2 by 1 respectively. Numbers 5 and 6 again come from taking reciprocals of 2 and 1.

Inverse Trig Functions

Since the trigonometric functions are periodic and are not one-to-one, we must limit their domains to principal values in order to obtain inverse functions. The restricted domain of the original trigonometric function becomes the range of the inverse function. The six inverse trigonometric so obtained are called arcsin, arccos, arctan, arccot, arcsec, and arccsc. Sometimes they are written as $\sin^{-1}$, $\cos^{-1}$, $\tan^{-1}$, $\cot^{-1}$, $\sec^{-1}$, and $\csc^{-1}$.

Principal values in the 1st and 4th quadrants are used for sine and tangent.

Arcsin is defined by y = arcsin x if and only if sin y = x, where the domain of x values is the closed interval [-1,1] and the range of y values is the closed interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

Arctan is defined by y = arctan x if and only if tan y = x, where the domain of x values is all real numbers and the range of y values is the open interval $(-\frac{\pi}{2}, \frac{\pi}{2})$.

Principal values in the 1st and 2nd quadrants are used for cosine and cotangent.

Arccos is defined by y = arccos x if and only if cos y = x, where the domain of x values is the closed interval [-1,1] and the range of y values is the closed interval $[0, \pi]$.

Arccot is defined by y = arccot x if and only if cot y = x, where the domain of x values is all real numbers and the range of y values is the open interval $(0, \pi)$.

The last two inverse trig functions are rarely used. Principal values for secant and cosecant are generally considered to be the same as for cosine and sine, as indicated below, but some authors list them as 1st and 3rd quadrants. $\cup$ is the symbol for union of sets.

Arcsec is defined by y = arcsec x if and only if sec y = x, where the domain of x values is $(-\infty, -1] \cup [1, \infty)$ and the range of y values is $[0, \frac{\pi}{2}) \cup (\frac{\pi}{2}, \pi]$.

Arccsc is defined by y = arccsc x if and only if csc y = x, where the domain of x values is $(-\infty, -1] \cup [1, \infty)$ and the range of y values is $[-\frac{\pi}{2}, 0) \cup (0, \frac{\pi}{2}]$.

Euler's Identity

Additional trigonometric formulas can be derived from an important identity, called **Euler's Identity**, which states the following: $\mathbf{e^{i\theta} = \cos\theta + i\sin\theta}$. The proof of this theorem is simple. If we perform the Taylor expansion of e^{ix}, we get

$$e^{ix} = 1 + ix - \frac{x^2}{2!} - \frac{ix^3}{3!} + \frac{x^4}{4!} + \frac{ix^5}{5!} - \frac{x^6}{6!} - \frac{ix^7}{7!} \ldots$$ where the terms of even degree form the Taylor expansion for cos x, and the terms of odd degree form the Taylor expansion for sin x.

Trig Formulas for the Sum of Two Angles

Now we can use Euler's Identity to derive angle sum formulas. Start with the equation $e^{i(A+B)} = (e^{iA})(e^{iB})$. Apply Euler's identity to both sides, which gives the following: $\cos(A + B) + i \sin(A + B) = (\cos A + i \sin A)(\cos B + i \sin B)$. Multiplying out the right side and combining the real parts and imaginary parts gives us:
$\cos(A + B) + i \sin(A + B) = \cos A\cos B - \sin A\sin B + i(\sin A\cos B + \cos A\sin B)$.

The formula for sine of a sum of two angles comes from equating the imaginary parts of both sides:

$$\mathbf{\sin(A + B) = \sin A\cos B + \cos A\sin B}$$

The formula for cosine of a sum comes from equating the real parts of both sides:

$$\mathbf{\cos(A + B) = \cos A\cos B - \sin A\sin B}$$

By dividing these two formulas, we obtain the formula for tangent of a sum.

$$\mathbf{\tan(A + B)} = \frac{\sin(A+B)}{\cos(A+B)} = \frac{\sin A\cos B + \cos A\sin B}{\cos A\cos B - \sin A\sin B} = \frac{\tan A + \tan B}{1 - \tan A\tan B}$$

Trig Formulas for the Difference of Two Angles

Angle difference formulas can be found simply by substituting (-B) for B in the above derivations.

Start with the equation $e^{i(A-B)} = (e^{iA})(e^{i(-B)})$. Apply Euler's identity to both sides, which gives: $\cos(A - B) + i \sin(A - B) = (\cos A + i \sin A)(\cos(-B) + i \sin(-B))$. Multiplying out the right side and combining the real parts and imaginary parts gives us $\cos(A - B) + i \sin(A - B) = \cos A\cos(-B) - \sin A\sin(-B) + i(\sin A\cos(-B) + \cos A\sin(-B))$. Now, since $\sin(-x) = -\sin(x)$ and $\cos(-x) = \cos(x)$, we substitute obtaining
$\cos(A - B) + i \sin(A - B) = \cos A\cos B + \sin A\sin B + i(\sin A\cos B - \cos A\sin B)$.

The formula for sine of a difference of two angles comes from equating the imaginary parts of both sides:

sin(A - B) = sinAcosB - cosAsinB

The formula for cosine of a difference comes from equating the real parts:

cos(A - B) = cosAcosB + sinAsinB

By dividing these two formulas, we obtain the formula for tangent of a difference.

$$\textbf{tan (A - B)} = \frac{\sin(A-B)}{\cos(A-B)} = \frac{\sin A\cos B - \cos A\sin B}{\cos A\cos B + \sin A\sin B} = \frac{\tan A - \tan B}{1+\tan A\tan B}$$

Double Angle Formulas

Double angle formulas can be derived by substituting A for B in the angle sum formulas.

sin(A + A) = sin (2A) = sinAcosA + cosAsinA = 2sinAcosA

cos(A + A) = cos (2A) = cosAcosA – sinAsinA = $\cos^2$A – $\sin^2$A

$$\tan(A + A) = \tan(2A) = \frac{\tan A + \tan A}{1-\tan A\tan A} = \frac{2\tan A}{1-\tan^2 A}.$$

Half Angle Formulas

The double angle cosine formula can be used to produce another set of formulas called half angle formulas. From the unit circle we know that $\cos^2$A + $\sin^2$A = 1. Thus, the double angle cosine formula can be written in two other ways: cos (2A) = 1 – 2$\sin^2$A = 2$\cos^2$A – 1. The first variation of the formula is used to obtain the half angle sine formula, the second variation is used to obtain the half angle cosine formula, and the half angle tangent formula is obtained by dividing.

$$\cos(2A) = 1 - 2\sin^2 A \Rightarrow \sin^2 A = \frac{1-\cos(2A)}{2}$$

$$\cos(2A) = 2\cos^2 A - 1 \Rightarrow \cos^2 A = \frac{1+\cos(2A)}{2}$$

The forms above are useful for integrating $\sin^2$A and $\cos^2$A.

Substituting $\frac{A}{2}$ for A, we obtain the half angle formulas:

$$\sin\left(\frac{A}{2}\right) = \pm\sqrt{\frac{1-\cos A}{2}} \quad \text{and} \quad \cos\left(\frac{A}{2}\right) = \pm\sqrt{\frac{1+\cos A}{2}}.$$

Dividing these last two results, $\tan\left(\frac{A}{2}\right) = \sqrt{\frac{1-\cos A}{1+\cos A}}$.

Law of Sines

Two other formulas which are useful to solving for the sides and angles of a triangle are the **Law of Sines** and the **Law of Cosines**. The Law of Sines states that the ratio of a side of a triangle to the sine of the opposite angle is constant for all three sides.

In other words, $\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}$

where a,b,c are the sides and A,B,C are the angles opposite a,b,c respectively.

To prove the Law of Sines, construct a triangle ABC, and draw an altitude h from angle C to its opposite side. The result will be two right triangles, each having one of the acute angles of the original triangle opposite the altitude h. From one of the triangles we can see that h = a sin B, from the other h = b sin A. Equating these parts gives us

$a \sin B = b \sin A \Rightarrow \frac{a}{\sin A} = \frac{b}{\sin B}$. Since the name of each side and angle is arbitrary, this equation holds true for all three sides and angles of any triangle.

