


Review Notes for IB Standard Level Math

© 2015-2016, Steve Muench
steve.muench@gmail.com
 @stevemuench

These notes are free of charge. If you paid to obtain them,
please send me an email to let me know about it.

Feel free to share the link to these notes
<http://bit.ly/ib-sl-maths-review-notes>
or my worked solutions to the November 2014 exam
<http://bit.ly/ib-sl-maths-nov-2014>
or my worked solutions to the May 2015 (Timezone 2) exam
<http://bit.ly/ib-sl-maths-may-2015-tz2>
or my worked solutions to the November 2015 exam
<https://bit.ly/ib-sl-maths-nov-2015>
with any student you believe might benefit from them.

If you downloaded these notes from a source other than
the bit.ly link above, please check there to make sure
you are reading the latest version. It may contain
additional content and important corrections!

November 4, 2016

Note: Additional IB Maths revision resources available at
<http://bit.ly/ib-maths-studynova>

Contents

1	Algebra	8
1.1	Rules of Basic Operations	8
1.2	Rules of Roots	8
1.3	Rules of Exponents	8
1.4	Allowed and Disallowed Calculator Functions During the Exam	8
1.5	Sequences and Series	9
1.6	Arithmetic Sequences and Series	9
1.7	Sum of Finite Arithmetic Series $(u_1 + \cdots + u_n)$	9
1.8	Partial Sum of Finite Arithmetic Series $(u_j + \cdots + u_n)$	10
1.9	Geometric Sequences and Series	10
1.10	Sum of Finite Geometric Series	11
1.11	Sum of Infinite Geometric Series	11
1.11.1	Example Involving Sum of Infinite Geometric Series	11
1.12	Sigma Notation	12
1.12.1	Sigma Notation for Arithmetic Series	13
1.12.2	Sigma Notation for Geometric Series	13
1.12.3	Sigma Notation for Infinite Geometric Series	14
1.12.4	Defining Functions Using Sigma Notation	14
1.13	Applications: Compound Interest	15
1.14	Applications: Population Growth	15
1.15	Logarithms	15
1.16	Using Logarithms to Solve Equations	16
1.17	Using Exponentiation to Solve Equations	17
1.18	Logarithm Facts Involving 0 and 1	17
1.19	Laws of Exponents and Logarithms	17
1.20	Change of Base	18
1.21	Powers of Binomials and Pascal's Triangle	18
1.22	Expansion of $(a + b)^n$	19
1.23	The Binomial Theorem	20
1.23.1	Using The Binomial Theorem for a Single Term	21
1.23.2	Example of Using Binomial Theorem	22
1.24	Solving Systems of Three Linear Equations Using Substitution	23
1.25	Solving Systems of Three Linear Equations Using Technology	24
2	Functions and Equations	25
2.1	Sets	25
2.2	Union and Intersection of Sets	25
2.3	Common Sets of Numbers	26
2.4	Intervals of Real Numbers	26
2.5	Concept of Function	27
2.6	Graph of a Function	27
2.7	Domain of a Function	28
2.8	Range of a Function	28
2.9	Composing One Function with Another	29
2.10	Identity Function	30
2.11	Inverse Function	30
2.12	Determining the Inverse Function as Reflection in Line $y = x$	30
2.13	Determining the Inverse Function Analytically	31
2.14	Drawing and Analyzing Graphs with Your Calculator	32
2.14.1	Drawing the Graph of a Function	32

2.14.2	Restricting the Domain of a Graph	32
2.14.3	Zooming Graph to See Exactly What You Want	33
2.14.4	Finding a Maximum Value in an Interval	33
2.14.5	Finding a Minimum Value Value in an Interval	34
2.14.6	Finding the x -Intercepts or “Zeros” of a Graph in an Interval	34
2.14.7	Finding the y -Intercept of a Graph	35
2.14.8	Vertical Asymptotes	35
2.14.9	Graphing Vertical Lines	36
2.14.10	Horizontal Asymptotes	36
2.14.11	Tips to Compute Horizontal Asymptotes of Rational Functions	37
2.14.12	Graphing Horizontal Lines	37
2.14.13	Symmetry: Odd Functions	37
2.14.14	Symmetry: Even Functions	38
2.14.15	Solving Equations Graphically	38
2.15	Transformations of Graphs	40
2.15.1	Horizontal and Vertical Translations	40
2.15.2	Vertical Reflection	41
2.15.3	Horizontal Reflection	42
2.15.4	Vertical Stretch	42
2.15.5	Horizontal Stretch	42
2.15.6	Order Matters When Doing Multiple Transformations in Sequence	43
2.15.7	Graphing the Result of a Sequence of Transformations	44
2.15.8	Determining Point Movement Under a Sequence of Transformations	45
2.15.9	Vector Notation for Function Translation	47
2.16	Quadratic Functions	48
2.16.1	Using the Quadratic Formula to Find Zeros of Quadratic Function	49
2.16.2	Finding the Vertex If You Know the Zeros	49
2.16.3	Graph and Axis of Symmetry	50
2.16.4	Computing the Vertex From the Coefficients	50
2.16.5	Using the Discriminant to Find the Number of Zeros	51
2.16.6	Y-Intercept Form	52
2.16.7	X-Intercept Form	53
2.16.8	Completing the Square to Get Binomial Squared Form	53
2.16.9	Vertex (h, k) Form	54
2.17	Reciprocal Functions	55
2.18	Rational Functions	56
2.19	Exponential Functions	56
2.20	Continuously Compounded Interest	57
2.21	Continuous Growth and Decay	57
2.22	Logarithmic Functions	57
3	Circular Functions and Trigonometry	59
3.1	Understanding Radians	59
3.1.1	Degrees Represent a Part of a Circular Path	59
3.1.2	Computing the Fraction of a Complete Revolution an Angle Represents	59
3.1.3	Attempting to Measure an Angle Using Distance	59
3.1.4	Arc Distance on the Unit Circle Uniquely Identifies an Angle θ	61
3.1.5	Computing the Fraction of a Complete Revolution for Angle in Radians	61
3.2	Converting Between Radians and Degrees	61
3.2.1	Converting from Degrees to Radians	61
3.2.2	Converting from Radians to Degrees	62

3.3	Length of an Arc Subtended by an Angle	62
3.4	Inscribed and Central Angles that Subtend the Same Arc	63
3.5	Area of a Sector	63
3.6	Definition of $\cos \theta$ and $\sin \theta$	64
3.7	Interpreting $\cos \theta$ and $\sin \theta$ on the Unit Circle	65
3.8	Radian Angle Measures Can Be Both Positive and Negative	65
3.9	Remembering the Exact Values of Key Angles on Unit Circle	66
3.10	The Pythagorean Identity	67
3.11	Double Angle Identities	68
3.12	Definition of $\tan \theta$	68
3.13	Using a Right Triangle to Solve Trigonometric Problems	68
3.13.1	Using Right Triangle with an Acute Angle	68
3.13.2	Using Right Triangle with an Obtuse Angle	69
3.14	Using Inverse Trigonometric Functions on Your Calculator	70
3.15	Circular Functions \sin , \cos , and \tan	71
3.15.1	The Graph of $\sin x$	71
3.15.2	The Graph of $\cos x$	71
3.15.3	The Graph of $\tan x$	72
3.15.4	Transformations of Circular Functions	72
3.15.5	Using Transformation to Highlight Additional Identities	73
3.15.6	Determining Period from Minimum and Maximum	74
3.16	Applications of the \sin Function: Tide Example	74
3.17	Applications of the \cos Function: Ferris Wheel Example	76
3.18	Solving Trigonometric Equations in a Finite Interval	78
3.19	Solving Quadratic Equations in \sin , \cos , and \tan	78
3.20	Solutions of Right Triangles	79
3.21	The Cosine Rule	80
3.22	The Sine Rule	80
3.23	Area of a Triangle	81
4	Vectors	82
4.1	Vectors as Displacements in the Plane	82
4.2	Vectors as Displacements in Three Dimensions	82
4.3	Terminology: Tip and Tail	82
4.4	Representation of Vectors	83
4.5	Magnitude of a Vector	83
4.6	Multiplication of a Vector by a Scalar	85
4.7	Negating a Vector	85
4.8	Sum of Vectors	86
4.9	Difference of Vectors	86
4.10	Unit Vectors	88
4.11	Scaling Any Vector to Produce a Parallel Unit Vector	88
4.12	Position Vectors	89
4.13	Determining Whether Vectors are Parallel	89
4.14	Finding Parallel Vector with Certain Fixed Length	90
4.15	Scalar (or “Dot”) Product of Two Vectors	90
4.16	Perpendicular Vectors	90
4.17	Base Vectors for Two Dimensions	90
4.18	Base Vectors for Three Dimensions	91
4.19	The Angle Between Two Vectors	91
4.20	Vector Equation of a Line in Two and Three Dimensions	92

4.21	Vector Equation of Line Passing Through Two Points	93
4.22	Finding the Cartesian Equation from a Vector Line	94
4.23	The Angle Between Two Vector Lines	94
4.24	Distinguishing Between Coincident and Parallel Lines	95
4.25	Finding the Point of Intersection of Two Lines	95
4.25.1	Finding Intersection Between a Line and an Axis	96
5	Statistics	98
5.1	Concepts	98
5.1.1	Population versus Sample	98
5.1.2	Discrete Data versus Continuous Data	98
5.2	Presentation of Data	98
5.2.1	Frequency Distribution Tables	98
5.2.2	Frequency Histograms with Equal Class Intervals	99
5.3	Mean	99
5.4	Median	100
5.5	Mode	101
5.6	Cumulative Frequency	101
5.6.1	Cumulative Frequency Table	101
5.6.2	Cumulative Frequency Graphs	101
5.7	Dispersion	103
5.7.1	Range	103
5.7.2	Quartiles	103
5.7.3	Interquartile Range	103
5.7.4	Box and Whiskers Plots	103
5.7.5	Percentiles	104
5.7.6	Variance	104
5.7.7	Standard Deviation	105
5.8	Use Cumulative Frequency Graph to Find Median and Quartiles	105
5.9	Computing Statistics Using Your Calculator	105
5.9.1	Calculating Statistics for a Single List	106
5.9.2	Calculating Statistics for a List with Frequency	106
5.10	Linear Correlation of Bivariate Data	107
5.10.1	Scatter Diagrams	107
5.10.2	Pearson's Product-Moment Correlation Coefficient r	108
5.10.3	Lines of Best Fit	108
5.10.4	Equation of the Regression Line	108
6	Probability	110
6.1	Probability Concepts	110
6.2	Probability of an Event	110
6.3	Probability of an Event's Not Occurring	110
6.4	Independent, Dependent, and Mutually Exclusive Events	110
6.5	Probability of A and B for Independent Events	111
6.6	Probability of A or B for Mutually Exclusive Events	112
6.7	Probability of Mutually Exclusive Events	112
6.8	Probability of A or B for Non-Mutually Exclusive Events	112
6.9	Lists and Tables of Outcomes	113
6.10	Conditional Probability of Dependent Events	113
6.11	Testing for Independent Events	113
6.12	Venn Diagrams	113
6.13	Tree Diagrams	114

6.14	Probabilities With and Without Replacement	115
6.15	Discrete Random Variables and Their Probability Distributions	115
6.15.1	Explicitly Listed Probabilities	115
6.15.2	Probability Distribution Given by a Function	116
6.15.3	Explicit Probability Distribution Involving an Unknown	116
6.16	Expected Value (Mean), $E(X)$ for Discrete Data	117
6.16.1	Expected Value for a “Fair” Game	118
6.17	Binomial Distribution	119
6.17.1	Using the Calculator for Binomial Distribution Problems	120
6.17.2	Tip for Calculators Whose <code>binomCdf</code> Does Not Have Lower Bound	121
6.18	Mean and Variance of the Binomial Distribution	122
6.18.1	Example of Mean and Variance of Binomial Distribution	122
6.19	Normal Distribution and Curves	123
6.20	Graphing the Normal Distribution and Computing Bounded Area	123
6.21	Standardizing Normal Variables to Get z -values (z -scores)	124
6.22	Using the Calculator for Normal Distribution Problems	125
6.22.1	Tip for Calculators Whose <code>normalCdf</code> Does Not Have Lower Bound	126
6.23	Using Inverse Normal Cumulative Density Function	126
6.24	Determining z Value and σ from the Probability	126

7 Calculus **128**

7.1	Overview of Concepts for Differential Calculus	128
7.1.1	Rate of Change in Distance	128
7.1.2	Derivative Gives Instantaneous Rate of Change Using Tangent Lines	130
7.1.3	Relationships Between Function and Derivative	131
7.1.4	Summary of Derivative Concepts	133
7.2	Equations of Tangents and Normals	133
7.3	Notation for Derivatives	134
7.4	Graphing Derivatives with Your Calculator	134
7.4.1	Graphing Derivatives on the TI-Nspire CX	135
7.4.2	Graphing Derivatives on the TI-84 Silver Edition	135
7.5	Computing Derivatives at a Point with Your Calculator	136
7.5.1	Computing Derivative at a Point on the TI-Nspire CX	136
7.5.2	Computing Derivative at a Point on the TI-84 Silver Edition	137
7.6	Rules for Computing Derivatives	137
7.6.1	Derivative of x^n	137
7.6.2	Derivative of $\sin x$	138
7.6.3	Derivative of $\cos x$	138
7.6.4	Derivative of $\tan x$	138
7.6.5	Derivative of e^x	138
7.6.6	Derivative of $\ln x$	138
7.7	Differentiating a Scalar Multiple	138
7.8	Differentiating a Sum or Difference	138
7.9	Example Using Derivation Rules	139
7.10	“Chain Rule”: Differentiating Composed Functions	139
7.11	“Product Rule”: Differentiating Product of Functions	141
7.12	“Quotient Rule”: Differentiating Quotient of Functions	144
7.13	Using the First Derivative to Find Local Maxima and Minima	146
7.14	Analyzing Zeros of Derivative Graph to Find Maxima and Minima	147
7.15	Using the Second Derivative to Determine Concavity and Points of Inflection	148

7.16	Overview of Concepts for Integral Calculus	148
7.16.1	Area Under Curve	149
7.16.2	Relationship Between Derivation and Integration	150
7.17	Anti-Differentiation to Compute Integrals	151
7.17.1	Indefinite Integration as Anti-Differentiation	152
7.18	Rules for Computing the Indefinite Integral (Antiderivative)	152
7.18.1	Indefinite Integral of x^n ($n \in \mathbb{Q}$)	153
7.18.2	Indefinite Integral of $\sin x$	153
7.18.3	Indefinite Integral of $\cos x$	154
7.18.4	Indefinite Integral of $\frac{1}{x}$	154
7.18.5	Indefinite Integral of e^x	154
7.19	Integrating Constant Multiple of a Function	154
7.20	Integrating Sums and Differences of Functions	155
7.21	Integration Using Substitution	155
7.21.1	Simple Example Involving Linear Factor	156
7.21.2	Example Involving Additional Factor of x	157
7.21.3	Example Involving Other Additional Factors	158
7.21.4	Summary of Strategy for Integration by Substitution	158
7.22	Using Additional Information to Determine Constant of Integration	158
7.23	Computing the Definite Integral	160
7.23.1	Computing the Definite Integral Analytically	160
7.23.2	Computing the Definite Integral Using Technology	161
7.23.3	Computing Normal Distribution Probability Using a Definite Integral	161
7.24	Displacement s , Velocity v , and Acceleration a	162
7.25	Determining Position Function from Acceleration	162
7.26	Total Area Under a Curve	164
7.27	Area Between Curves	164
7.27.1	Area Between Curves Using Calculator	164
7.27.2	Area Between Curves Analytically (Without the Calculator)	165
7.28	Net Change in Displacement versus Total Distance Traveled	165
7.28.1	Difference Between Distance and Displacement	165
7.28.2	Computing Total Distance Traveled	166

1 Algebra

1.1 Rules of Basic Operations

You need to be familiar with the rules of basic operations shown in Table 1.

$a(b + c) = ab + ac$	$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$
$\frac{a}{b} + \frac{c}{d} = \frac{a}{b} \cdot \frac{d}{d} + \frac{c}{d} \cdot \frac{b}{b} = \frac{ad+cb}{bd}$	$-(a + b) = -a - b$
$-(a - b) = b - a$	$(a + b)^2 = a^2 + 2ab + b^2$
$(a - b)^2 = a^2 - 2ab + b^2$	$(a + b)(a - b) = a^2 - b^2$
$(a + b)(c + d) = ac + ad + bc + bd$	

Table 1: Basic Operations

1.2 Rules of Roots

You need to be familiar with the rules of roots shown in Table 2.

$\sqrt{ab} = \sqrt{a}\sqrt{b}$	$\sqrt{\frac{a}{b}} = \frac{\sqrt{a}}{\sqrt{b}}$	$a^{\frac{1}{q}} = \sqrt[q]{a}$	$a^{\frac{p}{q}} = (\sqrt[q]{a})^p$
$a^{-\frac{p}{q}} = \frac{1}{(\sqrt[q]{a})^p}$	$\frac{p}{\sqrt{a}} \cdot \frac{\sqrt{a}}{\sqrt{a}} = \frac{p\sqrt{a}}{a}$	$(\sqrt{a})^2 = a$	$(\sqrt[q]{a})^q = a$

Table 2: Rules of Roots

1.3 Rules of Exponents

You need to be familiar with the rules of exponents shown in Table 3.

$a^0 = 1$	$a^1 = a$
$a^b \cdot a^c = a^{b+c}$	$(a^b)^c = a^{b \cdot c}$
$\left(\frac{a}{d}\right)^c = \frac{a^c}{d^c}$	$a^{(-c)} = \frac{1}{a^c}$
$\frac{a^b}{a^c} = a^{(b-c)}$	$\frac{a^b}{d^c} \cdot \frac{d^c}{a^f} = \frac{a^b}{a^f} \cdot \frac{d^c}{d^c} = a^{b-f} \cdot d^{c-e} = \frac{a^{b-f}}{d^{e-c}}$

Table 3: Rules of Exponents

1.4 Allowed and Disallowed Calculator Functions During the Exam

The International Baccalaureate documentation for examiners dated 18/11/2015, provides the following guidance about the functions of the TI NSpire calculator during an exam. The following options in “Press to Test” mode must be ticked and therefore **blocked**: *Limit geometry functions; Disable function and conic grab and move; Disable vector functions including eigenvectors and eigenvalues; Disable isPrime function; Disable differential equation graphing; Disable 3D graphing; Disable implicit graphing, conic templates, conic analysis, and geometric conics.* The following options in “Press to Test” mode must be unticked and therefore **allowed**: *Disable inequality graphing; Limit trigonometric functions; log_b x template and summation functions; Disable Polynomial Root Finder and Simultaneous Equation Solver.*

1.5 Sequences and Series

A *sequence* is an ordered list of numbers. The list can be finite or infinite. We name the sequence with a letter like u , and refer to an individual element (or “term”) of this ordered list using a subscript that represents the ordinal position the particular element occupies in the list. The first term in the sequence u is denoted by u_1 , the second term by u_2 , and so on. The n th term in the sequence is u_n .

You describe a sequence by giving a formula for its n th term. This could be something simple like the sequence each of whose terms is the number 1:

$$u_n = 1$$

More typically, the formula for the n th term involves the variable n in some way. It might be involved in calculating the value of the term like:

$$u_n = n^2 \qquad \text{or} \qquad u_n = \frac{1}{n}$$

The n can also appear in the subscript to reference the values of previous terms. For example, after defining the value of the first and second term, this sequence defines the n th term to be the sum of the previous two terms:

$$\begin{aligned} u_1 &= 1 \\ u_2 &= 1 \\ u_n &= u_{n-1} + u_{n-2} \end{aligned}$$

A *series* is the sum of the terms in a sequence.

1.6 Arithmetic Sequences and Series

A sequence is *arithmetic* if you compute its n th term by adding a constant value d to the previous element:

$$u_n = u_{n-1} + d$$

This value d is known as the *common difference* because if you take any term in the sequence and subtract the value of the preceding one, the difference is d . This is a result of rearranging the previous equation to get:

$$u_n - u_{n-1} = d$$

Given the above definition for u_n involving d , we can also represent the n th element of an arithmetic sequence as the first element to which the common difference d has been added $n - 1$ times.

$$u_n = u_1 + (n - 1)d$$

An *arithmetic series* is the sum of an arithmetic sequence.

1.7 Sum of Finite Arithmetic Series ($u_1 + \dots + u_n$)

It only makes sense to talk about the sum of a *finite* arithmetic sequence. Since the elements of an *infinite* arithmetic sequence continue growing by the common difference, we don't try to compute the sum of the corresponding infinite arithmetic series because its value would always be ∞ (or $-\infty$ if d is a negative number).

To calculate the sum of the first n terms of a *finite* arithmetic sequence we use the formula:

$$S_n = \frac{n}{2} (u_1 + u_n)$$

This formula comes from observing that the sum of the pair of $u_1 + u_n$ (the first and last terms) is the same as the sum of $u_2 + u_{n-1}$ (the second and penultimate terms), which is the same as the sum of $u_3 + u_{n-2}$ (the third and antepenultimate terms), etc. Since we're pairing up the terms, we end up with $\frac{n}{2}$ pairs having the same sum as $u_1 + u_n$. By multiplying this sum by the number of pairs, we get the sum of all of the terms of the series up to n . Since above we saw that $u_n = u_1 + (n - 1)d$, we can also write this formula as:

$$S_n = \frac{n}{2} (u_1 + (u_1 + (n - 1)d)) = \frac{n}{2} (2u_1 + (n - 1)d)$$

1.8 Partial Sum of Finite Arithmetic Series ($u_j + \dots + u_n$)

Sometimes you might need to calculate the sum of a series starting at a j th term other than the first term, that is where $j > 1$. These kind of problems are easy once you understand that the sum of the first n terms equals the sum of the first $j - 1$ terms plus the sum of the terms from the j th position to the n th. In other words,

$$\begin{aligned} S_n &= (u_1 + \dots + u_{j-1}) + (u_j + \dots + u_n) \\ S_n &= S_{j-1} + S_{j\dots n} \end{aligned}$$

Rearranging the equation we have:

$$S_{j\dots n} = S_n - S_{j-1}$$

Therefore, to find the sum of the series from the j th term through the n th term (inclusive), calculate S_n and subtract from it the value of S_{j-1} .

1.9 Geometric Sequences and Series

A sequence is *geometric* if you compute its n th term by multiplying the previous element by a constant value r :

$$u_n = u_{n-1} \cdot r$$

This value r is known as the *common ratio* because if you take any term in the sequence and divide by the value of the preceding one, the ratio is r . This is a result of rearranging the previous equation to get:

$$\frac{u_n}{u_{n-1}} = r$$

Given the above definition for u_n involving r , we can also represent the n th element of an arithmetic sequence as the first element by which the common ratio r has been multiplied $n - 1$ times.

$$u_n = u_1 \cdot r^{n-1}$$

A *geometric series* is the sum of a geometric sequence.

1.10 Sum of Finite Geometric Series

To calculate the sum of the first n terms of a geometric series, use the formula:

$$S_n = \frac{u_1(r^n - 1)}{r - 1}, \quad r \neq 1$$

1.11 Sum of Infinite Geometric Series

If the absolute value of the common ratio is less than one, then each time we compute the next term of a geometric sequence it is some fraction of the previous term. In other words, the new term is smaller in absolute value than the previous term. Another way to think of this is that the new term is closer to zero than the previous term was. This means we can calculate the sum of the *infinite* geometric series because the terms tend to zero as n gets larger and larger.