Law of Cosines

The Law of Cosines states that $a^2 = b^2 + c^2 - 2bc \cos A$. (a,b,c can be substituted for each other)

To prove this, construct a triangle ABC in the xy plane, where A = (0,0), B = (c,0) and C = (x,y) where x,y > 0. Since C has coordinates (x,y), and A has coordinates (0,0), the side b forms the hypotenuse of a right triangle, with the 3rd point being (x,0). It follows from this triangle that x = b cos A and y = b sin A. Using the distance formula, the length of side a is the distance from B to C, which is $\sqrt{(x-c)^2 + (y-0)^2}$.

Squaring both sides and substituting gives us $a^2 = (b \cos A - c)^2 + (b \sin A)^2$

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

$= b^2 (\sin^2 A + \cos^2 A) + c^2 - 2bc \cos A = b^2 + c^2 - 2bc \cos A.$

The first and last parts of this equation make the Law of Cosines.

Solving for Missing Parts of a Triangle Using the Law of Sines or Law of Cosines

Usually it is possible to use either the Law of Sines or the Law of Cosines to solve a problem, but not both of them. When all three sides of a triangle are known, it is impossible to solve for the angle measures using the Law of Sines because none of the ratios are known, but using the Law of Cosines it is easy to plug in sides a,b,c and solve for the angles.

Another case when it is impossible to use the Law of Sines is when there are two sides known, and only the angle between them is known. Here also, none of the required ratios for the Law of Sines is known, however the Law of Cosines can be used to first solve for the unknown side, then proceed to find the two missing angles as in the previous example.

When only one side of a triangle is known, the Law of Cosines cannot be used. However, when at least two angles are also known, the third angle can be found by subtraction, and thus the Law of Sines can be used because the angle opposite the known side can be used to compute the necessary ratio.

Rectangular and Polar Coordinates

The most common way to represent points on the plane is by using (x,y) coordinates, where the first number represents the units along the horizontal axis, and the second number represents the units along the vertical axis. This type of coordinates is known as **rectangular coordinates**.

There is another coordinate system which is often used in trigonometry and calculus, known as **polar coordinates**. With polar coordinates, the first coordinate represents the distance away from the origin, and the second coordinate represents the angle between the point and the positive horizontal axis. (This is the same way that angles are measured on the unit circle.)

Conversions between Rectangular and Polar Coordinates

There are many applications which cannot be done using rectangular coordinates, but can be done easily using polar coordinates, thus it is necessary to have an easy method of converting from one system to the other. Suppose we are given a point in rectangular coordinates, (x,y), which we want to convert into polar coordinates. To find the distance from the origin, simply use the Pythagorean Theorem to compute the hypotenuse of the inscribed right triangle, i.e. $r^2 = x^2 + y^2$, thus

$$\mathbf{r} = \sqrt{x^2 + y^2}$$

The x and y values correspond to the adjacent leg, and opposite leg, respectively, of the triangle. Since $\tan\theta = \frac{y}{x}$, $\theta =$ **Arctan** $\frac{y}{x}$**.**

Now suppose we are given a point in polar coordinates, (r, θ), which we want to convert into rectangular coordinates. Since the x coordinate is the adjacent leg, and $\cos\theta = \frac{\text{adjacent leg}}{\text{hypotenuse}} = \frac{x}{r}$, we can multiply both sides of this equation by r, obtaining

$$\mathbf{x = r \cos\theta}\,.$$

Since the y coordinate is the opposite leg, and $\sin\theta = \frac{\text{opposite leg}}{\text{hypotenuse}} = \frac{y}{r}$,

$$\mathbf{y = r \sin\theta}\,.$$

DeMoivre's Theorem

In the previous section we saw how to multiply complex numbers. DeMoivre's Theorem is a shortcut to taking powers and roots of complex numbers. The complex numbers must be expressed in polar form, where $r = \sqrt{a^2 + b^2}$ and $\theta = \text{Arctan}\ \frac{b}{a}$.

Powers of Complex Numbers

DeMoivre's Theorem states the following: If $z = r(\cos\theta + i \sin \theta)$ is a complex number and n is a positive integer, then we find z^n by taking the radius to the n^{th} power and multiplying θ by n.

$$z^n = (r(\cos \theta + i \sin \theta))^n = r^n(\cos n\theta + i \sin n\theta).$$

For example, calculate $(2 + 2\sqrt{3}\, i)^8$ using DeMoivre's Theorem. We convert $(2 + 2\sqrt{3}\, i)$ to polar form, $r = 4$, $\theta = \frac{\pi}{3}$, and $n = 8$. Then apply DeMoivre's Theorem,

$$(2 + 2\sqrt{3}\, i)^8 = (4(\cos\frac{\pi}{3} + i \sin \frac{\pi}{3}))^8 = 4^8(\cos(\frac{8\pi}{3}) + i \sin(\frac{8\pi}{3})) =$$

$$65536\,(-\frac{1}{2} + \frac{\sqrt{3}}{2} i) = -32768 + 32768\sqrt{3}\, i.$$

Roots of Complex Numbers

DeMoivre's Theorem states that if n is a positive integer, the complex number $z = r(\cos\theta + i \sin \theta)$ has exactly n distinct roots found by taking the positive n^{th} root of the radius and dividing $(\theta + 2\pi k)$ by n.

$$\sqrt[n]{z} = \sqrt[n]{r}\,(\cos \frac{\theta + 2\pi k}{n} + i \sin \frac{\theta + 2\pi k}{n}) \quad \text{where } k = 0, 1, 2, \ldots, n-1$$

For example, calculate $\sqrt{i}$ using DeMoivre's Theorem. We convert $\sqrt{i}$ to polar form, $r = 1$, $\theta = \frac{\pi}{2}$, and $n = 2$, so $\sqrt{i} =$

$$\sqrt{1}\,(\cos\frac{\theta}{2} + i \sin\frac{\theta}{2}) = \cos(\frac{\pi}{4}) + i \sin(\frac{\pi}{4}) = \frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}} \quad \text{and}$$

$$\sqrt{1}\,(\cos\frac{\theta + 2\pi}{2} + i \sin\frac{\theta + 2\pi}{2}) = \cos(\frac{5\pi}{4}) + i \sin(\frac{5\pi}{4}) = -\frac{1}{\sqrt{2}} - \frac{i}{\sqrt{2}}$$

Geometry Formulas

The following are several useful geometry formulas, several of which are derived from calculus.

Triangle Area Formulas

Triangle with sides a,b,c opposite angles A,B,C respectively:

$$\text{Area} = \frac{1}{2}ab\sin C$$

Equilateral Triangle with side s: $\text{Area} = \dfrac{s^2\sqrt{3}}{4}$

Triangle with sides a,b,c and semiperimeter $s = \dfrac{a+b+c}{2}$:

$$\text{Area} = \sqrt{s(s-a)(s-b)(s-c)}$$

Surface Area Formulas

Sphere with radius r: Surface area $= 4\pi r^2$

Circular cylinder with radius r and height h: Surface area $= 2\pi r^2 + 2\pi rh$

Circular cone with radius r and slant height l: Surface area $= \pi r^2 + \pi rl$

Volume Formulas

Prism with base area B and height h: Volume $= Bh$

Pyramid with base area B and height h: Volume $= \dfrac{Bh}{3}$

Circular cylinder with radius r and height h: Volume $= \pi r^2 h$

Circular cone with radius r and height h: Volume $= \dfrac{\pi r^2 h}{3}$

Sphere with radius r: Volume $= \dfrac{4\pi r^3}{3}$

Analytic Geometry

Equations of Lines in a Plane

There are three standard forms for expressing equations of lines in the xy plane. The first is called **slope-intercept** form, which is $y = mx + b$, where m is the **slope**, and b is the **y-intercept**. The slope can be defined as the change in y per unit change in x. The y-intercept is the value of the y-coordinate when $x = 0$, also the point at which the line crosses the vertical axis at the point (0,b).