To calculate the sum of an infinite geometric series, use the formula:

$$S_\infty = \frac{u_1}{1 - r}, \quad |r| < 1$$

1.11.1 Example Involving Sum of Infinite Geometric Series

As shown in Figure 1, the sides of an outermost square are 16 units. To create a smaller square, connect midpoints of the original square. Consider the sequence A_n of the areas of the shaded triangles in the lower-left corner of each square, formed when the next smaller square is drawn. Find the sum of the areas of these shaded triangles when this process continues infinitely.

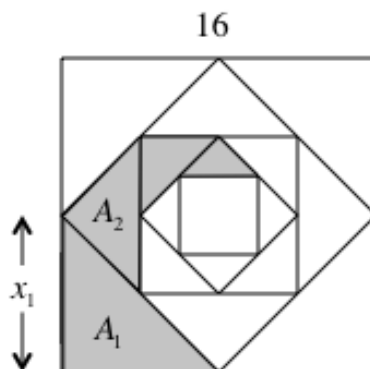


Figure 1: Progressively Smaller Triangle Areas

Since the original side length is 16, the legs of the first shaded triangle are 8, and the area $A_1 = \frac{1}{2} \cdot 8 \cdot 8 = 32$. Using Pythagorean Theorem, we can see that for a square of side-length a , its hypotenuse will always be $\sqrt{a^2 + a^2} = a\sqrt{2}$, so the hypotenuse of the first triangle is $8\sqrt{2}$. This means that the legs of the second triangle are half of this, or $4\sqrt{2}$. Therefore the area of the second triangle is $A_2 = \frac{1}{2} \cdot 4\sqrt{2} \cdot 4\sqrt{2} = 16$. The hypotenuse of the second shaded triangle is $4\sqrt{2}\sqrt{2} = 8$, so the legs of the third triangle are half of this: 4. Hence, the area of the third shaded triangle is $A_3 = \frac{1}{2} \cdot 4 \cdot 4 = 8$. The hypotenuse of the third shaded triangle is $4\sqrt{2}$, so the legs of the fourth shaded triangle are half of this: $2\sqrt{2}$. So, the area of the fourth shaded triangle is $A_4 = \frac{1}{2} \cdot 2\sqrt{2} \cdot 2\sqrt{2} = 4$. So, the pattern that is emerging is that $A = \{32, 16, 8, 4, \dots, 2^{6-n}, \dots\}$.

Considering the sequence A of the triangle areas, we see that each subsequent term is obtained by multiplying the previous term by $\frac{1}{2}$, so the sequence of the areas is a geometric sequence with common ratio $r = \frac{1}{2}$. Since the common ratio is less than one in absolute value, it makes

sense to compute the infinite sum, which is given by the formula in Section 1.11, “Sum of Infinite Geometric Series”:

$$\sum_{n=1}^{\infty} A_n = \frac{A_1}{1-r} = \frac{32}{1-\frac{1}{2}} = 64$$

1.12 Sigma Notation

Sigma notation is another convenient way to represent a series. Instead of saying an English phrase like “the sum of the first 5 terms of the sequence defined by $u_n = 2n^2 + 1$ ”, we can write the same expression using the much more compact formula:

$$\sum_{n=1}^5 2n^2 + 1$$

The \sum symbol is the capital Greek letter sigma, the first letter in the word “sum”. The $n = 1$ below the \sum indicates the variable whose value will be changing to compute the terms in the series, and it also indicates the starting value of this variable. The 5 above the \sum symbol provides the ending value of the variable n specified below the symbol. You compute each term in the sequence by substituting the current value of the changing variable into the expression that follows the \sum symbol. The series is the sum of these terms.

In the example above, the variable is n and its value starts at 1 and goes to 5, incrementing by 1 each time. This means that we can “expand” the compact sigma notation in the example into this expression:

$$\begin{aligned} \sum_{n=1}^5 2n^2 + 1 &= \left(\underbrace{2(1)^2 + 1}_{n=1} \right) + \left(\underbrace{2(2)^2 + 1}_{n=2} \right) + \left(\underbrace{2(3)^2 + 1}_{n=3} \right) + \left(\underbrace{2(4)^2 + 1}_{n=4} \right) + \left(\underbrace{2(5)^2 + 1}_{n=5} \right) \\ &= (2 \cdot 1 + 1) + (2 \cdot 4 + 1) + (2 \cdot 9 + 1) + (2 \cdot 16 + 1) + (2 \cdot 25 + 1) \\ &= (3) + (9) + (19) + (33) + (51) \\ &= 115 \end{aligned}$$

In the example above, by computing the difference between two pairs of consecutive terms, we can observe that the sequence (and hence the series) is not arithmetic because the terms do not share a common difference:

$$\begin{aligned} u_2 - u_1 &= 9 - 3 = 6 \\ u_3 - u_2 &= 19 - 9 = 10 \neq 6 \end{aligned}$$

By computing the ratio between two pairs of consecutive terms, we can also observe that the sequence (and hence the series) is not geometric because the terms do not share a common ratio:

$$\begin{aligned} \frac{u_2}{u_1} &= \frac{9}{3} = 3 \\ \frac{u_3}{u_2} &= \frac{19}{9} \neq 3 \end{aligned}$$

Of course there is nothing special about the letter n and the value of the variable being incremented from the start index to the end index need not start with the value 1. The sigma notation can be used with any letter for the variable and any integer values for the index start

and end. Other common variable letters that get used are i , j , k , and m . The following series are equivalent:

$$\sum_{n=3}^9 2n^2 + 1 = \sum_{i=3}^9 2i^2 + 1 = \sum_{j=3}^9 2j^2 + 1 = \sum_{m=3}^9 2m^2 + 1$$

1.12.1 Sigma Notation for Arithmetic Series

The series that the sigma notation represents might be an arithmetic series. Consider the example:

$$\begin{aligned} \sum_{n=1}^5 2n + 1 &= \left(\underbrace{2(1) + 1}_{n=1} \right) + \left(\underbrace{2(2) + 1}_{n=2} \right) + \left(\underbrace{2(3) + 1}_{n=3} \right) + \left(\underbrace{2(4) + 1}_{n=4} \right) + \left(\underbrace{2(5) + 1}_{n=5} \right) \\ &= (2 \cdot 1 + 1) + (2 \cdot 2 + 1) + (2 \cdot 3 + 1) + (2 \cdot 4 + 1) + (2 \cdot 5 + 1) \\ &= (3) + (5) + (7) + (9) + (11) \\ &= 35 \end{aligned}$$

If we compute the difference between two pairs of consecutive terms, we see that they both have a common difference of 2, so this series is arithmetic:

$$\begin{aligned} u_2 - u_1 &= 5 - 3 = 2 \\ u_3 - u_2 &= 7 - 5 = 2 \end{aligned}$$

Computing the ratio between two pairs of consecutive terms, we see that there is no common ratio, so the series is not geometric.

$$\begin{aligned} \frac{u_2}{u_1} &= \frac{5}{3} \\ \frac{u_3}{u_2} &= \frac{7}{5} \neq \frac{5}{3} \end{aligned}$$

1.12.2 Sigma Notation for Geometric Series

The series that the sigma notation represents might be a geometric series. Consider the example:

$$\begin{aligned} \sum_{k=1}^4 \frac{1}{3^{2k}} &= \left(\underbrace{\frac{1}{3^{2 \cdot 1}}}_{k=1} \right) + \left(\underbrace{\frac{1}{3^{2 \cdot 2}}}_{k=2} \right) + \left(\underbrace{\frac{1}{3^{2 \cdot 3}}}_{k=3} \right) + \left(\underbrace{\frac{1}{3^{2 \cdot 4}}}_{k=4} \right) \\ &= \frac{1}{9} + \frac{1}{81} + \frac{1}{729} + \frac{1}{6561} \\ &= \frac{820}{6561} \end{aligned}$$

If we compute the difference between two pairs of consecutive terms, we see that they do not have a common difference, so this series is not arithmetic:

$$\begin{aligned} u_2 - u_1 &= \frac{1}{81} - \frac{1}{9} = -\frac{8}{81} \\ u_3 - u_2 &= \frac{1}{729} - \frac{1}{81} = -\frac{8}{729} \neq -\frac{8}{81} \end{aligned}$$

Computing the ratio between two pairs of consecutive terms, we see that there is common ratio of $\frac{1}{9}$, so the series is geometric.

$$\frac{u_2}{u_1} = \frac{\frac{1}{81}}{\frac{1}{9}} = \frac{1}{81} \cdot \frac{9}{1} = \frac{1}{9}$$

$$\frac{u_3}{u_2} = \frac{\frac{1}{729}}{\frac{1}{81}} = \frac{1}{729} \cdot \frac{81}{1} = \frac{1}{9}$$

1.12.3 Sigma Notation for Infinite Geometric Series

The sigma notation can also represent an infinite sum by having ∞ as the end value of the variable. The infinite sum of the geometric series from the previous section would therefore be written like:

$$\sum_{k=1}^{\infty} \frac{1}{3^{2k}}$$

Since we determined above that this series is geometric with a common ratio of $r = \frac{1}{9}$ and since this respects the requirement for infinite series to have $|r| < 1$, we can use the formula from section 1.11 to calculate the infinite sum.

$$\sum_{k=1}^{\infty} \frac{1}{3^{2k}} = \frac{u_1}{1-r} = \frac{\frac{1}{3^{2 \cdot 1}}}{1 - \frac{1}{9}} = \frac{\frac{1}{9}}{\frac{8}{9}} = \frac{1}{9} \cdot \frac{9}{8} = \frac{1}{8}$$

1.12.4 Defining Functions Using Sigma Notation

You can define a function using sigma notation by including an additional variable in the expression other than the so-called “index variable” that is counting from the start to end value. For example, consider the following function $f(x)$:

$$f(x) = 1 + x + x^2 + x^3 + x^4$$

To write the function more compactly, you could use sigma notation like this:

$$f(x) = \sum_{k=0}^4 x^k$$

When expanding the sigma notation, each term replaces the value of the index variable with the current integer value (from 0 to 4 in this example). Any other variables, like the x here, remains as they are in the expression. It even makes sense to use this notation for functions defined by infinite sums. We could write:

$$f(x) = \sum_{k=0}^{\infty} x^k$$

This would be the function with an infinite number of terms:

$$f(x) = 1 + x + x^2 + x^3 + x^4 + x^5 + \dots$$

1.13 Applications: Compound Interest

Interest is a percentage paid on an original amount of money, the principal, after a certain interval of time. During subsequent periods, the interest earned in one period begins earning interest (as part of the new balance) during the next period. In this way, the interest compounds (or accumulates) period after period.

If P is the initial principal, r is the annual interest rate expressed as a decimal, and we are compounding annually, then after one year the amount of money will be $P + P \cdot r = P \cdot (1 + r)$. That 1 inside the parenthesis ensures that the original principal amount P is included in the new total, added to the fraction r of the principal that represents the interest. After the second year, the amount from the end of the first year is multiplied by the same $(1 + r)$ quantity to produce the new balance of $[P \cdot (1 + r)](1 + r) = P \cdot (1 + r)^2$. After three years, the balance is $P \cdot (1 + r)^3$. Observe that the yearly balance of the bank account is a geometric sequence with first (zeroth) term (u_0) equal to P and a common ratio of $(1 + r)$. The t -th term in the sequence represents the balance after t years of compounding annually.

$$P \cdot (1 + r)^t$$

The period of compounding can be more frequent than once per year. It could be twelve times a year (monthly), or fifty-two times a year (weekly), or 365 times a year (daily). In these cases, the annual interest percentage r is divided over the number of compounding periods a year, $\frac{r}{12}$, $\frac{r}{52}$, or $\frac{r}{365}$ respectively. The number of compounding periods will increase as well in these cases. In t years, there would be $12t$ monthly periods, $52t$ weekly periods, or $365t$ daily periods. In general, if there are n compounding periods per year, then the per-period fractional amount of the annual interest rate will be $\frac{r}{n}$ and the number of compounding periods will be $n \cdot t$. Therefore, the formula for calculating the balance of an initial investment P at r percent annual interest over t years, compounding n times a year is:

$$P \cdot \left(1 + \frac{r}{n}\right)^{n \cdot t}$$

1.14 Applications: Population Growth

The same formula for annual compounding of interest can be used to calculate population growth as well. If a population is initially P people, and it grows at the rate of r percent per year (expressed as a decimal), then the total number of people after t years is:

$$P \cdot (1 + r)^t$$

1.15 Logarithms

We know intuitively when we look at an equation like $5^x = 125$ what the value of x has to be to make the equality true. What is x here? It is the power to which we need to raise the base 5 to get the result 125. In this example, it must equal 3 since $5 \cdot 5 \cdot 5 = 5^3 = 125$. Similarly, if we see the equation $5^x = 25$ we know $x = 2$. But how do we solve an equation like $5^x = 100$? From our previous two observations, we understand that 5^2 is less than 100 and 5^3 is greater than 100. In other words, we know that $5^2 < 5^x < 5^3$. Recall that the value of an exponent is not limited to integer numbers like 2 or 3. Intuitively, the value of x must lie somewhere between 2 and 3. We could start guessing at the value by picking a number like 2.5 and using our calculator to compute $5^{2.5} = 55.9017$, which is still too small. Next we could try a value like $5^{2.9} = 106.417$, which is too large. So, now we know that $5^{2.5} < 5^x < 5^{2.9}$. Trying other values of x between 2.5 and 2.9 we could eventually converge on a value for x like 2.86135 after finding that:

$$\underbrace{5^{2.8613}}_{99.9915} < \underbrace{5^x}_{100} < \underbrace{5^{2.8614}}_{100.008}$$

In fact, $5^{2.86135} = 99.9995 \approx 100$. So, using this manual procedure and some intuition we've calculated an approximate value for the power to which we would need to raise the base 5 to get the value 100.

What if we were not happy with an approximation, but instead wanted the *exact* value of x that solves the equation $5^x = 100$? This is where logarithms come into play. They provide a mathematical notation to represent the *exact* value of the exponent required on a given base like 5 to equal a desired value like 100. So rather than being contented by saying that x is approximately equal to 2.86135, instead we can say that x equals "log base 5 of 100", which we write like this:

$$x = \log_5 100$$

The expression on the right hand side of the equals sign is an exact number (like π or $\sqrt{2}$) that has a place on the number line close to, but ever so slightly to the right of 2.86135. We conclude this because we saw that $5^{2.86135} = 99.9995$ was just a bit smaller than 100. What is the value of this expression?

$$5^{\log_5 100}$$

The number $\log_5 100$ is exactly the right number that, when used as a power on the base 5, gives the number 100. Here we are using it as a power on the base 5, so the result of this calculation is exactly 100.

$$5^{\log_5 100} = 100$$

So, we see that the action of "raising base 5 to a power" and "taking the log base 5" are inverse operations. One operation "undoes" the other. Consider the following expression. What does it equal?

$$\log_5 (5^3)$$

We know from above that $\log_5 (5^3)$ is a number that is exactly the right exponent to use on the base 5 to get the value 5^3 . So what is that exponent?

$$\log_5 (5^3) = 3$$

The exponent we need to use on the base 5 to get the result of 5^3 is just the number 3. So again, we see that the actions of "raising base 5 to a power" and "taking the log base 5" are inverse operations. One operation "undoes" the other.

1.16 Using Logarithms to Solve Equations

Given that the actions of "raising base b to a power" and "taking the log base b " are inverse operations, we can use these actions as tools when solving equations. If the unknown variable x is "trapped" in the position of the exponent like this:

$$9^{2x} = 6561$$

We can "take the log of both sides" to free the expression involving x from its trapped position up in the exponent. In order to have the inverse effect, we need to pick the correct base of the logarithm to match the base that is trapping our expression. In this case, the base is 9, so we'll take the \log_9 of both sides to keep the equation equal:

$$\begin{aligned} 9^{2x} &= 6561 \\ \log_9 (9^{2x}) &= \log_9 6561 \end{aligned}$$

The expression $\log_9(9^{2x})$ represents the power we need to put on the base 9 to get 9^{2x} , so that is just $2x$. As before the operations of “taking the log base 9” and “raising the base 9 to a power” are inverses, so they cancel or neutralize each other:

$$\begin{aligned}\log_9(9^{2x}) &= \log_9 6561 \\ 2x &= \log_9 6561 \\ x &= \frac{\log_9 6561}{2} = \frac{4}{2} = 2\end{aligned}$$

1.17 Using Exponentiation to Solve Equations

Given that the actions of “taking the log base b ” and “raising base b to a power” are inverse operations, we can use these actions as tools when solving equations. If the unknown variable x is “trapped” in the expression of a logarithm like this:

$$\log_9(3x) = 4$$

We can “raise a base to a power” on both sides to free the expression involving x from its trapped position inside the logarithm. In order to have the inverse effect, we need to pick the correct base of the logarithm to match the base that is trapping our expression. In this case, the base is 9, so we’ll use 9 as the base on both sides of the equals sign, and use the expression on both sides as the exponent on this base. Since those expressions were equal before, putting equal exponents on the same base 9 will produce results that are also equal:

$$\begin{aligned}\log_9(3x) &= 4 \\ 9^{\log_9(3x)} &= 9^4\end{aligned}$$

The expression $\log_9(3x)$ represents the power we need to put on the base 9 to get $3x$, and we have used it as power on exactly that base 9 here, so it must produce $3x$. As before the operations of “raising the base 9 to a power” and “taking the log base 9” are inverses, so they cancel or neutralize each other:

$$\begin{aligned}9^{\log_9(3x)} &= 9^4 \\ 3x &= 6561 \\ x &= \frac{6561}{3} = 2187\end{aligned}$$

1.18 Logarithm Facts Involving 0 and 1

Two facts that stem from the definition of logarithm are the following (for any value of $b > 0$):

$$\begin{array}{ll}\log_b 1 = 0 & \text{(since } b^0 = 1\text{)} \\ \log_b b = 1 & \text{(since } b^1 = b\text{)} \\ \log_b 0 = \textit{undefined} & \text{(since no value } x \text{ makes } b^x = 0\text{)}\end{array}$$

1.19 Laws of Exponents and Logarithms

There are three laws to remember to help simplify expressions involving logarithms:

$$\begin{aligned}\log_b a + \log_b c &= \log_b (a \cdot c) \\ \log_b a - \log_b c &= \log_b \left(\frac{a}{c}\right) \\ \log_b (a^c) &= c \cdot \log_b (a)\end{aligned}$$

1.20 Change of Base

A fourth law related to logarithms is the “Change of Base” law. This allows you to change a logarithm in one base p to a new base q .

$$\log_p a = \frac{\log_q a}{\log_q p}$$

Consider the expression $\log_{12}100$. To compute a decimal value for this using the change of base rule, you can re-express it in terms of base 10 as follows:

$$\log_{12}100 = \frac{\log_{10}100}{\log_{10}12} = \frac{2}{1.07918} = 1.85326$$

1.21 Powers of Binomials and Pascal’s Triangle

A *polynomial* is a mathematical expression involving a sum of powers of one or more variables multiplied by coefficients. For example, the following is a polynomial in variables x and y with integer coefficients:

$$15x^2y^3 + 7xy^2 - 6xy$$

The individual terms in the sum are called *monomials*. A polynomial that is the sum of two monomials is known as a *binomial*. For example, all of the following are examples of binomials:

$$\begin{aligned} 15x^2y^3 + 7xy^2 \\ 15x^2 - 7x \\ 15x + 7y \\ x + 1 \end{aligned}$$

When you multiply one binomial by another you expand the product one step at a time using the distributive property of multiplication. For example, consider:

$$\begin{aligned} (15x + 7y)(x + 2) &= (15x + 7y) \cdot x + (15x + 7y) \cdot 2 \\ &= 15x \cdot x + 7y \cdot x + 15x \cdot 2 + 7y \cdot 2 \\ &= 15x^2 + 7xy + 30x + 14y \end{aligned}$$

The same approach works, of course, when you multiply a binomial by itself:

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

Using shorthand notation, we see then that:

$$(x + 1)^2 = (x + 1)(x + 1) = x^2 + 2x + 1$$

We can use the same approach to multiply a binomial by itself to higher powers. For example, consider:

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

Continuing the pattern, we can calculate:

$$\begin{aligned}
 (x + 1)^4 &= (x + 1)^3(x + 1) = (x + 1)^3 \cdot x + (x + 1)^3 \cdot 1 \\
 &= (x^3 + 3x^2 + 3x + 1) \cdot x + (x^3 + 3x^2 + 3x + 1) \cdot 1 \\
 &= x^3 \cdot x + 3x^2 \cdot x + 3x \cdot x + 1 \cdot x + x^3 \cdot 1 + 3x^2 \cdot 1 + 3x \cdot 1 + 1 \cdot 1 \\
 &= x^4 + 4x^3 + 6x^2 + 4x + 1
 \end{aligned}$$

If you adopt the convention of always ordering the powers of x from highest to lowest going left to right, then as you continue expanding the binomial $(x + 1)$ to higher and higher powers, you would notice that the coefficients of the terms follow the pattern in Pascal's Triangle shown in Figure 2. If you study a row in the triangle like Row 6, you can see that each row begins and ends with 1 and that the coefficients in the inside positions are symmetric. Each of these "middle coefficients" in Row 6 is obtained by adding the coefficient above and to its left with the one above and to its right from the previous row. So $6 = 1 + 5$, $15 = 5 + 10$, and $20 = 10 + 10$. Knowing this pattern, it would be easy to extend the triangle to any number of rows.

Row 0:	1
Row 1:	1 1
Row 2:	1 2 1
Row 3:	1 3 3 1
Row 4:	1 4 6 4 1
Row 5:	1 5 10 10 5 1
Row 6:	1 6 15 20 15 6 1
Row 7:	1 7 21 35 35 21 7 1

Figure 2: Row n in Pascal's Triangle are coefficients of binomial expansion of $(x + 1)^n$

1.22 Expansion of $(a + b)^n$

In practice, the binomial raised to a power that you will need to expand will be more complex than the simple $(x + 1)^n$ we studied in the previous section. So it won't be possible, in general, to just read the numbers off of Pascal's Triangle and be done with it. For example, consider a slightly more complicated binomial expansion like this:

$$(3x + 2y)^6$$

We can still use Pascal's Triangle to find the coefficients, but we need to be more careful about following the pattern. Since we're raising the binomial to the sixth power, we'll use the coefficients from Row 6 in the triangle:

$$(3x + 2y)^6 = 1 \cdot (\dots) + 6 \cdot (\dots) + 15 \cdot (\dots) + 20 \cdot (\dots) + 15 \cdot (\dots) + 6 \cdot (\dots) + 1 \cdot (\dots)$$

From the simpler $(x + 1)^n$ examples in the previous section, we saw that the highest power of x started with the same power to which the binomial was being raised and then proceeded to get smaller by one in each successive term. The same pattern works when the first term in the binomial is $3x$...

$$\begin{aligned}
 (3x + 2y)^6 &= 1 \cdot (3x)^6(\dots) + 6 \cdot (3x)^5(\dots) + 15 \cdot (3x)^4(\dots) \\
 &\quad + 20 \cdot (3x)^3(\dots) + 15 \cdot (3x)^2(\dots) + 6 \cdot (3x)^1(\dots) + 1 \cdot (3x)^0(\dots)
 \end{aligned}$$

While it was hard to observe when the second term in the binomial was 1, there was a similar pattern occurring for the powers of the second term. However, the powers of the second term

start at 6 in the final term and count down to 0 in each preceding term (going right to left). Alternatively, you can think of them starting with 0 and counting up to 6 going left to right, like this...

$$(3x + 2y)^6 = 1 \cdot (3x)^6(2y)^0 + 6 \cdot (3x)^5(2y)^1 + 15 \cdot (3x)^4(2y)^2 + 20 \cdot (3x)^3(2y)^3 + 15 \cdot (3x)^2(2y)^4 + 6 \cdot (3x)^1(2y)^5 + 1 \cdot (3x)^0(2y)^6$$

Notice in each term of the expansion, the powers of the two monomials add up to 6, which is the power to which the original binomial is being raised. If we simplify the expression by raising each monomial to the indicated power, we get:

$$\begin{aligned} (3x + 2y)^6 &= 1 \cdot 3^6 \cdot x^6 \cdot 2^0 \cdot y^0 + 6 \cdot 3^5 \cdot x^5 \cdot 2^1 \cdot y^1 + 15 \cdot 3^4 \cdot x^4 \cdot 2^2 \cdot y^2 \\ &\quad + 20 \cdot 3^3 \cdot x^3 \cdot 2^3 \cdot y^3 + 15 \cdot 3^2 \cdot x^2 \cdot 2^4 \cdot y^4 + 6 \cdot 3^1 \cdot x^1 \cdot 2^5 \cdot y^5 \\ &\quad + 1 \cdot 3^0 \cdot x^0 \cdot 2^6 \cdot y^6 \\ &= 729x^6 + 6 \cdot 243x^5 \cdot 2y + 15 \cdot 81x^4 \cdot 4y^2 + 20 \cdot 27x^3 \cdot 8y^3 \\ &\quad + 15 \cdot 9x^2 \cdot 16y^4 + 6 \cdot 3x \cdot 32y^5 + 64y^6 \\ &= 729x^6 + 2916x^5y + 4860x^4y^2 + 4320x^3y^3 + 2160x^2y^4 + 576xy^5 + 64y^6 \end{aligned}$$

1.23 The Binomial Theorem

We have seen how to compute $(a + b)^n$ using the "brute force" procedure in the previous section, using the n -th row from Pascal's Triangle for the coefficients. However, this process would quickly become tedious if n were a larger number like 17 or 37, especially if you only need to calculate one term in the expansion.