Another standard form of a line is **point-slope form**. This form is useful to find the equation if two points of the line are known, or if the slope and one point are known. An equation in point-slope form is of the form $y - y_0 = m(x - x_0)$ where (x_0, y_0) is a point on the line. If two points are known, then the slope can be calculated by $\frac{y_1 - y_0}{x_1 - x_0}$, and either point can be used in the point-slope form. The slope-intercept form is actually a specific case of the point-slope form, where $x_0 = 0$, and y_0 is the y-intercept, which is b.

The third way to express the equation of a line is the **general form**, which is to put all variables and constants on one side of the equation and set it to zero in the form $Ax + By + C = 0$.

Conic Sections

A **conic section** is a curve which results from the intersection of a plane and a cone. The basic types of conic sections: the **ellipse**, the **circle**, the **hyperbola**, and the **parabola**.

Ellipses

An ellipse is the set of all points such that the sum of the distances from two distinct fixed points, called **foci**, is constant.

Ellipses are oval shaped, with the foci located inside the ellipse, and the line connecting the foci is known as the **major axis**. The line perpendicular to the major axis is the **minor axis**. The major axis intersects the ellipse at the vertices, and the midpoint between the two vertices (which is also the midpoint between the two foci) is known as the **center** and is located at the point (h,k).

The distance from the center to each vertex is defined as a, thus the major axis has a length of 2a. The distance from the center to each point where the minor axis crosses the ellipse is defined as b, thus the minor axis has a length of 2b. The distance from the center to each focus is defined as c.

Using these, we can derive the standard form of the equation of an ellipse. The ellipse was defined to be the set of all points such that the sum of the distances from the two foci is constant. For an ellipse with the major axis horizontal, the vertices are located at $(h \pm a, k)$ and the foci are located at $(h \pm c, k)$. The constant sum of the distances between a point and the foci can be found easily by using either vertex. It equals $(a + c) + (a - c)$, which simplifies to 2a.

Using the distance formula and any point (x,y) on the ellipse, we get the equation $\sqrt{[x-(h-c)]^2+(y-k)^2} + \sqrt{[x-(h+c)]^2+(y-k)^2} = 2a$. After simplifying this becomes $(a^2 - c^2)(x - h)^2 + a^2(y - k)^2 = a^2(a^2 - c^2)$. To simplify further, we need to make use of an identity relating a, b, and c. Construct a right triangle by connecting one of the foci to one of the endpoints of the minor axis. Since this point is equidistant from both foci, the length of the segment drawn must equal a, since it was shown earlier that the sum of the distance to each focus from any point on the ellipse is 2a. The legs of the triangle are b and c, thus in an ellipse, $a^2 = b^2 + c^2$ or $b^2 = a^2 - c^2$ by the Pythagorean Theorem.

Substituting in the previous expression yields $b^2(x - h)^2 + a^2(y - k)^2 = a^2b^2$. Dividing this equation by a^2b^2 gives the standard form for the equation of an ellipse with a horizontal major axis: $\mathbf{\dfrac{(x-h)^2}{a^2} + \dfrac{(y-k)^2}{b^2} = 1}$.

With the major axis vertical, the equation would be $\mathbf{\dfrac{(x-h)^2}{b^2} + \dfrac{(y-k)^2}{a^2} = 1}$.

In both cases, the major axis has length 2a, and the minor axis length 2b. The foci always lie on the major axis, c units on either side of the center.

The **eccentricity** of an ellipse measures its 'ovalness' or thinness. Eccentricity is defined as $\frac{c}{a}$. Hence a small eccentricity value (close to 0) means the ellipse is close to a circle in shape, i.e. the major axis is only slightly longer than the minor axis. A large eccentricity value (almost 1) means the major axis is substantially larger than the minor axis, thus the ellipse is very long and thin. The eccentricity of an ellipse must be between 0 and 1, because the focus points are always inside the ellipse, $c < a$.

Circles

A circle is the set of all points in a plane that are equidistant from one point, called the **center**. The standard form equation for a circle is $\mathbf{(x - h)^2 + (y - k)^2 = r^2}$ where (h,k) is the center, and r is the **radius**. Worth noting here is this equation is very similar to the equation for an ellipse, i.e. if we divided this equation by r^2, it would look exactly the same, except that a and b are equal. The circle is actually a special kind of ellipse with both foci located at the same point, the center. Hence the major and minor axes are the same, and the eccentricity is zero.

Hyperbolas

A hyperbola is the set of all points such that the difference of the distances from two distinct fixed points, called **foci**, is constant. Hyperbolas consist of two branches opening opposite each other, with one focus point inside each branch.

The line connecting the foci of a hyperbola is known as the **transverse axis**. The points where the transverse axis intersects the hyperbola are called **vertices**. The midpoint of the transverse axis is the **center**, denoted by (h,k). The axis perpendicular to the transverse axis is called the **conjugate axis**.

As the branches of a hyperbola become further from the center, they approach becoming straight lines. These lines are called the **asymptotes** of the hyperbola. Every hyperbola has two asymptotes, which intersect at the center. One end of each branch approaches one end of one of the asymptotes.

Distances in a hyperbola are similar to the ellipse:

a is the distance from the center to either vertex, c is the distance from the center to either focus,

b is the half width of a rectangle drawn between the branches of the hyperbola with length 2a and the asymptotes forming diagonals.

The Pythagorean relationship in a hyperbola is $a^2 + b^2 = c^2$.

A **rectangular hyperbola** is a hyperbola where a = b.

We can also derive the standard form of the equation of a hyperbola using the distance formula. As with the ellipse, the constant difference of the distances between a point on the hyperbola and the foci can be found easily by using either vertex. It equals (c + a) - (c – a) or 2a. For any point (x,y) on the hyperbola, we get an equation similar to that of the ellipse except the plus sign before the second term is changed to a minus sign.

The standard form of the equation of a hyperbola with transverse axis horizontal simplifies to $\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1$.

With the transverse axis vertical, the equation becomes $\frac{(y-k)^2}{a^2} - \frac{(x-h)^2}{b^2} = 1$.

The **asymptotes** for a hyperbola with center at (h,k) and transverse axis horizontal can be expressed by the equations $y = k \pm \frac{b}{a}(x-h)$. The asymptotes for a hyperbola with center at (h,k) and transverse axis vertical can be expressed by the equations $y = k \pm \frac{a}{b}(x-h)$. We find the asymptote equations by taking the equation of the hyperbola, changing the 1 to a 0, and solving for y.

The **eccentricity** of a hyperbola is also defined as $\frac{c}{a}$ and measures its flatness. The eccentricity of a hyperbola must be greater than 1, because c > a. If the eccentricity is close to 1, the hyperbola is rather flat and the branches are pointed. As the eccentricity becomes greater than 1, the hyperbola branches open wider.

Parabolas

A parabola is the set of all points in a plane that are equidistant from a fixed point called the **focus**, and a fixed line called the **directrix**. The **vertex** is defined as the midpoint between the focus and the directrix. The **axis** of a parabola is the line connecting the focus and the vertex. In most cases, the axis of the parabola is parallel to either the x-axis or the y-axis.

The equation of a parabola with vertical axis, vertex at (h,k) and directrix at $y = k - p$, in standard form is $\mathbf{(x - h)^2 = 4p(y - k)}$, where p represents the distance between the focus and the vertex. With horizontal axis, vertex at (h,k) and directrix at $x = h - p$, the equation is $\mathbf{(y - k)^2 = 4p(x - h)}$.

To derive the equation of a parabola, we use the point (x,y) on the parabola and the vertex (h,k). If we consider the case where the parabola opens upward, then the focus lies on the point (h, k + p) and the directrix is $y = k - p$. Using the distance formula, the distance from the focus to the point (x,y) equals the distance from the directrix to the point. $\sqrt{(x-h)^2 + [y-(k+p)]^2} = y - (k - p)$. Squaring both sides gives $(x - h)^2 + y^2 - 2y(k + p) + (k + p)^2 = y^2 - 2y(k - p) + (k - p)^2$. Simplifying this gives the formula $(x - h)^2 = 4p(y - k)$.