The *Binomial Theorem* gives a more concise formula to calculate any particular term in the expansion of $(a + b)^n$ for any value of n . It relies on the observation that the r -th entry in n -th row of Pascal's Triangle is precisely the number of ways there are to choose r items out of a set of n items (where the order of the items does not matter). This is a calculation from combinatorics that has a special notation named " n choose r " that looks like this:

$$\begin{aligned} \binom{n}{r} &= \frac{\text{Permutations of } r \text{ elements without repetition from set of } n \text{ elements}}{\text{Permutations of } r \text{ elements}} \\ &= \frac{n \cdot (n-1) \cdots (n-r+1)}{r!} \\ &= \frac{1}{r!} \cdot (n \cdot (n-1) \cdots (n-r+1)) \\ &= \frac{1}{r!} \cdot \frac{n \cdot (n-1) \cdots (n-r+1) \cdot \cancel{(n-r)} \cdot \cancel{(n-r-1)} \cdot \cancel{(n-r-2)} \cdots \cancel{2} \cdot \cancel{1}}{\cancel{(n-r)} \cdot \cancel{(n-r-1)} \cdot \cancel{(n-r-2)} \cdots \cancel{2} \cdot \cancel{1}} \\ &= \frac{1}{r!} \cdot \frac{n!}{(n-r)!} \\ &= \frac{n!}{(r!)(n-r)!} \end{aligned}$$

Here the exclamation point stands for the "factorial" of a number, which is the product of all natural numbers from n down to 1. For example, $6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$. If you have 6 items and you want to know how many different ways there are to pick 4 items out of that set (where order does not matter), then you can use this "6 choose 4" formula to calculate it like this:

$$\binom{6}{4} = \frac{6!}{(4!)(6-4)!} = \frac{6!}{(4!)(2)!} = \frac{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{(4 \cdot 3 \cdot 2 \cdot 1)(2 \cdot 1)} = \frac{6 \cdot 5}{2} = 15$$

The binomial theorem uses this “ n choose r ” notation and the sigma notation to write the expansion of the binomial $(a + b)$ to the n -th power:

$$(a + b)^n = \sum_{r=0}^n \binom{n}{r} (a)^{n-r} (b)^r$$

Using this notation we can recompute the example of $(3x + 2y)^6$ from the previous section:

$$\begin{aligned} (3x + 2y)^6 &= \sum_{r=0}^6 \binom{6}{r} (3x)^{6-r} (2y)^r \\ &= \binom{6}{0} (3x)^{6-0} (2y)^0 + \binom{6}{1} (3x)^{6-1} (2y)^1 + \binom{6}{2} (3x)^{6-2} (2y)^2 + \\ &\quad \binom{6}{3} (3x)^{6-3} (2y)^3 + \binom{6}{4} (3x)^{6-4} (2y)^4 + \binom{6}{5} (3x)^{6-5} (2y)^5 + \\ &\quad \binom{6}{6} (3x)^{6-6} (2y)^6 \end{aligned}$$

Using your calculator’s $\mathbf{nCr}(n, r)$ function (in the **Probability** category under **Combinations**) to calculate the value of each $\binom{6}{k}$ expression, expanding the power of each term, and simplifying you’ll end up with the same answer we calculated above.

$$\begin{aligned} (3x + 2y)^6 &= 1 \cdot (3x)^6 + 6 \cdot (3x)^5 (2y)^1 + 15 \cdot (3x)^4 (2y)^2 \\ &\quad 20(3x)^3 (2y)^3 + 15(3x)^2 (2y)^4 + 6(3x)^{6-5} (2y)^5 \\ &\quad 1 \cdot (2y)^6 \\ &= 729x^6 + 2916x^5y + 4860x^4y^2 + 4320x^3y^3 + 2160x^2y^4 + 576xy^5 + 64y^6 \end{aligned}$$

1.23.1 Using The Binomial Theorem for a Single Term

A question like, “What is the coefficient of the x^2 term in $(3x^2 + \frac{1}{x})^7$?” is easy to answer using the Binomial Theorem, without having to do all the work to expand every term.

It tells us that for any r between 0 and 7, the r -th term will be:

$$\binom{7}{r} (3x^2)^{7-r} \left(\frac{1}{x}\right)^r$$

We can rewrite this r -th term by simplifying the exponents and writing $\frac{1}{x}$ as x^{-1} like this:

$$\begin{aligned} \binom{7}{r} (3x^2)^{7-r} \left(\frac{1}{x}\right)^r &= \binom{7}{r} 3^{7-r} x^{2(7-r)} (x^{-1})^r \\ &= \binom{7}{r} 3^{7-r} x^{14-2r} (x)^{-r} \\ &= \binom{7}{r} (3)^{7-r} (x)^{14-3r} \end{aligned}$$

So, we can find which value of r will produce the term with x^2 in the expansion by solving the simple equation based on the exponent $14 - 3r$ in the r -th term:

$$\begin{aligned} 14 - 3r &= 2 \\ 12 &= 3r \\ r &= \frac{12}{3} = 4 \end{aligned}$$

Therefore, the $r = 4$ term will be the one that has x^2 . Substituting $r = 4$ into the Binomial Theorem, we can compute the desired term of the expansion. **N.B.** Since the first term has $r = 0$, the $r = 4$ term is the *fifth* monomial in the expansion:

$$\begin{aligned} \binom{7}{4} (3x^2)^{7-4} \left(\frac{1}{x}\right)^4 &= 35 \cdot (3x^2)^3 \left(\frac{1}{x}\right)^4 \\ &= 35 \cdot 3^3 \cdot x^6 \frac{1}{x^4} \\ &= 35 \cdot 27 \cdot x^2 \\ &= 945x^2 \end{aligned}$$

Therefore, without expanding all of the terms, we have determined that 945 is the answer to the question “What is the coefficient of the x^2 term in $(3x^2 + \frac{1}{x})^7$?”

1.23.2 Example of Using Binomial Theorem

Consider the following problem:

$$\text{Given that } \left(1 + \frac{2}{3}x\right)^n (3 + nx)^2 = 9 + 84x + \dots, \text{ find the value of } n$$

Since the right side of the equation only provides us information about the units (9) and the x term ($84x$), we will focus our attention only on the terms with no x and x to the first power. We can ignore terms with x^2 or higher powers of x .

Using the Binomial Theorem, the first ($r = 0$) term and second ($r = 1$) term of the expansion is:

$$\begin{aligned} \left(1 + \frac{2}{3}x\right)^n &= \binom{n}{0} (1)^{n-0} \left(\frac{2}{3}x\right)^0 + \binom{n}{1} (1)^{n-1} \left(\frac{2}{3}x\right)^1 + \dots \\ &= 1 \cdot 1 \cdot 1 + \binom{n}{1} \cdot 1 \cdot \frac{2}{3}x + \dots = 1 + \binom{n}{1} \cdot \frac{2}{3}x + \dots \end{aligned}$$

So, we can substitute this expansion into the left-hand side of the equation in the problem, and simplify. Any time we get a power of x^2 while multiplying we can ignore that term since we only care about the units and the x term:

$$\begin{aligned} \left(1 + \frac{2}{3}x\right)^n (3 + nx)^2 &= \left(1 + \binom{n}{1} \frac{2}{3}x + \dots\right) (3 + nx)^2 \\ &= 1 \cdot (3 + nx)^2 + \binom{n}{1} \frac{2}{3}x (3 + nx)^2 + \dots \\ &= (9 + 6nx + n^2x^2) + \binom{n}{1} \frac{2}{3}x (9 + 6nx + n^2x^2) + \dots \\ &= (9 + 6nx + \dots) + \binom{n}{1} \frac{2}{3}x (9 + 6nx + \dots) + \dots \\ &= 9 + 6nx + \binom{n}{1} \frac{2}{3}x \cdot 9 + \binom{n}{1} \frac{2}{3}x \cdot 6nx + \dots \\ &= 9 + 6nx + 6 \binom{n}{1} x + 4 \binom{n}{1} nx^2 + \dots \\ &= 9 + \left(6n + 6 \binom{n}{1}\right) x + \dots \end{aligned}$$

At this point, we can set the coefficient on x equal to 84 and solve the problem:

$$\begin{aligned}6n + 6\binom{n}{1} &= 84 \\6\left(n + \binom{n}{1}\right) &= 84 \\ \left(n + \binom{n}{1}\right) &= \frac{84}{6} = 14\end{aligned}$$

Since $\binom{n}{1} = \frac{n!}{(1!(n-1)!)} = \frac{n(n-1)!}{(1)(n-1)!} = n$, we have:

$$\begin{aligned}\left(n + \binom{n}{1}\right) &= \frac{84}{6} = 14 \\(n + n) &= \frac{84}{6} = 14 \\2n &= 14 \\n &= 7\end{aligned}$$

Out of curiosity, Mathematica tells me that the complete expansion works out to:

$$\frac{6272x^9}{2187} + \frac{23744x^8}{729} + \frac{39328x^7}{243} + \frac{4144x^6}{9} + \frac{22232x^5}{27} + \frac{8596x^4}{9} + 714x^3 + 329x^2 + 84x + 9$$

1.24 Solving Systems of Three Linear Equations Using Substitution

Consider the function $f(x) = px^3 + qx^2 + rx$. The graph passes through the origin and the points $A(-2, -8)$, $B(1, -2)$, and $C(2, 0)$. Find three linear equations in p , q , and r , then solve the system of linear equations to find the values of p , q , and r .

By substituting in the x and y values of the three points into $f(x)$, we produce the following system of three linear equations in three variables:

$$\begin{array}{ll}-8p + 4q - 2r = -8 & \text{Equation using } A(-2, -8) \\p + q + r = -2 & \text{Equation using } B(1, -2) \\8p + 4q + 2r = 0 & \text{Equation using } C(2, 0)\end{array}$$

One way to solve this problem is to use the substitution method. Adding the first and third equations together gives:

$$\begin{aligned}-8p + 4q - 2r &= -8 \\+ (8p + 4q + 2r) &= 0 \\8q &= -8 \\q &= -1\end{aligned}$$

Then, once finding the value of q , you can substitute it into any two of the three equations to end up with two equations in two variables p and r . Picking the first and second equations and substituting in our value of q we get:

$$\begin{aligned} -8p - 2r &= -4 \\ p + r &= -1 \end{aligned}$$

We can solve for p using the second equation to get $p = -1 - r$ and then substitute that into the other equation to solve for r :

$$\begin{aligned} -8(-1 - r) - 2r &= -4 \\ 8 + 8r - 2r &= -4 \\ 6r &= -12 \\ r &= -2 \end{aligned}$$

So, now we know that $q = -1$ and $r = -2$, we substitute those into any of the three equations to find the remaining value p . Using the second equation of the original three, and substituting in the values of q and r that we know, we have:

$$\begin{aligned} p - 1 - 2 &= -2 \\ p &= 1 \end{aligned}$$

So the answer is $p = 1$, $q = -1$, and $r = -2$.

1.25 Solving Systems of Three Linear Equations Using Technology

You can use the TI Nspire CX calculator to easily solve a system of three linear equations like the one in the previous section. To access this functionality, use the “Menu” button while in a calculator document to pick the “Algebra → Solve Systems of Linear Equations...” option. As shown in Figure 3, enter the number of equations and a comma-separated list of the variables involved, then press the “OK” button to continue.

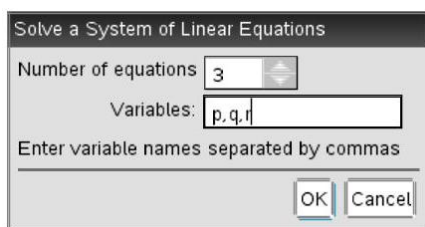


Figure 3: Entering Number of Equations and Variables Involved

Enter the three equations as shown in Figure 3, and press “Enter” to compute the solution.

$$\text{linSolve} \left(\begin{cases} -8p+4q-2r=-8 \\ p+q+r=-2 \\ 8p+4q+2r=0 \end{cases}, \{p,q,r\} \right)$$

Figure 4: Entering System of Three Linear Equations in p , q , and r variables

In the time it takes to type in the equations, your calculator produces the answer $\{1, -1, -2\}$. The order of the solutions matches the order of the variables you entered into the “Variables” field of the dialog box in the first step. Therefore, since you typed in p, q, r into that box earlier, the solution is $p = 1$, $q = -1$, and $r = -2$.

2 Functions and Equations

2.1 Sets

A set is a collection of elements. The set can contain no elements, a finite (i.e. countable) number of elements, or an infinite number of elements. If the set is empty, then we write:

$$S = \emptyset$$

If the set is finite, then its elements can be listed explicitly:

$$S = \{2, 3, 5, 7, 11\}$$

If there is an obvious pattern to the elements in the set, then rather than listing every one, you can provide an example of the list and use dots to show where the pattern continues. For example, the set T below contains all odd numbers between 1 and 101:

$$T = \{1, 3, 5, 7, 9, \dots, 101\}$$

If the set contains an infinite number of elements, the dots can indicate that the pattern goes on forever. For example, the set D below are all even numbers greater than or equal to 2:

$$D = \{2, 4, 6, 8, 10, \dots\}$$

A set can be described in words like the example below that defines the set P to be the set of all prime numbers:

$$P = \{n \in \mathbb{N} \mid n \text{ is prime}\}$$

We read this definition, “ P is the set of all natural numbers n such that n is prime.” To the left of the vertical “such that” bar, we provide a variable and indicate the set of numbers to which it belongs. On the right of the vertical bar, we provide a condition that indicates the “rule” that defines which elements of that set belong to the current set we are defining.

The vertical “such that” bar can be used with other kind of conditions to help define the elements of a set. For example, the set W below is defined to be the set of all real numbers between -2π and 2π (inclusive).

$$W = \{x \in \mathbb{R} \mid -2\pi \leq x \leq 2\pi\}$$

2.2 Union and Intersection of Sets

Addition and multiplication define two operations on numbers that we use to produce new numbers as a result. Analogously, we frequently use two operations on sets called union and intersection to produce new sets from existing ones. The *union* of two sets A and B is defined as the set that contains the elements that are either in A or in B (or in both). It is written using the “cup” symbol that looks like the letter u in the word **union**. For example:

$$\{1, 3, 5, 7\} \cup \{2, 3, 6\} = \{1, 2, 3, 5, 6, 7\}$$

Notice that even though the number 3 appears in both sets, it only occurs once in the set that is their union. Sets never contain repeated occurrences of the same element.

The *intersection* of two sets A and B is defined as the set that contains the elements that are in both A and B . It is written using the “cap” symbol that looks like an upside-down letter u . For example:

$$\{1, 3, 5, 7\} \cap \{2, 3, 6\} = \{3\}$$

If the sets contain no common elements, then their intersection is the empty set:

$$\{1, 3, 5, 7\} \cap \{2, 6\} = \emptyset$$

2.3 Common Sets of Numbers

In IB Maths Standard we work with four important sets of numbers. The set of natural numbers \mathbb{N} includes zero as well as the “counting numbers”, the positive whole numbers 1, 2, 3, etc. The set of integers \mathbb{Z} includes all of the natural numbers along with their negatives. The blackboard letter \mathbb{Z} used to denote this set comes from the German word for numbers *zahlen*. The set of rational numbers \mathbb{Q} contains elements formed by a *ratio* of two integers, that is, one number divided by another: in short, all fractions. By thinking of the fractions formed by an integer in the numerator and the number one (1) in the denominator, it’s easy to see that rationals also include all integers. The blackboard letter \mathbb{Q} naming this set comes from the word “quotient”, another word for one number divided by another. Finally, the set of real numbers \mathbb{R} includes all of the rational numbers, along with *irrational* numbers like π and $\sqrt{2}$ which cannot be represented exactly as a fraction.

When all the elements of one set A are contained in another set B , we say that A is a *subset* of B and write $A \subset B$. Given the definitions above, we have the following subset relationships between the common sets of numbers we will work with:

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$$

2.4 Intervals of Real Numbers

When working with the real numbers \mathbb{R} we often will work with subsets defined by the set of reals contained between a lower bound and an upper bound. For example, the set W of all real numbers between 0 and $\frac{\pi}{4}$, which above we learned could be written as:

$$W = \left\{ x \in \mathbb{R} \mid 0 \leq x \leq \frac{\pi}{4} \right\}$$

If we change the \leq symbol in the set’s definition to use $<$ instead, then we can define a subset J of W that does not include the lower bound 0 and does not include the upper bound $\frac{\pi}{4}$:

$$\left\{ x \in \mathbb{R} \mid 0 < x < \frac{\pi}{4} \right\} = J \subset W$$

A set defined by all real numbers between an upper bound and a lower bound is called an *interval*. The set W defined above is known as a *closed* interval because it contains its upper and lower bounds, while the set J is known as an *open* interval because it omits (excludes) its bounds. This idea of interval of real numbers is used so frequently that we have a more compact notation for it. We can rewrite the sets W and J using this interval notation like this:

$$W = \left[0, \frac{\pi}{4} \right] \quad J = \left(0, \frac{\pi}{4} \right)$$

The square brackets indicate a closed interval that includes its bounds, while the round brackets (parentheses) denote an open interval that excludes its bounds. Since the set of real numbers is often represented as a number line, the upper and lower bounds of an interval are usually called *endpoints* of the interval because they define the endpoints of a segment of the real number line. Using this terminology, a closed interval includes its endpoints while an open interval excludes them.

Of course, an interval can independently include or exclude its upper and lower endpoint, so we can write intervals using a mixture of the square brackets and round brackets like the example below. The interval C include its lower endpoint 0 but excludes its upper endpoint

$\frac{\pi}{4}$. Conversely, the interval D excludes its lower endpoint 0 but includes its upper endpoint $\frac{\pi}{4}$. These intervals are known as *semi-open*:

$$C = \left[0, \frac{\pi}{4}\right) \quad D = \left(0, \frac{\pi}{4}\right]$$

If an interval continues forever in the direction toward positive infinity or toward negative infinity, then we always write the side of the interval containing an infinity as an open interval end like this:

$$\begin{aligned} [-5, \infty) &= \{x \in \mathbb{R} \mid x \geq -5\} \\ (7, \infty) &= \{x \in \mathbb{R} \mid x > 7\} \\ (-\infty, \sqrt{2}] &= \{x \in \mathbb{R} \mid x \leq \sqrt{2}\} \\ (-\infty, \pi) &= \{x \in \mathbb{R} \mid x < \pi\} \\ (-\infty, \infty) &= \{x \in \mathbb{R}\} \end{aligned}$$

2.5 Concept of Function

A *function* is an operation that uniquely associates each element of an “input” set with a member of another “output” set. In IB Standard Level Maths, both the input and output sets are typically the set of all real numbers \mathbb{R} , or some subset of it like the positive real numbers or the interval $[0, 2\pi]$, for example.

We give each function a name like f or g and say that the function “maps” a value in the input set to a value in the output set. Using a variable like x to represent an arbitrary element of the input set, we can describe a function by the operation it performs on this arbitrary element. For example, a function named f that takes an arbitrary value x from its input set and raises the base 2 to the power x would be written:

$$f(x) = 2^x$$

We read the expression $f(x)$ as “ f of x ”. Figure 5 shows the function, its input set, and its output set in a pictorial form.

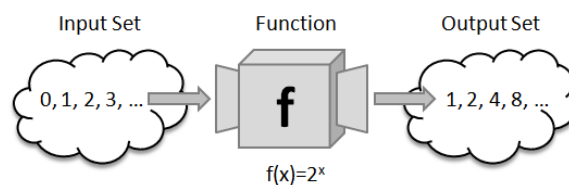


Figure 5: Function f associates elements of input set with output set

2.6 Graph of a Function

For functions that map elements of the real numbers to other real numbers ($\mathbb{R} \rightarrow \mathbb{R}$), we can visualize the operation the function performs by considering the set of all points (x, y) in the Cartesian plane (written in shorthand as $\mathbb{R} \times \mathbb{R}$ or \mathbb{R}^2) where the x -coordinate is in the input set of the function and the y -coordinate is the value $f(x)$ to which the function maps x . Consider the function $f(x) = \frac{\sqrt{x+4}}{x-4} + 1$. Figure 6 shows the visualization the Texas Instruments Nspire-CX calculator produces for the set of points defined by:

$$\left\{ (x, y) \in \mathbb{R}^2 \mid y = \frac{\sqrt{x+4}}{x-4} + 1 \right\}$$

This visualization is known as the *graph* of the function $f(x)$. Since the calculator supports graphing multiple functions on the same graph, it names the first function $f1(x)$, the second function $f2(x)$, and so on.

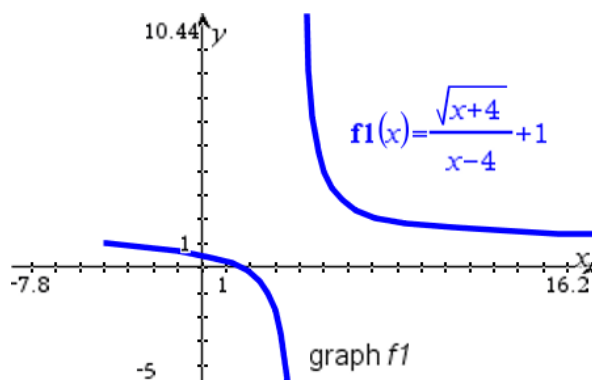


Figure 6: Graph of $f(x) = \frac{\sqrt{x+4}}{x-4} + 1$

2.7 Domain of a Function

A function f 's input set is called its *domain*. On some occasions, a function's domain may be limited on purpose by the problem statement to focus your attention on a particular interval such as $\{x \in \mathbb{R} \mid -2\pi \leq x \leq 2\pi\}$. However, more typically a function's domain will consist of the set of all real numbers \mathbb{R} , except where the function is not defined. Consider again the function in Figure 6 $f(x) = \frac{\sqrt{x+4}}{x-4} + 1$. The value 4 would cause the $x-4$ in the denominator to be zero, which would cause trouble for the function f because we cannot divide by zero. Furthermore, the numerator is only well-defined if the value under the square root is non-negative. In other words, it must be the case that $x+4 \geq 0$, so $x \geq -4$. So the domain of $f(x)$ is the set $\{x \in \mathbb{R} \mid x \geq -4 \text{ and } x \neq 4\}$. Another way of writing this using interval notation is $[-4, 4) \cup (4, \infty)$. Notice in Figure 6 that the graph of this function does not exist to the left of $x = -4$ and that it is undefined at $x = 4$ where there is a vertical asymptote. We can see the domain of the function f in its graph by scanning across the x -axis and noticing which values of x the function takes on along its graph.

2.8 Range of a Function

A function f 's output set is called its *range*. It consists of all elements that result by applying the function $f(x)$ to all elements x in the domain. We can see the range of the function f in its graph by scanning up and down the y -axis and noticing which values of y the function takes on along its graph. Looking again at Figure 6, we see that around the vertical asymptote at $x = 4$, the function's value goes down to negative infinity ($-\infty$) on the left of the asymptote and up to positive infinity (∞) on the right of the asymptote. So, by choosing a value for x that is appropriately close (but not equal to) 4, we can produce function values that are as large as we want in either positive or negative direction. So, it would appear that the range is all of the real numbers (\mathbb{R}). However, the function has a horizontal asymptote at $y = 1$. As the value of x gets larger and larger in the positive direction, the value of the function gets closer and closer to 1 from above, but never actually reaches the value 1. So, on second thought it would appear that the range of the function f from Figure 6 is all real numbers except 1, or written in set notation $\{y \in \mathbb{R} \mid y \neq 1\}$. However, notice that the function is defined at $x = -4$ and that $f(-4) = \frac{\sqrt{-4+4}}{-4-4} + 1 = \frac{0}{-8} + 1 = 1$. Therefore, even though the function never takes on the value 1 as the value of x gets larger and larger, it does equal 1 right at $x = -4$ which is part of

the domain. Therefore, the range is all real numbers (\mathbb{R}).

Consider another example of the function $f(x) = x^2$ shown in Figure 7.

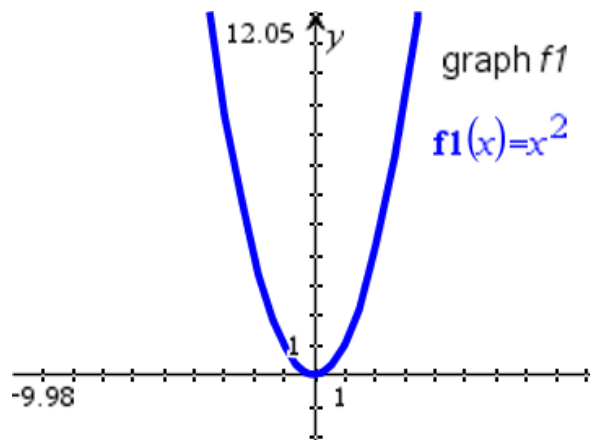


Figure 7: Graph of $f(x) = x^2$

Here there are no values in \mathbb{R} where the function is undefined, so the domain is all real numbers (\mathbb{R}). Scanning up and down along the y -axis, we see that the function's y values never go below zero. When the function is applied to any negative numbers in the domain, the result is a positive number. So, the range of this function is the set of all non-negative real numbers $\{y \in \mathbb{R} \mid y \geq 0\}$.

2.9 Composing One Function with Another

Given a value x in the domain of a function f , if you take the result of applying f to x and then apply another function g to that result, you end up with $g(f(x))$, which we read “ g of f of x ”. We call this “chaining” of functions *composition*. Consider two functions f and g defined as follows:

$$f(x) = 2^x \quad g(x) = x^2 + 1$$

Figure 8 shows what happens if we compose function g with function f , which means first applying $f(x)$ and then “chaining” that output into the input of $g(x)$ to produce $g(f(x))$. We write the function “ g composed with f ” using a circle symbol between them: $g \circ f$. We see that the result of composing g with f is the function:

$$g \circ f(x) = (2^x)^2 + 1$$

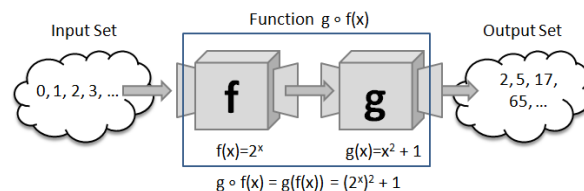


Figure 8: Composite function $g \circ f(x)$

On the other hand, as shown in Figure 9 if we first apply g to x and then “chain” that output $g(x)$ into the input of the function f , we produce a different resulting composed function $f \circ g(x)$:

$$f \circ g(x) = 2^{x^2+1}$$

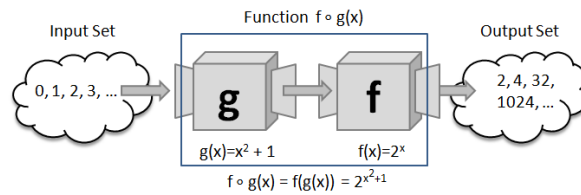


Figure 9: Composite function $f \circ g(x)$

So we see from this simple example that by changing the order of function composition, we produce different resulting functions. This tells us that the order of composition is not commutative like addition or multiplication is. In other words, in most cases:

$$f \circ g(x) \neq g \circ f(x)$$

2.10 Identity Function

The identity function maps a value in the domain to itself and is represented by the function $f(x) = x$.

2.11 Inverse Function

Given a function f , its inverse function f^{-1} “reverses” or “inverts” the mapping that f performs on its input x . If you first put a value x into the function f and then put that result into the function f^{-1} , then you get back the original value x . In other words, f composed with f^{-1} is the identity function.

$$f \circ f^{-1}(x) = f^{-1} \circ f(x) = x$$

2.12 Determining the Inverse Function as Reflection in Line $y = x$

Figure 10 shows how the e^x and $\ln x$ functions, which are inverses of each other, are a reflection in the line $y = x$. If you only need to graph the inverse, simply draw its reflection in the line $y = x$.

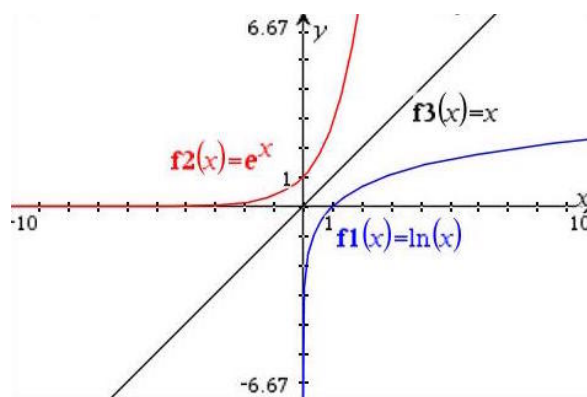


Figure 10: Inverse Functions e^x and $\ln x$ are Reflections in $y = x$

Another important fact to remember about inverse functions is that if the point (p, q) lies on the curve of function $f(x)$, then the point (q, p) lies on the curve of the inverse $f^{-1}(x)$. This is a direct corollary of the fact that any inverses are reflections in the line $y = x$.

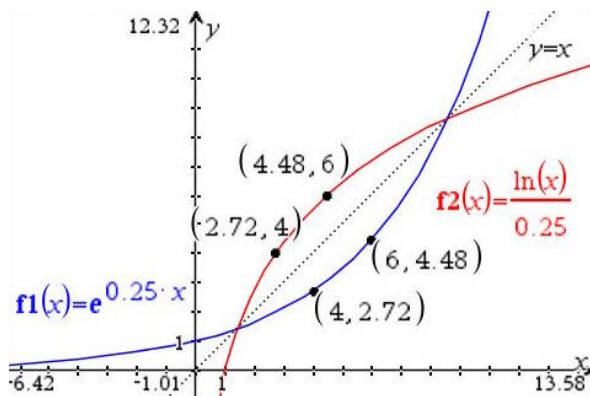


Figure 11: If (p, q) Lies on Graph of $f(x)$, then (q, p) Lies on Graph of Inverse $f^{-1}(x)$

Figure 11 shows the graphs of two functions which are inverses $f_1(x) = e^{0.25x}$ and $f_2(x) = \frac{\ln x}{0.25}$. Notice how, as expected, they are each others reflection in the line $y = x$. But also notice how the points $(4, 2.72)$ and $(6, 4.48)$ lie on the curve $f_1(x)$ and the points with inverted x - and y -coordinates $(2.72, 4)$ and $(4.48, 6)$ lie on the inverse curve $f_2(x)$. This makes sense because the definition of an inverse function is that it “undoes” the mapping that the other function did. If $f_1(x)$ maps the number 4 in the domain (x -axis) to the number 2.72 in the range (y -axis) – which is what saying “the point $(4, 2.72)$ is on curve f_1 ” means – then the inverse function must map 2.72 back to 4 in order that the original operation be “undone” to get back to the point you started with.

2.13 Determining the Inverse Function Analytically

Given a function $f(x)$ to determine its inverse function $f^{-1}(x)$ analytically, follow these steps:

1. Write the function using $y = \dots$ instead of $f(x) = \dots$
2. Switch all occurrences of x and y
3. Solve the new equation for y
4. Write the function using $f^{-1}(x) = \dots$ instead of $y = \dots$

For example, assume that $f(x) = \frac{x+1}{x+2}$ for $x > -2$. To find the inverse $f^{-1}(x)$, do these steps. Start by replacing $f(x)$ with y :

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

Then swap the x and y variables and solve for y :

$$\begin{aligned} x &= \frac{y+1}{y+2} \\ x(y+2) &= y+1 \\ xy+2x &= y+1 \\ 2x-1 &= y-xy \\ 2x-1 &= y(1-x) \\ \frac{2x-1}{1-x} &= y \end{aligned}$$

Finally, replace the y with $f^{-1}(x)$.

$$f^{-1}(x) = \frac{2x-1}{1-x}$$

Figure 12 shows the graphs of f and f^{-1} . Notice again that the graphs of inverse functions are reflections in the line $y = x$.

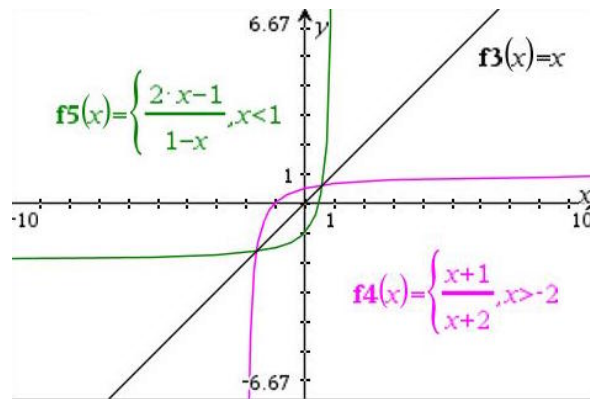


Figure 12: Inverse Functions are Reflections in $y = x$

2.14 Drawing and Analyzing Graphs with Your Calculator

Your graphing calculator can graph multiple functions and analyze the graphs to find minimum, maximum, x -intercept, y -intercept, where functions intersect, and where the function has “zeros” (which means where it intersects the x -axis, in other words, the line $y = 0$).

2.14.1 Drawing the Graph of a Function

On the TI-Nspire calculator, to draw a graph create a new document of type “Graph”. You can type one or more functions that the calculator names $f1(x)$, $f2(x)$, etc. The checkbox next to the function name indicates whether you want to show or hide the function in question. When entering the equation, it is best to use the dedicated buttons on the calculator to enter common constructs like $\ln(x)$, e^x , or $\sin x$ to ensure the calculator understands what you mean.

2.14.2 Restricting the Domain of a Graph

When typing in a function to graph on the TI-Nspire, by using the “such that” symbol (vertical bar), you can restrict the domain to limit the values of x for which the graph will be drawn. Figure 13 shows entering the function $x^2 - x - 1$ but restricting the domain to only between -2 and 2.

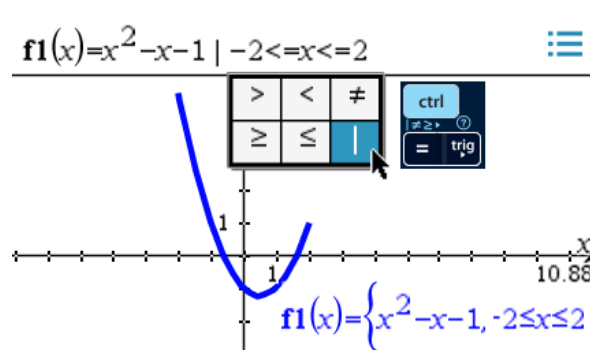


Figure 13: Restricting the Domain of a Graph Using the “Such That” Vertical Bar

2.14.3 Zooming Graph to See Exactly What You Want

After you've graphed a function, use the "Window/Zoom" option to scale the graph to see the portion you care to work with. The easiest option to use is the combination of "Zoom - Out" and then "Zoom - Box". Zoom out until you see all of the graph you care about, then zoom to the box that you want to fill the screen. You click in the upper left and lower right corners to define the bounding box of what you'd like to see in the window. Figure 14 shows the graph of $\frac{x^3 - 16x}{x^2 + 1}$ after first zooming out, and then zooming to a box to define the exact area to display.

NOTE: If you select a *rectangular* box with different width and height, then the calculator will scale the axes differently to make it fit the window. For this reason, if possible, it's best to select a box that is roughly square in shape so the axes will maintain the same scale.

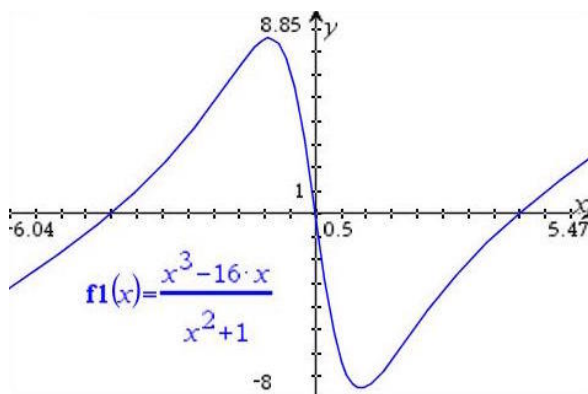


Figure 14: Graph After Zooming Out and then Zooming to Box

2.14.4 Finding a Maximum Value in an Interval

To find the maximum value of a graph in a given interval, use the "Menu" button to pick the "Analyze Graph → Maximum" option. You need to pick a lower bound and a higher bound of an interval in which the calculator should determine the maximum value. When the calculator prompts you to pick a "lower bound?" you click to position the slider just to the left of the point that visually appears to be the local minimum. Next, as shown in Figure 15, when it prompts you to pick an "upper bound?" you click to position the slider just to the right of that minimum. Once you've done this, the calculator adds the (x, y) coordinates of the maximum point in that interval to the graph.

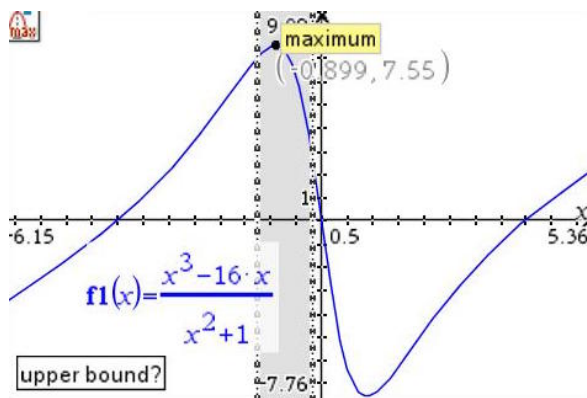


Figure 15: Selecting the Upper Bound in Which to Determine the Maximum

2.14.5 Finding a Minimum Value Value in an Interval

To find the minimum value of a graph in a given interval, use the “Menu” button to pick the “Analyze Graph →Minimum” option. The steps after that are similar to those in Section 2.14.4, “Finding a Maximum Value in an Interval”. Figure 16 shows the graph from the previous section after analyzing it to find both the maximum and the minimum values

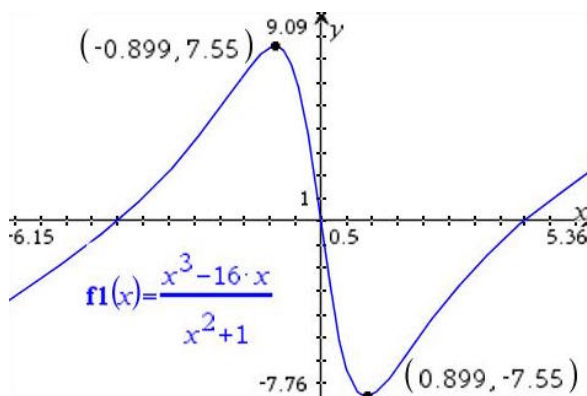


Figure 16: Graph Showing Minimum and Maximum Points

2.14.6 Finding the x -Intercepts or “Zeros” of a Graph in an Interval

To find the points where a graph crosses (or *intercepts*) the x -axis, for each x -intercept you see in the graph, repeat the following steps. Use the “Menu” button to pick the “Analyze Graph →Zero” option. Then click to define the lower and upper bounds of the interval around the zero you want to identify. The calculator will add that zero to the graph. Figure 17 shows the graph from the previous section after having repeated the “Analyze Graph →Zero” option three times.

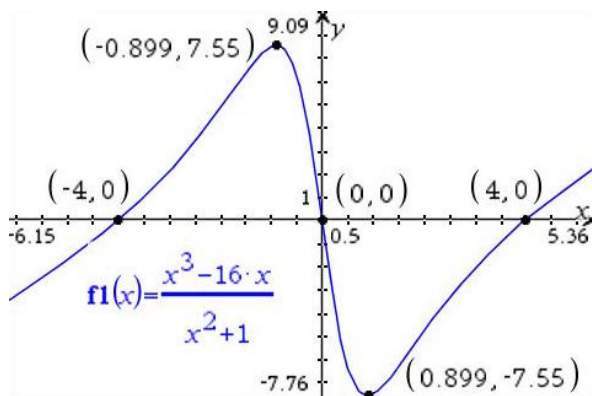


Figure 17: Graph Showing Zero Points

2.14.7 Finding the y -Intercept of a Graph

To find the points where a graph crosses (or *intercepts*) the y -axis, use the “Trace \rightarrow Graph Trace” option. You could use the arrow keys and press [enter] to place points on the graph, but it’s even easier to just type the value of x whose function value you want to see in the trace. Since we are looking for the y -intercept, that will be when the value of $x = 0$. Therefore, after picking the trace option, type the [0] key and press [enter] twice, followed by [esc] to signal you’re done adding trace points. As shown in Figure 18, which uses a slightly different equation for the graph to have a more interesting y -intercept, it places the (x, y) coordinate of the y -intercept on the graph. In your own work, when using the “Graph Trace” feature to add points to a curve, keep in mind that if you have multiple functions on your graph that you can use the up and down arrow keys to select which graph you want to trace before entering a particular x value for which to trace/plot a point.

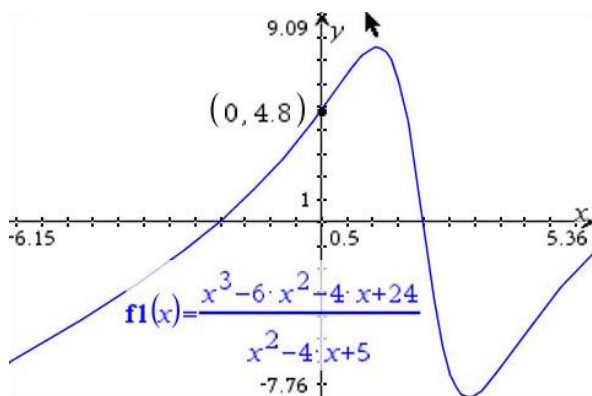


Figure 18: Graph Showing Y-Intercept Point

2.14.8 Vertical Asymptotes

To find vertical asymptotes, analyze where the function would be undefined. Typically this will be the value(s) of x that will make the expression in the denominator equal zero.

Consider the function $\frac{2x^2 - 3x - 6}{x^2 - 4}$. To find the values of x that make the denominator zero and cause the function to be undefined, solve the equation:

$$x^2 - 4 = 0$$

$$(x + 2)(x - 2) = 0$$

So $x = 2$ or $x = -2$. As shown in Figure 19 the graph of this function has vertical asymptotes as $x = 2$ and $x = -2$.

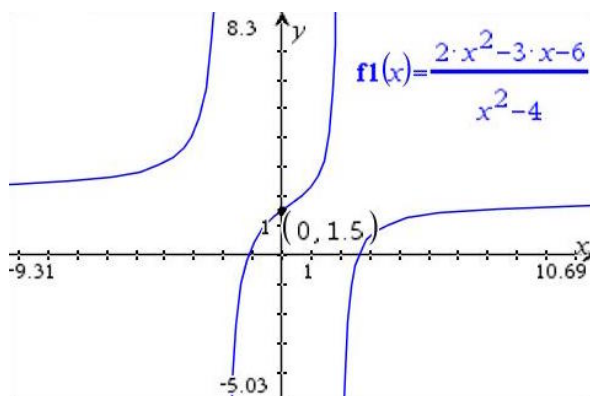


Figure 19: Graph with Two Vertical Asymptotes at -2 and 2

2.14.9 Graphing Vertical Lines

If you want to add a vertical line to a graph (perhaps to illustrate a vertical asymptote), the way you do it on the Nspire calculator is to choose:

“Graph Entry/Edit → Equation → Line → Vertical Line $x = c$ ”

Then you enter the value of x that you want the vertical line to pass through. After repeating this for the values $x = 2$ and $x = -2$ and using the graph attributes to make the line dotted, you end up with the graph in Figure 20.

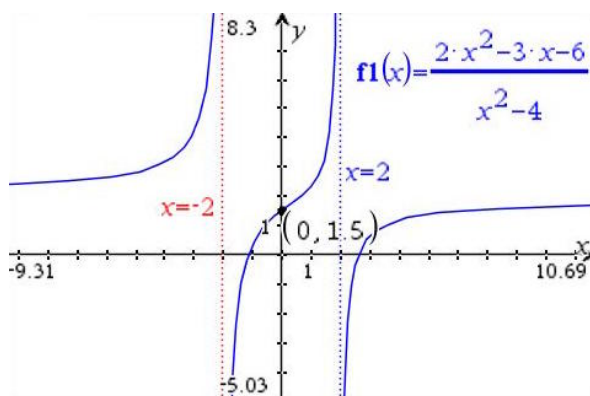


Figure 20: Dotted Vertical Lines Showing Vertical Asymptotes of Another Graph

2.14.10 Horizontal Asymptotes

Horizontal asymptotes occur when the function value approaches a fixed number as the value of x increases without bound, or decreases without bound. Said another way, as $x \rightarrow \infty$ if $f(x)$ gets closer and closer to some value a but never reaches it, then say that the function $f(x)$ has a horizontal asymptote at $y = a$. The way we determine whether such a value a exists is to consider the limit of the function as $x \rightarrow \infty$ and as $x \rightarrow -\infty$.