The **eccentricity** of a parabola can be defined as the ratio of the distance from the focus to the distance from the directrix for any point on the parabola. The eccentricity of a parabola is always 1.

General Second Degree Equation and Rotation of Axes

When the axes of a conic section are not parallel to the x and y axes, a general equation of the form $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$ results. If the B term in the general equation is nonzero, and we must **rotate the axes** to identify the conic. Rotation of axes involves eliminating the B term by creating a new set of axes, x' and y', which are parallel to the axes of the conic and which form the x'y' plane. A new equation for the conic in the x'y' plane has new coefficients, $A'(x')^2 + C'(y')^2 + D'x' + E'y' + F' = 0$. B' will be zero if we use as an angle of rotation, θ, found by solving $\cot 2\theta = \dfrac{A-C}{B}$.

The new coefficients are then found by substituting into the following equations:

$$A' = A\cos^2\theta + B\cos\theta\sin\theta + C\sin^2\theta$$

$$C' = A\sin^2\theta - B\cos\theta\sin\theta + C\cos^2\theta$$

$$D' = D\cos\theta + E\sin\theta$$

$E' = -D \sin\theta + E \cos\theta$

$F' = F.$

After the new coefficients are found, with $B' = 0$, the conic can be identified and written in one of the standard forms described in the previous pages.

Discriminant of a General Conic

A quick way to identify a conic without rotating the axes is by evaluating the **discriminant,** defined as $B^2 - 4AC$. When $B = 0$, the type of conic depends entirely on A and C. We can easily see that if A or C (not both) are zero, the conic is a parabola, and $B^2 - 4AC$ is zero. If A and C have opposite signs, then the conic is a hyperbola, and $B^2 - 4AC$ is positive. If A and C have the same sign, then the conic is an ellipse, and $B^2 - 4AC$ is negative. If A and C are equal, then the ellipse is also a circle.

When the axes are rotated, $(B')^2 - 4A'C' = B^2 - 4AC$. Thus we can remember the rule from the simple case when there is no xy term and use the rule when there is an xy term. The graph of $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$ can be identified as follows:

$B^2 - 4AC = 0 \longrightarrow$ parabola

$B^2 - 4AC > 0 \longrightarrow$ hyperbola

$B^2 - 4AC < 0 \longrightarrow$ ellipse or circle.

Vectors

A **vector** is a line segment that has a length, called **magnitude**, and a **direction**. Every vector has an **initial point** and a **terminal point**. To compute the magnitude of a vector **v** (denoted $\| \mathbf{v} \|$) given an initial point (x_1, y_1) and a terminal point (x_2, y_2), use the distance formula. For example, a vector with an initial point (2 , 2) and a terminal point (5 , 6) has a magnitude of $\sqrt{[(5-2)^2 + (6-2)^2]} = 5$.

The **component form** of **v** is expressed by $< x_2 - x_1 , y_2 - y_1 >$. In our example, the component form would be $< 3 , 4 >$.

Two vectors with the same magnitude and same direction are said to be **equal**. A vector with initial point (1 , -3) and terminal point (4 , 1) is equal to the vector with initial point (2 , 2) and terminal point (5 , 6) because each of them has component form $<3 , 4>$.

A vector with magnitude 1 is called a **unit vector**. Every vector **v** has exactly one unit vector which points in the same direction. It is computed by dividing each component by the magnitude. $< 3 , 4 >$ has magnitude of 5, therefore its unit vector is $< \frac{3}{5} , \frac{4}{5} >$.

Vector Addition, Subtraction, and Scalar Multiplication

The sum of vectors **p** and **q**, denoted by $<p_1 , p_2>$ and $<q_1 , q_2>$, is $< p_1 + q_1 , p_2 + q_2>$. This can also be seen graphically by placing the initial point of **p** at the origin, and placing the initial point of **q** at the terminal point of **p**. The sum of vectors **p** and **q** is the vector with initial point at the origin and terminal point at the terminal point of **q**.

The difference of vectors **p** and **q** is similar to addition, but the corresponding components are subtracted. Graphically, we place the initial points of both **p** and **q** at the origin. The vector **p - q** will go from the terminal point of **q** to the terminal point of **p**. We could also draw vector **q** going in the opposite direction as **-q** and add it to vector **p**.

The scalar product of a positive real number k and a vector **v** has the same direction as the vector, and each component and the magnitude are multiplied by the real number. If k is negative, then the product has the opposite direction. Vectors which have the same direction but different magnitudes are called **parallel** vectors. If two vectors are parallel, then the ratio between any component of the first and the corresponding component of the second is constant.

Dot Product of Vectors

The **dot product** is the multiplication of two vectors which results in a scalar. It is sometimes called the **inner product**. The dot product of vectors **p** and **q**, denoted by $<p_1 , p_2>$ and $<q_1 , q_2>$, is a real number equal to $p_1q_1 + p_2q_2$. For example, the dot product of $< 3 , 4 >$ and $< -1 , 2 >$ is $(3)(-1) + (4)(2) = 5$.

Angle between Two Vectors

The dot product is useful to find the angle θ between two vectors **x** and **y**. If we graph **x** and **y** with initial points at the origin, we can construct a triangle with one side from the origin to the endpoints of each vector. The third side connecting the two endpoints will be the difference **x** - **y**. By applying the Law of Cosines to this triangle, we obtain $\|\mathbf{x}-\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\,\|\mathbf{x}\|\,\|\mathbf{y}\|\cos\theta$. Since $\|\mathbf{x}-\mathbf{y}\|^2 = (x_1 - y_1)^2 + (x_2 - y_2)^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2(x_1y_1 + x_2y_2)$, we can simplify this to $x_1y_1 + x_2y_2 = \|\mathbf{x}\|\,\|\mathbf{y}\|\cos\theta$.

This becomes the formula: $\cos\theta = \dfrac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}$.

Two vectors whose dot product is zero are perpendicular, and are called **orthogonal** vectors.

Projection of One Vector onto Another

We wish to project vector **u** onto vector **b**, written as $\mathrm{proj}_{\mathbf{b}}\mathbf{u}$. If these vectors are placed with their initial points together, by trigonometry, the cosine of the acute angle between them = $\dfrac{\|\mathrm{proj}_{\mathbf{b}}\mathbf{u}\|}{\|\mathbf{u}\|}$. But by the previous formula, the cosine of this angle = $\dfrac{\mathbf{u}\cdot\mathbf{b}}{\|\mathbf{u}\|\|\mathbf{b}\|}$. Thus the magnitude: $\|\mathrm{proj}_{\mathbf{b}}\mathbf{u}\| = \dfrac{|\mathbf{u}\cdot\mathbf{b}|}{\|\mathbf{b}\|}$. Multiplying this by a unit vector in the direction of **b**, gives the vector projection: $\left(\dfrac{\mathbf{u}\cdot\mathbf{b}}{\|\mathbf{b}\|^2}\right)(\mathbf{b})$.

Distance from a Point to a Line

From these formulas it is possible to find the distance from a point to a line in a plane. The perpendicular distance is the shortest distance from a line $ax + by + c = 0$ to a point (x_0, y_0) not on the line. First we make a vector from any point (x, y) on the line to (x_0, y_0): $\langle x_0-x, y_0-y\rangle$. Now take the length of the projection of that vector onto a vector perpendicular to the line, $n = ai + bj$, simplify, and we come up with this formula.

The magnitude of the projection, the perpendicular distance = $\dfrac{|a(x_0 - x) + b(y_0 - y)|}{\|n\|}$.

The distance from a point to a line $= \dfrac{|ax_0 + by_0 + c|}{\sqrt{(a^2 + b^2)}}$.

Distance from a Point to a Plane

The formula for the distance to the plane $ax + by + cz + d = 0$ from a point (x_0, y_0, z_0) not on the plane is derived in a similar way.

$$\text{The distance from a point to a plane} = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{(a^2 + b^2 + c^2)}}.$$

Equations of Lines in Space

A line in space can be determined uniquely by specifying a point on the line and a direction vector parallel to it. Given a point (x_0, y_0) on a line, and direction vector **v** = <a,b> , the line in two-space has parametric equations $x = x_0 + at$, and $y = y_0 + bt$. Of course, in two-space, we could eliminate the parameter and have the familiar equations at the beginning of this section.