Consider the function in Figure 20, $\frac{2x^2 - 3x - 6}{x^2 - 4}$. When we take the limit of this function as $x \rightarrow \infty$, the highest powers of x in the numerator and denominator make a difference. So,

we can ignore the lower powers of x when thinking about the limit.

$$\lim_{x \rightarrow \infty} \frac{2x^2 - 3x - 6}{x^2 - 4} = \lim_{x \rightarrow \infty} \frac{2x^2}{x^2} = 2$$

The same limit occurs if we let $x \rightarrow -\infty$ because the even powers of x cause the result to be positive like before. So in both cases, the limit as $x \rightarrow \pm\infty$ is 2. So the graph of this function has a horizontal asymptote at $y = 2$.

2.14.11 Tips to Compute Horizontal Asymptotes of Rational Functions

While calculating the limit for $x \rightarrow \pm\infty$ in Section 2.14.10, “Horizontal Asymptotes”, we looked only at the highest power of x in the numerator and the denominator, since its value will dominate as x gets larger and larger. In that section’s example, the highest power in the numerator and denominator was the same (i.e. 2) so these canceled out and the value of the limit was given by the ratio of the coefficients. However, if the numerator’s highest power of x is larger than the denominator’s highest power of x , then there is no limit. That is, the limit will be ∞ because the numerator gets larger faster than the denominator. In contrast, if the numerator’s highest power of x is smaller than the denominator’s highest power of x , then the limit is zero (0) since the denominator gets larger faster than the numerator.

2.14.12 Graphing Horizontal Lines

If you want to add a horizontal line to a graph (perhaps to illustrate a horizontal asymptote), the way you do it on the Nspire calculator is to choose:

“Graph Entry/Edit → Equation → Line → Line Slope Intercept $x = c$ ”

Then you enter the value of zero (0) in the first box as the coefficient for x and the value a that you want the y -intercept to be. After doing this for the values $y = 0 \cdot x + 2$ and using the graph attributes to make the line dotted, you end up with the graph in Figure 21.

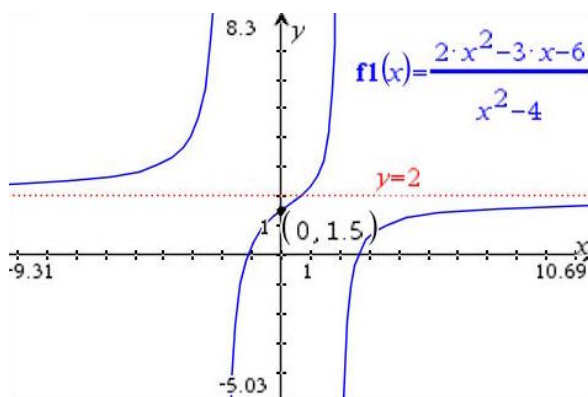


Figure 21: Dotted Vertical Lines Showing Vertical Asymptotes of Another Graph

2.14.13 Symmetry: Odd Functions

A function f that is symmetric about the origin is called an *odd function*. For each x in the domain of f , an odd function satisfies the condition that $f(-x) = -f(x)$. That is, if you put in the value -5 for example, you get the opposite (i.e. negative) of the value you would get if you put positive 5 into the function. For an odd function, this happens for any value of x .

The functions are called “odd” because this effect occurs for some functions where the highest

power of x is an odd number. Not all such equations are odd functions, however. A line that passes through the origin is an odd function. The power on the x in a linear equation is the odd number 1. A cubic equation that passes through the origin is also an odd function. The power on the x in a linear equation is the odd number 3. Figure 22 shows examples of a linear and cubic equation that are odd functions because they are symmetric about the origin.

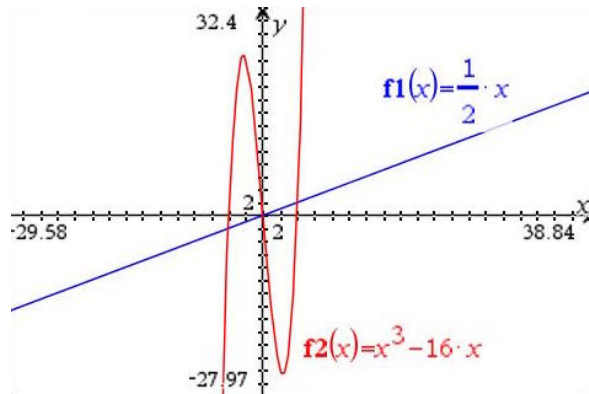


Figure 22: Linear and Cubic Equations Through the Origin are Odd Functions

2.14.14 Symmetry: Even Functions

A function f that is symmetric about the y -axis is called an *even function*. For each x in the domain of f , an even function satisfies the condition that $f(-x) = f(x)$. That is, if you put in the value -5 for example, you get the same value you would get as if you put positive 5 into the function. For an even function, this happens for any value of x .

The functions are called “even” because this symmetry effect occurs for some functions where highest power of x is an even number. Not all such equations are even functions, however. An absolute value function can be an even function, as well as a quadratic equation if its vertex lies on the y -axis. Figure 23 shows examples of a linear and cubic equation that are odd functions because they are symmetric about the origin.

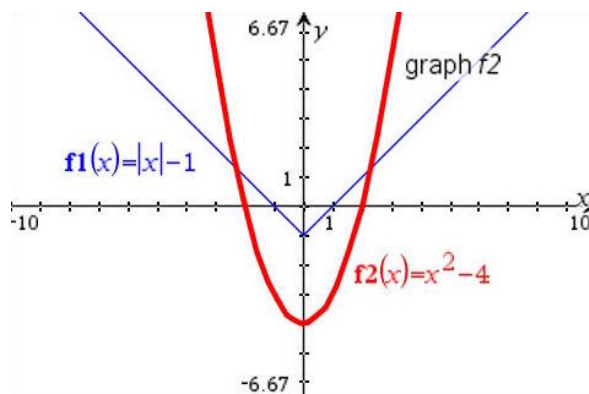


Figure 23: Absolute Value and Quadratic Examples of Even Functions

2.14.15 Solving Equations Graphically

Very often we are presented with two functions and asked to find the value x that makes both functions equal. Of course, for some situations, we have existing techniques to find the solution. If both functions are lines, we can use substitution to solve the system of linear equations. Even

if both functions are quadratic like:

$$\begin{aligned}f1(x) &= x^2 - x - 4 \\f2(x) &= 6x^2 + x - 10\end{aligned}$$

we can create an equation by setting the function $f1(x)$ equal to the function $f2(x)$ and then solve this equation like any quadratic we're familiar with:

$$\begin{aligned}x^2 - x - 4 &= 6x^2 + x - 10 \\0 &= 5x^2 + 2x - 6\end{aligned}$$

Then we use the quadratic formula (with $a = 5$, $b = 2$, and $c = -6$) to find the two values of x that solve the equation:

$$\begin{aligned}x &= \frac{-2 \pm \sqrt{2^2 - 4 \cdot 5 \cdot (-6)}}{2 \cdot 5} \\&= \frac{-2 \pm \sqrt{124}}{10} \\&= \frac{-2 \pm 11.136}{10} \\&= -1.314 \text{ or } 0.914\end{aligned}$$

However, if you don't have an easy analytic way to solve the equation involving the two functions, then fall back to using a surefire visual technique using your graphing calculator to get the solution. When presented with two functions $f1(x)$ and $f2(x)$, of course their graphs may or may not intersect. If they *do* intersect, then this means that those points of intersection belong to both graphs. In other words, when function $f1(x)$ and $f2(x)$ intersect at the point (p, q) , that means that $f1(p) = q = f2(p)$. Therefore, finding where the two functions intersect is a graphical way to solve the equation $f1(x) = f2(x)$.

Consider the two non-trivial equations below:

$$\begin{aligned}f1(x) &= \frac{1}{2}x^5 - 4x^3 + x + 7 \\f2(x) &= e^x + 5\end{aligned}$$

Even if we set these two equations equal, we don't know any analytical techniques to solve for x . Here, the graphing calculator comes to the rescue. We simply graph $f1(x)$ and $f2(x)$ and use the calculator's "Analyze Graph \rightarrow Intersection" feature to compute the coordinates of the points where the two functions intersect. The x coordinate values of these point we can find visually in this way are the solutions to the equation $f1(x) = f2(x)$. The feature works similarly to finding the minimum, maximum, or zero points. The calculator asks you to indicate the lower bound and upper bound of an interval in which you want its help in computing the intersection. You can repeat the process for each of the places that your eyes see the graphs cross. Figure 24 shows the result of following this procedure to compute the coordinates of the three places where $f1(x) = f2(x)$. The solutions to this equation are the x values from the three coordinates: $x \in \{-2.81, 0.59, 3.02\}$.

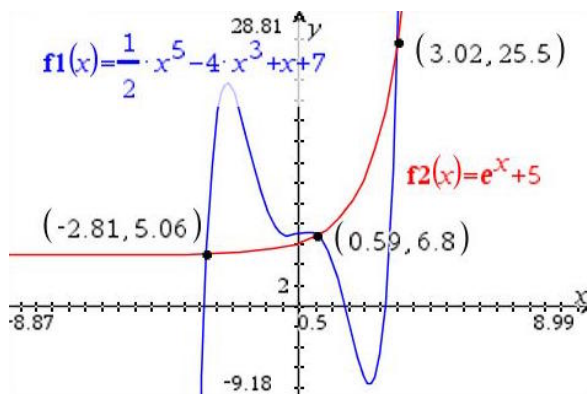


Figure 24: Finding Points where $f_1(x) = f_2(x)$ Graphically

NOTE: If your two curves intersect in multiple points, instead of repeatedly using “Analyze Graph → Intersection” try the “Geometry → Points & Lines → Intersection Point(s)” instead. After clicking on two curves on your graph, this alternative finds all their points of intersection in a single command.

2.15 Transformations of Graphs

There are four main kinds of transformations you can do to a graph:

1. *Translate* it a given number of units horizontally and/or vertically
2. *Reflect* it in the x -axis or the y -axis
3. *Vertically Stretch* it by a given factor
4. *Horizontally Stretch* is by a given factor

Each of these transformations corresponds to an analytical change that you make to the original function’s definition to achieve the modification to the graph. The following sections describe each one in turn.

2.15.1 Horizontal and Vertical Translations

Given a function $f(x)$, to translate it b units to the right, use $f(x - b)$. To translate it b units to the left, use $f(x - (-b)) = f(x + b)$. Figure 25 shows these two kinds of horizontal translation.

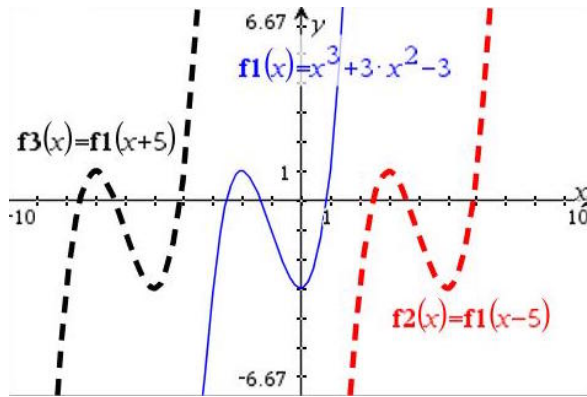


Figure 25: Horizontal Translation of $f_1(x)$ to Left and Right

Given a function $f(x)$, to translate it b units up, use $f(x) + b$. To translate it b units down, use $f(x) + (-b) = f(x) - b$. Figure 26 shows both kinds of this vertical translation.

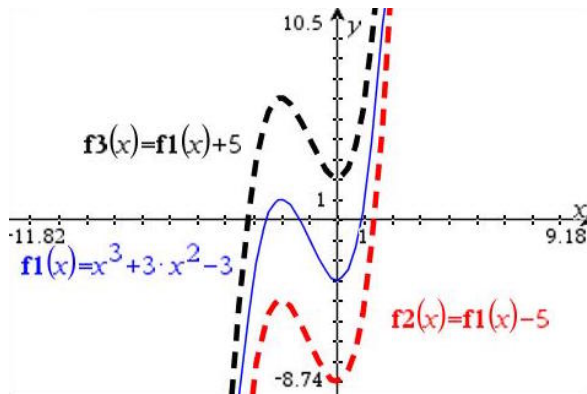


Figure 26: Vertical Translation of $f_1(x)$ Up and Down

2.15.2 Vertical Reflection

Given a function $f(x)$, to reflect it vertically – causing what was “up” to become “down” – multiply the function by negative one. In other words, use $-f(x)$. The line of symmetry around which you reflect the function will be a *horizontal* line of symmetry, parallel (or in some cases, exactly equal to) the x -axis.

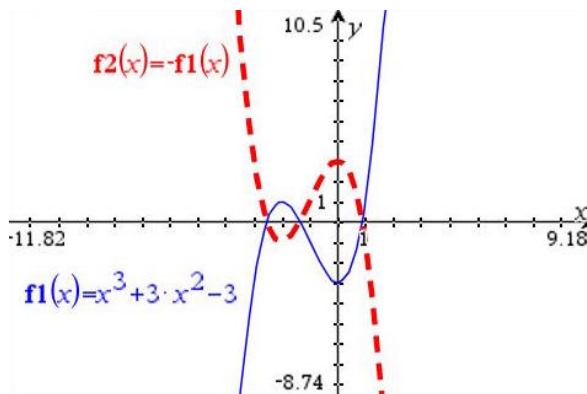


Figure 27: Reflection of $f_1(x)$ in the x -axis

2.15.3 Horizontal Reflection

Given a function $f(x)$, to reflect it horizontally – causing what was “left” to become “right” – multiply the input to the function by negative one. In other words, use $f(-x)$. The line of symmetry around which you reflect the function will be a *vertical* line of symmetry, parallel (or in some cases, exactly equal to) the y -axis.

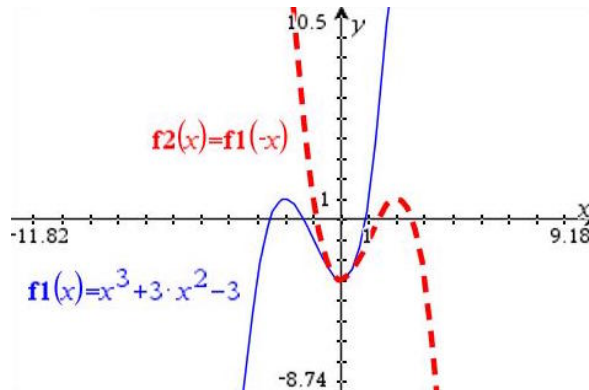


Figure 28: Reflection of $f1(x)$ in the y -axis

2.15.4 Vertical Stretch

Given a function $f(x)$ to stretch it by scale factor b in the vertical direction, use $b \cdot f(x)$. Notice in the example in Figure 29 that if $|b| > 1$ then the deformation is a stretch. If $|b| < 1$ then the deformation is a shrink (or squash).

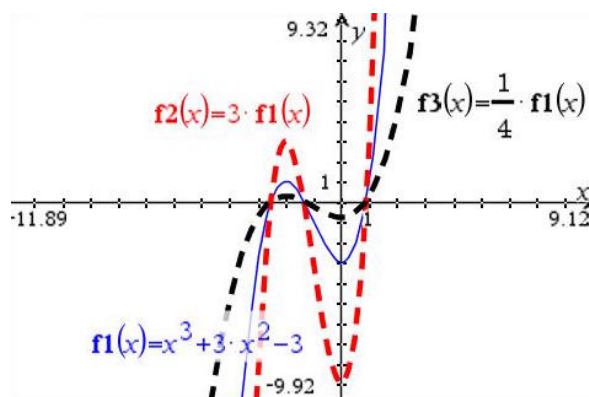


Figure 29: Vertical Stretch of $f1(x)$ by factor of 3 and $\frac{1}{4}$

2.15.5 Horizontal Stretch

Given a function $f(x)$ to stretch it by scale factor b in the horizontal direction, use $f(\frac{1}{b} \cdot x)$.

Notice in the example in Figure 30 that if $|b| > 1$ (meaning $|\frac{1}{b}| < 1$) the deformation is a stretch. If $|b| < 1$ (meaning $|\frac{1}{b}| > 1$) the deformation is a shrink (or squash).

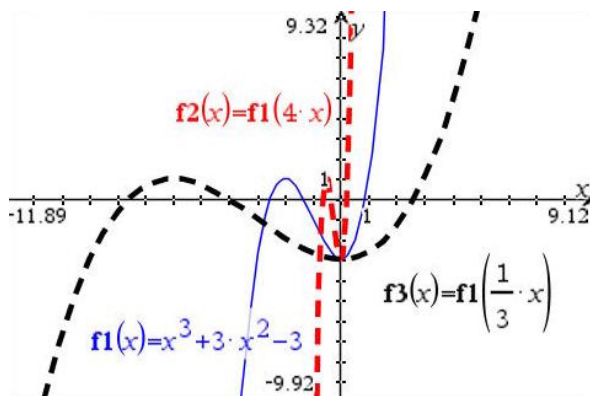


Figure 30: Horizontal Stretch of $f_1(x)$ by factor of 3 and $\frac{1}{4}$

2.15.6 Order Matters When Doing Multiple Transformations in Sequence

You can apply multiple transformation to a graph by applying the operations one after another. As we'll see, even in this very simple example, the order in which you carry out the operations is important because changing the order can produce different results. For this example, we'll start with a function $f(x) = x^3 + 3x^2 - 3$ and try two different experiments:

1. Translate $f(x)$ up by 5 units, then reflect the result in the x -axis
2. Reflect $f(x)$ in the x -axis, then translate the result up by 5 units

To carry out the first experiment, use your graphing calculator to define $f_1(x) = x^3 + 3x^2 - 3$. Then define $f_2(x) = f_1(x) + 5$. This represents a vertical translation by 5 units and the new function $f_2(x)$ represents the result of this transformation operation. Finally, define $f_3(x) = -f_2(x)$. This reflects the result of the first transformation in the x -axis. Figure 31 shows the $f_3(x)$ graph that results from these two consecutive transformations in this order.

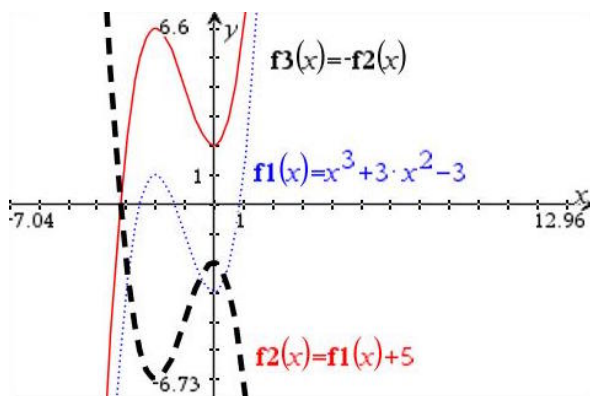


Figure 31: First Translate Up by 5, Then Reflect in x -axis

To carry out the second experiment, use your graphing calculator to change $f_2(x) = -f_1(x)$. This reflects the initial $f_1(x)$ function in the x -axis as the first transformation. Next, change $f_3(x) = f_2(x) + 5$. This translates the result of the first transformation up by 5 units. Figure 32 shows the different resulting $f_3(x)$ function graph obtained by these two consecutive transformations in this different order.

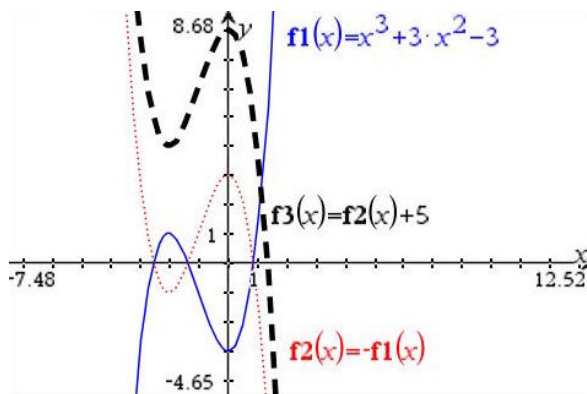


Figure 32: First Reflect in x -axis, Then Translate Up by 5

2.15.7 Graphing the Result of a Sequence of Transformations

In this section we study an example of doing six operations in sequence to understand how to graph the result of this combined transformation. Starting with the original function $f(x) = \frac{1}{2}x^2 - 4x + 6$, we'll do the following six transformation operations in this order:

1. Translate left by 1 unit
2. Reflect in the y -axis
3. Stretch horizontally by factor of 2
4. Translate down by 2 units
5. Reflect in the x -axis
6. Stretch vertically by factor of $\frac{1}{3}$ (resulting in a shrink/squash)

To carry out the work, we start by graphing the function $f_1(x) = \frac{1}{2}x^2 - 4x + 6$, then we define the following additional functions to correspond to the six operations in order:

- | | |
|---------------------------------|--|
| 1. $f_2(x) = f_1(x + 1)$ | Translate left by 1 unit |
| 2. $f_3(x) = f_2(-x)$ | Reflect in the y -axis |
| 3. $f_4(x) = f_3(\frac{1}{2}x)$ | Stretch horizontally by factor of 2 |
| 4. $f_5(x) = f_4(x) - 2$ | Translate down by 2 units |
| 5. $f_6(x) = -f_5(x)$ | Reflect in the x -axis |
| 6. $f_7(x) = \frac{1}{3}f_6(x)$ | Stretch vertically by factor of $\frac{1}{3}$ (resulting in a shrink/squash) |

As you enter each equation, by default the calculator will graph them, but you can hide the graphs of the interim functions f_2 , f_3 , f_4 , f_5 , and f_6 to leave only the original graph $f_1(x)$ and the final $f_7(x)$. The easiest way to hide the interim graphs is to revisit the function entry area and uncheck the ones you don't want to see. Figure 33 shows the result of graphing the initial $f_1(x)$ function as well as the final f_7 function.

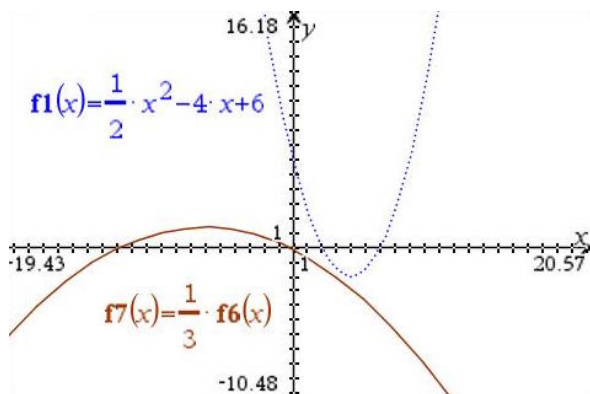


Figure 33: Function $f7(x)$ Results After Six Transformation Operations to Function $f1(x)$

If you are required to state the resulting, simplified equation for the final transformed graph, then you need to work backwards and in each step substitute the function with its definition until you get back to an expression in x that you can simplify completely. The steps involved to get the transformed equation in the above example go like this:

$$\begin{aligned}
 f7(x) &= \frac{1}{3}f6(x) \\
 &= \frac{1}{3}(-f5(x)) \\
 &= \frac{1}{3}(-(f4(x) - 2)) \\
 &= \frac{1}{3}\left(-\left(f3\left(\frac{1}{2}x\right) - 2\right)\right) \\
 &= \frac{1}{3}\left(-\left(f2\left(-\frac{1}{2}x\right) - 2\right)\right) \\
 &= \frac{1}{3}\left(-\left(f1\left(-\frac{1}{2}x + 1\right) - 2\right)\right) \\
 &= \frac{1}{3}\left(-\left(\frac{1}{2}\left(-\frac{1}{2}x + 1\right)^2 - 4\left(-\frac{1}{2}x + 1\right) + 6 - 2\right)\right) \\
 &= \frac{1}{3}\left(-\left(\frac{1}{8}x^2 - \frac{x}{2} + \frac{1}{2} + 2x\right)\right) \\
 f7(x) &= -\frac{1}{24}x^2 + \frac{x}{2} - \frac{1}{6}
 \end{aligned}$$

2.15.8 Determining Point Movement Under a Sequence of Transformations

In this section we continue with the example from Section 2.15.7, “Graphing the Result of a Sequence of Transformations”. To determine exactly where a sequence of transformations moves a *particular* point on the original graph, we cannot simply take the x -coordinate of that original point and put it in into the resulting, new function. To see why that does not work, let’s look at an example. Figure 34 shows the vertex $(4, 2)$ of the original $f1(x)$ function. Intuitively by translating and reflecting and stretching, the vertex of the original parabola should have moved to the vertex of the resulting parabola. However, if we use the calculator’s “Graph Trace” to add a point corresponding to $(4, f7(4))$, we see that the point on the resulting curve with x -coordinate of 4 is not the new curve’s vertex. So, we need a different technique to get the right answer.