Given a point (x_0, y_0, z_0) in a line, and direction vector **v** = <a,b,c>, the line in three-space has parametric equations $x = x_0 + at$, $y = y_0 + bt$, and $z = z_0 + ct$. If we eliminate the parameter in three-space, we get a three-part symmetric equation

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}.$$

Equations of Planes

An equation for a plane can be written given a point (x_0, y_0, z_0) on the plane and a vector <a,b,c> which is normal or perpendicular to the plane. If (x, y, z) is any other point in the plane, the vector $\langle x - x_0, y - y_0, z - z_0 \rangle$ is orthogonal to vector <a,b,c> and so the dot product will equal zero. This gives the equation for a plane as:

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0,$$

which simplifies to: $\qquad ax + by + cz + d = 0.$

Arithmetic Series & Geometric Series

Arithmetic Sequence

An **Arithmetic Sequence** is a sequence of numbers, which has the condition that the difference between any two consecutive terms is constant. The first term of an arithmetic sequence is generally labeled a or a_1 and the common difference called d. The ith term in the sequence is given the label a_i.

Finding Terms of an Arithmetic Sequence

Given any term in the sequence, and the common difference, it is possible to find any other term in the sequence. For example, if a_1 is 8, and $d = -3$, we can find a_7 by adding 7-1=6 multiples of −3, obtaining an answer of −10. If we are told the ith and jth terms in the sequence, we can find any other term. First find the common difference, $\mathbf{d = \frac{a_j - a_i}{j - i}}$. Then any other term in the sequence can be found by adding or subtracting multiples of d. For example, if a_5 is 13, and a_8 is 22, $d = \frac{22-13}{8-5} = 3$. Any other term can be found by adding or subtracting multiples of d.

Sum of an Arithmetic Sequence

It is also possible to compute the sum of the first n terms of an arithmetic sequence. For a sequence of length n, pairs of terms which have a common sum can be grouped together to make the additions easier. $a_1 + a_n$, $a_2 + a_{n-1}$, $a_3 + a_{n-2}$, and so on, all have sums of $a_1 + a_n$ or $2a_1 + (n-1)d$. There are $\frac{n}{2}$ such pairs. Hence the sum of all terms in the sequence is $\mathbf{\frac{n}{2}(a_1 + a_n)}$ or $\mathbf{\frac{n}{2}(2a_1 + (n-1)d)}$.

To compute the sum of a group of terms which does not include the first, simply treat the first term in that group as though it were the first in the sequence.

Geometric Sequence

A **Geometric Sequence** is a sequence of numbers, which has the condition that the ratio between any two consecutive terms is constant. The first term of a geometric sequence is generally labeled a or a_1 and the common ratio called r.

Finding Terms of a Geometric Sequence

Given any term in the sequence, and the common ratio, it is possible to find any other term in the sequence. For example, if a_1 is 5, and the common ratio is 2, we can find a_7 by multiplying 5 by 7-1=6 factors of 2, obtaining an answer of 320. If we are told the ith and jth terms in the sequence, we can find any other term, as well as the common ratio, $\mathbf{r} = \sqrt[(j-i)]{\frac{\mathbf{a}_j}{\mathbf{a}_i}}$. For example, if $a_3 = 13$, and $a_7 = 117$, the common ratio $= \sqrt[(7-3)]{\frac{117}{13}} = \sqrt[4]{9} = \sqrt{3}$. Any other term can be found by multiplying or dividing by factors of $\sqrt{3}$.

Sum of a Geometric Sequence

It is also possible to compute the sum of the first n terms of a geometric sequence, or all terms of an infinite sequence with $| r | < 1$.

For an infinite sequence, represent the sum S as $S = a + ar + ar^2 + ar^3 + \ldots$. Multiplying both sides of this equation by r gives us $rS = ar + ar^2 + ar^3 + ar^4 + \ldots$. Subtracting the 2nd equation from the first gives us $S - rS = a$ or $\mathbf{S} = \frac{\mathbf{a}}{\mathbf{1-r}}$. This is the formula for the sum of an infinite geometric series.

For a finite sequence of length n, the only difference is that when the 2nd equation above is subtracted from the first, the result will be $S - rS = a - ar^{n+1}$, because here the last term does not cancel out. Hence this formula is $\mathbf{S} = \frac{\mathbf{a - ar^{n+1}}}{\mathbf{1-r}}$.

To compute the sum of a group of terms which does not include the first, simply treat the first term in that group as though it were the first in the sequence.

Binomial Theorem & Pascal's Triangle

Patterns in Powers of Binomials

Suppose we want to raise a binomial of the form $(x + y)$ to a large power. This might appear to be extremely difficult or impossible without a computer. First we can raise the binomial to smaller powers and try to observe a pattern.

$(x + y)^0 = 1$

$(x + y)^1 = x + y$

$(x + y)^2 = x^2 + 2xy + y^2$

$(x + y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$

$(x + y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$

$(x + y)^5 = x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5$

and so on .. a few observations about these expressions. The terms of each expression of $(x + y)^n$ are always arranged in order such that the power of x always decreases, and the power of y always increases, and their sum equals n. In each, there are always n + 1 terms, and the coefficients of the terms were symmetric about the middle. Furthermore, the sum of the coefficients in the expression of $(x + y)^n$ equals 2^n. The coefficients of the first term and last term of each expression are always 1, while the coefficients of the second term and second to the last term of $(x + y)^n$ always equal n.

Binomial Theorem

However the rest of the coefficients in between are not nearly as obvious. To compute them we use the **Binomial Theorem**, which states that the mth term of the expansion of $(x + y)^n$ is ${}_nC_m\, x^{n-m}\, y^m$, where the coefficient ${}_nC_m$ is equal to $\dfrac{n!}{(n-m)!\,m!}$.

There will be n+1 terms with m going from zero to n.

Pascal's Triangle

Another feature of the expressions given above is that the sum of the coefficients ${}_nC_m$ and ${}_nC_{m+1}$ always equals the coefficient ${}_{n+1}C_{m+1}$. This feature is easy to see in **Pascal's Triangle** which is an infinite isosceles triangle of numbers where each row contains the coefficients in the expansion of $(x + y)^n$.

$$\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 1 & & 1 & & & & & \\
 & & & & 1 & & 2 & & 1 & & & & \\
 & & & 1 & & 3 & & 3 & & 1 & & & \\
 & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
 & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1
\end{array}$$

The first row (technically row 0) contains the number 1, row 1 contains 1 1, row 2 contains 1 2 1, row 3 contains 1 3 3 1, row 4 contains 1 4 6 4 1, row 5 contains 1 5 10 10 5 1, etc. Each element of each row can be found either from the binomial theorem or by adding the 2 elements directly above it.

The coefficients of a binomial expansion can be found by reading across the coefficients of a row in Pascal's Triangle. For example, the binomial $(x + 3)^5 = x^5 + 5x^4(3)^1 + 10x^3(3)^2 + 10x^2(3)^3 + 5x^1(3)^4 + (3)^5 = x^5 + 15x^4 + 90x^3 + 270x^2 + 405x + 243$. It is also possible to search for only one or more terms from a row of the binomial expansion. For example, to find the coefficient of x^2y^2 in the expansion of $(3x + 2y)^4$, we would obtain the coefficient ${}_4C_2$ from Pascal's Triangle, which is 6. Multiplying $6(3x)^2(2y)^2$ gives us $216x^2y^2$.

Combinations and Permutations

Another use of the term ${}_nC_m$ is to represent the number of possible **combinations** of m items chosen from a group of n items, or subsets of n of length m. We just stated that ${}_4C_2 = 6$, meaning that there are 6 ways to select 2 items from a group of 4. If we have 4 items numbered 1 through 4, the sets of 2 are {1,2} , {1,3} , {1,4} , {2,3} , {2,4} , {3,4}.

The number of **permutations** of m items chosen from a group of n items, denoted ${}_nP_m$, is the number of ordered subsets of n of length m. To calculate the number of permutations, consider selecting m items, one at a time, from a group of n items. There are n possible choices for the first item. For the second item, there are only n - 1 choices because one item has already been used. For the third item, there are only n - 2 choices because two items have already been used, and so on. For the m^{th} item, there are $n - (m - 1) = n - m + 1$ choices, since m - 1 items have been used. Hence, the total number of permutations ${}_nP_m$ can be expressed as a product $(n)(n - 1)(n - 2) \ldots (n - m + 1)$, which can be simplified to $\frac{n!}{(n-m)!}$. For example, the number of permutations of 2 items chosen from 4, denoted by ${}_4P_2$, is $\frac{4!}{(4-2)!} = \frac{24}{2} = 12$.

A formula for **combinations** ${}_nC_m$, as mentioned earlier, is $\frac{n!}{(n-m)!\,m!}$. This is the same as the formula for ${}_nP_m$ divided by m!. The reason for this is because there are m! ways to order m items. Since the order does not matter for ${}_nC_m$, there will always be m! permutations for every one combination. In our example, there are twice as many permutations as combinations, because each combination listed can be reversed, thereby producing a different permutation.

Probability

Probability

The probability of an event occurring is the chance or likelihood that it will happen. In mathematical terms, **probability can be defined as the total number of favorable outcomes divided by the total number of possible outcomes**.

As a very simple example, when tossing a coin, the probability of the coin coming up heads is 1/2. This is obtained by dividing the total number of favorable outcomes {heads} = 1 by the total number of possible outcomes {heads, tails} = 2.

The probability of an event A occurring is denoted by P(A). The highest possible probability for any event to occur is 1, that happens when all of the possible outcomes are favorable. In other words, the event is guaranteed to occur. The lowest possible probability for an event to occur is 0, when none of the possible outcomes are favorable. In other words, the event is impossible. In all other cases, the probability of an event is a real number between 0 and 1.

If the probability of an event occurring is p, then **the probability of the event not occurring**, or in other words the **complement** of p, is 1-p.

Independent, Dependent, and Mutually Exclusive Events

Two events A and B are said to be **independent** if the outcome of event A has no effect at all on the outcome of event B, and vice-versa. For example, tossing a coin twice represents two independent events because the outcome of the first toss has no effect on the outcome of the second toss.

If the outcome of an event A has an effect on the outcome of a future event B, we say that event B is **dependent** on event A. An example of this is drawing 2 cards from a standard deck of playing cards, without replacement. The probability that the first card drawn is a king is $\frac{4}{52} = \frac{1}{13}$, because there are 4 possible favorable outcomes (kings in the deck) and 52 total possible outcomes (total cards in the deck). The probability that the second card drawn is also a king depends on the first outcome. In either case, the total number of possible outcomes is 51, since there are now 51 cards in the deck. If the first card drawn was a king, then there are 3 kings left, hence the probability of a second king is $\frac{3}{51}$. However if the first card drawn was not a king, then there are 4 kings left, hence the probability of getting a king would be $\frac{4}{51}$.

Two events are said to be **mutually exclusive** if at most one of them can occur. For example, if A represents drawing a king from a deck, and B represents drawing a queen from a deck, then A and B are mutually exclusive because it is not possible for a card to be both a king and a queen.

Probability of Event A OR Event B

For two independent events A and B, the probability of either A or B occurring, i.e. **P(A or B) = P(A) + P(B) – P(A and B).** Note here that if A and B are mutually exclusive then P(A and B) = 0, so P(A or B) becomes simply P(A) + P(B). The reason for this can be seen from the previous example, P(drawing a king) = $\frac{4}{52}$, and P(drawing a queen) = $\frac{4}{52}$, so P(drawing a king or queen) = $\frac{4}{52} + \frac{4}{52} = \frac{8}{52}$.

Probability of Event A AND Event B

For two independent events A and B, the probability of both A and B occurring, i.e. **P(A and B) = P(A) * P(B).** For example, tossing two fair coins, the possible outcomes are {HH, HT, TH, TT}. The probability of getting a head on any one toss is $\frac{1}{2}$, so the probability of doing it twice in a row, using this formula, is $\frac{1}{2} * \frac{1}{2} = \frac{1}{4}$. This agrees with the result obtained directly by using the set of possible combined outcomes of the two rolls.

For two dependent events A and B, the probability of both A and B occurring, i.e. **P(A and B) = P(A) * P(B given A).** Using the example of drawing 2 cards from a standard deck of playing cards, without replacement, the probability that both cards will be kings, by this formula, is $\frac{4}{52} * \frac{3}{51} = \frac{1}{221}$.

Independent Binomial Events

The Binomial Theorem and Pascal's Triangle also are useful in determining probabilities. The probability of an event A with P(A) = 1/2 occurring exactly m times out of a possible n independent trials can be read directly from Pascal's Triangle. It is ${}_nC_m$ (the [m + 1]th term in row n) divided by 2^n.

If P(A) = p, but p is not necessarily 1/2, the probability of event A occurring exactly m times out of a possible n independent trials is ${}_nC_m\, p^m\, (1-p)^{n-m}$.

Divisibility Rules and Prime Numbers

Divisibility

An integer is **divisible** by another integer if when divided the remainder is zero. When the divisor is very large, it is often possible to determine only by a calculator or computer, but with relatively small numbers there are easier methods.

Divisibility by Powers of Two

Suppose we want to test if a number is divisible by 2. Since 10 is a multiple of 2, we need only to look at the last digit. If the last digit is divisible by 2 (0,2,4,6,8) then the number is divisible by 2, otherwise it is not. The rest of the digits are not important because we can add or subtract any number of multiples of 10 without affecting its divisibility by 2.

To test for divisibility by 4 is similar to testing for divisibility by 2, except here we must consider the last 2 digits. Since 100 is divisible by 4, we can add or subtract any number of multiples of 100 from a number without affecting its divisibility by 4. Hence we can conclude that a number is divisible by 4 if and only if its last 2 digits are divisible by 4.

Using similar reasoning, we can come up with divisibility rules for any power of 2. The smallest power of 10 which is evenly divisible by 2^n is 10^n. Hence, when testing for divisibility by 2^n, any number of multiples of 10^n can be added or subtracted without affecting divisibility by 2^n. This means we must test only the remainder when the given number is divided by 10^n. Therefore, **a number is divisible by 2^n if and only if its last n digits are divisible by 2^n**.

Divisibility by Powers of Five

Five is also a factor of 10, hence the same idea used for divisibility by powers of 2 will work for powers of 5. Since 10 is a multiple of 5, we again need to look at only the last digit. If the last digit is divisible by 5, i.e. 0 or 5, then the number is divisible by 5, otherwise it is not. The rest of the digits do not matter because any number of multiples of 10 can be added or subtracted without affecting its divisibility by 5.

For any power of 5, the smallest power of 10 which is evenly divisible by 5^n is 10^n. Hence, when testing for divisibility by 5^n, any number of multiples of 10^n can be added or subtracted without affecting divisibility by 5^n. Therefore, **a number is divisible by 5^n if and only if its last n digits are divisible by 5^n**.

Unfortunately, for values of n much greater than 2 it will be difficult to test the last n digits for divisibility by 5^n, so this method is not terribly useful except for very small values of n.

Divisibility by Three and by Nine

As would be expected, the tests for divisibility by numbers other than powers of 2 and powers of 5 are quite different, because 2 and 5 are the only prime factors of 10.

Testing divisibility by 3, we consider the sum of the digits of the number. This works because whenever 3 is added to a number, one of two things must happen.

A) If the units digit is less than 7 it will increase by 3, thereby increasing the sum of the digits by 3, which does not affect whether the sum of the digits is divisible by 3.

B) If there is a carry in the units digit, the units digit will decrease by 7, and the tens digit will increase by 1, thereby decreasing the sum of the digits by 6, which does not affect whether the sum of the digits is divisible by 3. It is also possible that a carry will result in more than just the units digit. But each time any digit carries as a result of adding 3, that digit is decreased by 7, and the next highest digit increases by 1. Divisibility by 3 remains unaffected.