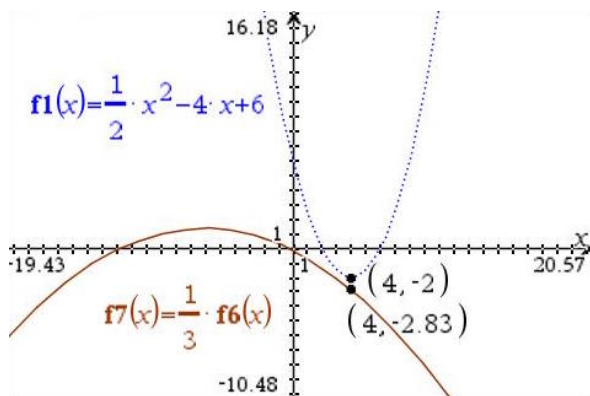


Figure 34: Visualizing Why Plugging in Original Point's x -Coordinate Doesn't Work

The correct way to determine where a particular point moves under sequence of transformations is to consider what each transformation in the sequence does to the x -coordinate or the y -coordinate of that point. The approach works for any point, but let's use our original parabola's vertex $(4, -2)$ as the example whose progress we follow through each step of the sequence of transformations. In each step below, we adjust either the x -coordinate or the y -coordinate based on what the current transformation would do to it. Horizontal translations or stretches only affect the x -coordinate. Vertical translations or stretches only affect the y -coordinate. A reflection in the x -axis flips the sign of the y -coordinate. A reflection in the y -axis flips the sign of the x -coordinate.

Starting with our vertex point $(4, -2)$, we proceed to record that each transformation step does to its coordinates:

1. $(4, -2) \rightarrow (\boxed{3}, -2)$ Translate left by 1 unit
2. $(3, -2) \rightarrow (\boxed{-3}, -2)$ Reflect in the y -axis
3. $(-3, -2) \rightarrow (\boxed{-6}, -2)$ Stretch horizontally by factor of 2
4. $(-6, -2) \rightarrow (-6, \boxed{-4})$ Translate down by 2 units
5. $(-6, -4) \rightarrow (-6, \boxed{4})$ Reflect in the x -axis
6. $(-6, 4) \rightarrow (-6, \boxed{\frac{4}{3}})$ Stretch vertically by factor of $\frac{1}{3}$ (resulting in a squash)

After repeating the same steps for the graph's y -intercept point $(0, 6)$ and the graph's two zeros at $(2, 0)$ and $(6, 0)$, Figure 35 shows the result which looks a lot more intuitive. The original y -intercept of $f1$ at $(0, 6)$ is mapped by the six-step transformation to the point $(2, -\frac{4}{3})$ on the resulting $f7$ curve. The original zeros of $f1$ at $(2, 0)$ and $(6, 0)$ are mapped to points on the resulting curve that are no longer on the x -axis.

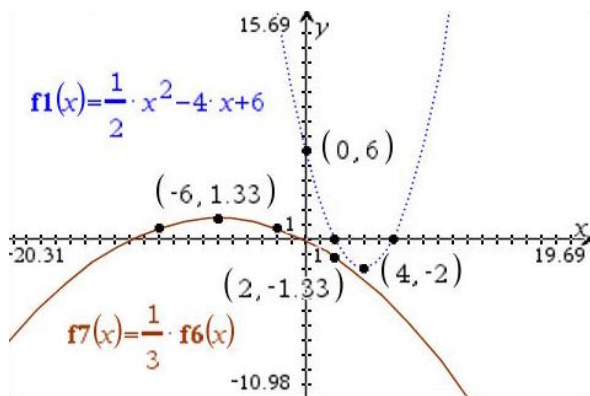


Figure 35: Vertex, y -intercept, and Zeros of f_1 Plotted on Transformed f_7

2.15.9 Vector Notation for Function Translation

As an alternative to expressing a translation in words like “function f is transformed p units to the right and q units up” sometimes you’ll see the translation expressed in vector notation like this:

$$\text{function } f \text{ undergoes translation } \begin{pmatrix} p \\ q \end{pmatrix}$$

The p is in the x -coordinate of the vector, so it represents the horizontal translation. A positive number p in this notation means “translated p units to the right”. The q is in the y -coordinate of the vector, so it represents the vertical translation. A positive number q in this notation means “translated q units up”.

Let’s consider an example that uses this vector notation and asks to follow a specific point through a transformation of functions. Let f and g be functions such that $g(x) = 2f(x+1) + 5$. The graph of f is mapped to the graph of g under the following transformations:

$$\text{vertical stretch by a factor of } k, \text{ followed by a translation } \begin{pmatrix} p \\ q \end{pmatrix}.$$

First, write down the values of k , p , and q . Then, consider the function $h(x) = -g(3x)$. The point $A(6, 5)$ on the graph of g is mapped to the point A' on the graph of h . Find A' .

Looking at the function $g(x)$, we see that starting with the original $f(x)$ that it has been:

1. translated horizontally left by 1 $\rightarrow f(x - (-1))$
2. stretched vertically by 2 $\rightarrow 2f(x - (-1))$
3. translated vertically up by 5 $\rightarrow 2f(x - (-1)) + 5$

So this tells us that $k = 2$, $p = -1$, and $q = 5$. Next, we look closely at the definition of $h(x)$ in terms of $g(x)$. We see that it has been:

1. $g(3x)$ \rightarrow stretched horizontally by $\frac{1}{3}$
2. $-g(3x)$ \rightarrow reflected in the x -axis

NOTE: In step 1 above, recall from Section 2.15.5, “Horizontal Stretch” that to stretch $f(x)$ horizontally by a factor of b , we multiply the argument of the function by the *reciprocal* of b . That is, $f(\frac{1}{b}x)$. Therefore if we see $h(x) = g(3x)$, this means that $\frac{1}{b} = 3$ and therefore $b = \frac{1}{3}$. In other words, if the function $h(x) = g(3x)$ then in words we say that “ $h(x)$ has horizontally stretched $g(x)$ by a factor of $\frac{1}{3}$ ”. When you “stretch” something by a factor less than one, it is sometimes called a “squash”.

We now use this information to follow what happens to our original point $A(6, 5)$ on $g(x)$ under the ordered steps of this transformation that produces the new function $h(x)$.

1. $(6, 5) \rightarrow (\boxed{2}, 5)$ stretched horizontally by $\frac{1}{3}$
2. $(2, 5) \rightarrow (2, \boxed{-5})$ reflected in the x -axis

Figure 36 illustrates three possible functions f , g , and h that satisfy this example, calling them f_1 , f_2 , and f_3 , respectively. It also depicts the original point A and the resulting point A' .

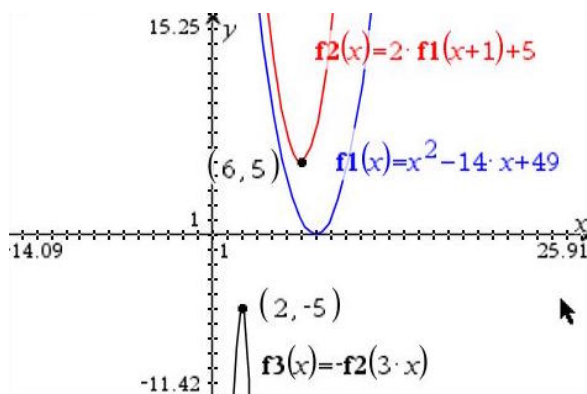


Figure 36: Original Function f_1 and Transformed Functions f_2 and f_3

2.16 Quadratic Functions

Quadratic functions have the form $f(x) = ax^2 + bx + c$ where a , b , and c are real number constants. The coefficient of the x^2 term needs to be non-zero or else the function would be a linear function, not quadratic. Figure 37 shows the graph of the quadratic function $\frac{1}{2}x^2 - 4x + 6$.

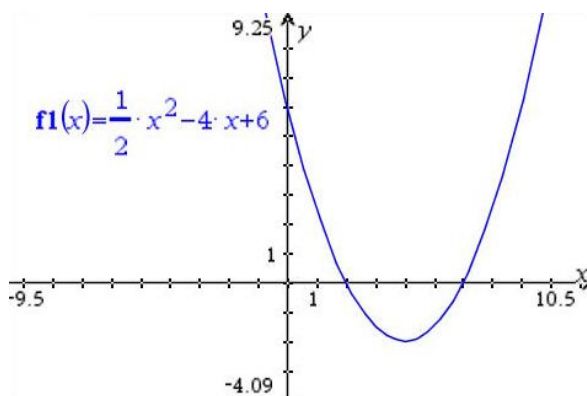


Figure 37: Example of a Quadratic Function

2.16.1 Using the Quadratic Formula to Find Zeros of Quadratic Function

The quadratic formula gives an explicit way to compute the zeros of (i.e. solutions to) any quadratic equation $ax^2 + bx + c = 0$. The zeros are always a formula of the coefficients a , b , and c as follows:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

In the example function shown in Figure 38, we have $a = \frac{1}{2}$, $b = -4$, and $c = 6$, so the values of its zeros using the quadratic formula are:

$$\begin{aligned} x &= \frac{-(-4) \pm \sqrt{(-4)^2 - 4 \cdot \frac{1}{2} \cdot 6}}{2 \cdot \frac{1}{2}} \\ &= \frac{4 \pm \sqrt{16 - 12}}{1} \\ &= 4 \pm \sqrt{4} \\ &= 4 \pm 2 \end{aligned}$$

Therefore, $x = 4 - 2 = 2$ or $x = 4 + 2 = 6$.

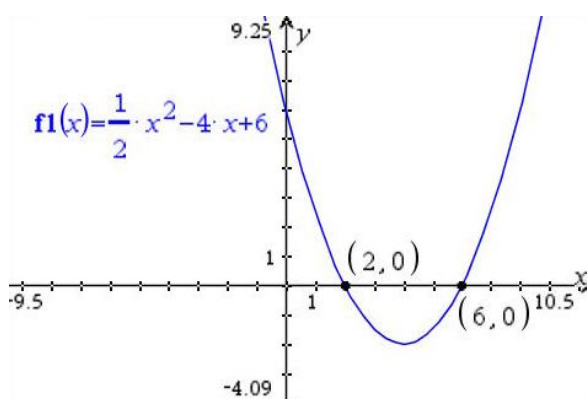


Figure 38: Two Distinct Zeros of Quadratic Function Found Using Quadratic Formula

2.16.2 Finding the Vertex If You Know the Zeros

Suppose that you already know the zeros x_1 and x_2 of a quadratic equation, by any of the following means:

- You factored it and determined which values of x make each factor zero
- You used the quadratic equation
- You read off the values because the function was already in x -intercept form

Then, it is very easy to find the x -coordinate of the vertex.

- If there is only a single root (also known as “two equal roots”), then that root is the x -coordinate of the vertex.
- If there are two distinct roots, then the vertex is halfway between the two roots. To find the point that is halfway between the roots, just take the average of the roots: add them together and divide by two.

Once you have the x -coordinate of the vertex, call it h , compute the y -coordinate k of the vertex by computing $f(h)$.

2.16.3 Graph and Axis of Symmetry

The graph of a quadratic function forms a parabola, and it is symmetric about the vertical line $x = h$ that passes through its vertex (h, k) . We can derive information about the symmetry of the function by rearranging the quadratic formula from Section 2.16.1, “Using the Quadratic Formula to Find Zeros of Quadratic Function”. If we separate the two terms in the numerator by writing each as a fraction over the common denominator of $2a$, we get:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$
$$x = \frac{-b}{2a} \pm \frac{\sqrt{b^2 - 4ac}}{2a}$$

This tells us that the zeros are always computed by taking the quantity $\frac{-b}{2a}$ and adding to it or subtracting from it the same quantity $\frac{\sqrt{b^2 - 4ac}}{2a}$. This means that the axis of symmetry for any quadratic function will always be the line $x = \frac{-b}{2a}$. Figure 39 shows the axis of symmetry of the function $\frac{1}{2}x^2 - 4x + 6$ lies on the line:

$$x = \frac{-b}{2a} = \frac{-(-4)}{2(\frac{1}{2})} = \frac{4}{1}$$
$$x = 4$$

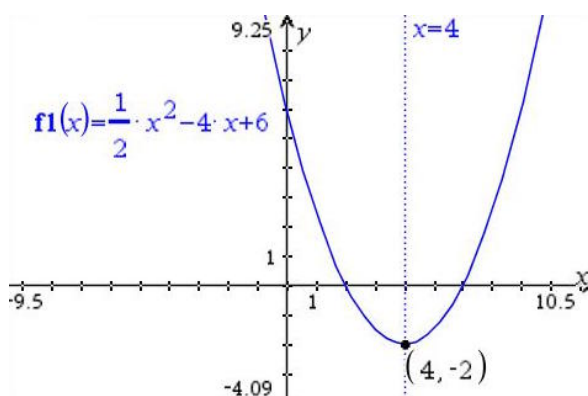


Figure 39: Quadratic Function’s Axis of Symmetry $x = \frac{-b}{2a}$

2.16.4 Computing the Vertex From the Coefficients

Since the parabola is symmetric about its vertex, and in Section 2.16.3, “Graph and Axis of Symmetry” we learned that a quadratic function’s axis of symmetry is always the line $x = \frac{-b}{2a}$, we can infer an easy way to calculate the coordinates of its vertex. The vertex lies on the line of symmetry, so its x -coordinate must be $\frac{-b}{2a}$. But the vertex lies on the function $f(x)$ ’s graph, so that means that the y -coordinate of the vertex must be the function f ’s value at this point $x = \frac{-b}{2a}$.

Therefore, for any quadratic function $f(x)$, its vertex is at $\left(\frac{-b}{2a}, f\left(\frac{-b}{2a}\right)\right)$.

Since for any quadratic function we have $f(x) = ax^2 + bx + c$, if we evaluate $f\left(-\frac{b}{2a}\right)$, we get:

$$\begin{aligned} f\left(-\frac{b}{2a}\right) &= a\left(-\frac{b}{2a}\right)^2 + b\cdot\left(-\frac{b}{2a}\right) + c \\ &= a\cdot\frac{b^2}{4a^2} - \frac{b^2}{2a} + c \\ &= \frac{b^2}{4a} - \frac{b^2}{2a} + c \\ &= \frac{b^2 - 2b^2}{4a} + c \\ &= c - \frac{b^2}{4a} \end{aligned}$$

So, the vertex of any quadratic function is at the coordinates $\left(-\frac{b}{2a}, c - \frac{b^2}{4a}\right)$.

2.16.5 Using the Discriminant to Find the Number of Zeros

In Section 2.16.3 we rewrote the quadratic equation from Section 2.16.1, “Using the Quadratic Formula to Find Zeros of Quadratic Function” in the equivalent form:

$$x = \frac{-b}{2a} \pm \frac{\sqrt{b^2 - 4ac}}{2a}$$

The second term of this expression (i.e. after the \pm sign) includes a square root with the expression $b^2 - 4ac$ under it. Since you can't take the square root of a negative number, we know that the square root function is only well-defined when $b^2 - 4ac \geq 0$. This quantity under the square root of the quadratic formula is called the *discriminant*, because as we'll see in this section, its value helps us discriminate between three possible situations for the zeros of a quadratic function:

1. There are two *distinct* zeros
2. There is a single zero (vertex lies on x -axis), also known as having “two *equal* zeros”
3. There is no zero (parabola never intersects x -axis)

If the discriminant ($b^2 - 4ac$) is greater than zero, then the second term of the expression above is non-zero, and therefore the quadratic function has two distinct zeros that are obtained by taking the quantity $\frac{-b}{2a}$ and adding to it and subtracting from it, the non-zero quantity $\frac{\sqrt{b^2 - 4ac}}{2a}$. This is the case in Figure 38.

If the discriminant ($b^2 - 4ac$) equals zero, then the second term of this expression evaluates to zero, which means then that the zeros of the function are:

$$\begin{aligned} x &= \frac{-b}{2a} \pm 0 \\ x &= \frac{-b}{2a} \end{aligned}$$

In this case, the parabola only has a single zero. Another way you will see this stated is that that parabola has “two equal zeros” or “two equal roots”. This corresponds to the graph where

the vertex lies right on the x -axis. Figure 40 shows an example of a quadratic function with a single zero. Its discriminant is:

$$b^2 - 4ac = 18^2 - 4 \cdot 1 \cdot 81 = 324 - 324 = 0$$

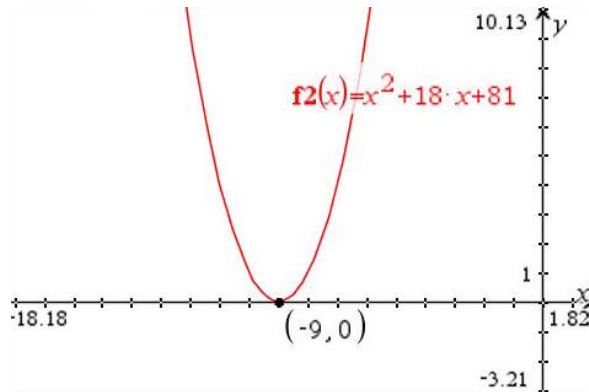


Figure 40: Quadratic with Discriminant = 0 has Single Zero (or “Two Equal Zeros/Roots”)

Finally, if the discriminant ($b^2 - 4ac$) is less than zero, then the square root in the second term of the expression above is undefined, so there are no zeros. As shown in Figure 41, its graph never intersects the x -axis.

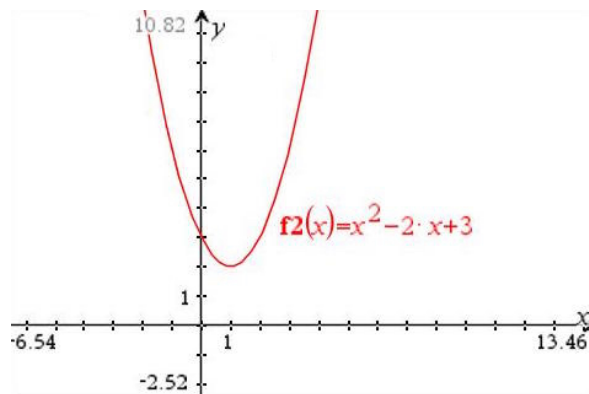


Figure 41: Quadratic Function with Negative Discriminant has No Zeros

2.16.6 Y-Intercept Form

The standard way of writing a quadratic function, $f(x) = ax^2 + bx + c$, is called the “Y-Intercept Form” because in this form you can easily read off the value of y where the function intercepts (i.e. crosses) the y -axis. Since the y intercept is the point where $x = 0$, if we evaluate $f(0) = a \cdot 0^2 + b \cdot 0 + c = c$ we see that the y intercept is just the coefficient c of the units term. By following the steps explained in Section 2.14.7, “Finding the y -Intercept of a Graph”, we can easily plot the y -intercept of the example quadratic function in Figure 42. We see that the y -coordinate of the y -intercept is just the coefficient 6 from its units term, as expected.

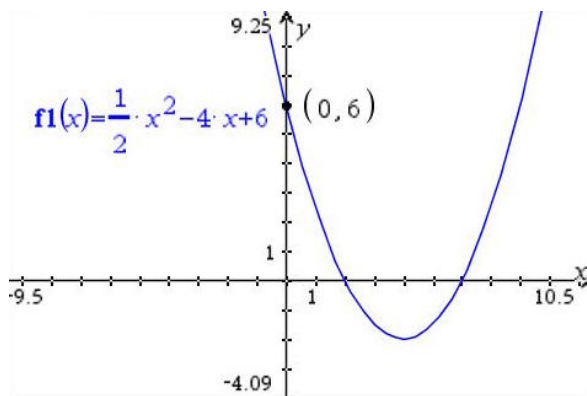


Figure 42: Quadratic Function with y Intercept Point

2.16.7 X-Intercept Form

In order to put a quadratic equation into a form that makes it easy to read off the x -intercepts of the function, we need to factor the original expression $ax^2 + bx + c$ into a form that looks like $a(x - p)(x - q)$, where p and q are the two roots (i.e. “zeros” or solutions) of the quadratic equation. We can do this mechanically by first using the quadratic formula to find the roots p and q , and then rewriting the original quadratic equation as $a(x - p)(x - q)$. Another possibility is to factor out the constant a that is the coefficient of the x^2 term, and then try to factor the remaining quadratic expression that will now start with x^2 .

The reverse is of course also true. If you see a quadratic equation that is already in the x -intercept form, you can just read off the values of the zeros p and q . For example, if $f(x) = 17(x - 10)(x + 3)$, then to find the zeros consider $17(x - 10)(x + 3) = 0$ which can only be zero if either $(x - 10) = 0$ or else if $(x + 3) = 0$. Therefore, $x = 10$ or $x = -3$. Those were numbers you can immediately see when the quadratic expression is written in the x -intercept format.

Given the example function $f(x) = \frac{1}{2}x^2 - 4x + 6$ from Figure 42, let’s rewrite it in x -intercept form for practice:

$$\begin{aligned} \frac{1}{2}x^2 - 4x + 6 &= \frac{1}{2}(x^2 - 8x + 12) && \text{Start by factoring out the coefficient of the } x^2 \text{ term} \\ &= \frac{1}{2}(x - 6)(x - 2) && \text{Factoring is straightforward, so we don't need quadratic formula} \end{aligned}$$

So, the x intercepts are 6 and 2 and the equivalent x -intercept form of the original function is therefore:

$$\underbrace{\frac{1}{2}x^2 - 4x + 6}_{y\text{-Intercept Form}} = \underbrace{\frac{1}{2}(x - 6)(x - 2)}_{x\text{-Intercept Form}}$$

When written in this x -intercept format, we can easily read off that the zeros of the function are 6 and 2, which matches the results we found using the quadratic formula as shown in Figure 38.

2.16.8 Completing the Square to Get Binomial Squared Form

“Completing the square” is a procedure that you can follow to rewrite any quadratic function in a form that includes a squared binomial $(x - h)^2$. The procedure goes like this:

$ax^2 + bx + c$	Start with the equation in standard y -intercept form
$(ax^2 + bx) + c$	Group the x^2 and x terms with parentheses
$a\left(x^2 + \frac{b}{a}x\right) + c$	Factor out a from the group
$a\left(x^2 + \frac{b}{a}x + \left(\frac{b}{2a}\right)^2\right) + c - a\left(\frac{b}{2a}\right)^2$	Add constant inside, subtract constant outside
$a\left(x + \frac{b}{2a}\right)^2 + \left(c - \frac{b^2}{4a}\right)$	Factor into square binomial and simplify
$a\left(x - \left(-\frac{b}{2a}\right)\right)^2 + \left(c - \frac{b^2}{4a}\right)$	Rewrite to see $x = -\frac{b}{2a}$ coordinate of vertex

In the fourth step above, the constant $\left(\frac{b}{2a}\right)^2$ that we add inside of the x^2 and x group is the square of one half of the coefficient of the x term $\frac{b}{a}$. We add this strategic value because in general, when you square the binomial $(x + q)$, you get the expression $(x + q)^2 = x^2 + 2q \cdot x + q^2$. The units term q^2 is the square of one half of the coefficient of the x term ($= 2q$). So the number we add inside the group is the one that *must* be there in order to use this squaring-a-binomial rule in reverse to factor it back to the binomial square form. However, if we introduce a term being added somewhere in the original expression, we also have to subtract that same quantity somewhere else in the expression to avoid inadvertently changing the meaning of the original expression. Notice that the a has been factored out of the x^2 and x group parenthesis so anything added inside the parenthesis is really being multiplied by a as well. This is why we have to subtract the quantity $a\left(\frac{b}{2a}\right)^2$ from the constant c outside of the group to avoid changing the original expression.

So, for any quadratic equation, we have the following equality with standard y -intercept form on the left of the equals sign and “binomial squared” form on the right:

$$ax^2 + bx + c = a\left(x - \left(-\frac{b}{2a}\right)\right)^2 + \left(c - \frac{b^2}{4a}\right)$$

2.16.9 Vertex (h, k) Form

We saw in Section 2.16.4, “Computing the Vertex From the Coefficients” that the vertex of a quadratic function was at the point $\left(-\frac{b}{2a}, c - \frac{b^2}{4a}\right)$. We also saw in Section 2.16.8, “Completing the Square to Get Binomial Squared Form” that a quadratic function in standard y -intercept form can be written as:

$$a\left(x - \left(-\frac{b}{2a}\right)\right)^2 + \left(c - \frac{b^2}{4a}\right)$$

The quantity being subtracted from x in the expression above $-\frac{b}{2a}$ is none other than the x -coordinate of the vertex we see just above. Likewise, the quantity being added to the binomial

squared term $\left(c - \frac{b^2}{4a}\right)$ is exactly the y -coordinate of the vertex. So, if we let:

$$h = -\frac{b}{2a}$$

$$k = c - \frac{b^2}{4a}$$

then we can write any quadratic equation in the following form, where h and k are the x - and y -coordinates respectively of the function's vertex:

$$ax^2 + bx + c = a(x - h)^2 + k$$

This form is called the *vertex form* of a quadratic function.

Given the example function $f(x) = \frac{1}{2}x^2 - 4x + 6$ from Figure 42, let's rewrite it in vertex form for practice. Start by identifying the coefficients:

$$a = \frac{1}{2}$$

$$b = -4$$

$$c = 6$$

Next we compute the h - and k -coordinates of the vertex using the formula we learned above:

$$h = -\frac{b}{2a} = -\frac{-4}{2 \cdot \frac{1}{2}} = 4$$

$$k = c - \frac{b^2}{4a} = 6 - \frac{(-4)^2}{4 \cdot \frac{1}{2}} = 6 - \frac{16}{2} = 6 - 8 = -2$$

So, the vertex is at point $(4, -2)$ and the equivalent vertex form of the original function is therefore:

$$\underbrace{\frac{1}{2}x^2 - 4x + 6}_{y\text{-Intercept Form}} = \underbrace{\frac{1}{2}(x - 4)^2 - 2}_{\text{Vertex Form}}$$

2.17 Reciprocal Functions

Reciprocal functions are of the form $f(x) = \frac{k}{x}$ where k is some constant. Their graph is a hyperbola with two curving branches having both a horizontal and vertical asymptote. They are symmetric around the line $y = x$, which makes them their own inverse function. If $k > 0$ then the curving branches of the function are in quadrants I and III as shown in Figure 43. If $k < 0$, then they are instead in quadrants II and IV.

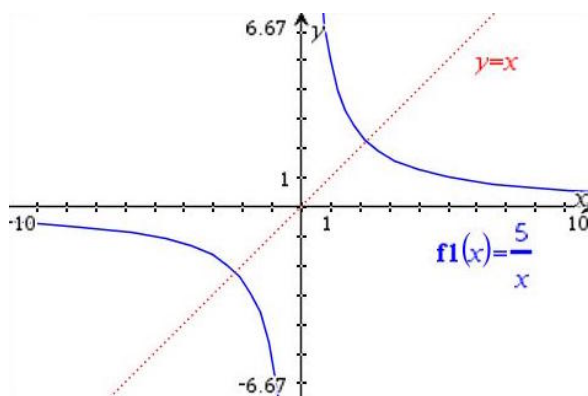


Figure 43: Reciprocal Function is a Hyperbola

2.18 Rational Functions

Rational functions are the quotient of two linear functions and have the form $\frac{ax + b}{cx + d}$. Like reciprocal functions, they have a horizontal asymptote and a vertical asymptote. Following the rules we learned in Section 2.14.8, “Vertical Asymptotes”, we can compute that the vertical asymptote must exist where the denominator $cx + d = 0$. Therefore the vertical asymptote is the line $x = -\frac{d}{c}$. Following the rules we learned in Section 2.14.10, “Horizontal Asymptotes”, we can determine that the a horizontal asymptote must be the line $y = \frac{a}{c}$. The x intercept will occur when the numerator $ax + b = 0$, so we can compute that the x intercept is at $x = -\frac{b}{a}$. By evaluating the function at $x = 0$ we can find that the y -intercept is $\frac{b}{d}$. Figure 44 shows an example of a rational function.

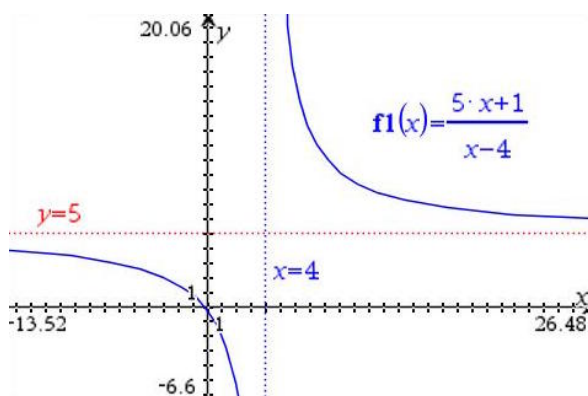


Figure 44: Rational Function is a Hyperbola with Asymptotes Away from the Axes

2.19 Exponential Functions

Exponential functions have the form $f(x) = ka^{mx}$ for some constants a (the base) and coefficients k and m . Very frequently the base e is used, since it occurs in formulas relating to the growth and decay of processes in nature and in the business world. As shown in Figure 45, for positive values of m , the graph models growth. The values of the function get larger and larger as x gets larger. When $m > 1$, the larger the value, the more steep the growth. When values $0 < m < 1$, then the function shows slower, flatter growth. When the value of m is negative, this models decay instead of growth since the values of the graph get smaller and smaller as x increases.

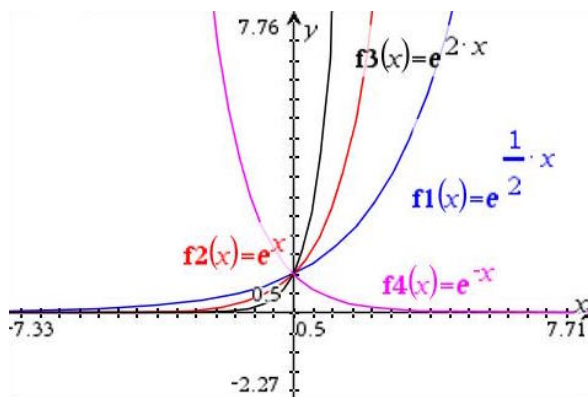


Figure 45: Examples of Exponential Functions Having Base e

Recall from Section 1.16, “Using Logarithms to Solve Equations” that we can exploit the fact that these the e^x and \ln functions are inverses to help solve problems. Consider an example equation like:

$$1234e^{0.19x} = 20000$$

The variable x whose value we seek to solve for is “trapped” up in the exponent of the base e . We can free it from that position by applying the appropriate inverse function, the natural logarithm $\ln(x)$ in this case, to both sides of the equation:

$$\begin{aligned} 1234e^{0.19x} &= 20000 \\ e^{0.19x} &= \frac{20000}{1234} \approx 16.2075 \\ \ln(e^{0.19x}) &= \ln 16.2075 \\ 0.19x &= 2.7855 \\ x &= 14.6604 \end{aligned}$$

2.20 Continuously Compounded Interest

In section Section 1.13, “Applications: Compound Interest” you saw how to calculate compound interest $A(t)$ after t years, at a rate of r for a discrete number of periods n in a year:

$$A(t) = P \cdot \left(1 + \frac{r}{n}\right)^{n \cdot t}$$

If you let the number of compounding periods n in a year increase without bound, you arrive at the concept of *continuously compounded* interest. In this model, your money is continuously earning tiny amounts of additional interest as time moves forward. These types of continually compounded growth (or decay, in the case of losing money) are modeled by exponential functions using the base of the natural logarithm e . The formula for what sounds like an even more complex concept, continuously compounded interest, is actually even simpler than the formula for the discrete compounding case. It is:

$$A(t) = P \cdot e^{rt}$$

where r is again the rate of interest and t is the time in years.

2.21 Continuous Growth and Decay

Many natural processes exhibit continuous growth or decay. Two examples are population growth and the decay of radioactivity in a given material. These processes of continuous growth or decay are modeled using essentially the same exponential equation as for continuously compounded interest. You start out with some initial quantity C of some entity (people, radioactive particles, etc.), and a rate r , then the amount of the entity after t units of time is given by:

$$A(t) = C \cdot e^{rt}$$

2.22 Logarithmic Functions

Logarithmic functions have the form $f(x) = k \cdot \log_b x$. For a given base b , the function $f(x) = \log_b x$ is the inverse function of the exponential function $f(x) = b^x$. So, $y = \log_b x$ if and only if $b^y = x$. Many natural processes model their continuous growth or decay based on the exponential function in Section 2.21, “Continuous Growth and Decay”. Since this uses the base

e, the natural logarithm ($\log_e x$ usually written $\ln x$) is its inverse. Figure 46 shows how the graphs of e^x and $\ln(x)$ are reflections in the line $y = x$, which means they are inverse functions.

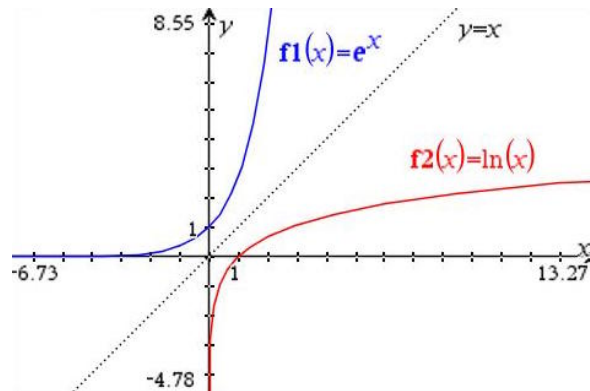


Figure 46: Exponential Base e^x and Natural Logarithm $\ln(x)$ are Inverse Functions

Recall from Section 1.17, “Using Exponentiation to Solve Equations” that if we have an equation involving a logarithm, where the variable x whose value we seek is “trapped” inside the log function, then we can free it from that position by raising an appropriate base to the power represented by each side of the equation. If the two sides were equal before, then so will be the same base raised to equal powers. Consider the example:

$$25 \ln(0.35x) = 200$$

First we isolate the $\ln(0.35x)$ by dividing both sides by 25:

$$\begin{aligned} 25 \ln(0.35x) &= 200 \\ \ln(0.35x) &= \frac{200}{25} = 8 \end{aligned}$$

Next, we write a new equation involving the base e of the natural logarithm on both sides:

$$e^{\boxed{\dots}} = e^{\boxed{\dots}}$$

Since we know that the two expressions that we will write into the boxes as exponents on the base e are equal already, we can conclude that result of raising e to those equal powers will also be equal. Therefore, we can write:

$$\begin{aligned} e^{\ln(0.35x)} &= e^8 \\ 0.35x &= e^8 \approx 2980.958 \\ x &= \frac{2980.958}{0.35} = 8517.023 \end{aligned}$$

3 Circular Functions and Trigonometry

3.1 Understanding Radians

This section motivates how radians provide an alternative to degrees in measuring the part of a circular path that an angle represents. It explains how to compute the fraction of the complete circular revolution to which this part corresponds. Finally, it shows how to convert back and forth between radians and degrees.

3.1.1 Degrees Represent a Part of a Circular Path

The angle θ in Figure 47 represents the amount of rotation that would be required to bring one intersecting line (or segment) in line with the other. About 2400 BC, the Sumerians observed the planets and noticed that the Sun’s path across the sky took about 360 days to make one year’s cycle. So, they divided this circular path into 360 degrees to track each day’s passage of the Sun’s whole journey. These 360 degrees (written 360°) are still in use today as one way to measure the amount of rotation in an angle. The degrees measuring the angle θ ’s rotation are some *part* of the total journey around the circle.

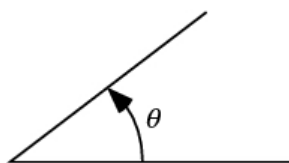


Figure 47: Angle is Rotation to Bring One Intersecting Segment in Line with the Other

3.1.2 Computing the Fraction of a Complete Revolution an Angle Represents

Knowing the total number of units that represent a complete circular trip, for example 360 degrees, you can use the part of the total journey divided by the total to calculate the fraction of the overall journey the angle represents. For example, if θ measures 45° , then the fraction of a total revolution that this rotation represents is $\frac{\text{part}}{\text{total}} = \frac{45^\circ}{360^\circ} = \frac{1}{8}$ of the circular journey. Put another way, after 45 days the Sumerian’s knew that their 360-day year was $\frac{1}{8}$ over. So, while the number 360 is a historical quantity, the key concept involved is that the angle represents a **part of a circular path** and you can use this to calculate the **fraction of a complete revolution**.

3.1.3 Attempting to Measure an Angle Using Distance

Another way to describe the angle θ would be to explain how far someone would have to walk along a circular path to “push” one segment to be in line with another. Figure 48 shows this idea. Imagine the same angle θ from Figure 47 with its vertex sitting at the center of a circle and a man standing some small distance r_s away from the vertex. Let this distance r_s be the radius of the circle. Suppose the segments of the angle are fixed with a pin at the center of this circle that allows them to move like hands of a clock. If the man’s goal is to push the segment closest to him to be in line with the other segment forming the angle, how far would he need to walk around the circumference of this circular path? The answer is, “It depends on the radius of the circle!” We know the formula for the circumference of a circle of radius r_s is $2\pi r_s$, so a walk around the entire circumference would be $2\pi r_s$ units. But clearly the man wouldn’t need to walk all the way around the circle to line up the segments, so he only needs to walk some **part of the circular path** to get the job done.

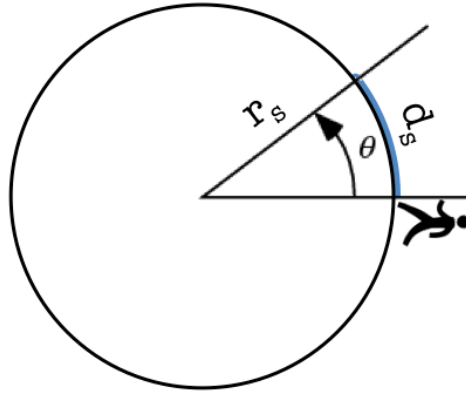


Figure 48: Rotation θ in Terms of Distance Walked Around a Circumference

If the man measures the distance d_s that he walks, how can he determine what fraction of the circular path he has traveled? A fraction always needs a *part* quantity in the numerator and a *total* or “whole” amount in the denominator. In this example, the part value for the numerator is the distance d_s the man walked. The quantity for the denominator’s total distance is the full circumference of the circle $2\pi r_s$. Therefore, the fraction of the circular path that the man has traveled is $\frac{\text{part}}{\text{total}} = \frac{d_s}{2\pi r_s}$.

What happens if the man moves a bigger distance r_b (for “bigger”) away from the center of the circle and repeats his walk? This time he will have to walk a greater distance d_b along the circumference of the larger circle of radius r_b . Using the same formula as above, we would calculate the fraction of the circular path to be $\frac{\text{part}}{\text{total}} = \frac{d_b}{2\pi r_b}$. Figure 49 shows the man’s longer walk on his second attempt.

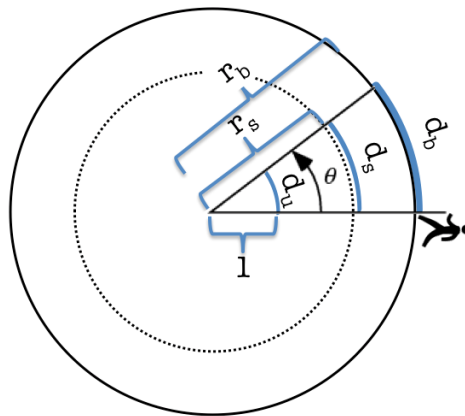


Figure 49: Man Walks Longer Along Circle with Larger Radius

Figure 49 also depicts his third attempt. That time he positions himself exactly one unit away from the vertex, and walks a distance of d_u (for “unit”). On this third attempt, the fraction of the circular path would be $\frac{\text{part}}{\text{total}} = \frac{d_u}{2\pi \cdot 1} = \frac{d_u}{2\pi}$.

Therefore, using an arc distance alone is an ambiguous way to represent an angle because the same angle θ can correspond to many different arc distances (d_u , d_s , d_b), depending on how far away from the vertex you measured the “walk” along the circular path.

3.1.4 Arc Distance on the Unit Circle Uniquely Identifies an Angle θ

In Figure 49, since the angle θ hasn't changed between the first walk of length d_s , the longer walk of length d_b , and the third walk of length d_u , the **fraction of the circular path** is still the same as it was before. Therefore, we must have:

$$\frac{d_s}{2\pi r_s} = \frac{d_b}{2\pi r_b} = \frac{d_u}{2\pi \cdot 1}$$

We can multiply through by 2π to eliminate the common factor in all the denominators, in order to reveal a very interesting observation.

$$\begin{aligned} 2\pi \cdot \frac{d_s}{2\pi r} &= 2\pi \cdot \frac{d_b}{2\pi r_b} = 2\pi \cdot \frac{d_u}{2\pi} \\ 2\pi \cdot \frac{d_s}{\cancel{2\pi}r} &= 2\pi \cdot \frac{d_b}{\cancel{2\pi}r_b} = 2\pi \cdot \frac{d_u}{\cancel{2\pi}} \\ \frac{d_s}{r_s} &= \frac{d_b}{r_b} = d_u \end{aligned}$$

In other words, for any radius other than 1, the *ratio* of the arc distance to the radius of its circular path is equal to the arc distance d_u on the unit circle. This means that if we fix our sights on only using the unit circle to measure angles, then the arc distance itself uniquely identifies the angle θ . We define the *radian measure* of any angle θ to be this unambiguous arc distance **on the unit circle**.

3.1.5 Computing the Fraction of a Complete Revolution for Angle in Radians

Radians measure the **part of a circular path** – specifically of the *unit* circular path – just like degrees do. To compute the fraction of the complete revolution this represents, just divide the number of radians by 2π , which represents the total distance around the circumference of the unit circle. If you are given an arc distance for an angle and the radius is not equal to one, then in order to compute the radian measure of the angle just divide the arc distance by the radius. This normalizes the arc distance by scaling it to correspond to the unambiguous distance on the *unit* circle.

3.2 Converting Between Radians and Degrees

For a given angle θ , θ_{rad} radians and θ_{deg} degrees are different numbers that each corresponds to the same part of a unit circular path. Correspondingly, $\frac{\theta_{rad}}{2\pi}$ and $\frac{\theta_{deg}}{360}$ both represent the same *fraction* of a complete revolution that the angle represents. We can use this fact to convert between them.

3.2.1 Converting from Degrees to Radians

If an angle θ measures θ_{deg} degrees, to convert it to radians:

- Determine the fraction of a complete revolution that θ_{deg} represents by dividing it by 360.
- The measure in radians will be the *same fraction* of the total unit circle distance of 2π , so multiply the fraction determined in the first step by 2π .

Therefore to convert from degrees to radians, we have:

$$\theta_{rad} = \underbrace{\frac{\theta_{deg}}{360}}_{\substack{\text{Fraction of} \\ \text{complete} \\ \text{revolution}}} \times \underbrace{2\pi}_{\substack{\text{Circumference of} \\ \text{complete revolu-} \\ \text{tion of unit circle}}} = \theta_{deg} \cdot \frac{2\pi}{360} = \underbrace{\theta_{deg} \cdot \frac{\pi}{180}}_{\substack{\text{Radian measure} \\ \text{of partial distance} \\ \text{around unit circle}}}$$

In short, to convert an angle θ in degrees to radians, multiply it by $\frac{\pi}{180}$. To remember the right fraction to multiply by, think about the fact that you are “introducing π ” when going from degrees to radians, so the π goes in the numerator.

3.2.2 Converting from Radians to Degrees

If an angle θ measures θ_{rad} radians, to convert it to degrees:

- Determine the fraction of a complete revolution that θ_{rad} represents by dividing it by 2π .
- The measure in radians will be the *same fraction* of the total number of degrees 360, so multiply the fraction determined in the first step by 360.

Therefore to convert from degrees to radians, we have:

$$\theta_{deg} = \underbrace{\frac{\theta_{rad}}{2\pi}}_{\substack{\text{Fraction of} \\ \text{complete} \\ \text{revolution}}} \times \underbrace{360}_{\substack{\text{Total degrees in} \\ \text{complete revolu-} \\ \text{tion of unit circle}}} = \theta_{rad} \cdot \frac{360}{2\pi} = \underbrace{\theta_{rad} \cdot \frac{180}{\pi}}_{\substack{\text{Degree measure} \\ \text{of partial rotation} \\ \text{around unit circle}}}$$

In short, to convert an angle θ in radians to degrees, multiply it by $\frac{180}{\pi}$. To remember the right fraction to multiply by, think about the fact that you are “removing π ” when going from radians to degrees, so the π goes in the denominator.

3.3 Length of an Arc Subtended by an Angle

When an angle θ has its vertex at the center of a circle, the angle determines (or “subtends”) an arc on the circle as shown in Figure 50.

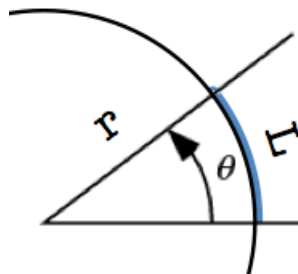


Figure 50: If Angle θ is in radians, Arc Length $L = \theta \cdot r$

Given a circle of radius r and an angle θ in radians, to compute the length of the arc subtended

by the angle, do the following:

- Determine the fraction of a complete revolution that θ_{rad} represents by dividing it by 2π .
- The arc length L will be the *same fraction* of the circle's total circumference, so multiply the fraction determined in the first step by the circumference $2\pi r$. This gives:

$$L = \underbrace{\frac{\theta_{rad}}{2\pi}}_{\substack{\text{Fraction of} \\ \text{complete} \\ \text{revolution}}} \times \underbrace{2\pi r}_{\substack{\text{Circumference} \\ \text{of the circle}}} = \theta_{rad} \cdot \frac{2\pi \cdot r}{2\pi} = \underbrace{\theta_{rad} \cdot r}_{\text{Arc length}}$$

In short, given a circle of radius r and an angle θ in radians, the length of the arc subtended by the angle is θr . If the angle θ is in degrees, then first convert it to radians by multiplying by $\frac{\pi}{180}$ then proceed to multiply that by the radius to calculate the arc length.