Hence we conclude that **a number is divisible by 3 if and only if the sum of its digits is divisible by 3**.

A similar argument can be made for testing divisibility by 9. Whenever 9 is added to a number, either

A) If the units digit is 0, it will become 9, thereby increasing the sum of the digits by 9, which does not affect whether the sum of the digits is divisible by 9.

B) If the units digit is not zero, a carry will result when 9 is added. The units digit will decrease by 1, and the tens digit will increase by 1, thereby leaving the sum of the digits unchanged. Just as in the previous case, it is possible that a carry will result in more than just the units digit. But each time any digit carries as a result of adding 9, that digit is decreased by 1, and the next highest digit increases by 1, thereby leaving the sum of the digits unchanged.

Hence we conclude that **a number is divisible by 9 if and only if the sum of its digits is divisible by 9**.

Divisibility by Larger Numbers

Divisibility tests for prime numbers greater than 5 exist, but are more complicated. These rules are all of the following form: A number is divisible by x if and only if x_1 times the units digit plus x_2 times the tens digit plus x_3 times the hundreds digit plus ..... is divisible by x. The x constants are determined by taking remainders when successive powers of 10 are divided by the number.

For example, let us derive the **formula for divisibility by 7**.

10^0 divided by 7 gives a remainder of 1,
10^1 divided by 7 gives a remainder of 3,
10^2 divided by 7 gives a remainder of 2,
10^3 divided by 7 gives a remainder of 6,
10^4 divided by 7 gives a remainder of 4,
10^5 divided by 7 gives a remainder of 5,
10^6 divided by 7 gives a remainder of 1.

Once a remainder has repeated, the calculations can stop because the remainders will all repeat. There are 6 remainders in this sequence. This shows that a 6 digit number is divisible by 7 if and only if 1 times the 10^0 digit plus 3 times the 10^1 digit plus 2 times the 10^2 digit plus 6 times the 10^3 digit plus 4 times the 10^4 digit plus 5 times the 10^5 digit is divisible by 7.

For numbers longer than 6 digits, repeat the pattern.

A number is **divisible by 7** if and only if

1 times all of the 10^{6k} digits
plus 3 times the 10^{6k+1} digits
plus 2 times the 10^{6k+2} digits

plus 6 times (or minus 1 times) the 10^{6k+3} digits

plus 4 times (or minus 3 times) the 10^{6k+4} digits

plus 5 times (or minus 2 times) the 10^{6k+5} digits

is divisible by 7, where k = any nonnegative integer.

It is possible to add or subtract any multiple of 7 from any of the coefficients above, in order to facilitate calculations. It would probably be easier to remember if the coefficients of the 6k + 3, 6k + 4, and 6k + 5 digits all were lessened by 7, becoming −1, -3, and −2 respectively, as 1,3,2,-1,-3,-2 is an easier sequence to remember than is 1,3,2,6,4,5.

Using similar methods, divisibility formulas for many larger numbers can be derived. Most are too long to mention in this paper, but here are a few of the simpler ones:

Divisibility by 11 if and only if

1 times 10^{2k} digits

minus 1 times 10^{2k+1} digits is divisible by 11.

Divisibility by 13 if and only if

1 times 10^{6k} digits

plus 10 times 10^{6k+1} digits

plus 9 times 10^{6k+2} digits

minus 1 times 10^{6k+3} digits

minus 10 times 10^{6k+4} digits

minus 9 times 10^{6k+5} digits is divisible by 13.

Divisibility by 27 if and only if

1 times 10^{3k} digits

plus 10 times 10^{3k+1} digits

plus 19 times 10^{3k+2} digits is divisible by 27.

Divisibility by 37 if and only if

1 times 10^{3k} digits

plus 10 times 10^{3k+1} digits

plus 26 times 10^{3k+2} digits is divisible by 37.

Divisibility by Composite Numbers

All of the above rules either test for divisibility by primes, or divisibility by powers of primes. These rules can be combined to test for divisibility by composite numbers which are the product of more than one distinct prime factor. For example, divisibility by both 2 and 3 implies divisibility by 6, and vice-versa. In fact, divisibility by x and divisibility by y is a sufficient test for divisibility by (xy) if x and y are relatively prime. (see next section)

Prime and Composite Numbers

A **prime number** is any positive integer that is not evenly divisible by any positive integer other than 1 or itself. For example, 7 is a prime number, but 6 is not prime, because it is divisible by both 2 and 3. A number that is not prime is said to be **composite**. One is technically considered neither prime nor composite, but every number greater than 2 is either prime or composite. The smallest prime number is 2, the smallest composite number is 4.

The Fundamental Theorem of Arithmetic states that all composite numbers can be expressed as the product of prime numbers in a unique way, this is known as its **prime factorization**. All positive numbers have a unique square root. Whenever a number can be expressed as the product of exactly two factors, one of these factors must be greater than or equal to the square root, the other factor must be less than or equal to the square root. Hence, to check if a number is prime, it is necessary to test whether the number is divisible by all numbers up to the largest integer less than or equal to its square root. If no factors are found, then the number is prime.

To determine the prime factorization of a number, first find its smallest prime factor. Then divide the number by the prime factor, and find the smallest prime factor of the quotient, and divide again. Repeat this process until the quotient is 1, then all of the prime factors have been found and the factorization is complete. For an example of this, we shall find the prime factorization of 60. The smallest prime factor of 60 is 2, divide 60/2 = 30. The smallest prime factor of 30 is 2, divide 30/2 = 15. The smallest prime factor of 15 is 3, divide 15/3 = 5. The smallest prime factor of 5 is 5, divide 5/5 = 1. The prime factors of 60 are 2,2,3,5. This method works for any number, regardless how large. The prime factorization of a prime number consists of only one factor, the number itself.

Two or more numbers are **relatively prime** if the greatest common factor among them is 1, i.e. they have no common prime factors. To test if all numbers in a set are relatively prime, it is necessary to first find the prime factorization of each number. If there are any prime factors common to all numbers in the set, then the product of those numbers is the greatest common factor, and the numbers are not relatively prime. If there are no prime factors common to all numbers in the set, then the numbers are relatively prime. For example, consider the numbers 28 and 35. The prime factors of 28 are 2,2, and 7. The prime factors of 35 are 5 and 7. Since 7 is common to both, 28 and 35 are not relatively prime. The numbers 28 and 45 are relatively prime, because the prime factors of 45 are 3,3, and 5, none common with the prime factors of 28.

One well-known fact about prime numbers is that **there is no largest prime number**, i.e. there are infinitely many of them. This is actually very easy to prove. Start by assuming that a number x is the largest prime number. Consider the product of all prime numbers less than x. Adding 1 to this number will either 1) produce a new prime number, or 2) be a composite number having only prime factors greater than x. (It cannot be divisible by any prime less than x since 1 was added to the product.)

Number Bases

Base Ten Numbers

The number system we use is called base 10. There are 10 digits, from 0 through 9. A number is made up of a set of digits, each to be multiplied by a power of 10 and summed together. For example, the number $1234 = 1(10^3) + 2(10^2) + 3(10^1) + 4(10^0)$.

Other Numbers as Bases

Other number bases exist as well, and they function the same way except the digits are multiplied by powers of a number other than 10. It is impossible to define a number system with only one digit, as then the only existing number would be zero. However all integer number base greater than one are defined. Each base has a different set of characters to represent digits. Number bases less than or equal to 10 use digits from 0 through one less than the base number. Bases between 11 and 36 use digits 0 through 9, followed by letters of the alphabet starting with A, i.e. A=10, B=11, and so forth. Numbers in bases other than base 10 are indicated with a subscript, i.e. 101_2 represents a number in base 2, or binary.

Converting from One Number Base to Another

Numbers can be converted from one base into another. Usually the easiest way to convert a number from base x to base y is to first convert from base x to base 10, then convert the base 10 number to base y.

To convert from another base to base 10, simply multiply each digit by the base raised to the exponent of the place it's in, and sum them together in base 10. For example, if we want to convert the number 123_8 to base 10, we compute $3(8^0) + 2(8^1) + 1(8^2) = 83$. If the base from which we are converting is greater than 10, then we simply substitute the appropriate numbers for the letters. For example, converting the number $A35D_{16}$ to base 10, we substitute 10 for A, and 13 for D, obtaining $13(16)^0 + 5(16^1) + 3(16)^2 + 10(16)^3 = 41821$.

Converting a number k in base 10 to another number base (base y) can be done in two different ways.

The immediately most obvious way would be to find the greatest power of y less than k, divide k by that power. The quotient becomes the leftmost digit, and the remainder gets divided by y raised to the one less power. Repeat the process then until the remainder becomes zero, then all further digits are zero. For example, converting the number 83_{10} to base 8. The largest power of 8 less than 83 is $8^2 = 64$. $\frac{83}{64}$ equals 1, with remainder 19. The leftmost digit will be 1, and 19 is divided by 8^1. The quotient of $\frac{19}{8}$ is 2, with remainder 3. The next digit will be 2, and 3 is divided by 8^0. The quotient of $\frac{3}{1}$ is 3, with remainder 0. Hence we are finished, the answer is $83_{10} = 123_8$.

While this method works well enough, there is a shortcut method. Start with the original number to be converted, 83, and divide by the base to which we are converting. The remainder becomes the rightmost digit of the new number. Divide the quotient by the base. The remainder becomes the next rightmost digit of the new number. Repeat this process until the quotient is zero. Doing the same conversion by this method, $\frac{83}{8}$ has a remainder of 3 and a quotient of 10. The rightmost digit of the new base 8 number will be 3, while 10 is divided by 8. $\frac{10}{8}$ has a remainder of 2 and a quotient of 1. The next rightmost digit of the new base 8 number will be 2, while 1 is divided by 8. $\frac{1}{8}$ has a remainder of 1. The leftmost digit of the new base 8 number, and since the quotient is 0, we are finished. The new number produced is the solution; it is 123_8. The same results can be obtained by both methods, but practically this second method is quicker.

Fast Conversions without Going through Base 10

Larger Base is Power of Smaller Base

There is one instance when there is a much faster way to convert between 2 number bases, x and y, than going through base 10. If x is an integer power of y (or vice-versa) then the direct conversion can be done very simply. Let z be the base y logarithm of x (i.e. the power y must be raised to in order to get x). Each group of z digits of y corresponds to one digit of x, and can be converted independently from the other digits in the number.

Larger Base to Smaller Base

For example, suppose we want to convert the number 571_8 to base 2. $8 = 2^3$, therefore each digit in the base 8 number corresponds to 3 digits in the base 2 number. Convert each digit into 3 base 2 digits, starting with the rightmost. $1_8 = 001_2$. These become the rightmost digits of the new base 2 number. $7_8 = 111_2$. These become the next rightmost digits of the new base 2 number. $5_8 = 101_2$. These become the next rightmost digits of the new base 2 number, and since there are no more digits left, the conversion is done. Hence the complete answer is $571_8 = 101111001_2$.

Smaller Base to Larger Base

Converting from the smaller base to the larger base is equally as simple. For example, we can convert 111010011_2 from to base 16. Since $2^4 = 16$, each group of 4 digits of the base 2 number will become 1 digit of the new base 16 number. Always start from the right side. The first 4 digits, $0011_2 = 3_{16}$. 3 becomes the rightmost digit of the base 16 number. The next 4 digits, $1101_2 = D_{16}$ (D=13) become the next rightmost digit of the base 16 number. Since there are fewer than 4 digits remaining in the base 2 number, assume zeroes for the others. $0001_2 = 1_{16}$. So 1 becomes the next rightmost digit of the base 16 number. Since there are no more digits, the problem is done and the final answer is $111010011_2 = 1D3_{16}$.

Larger and Smaller Bases Are Both Powers of the Same Integer

If x and y are both integral powers of a lesser number w, then the problem can be solved by first converting from base x to base w, then from base w to base y. For example, 8 and 16 are both powers of 2. To convert a base 8 number to base 16, first convert it from base 8 to base 2, then from base 2 to base 16.

Arithmetic in Other Number Bases

Addition and multiplication in other number bases is done the same way as in base 10. For example, suppose we want to add the numbers 123_8 and 456_8. In base 8 there doesn't exist a ones place, a tens place, and a hundreds place, instead there exists a 8^0 place, 8^1 place, 8^2 place, etc. We start by adding the rightmost digit, which is the 8^0 place. $3 + 6 = 9$ in base 10, but the digit 9 does not exist in base 8. $3_8 + 6_8 = 11_8$. Following the rules of addition, the number in the 8^0 place gets written, and the number in the 8^1 place gets carried. We then get the addition $1 + 2 + 5$ in the 8^1 place. $1_8 + 2_8 + 5_8 = 10_8$. The 0 gets written in the 8^1 place and the 1 gets carried over to the 8^2 place. $1 + 1 + 4 = 6$, so the answer to the problem is $123_8 + 456_8 = 601_8$.

Multiplication is also done in the same way. All of the partial products must be done in the appropriate base, and then added together for the final product in that same base.

Index

Bibliography

[1] Anton, Howard. Calculus with Analytic Geometry. Fourth Ed. New York: John Wiley & Sons, 1992.

[2] Byrd, Robert C. "A Failure to Produce Better Students." Congressional Record: S5393. Online. http://www.eurekanet.com/~brendonnell/byrdmath.html Jun. 9, 1997.

[3] Charles, Randall I., and Edward A. Silver. The Teaching and Assessing of Mathematical Problem Solving. Reston, VA: The National Council of Teachers of Mathematics, 1989.

[4] ETS Research. "Student Math Skills Improving Nationwide." Online. http://www.ets.org/research/pic/pr/reach.html May 25, 1995.

[5] Geary, David C., and Carmen O. Hamson. "Inproving the Mathematics and Science Achievement of American Children: Psychology's Role." Online. http://www.apa.org/ed/geary.html Jul. 30, 1998.

[6] James, Glenn, and Robert C. James, Ed. Mathematics Dictionary. Princeton, NJ: D. VanNostrand Company, 1976.

[7] Jennings, Marianne M. "Why Our Kids Can't Do Math." Reader's Digest Nov. 1997: 181-182. Reprinted from The Wall Street Journal Dec. 17. 1996.

[8] Larson, Roland E., and Robert P. Hostetler. Precalculus. Second Ed. Lexington, MA: D. C. Heath, 1989.

[9] Leo, John. "Hey, We're No. 19!" U.S. News & World Report. Mar. 9, 1998: 14.

[10] ---. "That So-Called Pythagoras." U.S. News & World Report. May 26, 1997: 14.

[11] Lester, Frank K. and Joe Garofalo, Ed. Mathematical Problem Solving Issues In Research. Philadelphia, PA: The Franklin Institute Press, 1982.

[12] Mathematical Sciences Education Board. Reshaping School Mathematics. Washington, D.C.: National Academy Press, 1990.

[13] Ratnesar, Romesh. "This Is Math?" Time. Aug. 25, 1997: 66-67.

[14] Schwartz, Martha. "Mathematically Correct." Testimony to the United States House of Representatives Committee of Education and the Workforce, Subcommittee on Early Childhood, Youth and Families. Online. http://mathematicallycorrect.com/schwartz.htm Jun. 11, 1998.

[15] Shumway, Richard J. Research in Mathematics Education. Reston, VA: The National Council of Teachers of Mathematics, 1980.

[16] U.S. National Research Center. "Summary of Findings, TIMSS United States." Online. http://ustimss.msu.edu/summary.htm Oct. 15, 1996.

[17] ---. "The Third International Mathematics and Science Study (TIMSS)." Online. http://ustimss.msu.edu/info.html Jul. 29, 1998.

[18] Westbury, Ian, et al. In Search of More Effective Mathematics Education. Norwood, NJ: Ablex Publishing Corp., 1994.