3.4 Inscribed and Central Angles that Subtend the Same Arc

The “Central Angle Theorem” is an important geometrical fact worth remembering that pertains to circles, angles, and arcs subtended by a central angle. Figure 51 illustrates that if a central angle $\angle AOB$ subtends an arc AB , then its measure θ is double the angle of any inscribed angle $\angle ACB$ or $\angle ADB$ that subtends the same arc. Or, vice versa, the inscribed angle's measure is half of the measure of the central angle when they subtend the same arc of the circle.

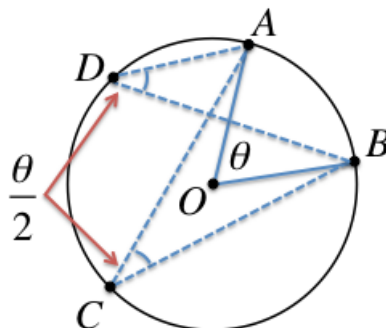


Figure 51: Central Angle is Double the Inscribed Angle that Subtends the Same Arc

3.5 Area of a Sector

When an angle θ has its vertex at the center of a circle, the angle's segments and the arc it subtends form a sector as shown in Figure 52.

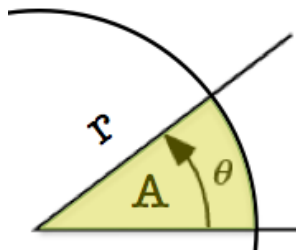


Figure 52: If Angle θ is in radians, Area of Sector $A = \frac{1}{2}\theta \cdot r^2$

Given a circle of radius r and an angle θ in radians, to compute the area of the sector formed by the angle, do the following:

- Determine the fraction of a complete revolution that θ_{rad} represents by dividing it by 2π .
- The area A of the sector will be the *same fraction* of the circle's total area, so multiply the fraction determined in the first step by the area πr^2 . This gives:

$$A = \underbrace{\frac{\theta_{rad}}{2\pi}}_{\substack{\text{Fraction of} \\ \text{complete} \\ \text{revolution}}} \times \underbrace{\pi r^2}_{\substack{\text{Area of the circle}}} = \theta_{rad} \cdot \frac{\pi \cdot r^2}{2\pi} = \underbrace{\frac{1}{2} \theta_{rad} \cdot r^2}_{\substack{\text{Sector area}}}$$

In short, given a circle of radius r and an angle θ in radians, the area of the sector formed by the angle is $\frac{1}{2}\theta r^2$. If the angle θ is in degrees, then first convert it to radians by multiplying by $\frac{\pi}{180}$ then proceed to multiply that by one half of the radius squared to calculate the arc length.

3.6 Definition of $\cos \theta$ and $\sin \theta$

Given the right triangle in Figure 53 with angles A and B and corresponding opposite side lengths a and b , we define the *sine* function of an angle θ (written $\sin \theta$) as the ratio of the length of the side *opposite* the angle to the hypotenuse and the *cosine* function of an angle θ (written $\cos \theta$) as the ratio of the length of the side *adjacent* the angle to the hypotenuse.

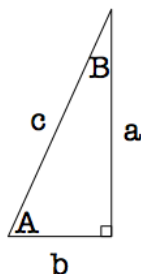


Figure 53: Right Triangle with Angles A and B

Given this definition and the angles in Figure 53, we have:

$$\begin{aligned} \sin A &= \frac{\text{opp}}{\text{hyp}} = \frac{a}{c} & \cos A &= \frac{\text{adj}}{\text{hyp}} = \frac{b}{c} \\ \sin B &= \frac{\text{opp}}{\text{hyp}} = \frac{b}{c} & \cos B &= \frac{\text{adj}}{\text{hyp}} = \frac{a}{c} \end{aligned}$$

These \sin and \cos functions are two of the three so-called *trigonometric* functions that we'll cover in this program. The word "trigonometry" comes from the Latin words "trigon" for three-sided figure and "metry" for measurement. Trigonometric functions help measure interesting relationships between the sides of a three-sided figure, the triangle. We'll define the third trigonometric function in Section 3.12, "Definition of $\tan \theta$ ".

3.7 Interpreting $\cos \theta$ and $\sin \theta$ on the Unit Circle

If we imagine the unit circle centered at the origin of the Cartesian plane, at any point (x, y) on the circle we can drop a perpendicular line to the x -axis to form a right triangle. As shown in Figure 54, the hypotenuse of this right triangle is always 1 unit because it is always a radius of our unit circle. Notice that the two legs of the triangle are exactly the distances x and y that are given by the point's (x, y) coordinates. From angle θ 's point of view, the side of the triangle that sits opposite it has length y , and the side that sits adjacent to it has length x . Using the definition of the sine function from Section 3.6, putting these facts together we can derive the following fact about the angle theta related to the unit circle:

$$\sin \theta = \frac{\text{opp}}{\text{hyp}} = \frac{y}{1} = y \qquad \cos \theta = \frac{\text{adj}}{\text{hyp}} = \frac{x}{1} = x$$

So as shown in Figure 54, any point on the unit circle has coordinates $(\cos \theta, \sin \theta)$.

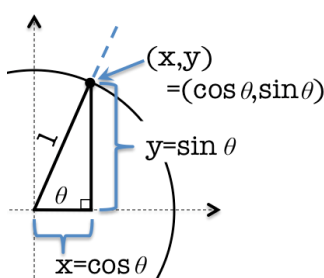


Figure 54: Right Triangle in Unit Circle with Angle θ

3.8 Radian Angle Measures Can Be Both Positive and Negative

From Section 3.1.4, “Arc Distance on the Unit Circle Uniquely Identifies an Angle θ ”, we know that radians are a measure of arc distance on the unit circle that corresponds to an amount of rotation. By convention, a positive radian angle measure is a distance measured starting at the point $(1, 0)$ on the unit circle (and on the x -axis) and traveling in a *counter-clockwise* direction. If instead the rotation is in a clockwise direction, then the radian angle measure is a negative distance. Figure 55 shows a θ_{pos} angle with a positive radian measure representing a certain amount of counter-clockwise rotation. It also shows a θ_{neg} angle with a negative radian measure representing a different amount of clockwise rotation.

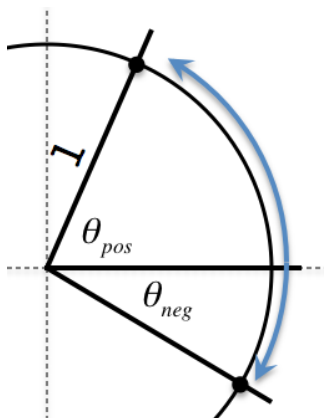


Figure 55: Right Triangle in Unit Circle with Angle θ

3.9 Remembering the Exact Values of Key Angles on Unit Circle

It's important to remember the cos and sin values for a number of key angles on the unit circle because they are not in the formula booklet and you are expected to know them on the non-calculator part of your exam. Figure 56 shows a visual approach using your left hand that you can use to memorize the cos and sin values for five key angles in the first quadrant. First, remember that the values are listed as cos, sin. That is, for each pair of values the cosine value of the angle is first and the sine value of the angle is second. To keep it simple, remember that all values have 2 in the denominator. The values in the numerator range from 0 to 4 and they are all under a square root. Going in counter-clockwise order, the values for the cosine numerators count down ($\sqrt{4}, \sqrt{3}, \sqrt{2}, \sqrt{1}, \sqrt{0}$). The values for the sine numerators count up ($\sqrt{0}, \sqrt{1}, \sqrt{2}, \sqrt{3}, \sqrt{4}$). Of course, some of the values like $\sqrt{4}$, $\sqrt{1}$, and $\sqrt{0}$ simplify to 2, 1, and 0 respectively, but remembering the counting up and down pattern makes it easier to remember.

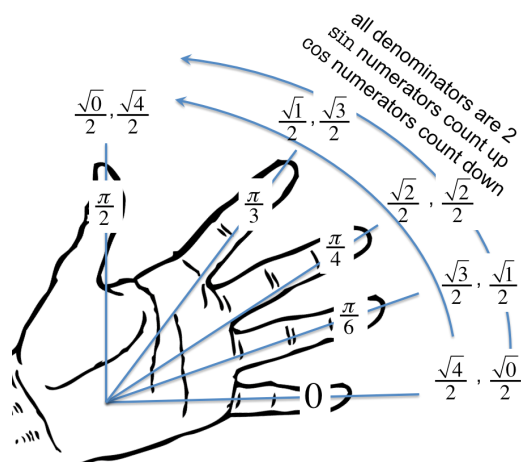


Figure 56: Using Your Hand to Remember Key cos and sin Values in First Quadrant

Once you know the values of the cos and sin functions for the five key angles in the first quadrant, you can easily remember the values of the other key angles in the other three quadrants because their cos and sin values are the same except for a strategically placed minus sign. Figure 57 shows the pattern of where the cos and sin function values are positive and negative. To calculate the cos and sin values of key angles in quadrants II, III, and IV, just determine which angle in the first quadrant should be its reflection in the x -axis or y -axis, remember the value from the "hand" and then apply the minus sign where appropriate based on which quadrant the angle falls in.

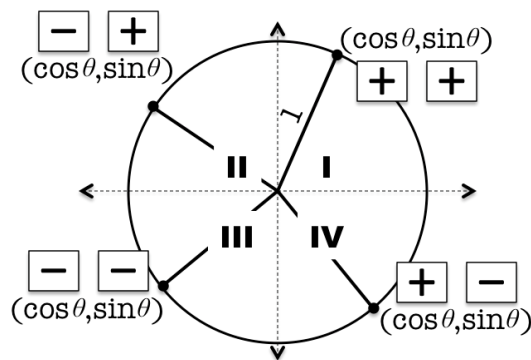


Figure 57: Positive/Negative Sign of cos and sin Values in Four Quadrants

Figure 58 shows cos and sin values for all key angles. Using the "hand" and "quadrant signs" makes remembering all of the values as easy as counting on your fingers.

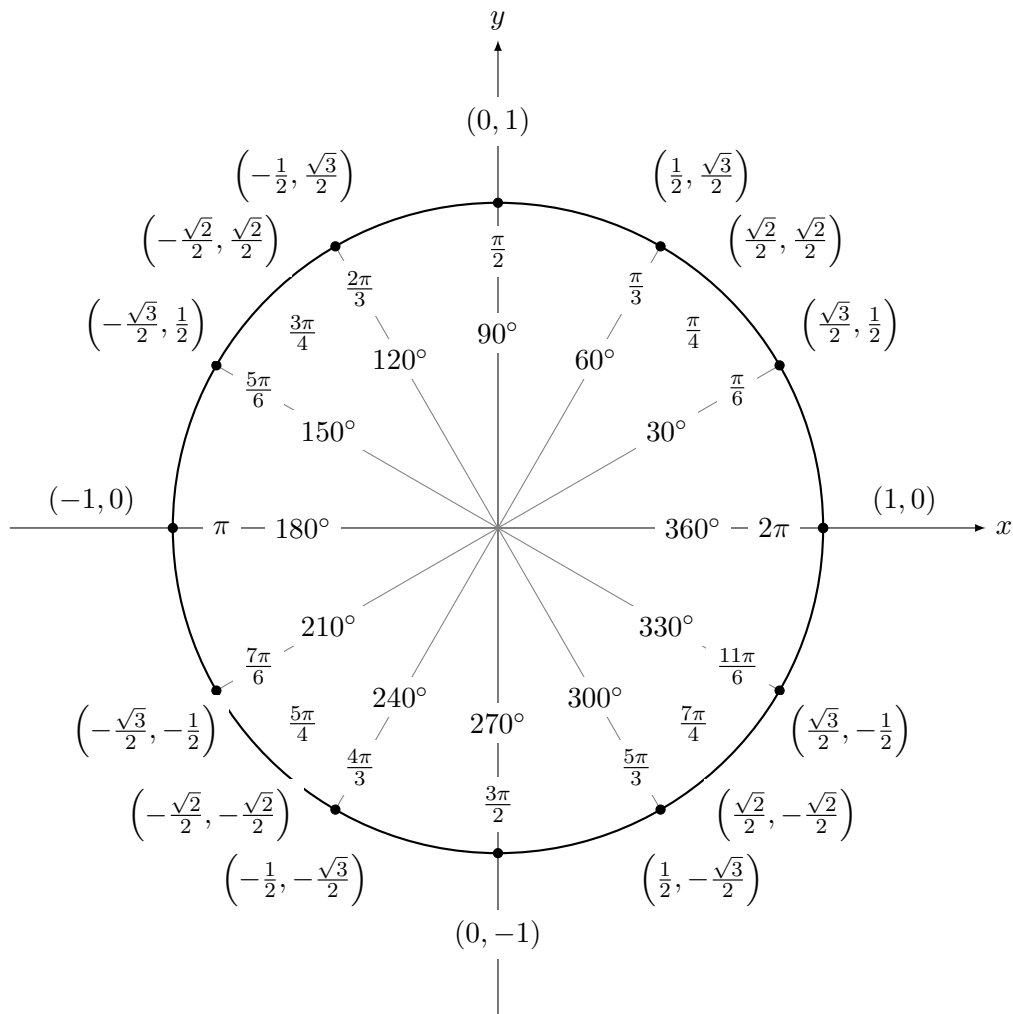


Figure 58: Values of cos and sin for All Key Unit Circle Angles

3.10 The Pythagorean Identity

As shown in Figure 59, the right triangle formed in the unit circle with left-hand angle θ has legs that measure $\cos \theta$ and $\sin \theta$. Since it's a right triangle, Pythagorean Theorem tells us that the sum of the squares of the legs equals the square of the hypotenuse. This gives rise to the *Pythagorean Identity* for cos and sin that $\cos^2 \theta + \sin^2 \theta = 1$.

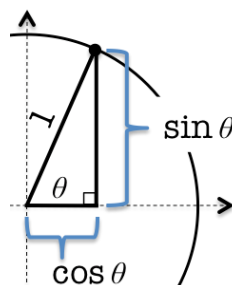


Figure 59: Pythagorean Identity for cos and sin

3.11 Double Angle Identities

The formula booklet lists two additional handy identities that can be used to calculate the cos or sin of a “double angle” 2θ if you already know the $\cos \theta$ and $\sin \theta$. You will use these identities to substitute one expression for another while solving equations involving sin and cos functions.

$$\begin{aligned}\cos 2\theta &= \cos^2 \theta - \sin^2 \theta = 2 \cos^2 \theta - 1 = 1 - 2 \sin^2 \theta \\ \sin 2\theta &= 2 \cos \theta \sin \theta\end{aligned}$$

3.12 Definition of $\tan \theta$

In addition to the two functions defined in Section 3.6, “Definition of $\cos \theta$ and $\sin \theta$ ”, the *tangent* function is a third important trigonometric ratio that captures interesting information about an angle θ in a right triangle. It is written $\tan \theta$ and is defined as the ratio between the length of the side opposite the angle and the length of the side adjacent to it.

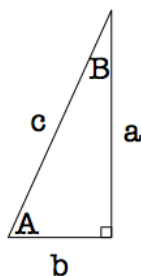


Figure 60: Right Triangle with Angles A and B

Given this definition and the angles in Figure 60, we can derive the following information. Note that in addition to being the ratio of opposite to adjacent sides, the $\tan \theta$ is also the ratio of $\sin \theta$ to $\cos \theta$.

$$\begin{aligned}\tan A &= \frac{\text{opp}}{\text{adj}} = \frac{a}{b} = \frac{\frac{a}{c}}{\frac{b}{c}} = \frac{\sin A}{\cos A} \\ \tan B &= \frac{\text{opp}}{\text{adj}} = \frac{b}{a} = \frac{\frac{b}{c}}{\frac{a}{c}} = \frac{\sin B}{\cos B}\end{aligned}$$

3.13 Using a Right Triangle to Solve Trigonometric Problems

We have talked in earlier sections about sin, cos, and tan in the context of the unit circle, but their definitions do not depend on having a hypotenuse of unit length. Any right triangle will do. The values of sin and cos range from -1 to 1 , so any value in that range is a legal value for the sin or cos of some angle θ . The value of the tan function can be any real number, so any value is legal to be the tan of some angle θ .

If a problem tells you the value of the sin, cos, or tan of an angle θ and that value does not correspond to one of the familiar “key angles” from the unit circle, then draw a right triangle and label the sides with the information you know. Using the Pythagorean Theorem, you can then figure out the remaining side, and then subsequently you can read off the values of the other trigonometric functions you may need to calculate.

3.13.1 Using Right Triangle with an Acute Angle

For example, a typical problem of this kind might take this form:

- Given $\cos x = \frac{3}{4}$ and that x is acute ($0 < x < \frac{\pi}{2}$), find $\sin 2x$

We know that x is an angle whose cosine is $\frac{3}{4}$. This value does not look like any of the small set of values ($0, 1, \frac{1}{2}, \frac{\sqrt{2}}{2}$, and $\frac{\sqrt{3}}{2}$) for the key angles of the unit circle, so we cannot immediately know the measure of the angle x in radians. Using the triangle technique here, we'll see that to solve this particular problem we don't really even need to find the value of x to get the final answer we're being asked to calculate.

Since the angle x is not one of the key angles, we draw a triangle with the information we know. As shown in Figure 61, we place the angle x in one corner. Then, since the $\cos x = \frac{3}{4}$ we know from the definition of cosine $\left(\frac{\text{adjacent}}{\text{hypotenuse}}\right)$ that 3 can be the length of the adjacent side while 4 is the hypotenuse. This is by definition what it means that $\cos x = \frac{3}{4}$.

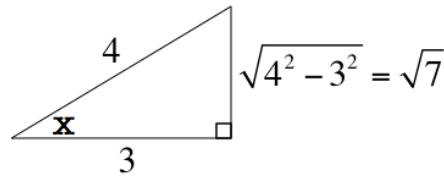


Figure 61: Using Right Triangle and Pythagorean Theorem for Trig Problems

Recall from Section 3.11, “Double Angle Identities”, that the $\sin 2x$ we've been asked to find can be rewritten via an identity as:

$$\sin 2x = 2 \cos x \sin x$$

We already know the value of $\cos x$ since the problem statement gave that to us. To complete the problem we need to determine the value of $\sin x$. From the definition of the sine function, its value is the ratio of the opposite side to the hypotenuse. The problem didn't provide us the length of the opposite side, but now that we've drawn a right triangle, we can use the Pythagorean Theorem to figure it out. As shown in the figure, the value of the opposite side is $\sqrt{4^2 - 3^2} = \sqrt{16 - 9} = \sqrt{7}$. So, the ratio of the opposite side to the hypotenuse that provides the value of $\sin x = \frac{\sqrt{7}}{4}$. Finally, to calculate the value of $\sin 2x$ we just need to substitute the values we now know for both $\cos x$ and $\sin x$ into the above double-angle formula:

$$\begin{aligned} \sin 2x &= 2 \cos x \sin x \\ &= 2 \cdot \frac{3}{4} \cdot \frac{\sqrt{7}}{4} = 2 \frac{3\sqrt{7}}{16} = \frac{3\sqrt{7}}{8} \end{aligned}$$

3.13.2 Using Right Triangle with an Obtuse Angle

Another typical problem will involve an obtuse angle. Consider this question:

- Given $\sin x = \frac{2}{3}$ and that x is obtuse ($\frac{\pi}{2} < x < \pi$), find $\cos x$ and $\cos 2x$

As shown in Figure 62, we setup our triangle with the obtuse angle x , putting the 2 for the opposite side and 3 for the hypotenuse since the definition of the sin function is the ratio of the opposite side to the hypotenuse.

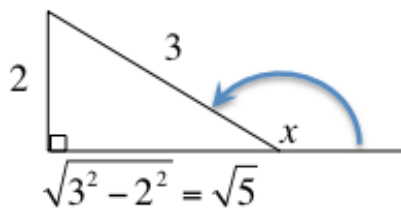


Figure 62: Using Right Triangle and Pythagorean Theorem for Obtuse Angle Problem

Using the Pythagorean Theorem, we determine that the adjacent side is $\sqrt{5}$, so the $\cos x$ should be the ratio of the adjacent side to the hypotenuse $\frac{\sqrt{5}}{3}$. However, given that this obtuse angle x is in the second quadrant, recall from Section 3.9, “Remembering the Exact Values of Key Angles on Unit Circle” that the value of its cosine should be a negative number. So the correct answer is $\cos x = -\frac{\sqrt{5}}{3}$.

To find the value of $\cos 2x$, we can use any of the identities from Section 3.11, “Double Angle Identities” and then fill in the values of the $\sin x$ and $\cos x$ that we already have. The three equations below show how all three identities allow us to conclude that $\cos 2x = \frac{1}{9}$:

$$\begin{aligned} \cos 2x &= 1 - 2\sin^2 x = 1 - 2 \cdot \left(\frac{2}{3}\right)^2 = 1 - 2 \cdot \frac{4}{9} = 1 - \frac{8}{9} = \frac{1}{9} \\ &= 2\cos^2 x - 1 = 2 \left(-\frac{\sqrt{5}}{3}\right)^2 - 1 = 2 \cdot \frac{5}{9} - 1 = \frac{10}{9} - 1 = \frac{1}{9} \\ &= \cos^2 x - \sin^2 x = \left(-\frac{\sqrt{5}}{3}\right)^2 - \left(\frac{2}{3}\right)^2 = \frac{5}{9} - \frac{4}{9} = \frac{1}{9} \end{aligned}$$

3.14 Using Inverse Trigonometric Functions on Your Calculator

Suppose a problem tells you, or as part of your work you conclude, that $\sin x = 0.156$. Since the value is not one of the “key angles” from the unit circle, you need to use the *inverse sine* function (\sin^{-1}) on your calculator to solve for x . Since the functions are periodic, however, there are infinitely many angles x that would have a sin value of 0.156. For this reason, the range of the \sin^{-1} function is restricted to the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$ so that there is an unambiguous answer when you type it into your calculator. This interval represents quadrants *I* and *IV*. Depending on the problem you are solving, you may have to add some integer multiple of 2π to the answer your calculator gives you.

Similar inverse functions exist for cosine and tangent, and are written \cos^{-1} and \tan^{-1} . For similar reasons as noted above for the inverse sine function, the range of the \cos^{-1} is restricted to the interval $[0, \pi]$. This interval represents quadrants *I* and *II*. Depending on the problem you are solving, you may have to add some integer multiple of 2π to the answer your calculator gives you.

The \tan^{-1} function’s domain accepts any real number, but its range is like sin’s $(-\frac{\pi}{2}, \frac{\pi}{2})$ but omitting the endpoints. Depending on the problem you are solving, you may have to add some integer multiple of π to the answer your calculator gives you.

3.15 Circular Functions sin, cos, and tan

We've seen the trigonometric functions sin, cos, and tan in earlier sections. In addition to being related to right-triangles, since the functions also relate to the unit circle they are sometimes know synonymously as *circular functions*. In this section we'll study the graphs of the circular functions sin, cos, and tan. These are periodic functions whose values repeat after a given *period*. The value of the sin and cos functions cycle back and forth between a minimum value and a maximum value, whereas the value of tan is unbounded. The *amplitude* of a periodic function is one half of the difference between its maximum value and its minimum value. For each graph we'll list the domain, range, amplitude, period, minimum, maximum, and zeros.

3.15.1 The Graph of $\sin x$

Figure 63 shows a portion of the graph of the $\sin x$ function. The domain of the function is all real numbers (\mathbb{R}), and its range is $\{x \in \mathbb{R} \mid -1 \leq x \leq 1\}$. It cycles smoothly between its maximum value at 1 and its minimum value at -1 over a period of 2π . That is, every positive and negative multiple of 2π the graph repeats itself. Its maximum value occurs at $x = \frac{\pi}{2}$, where it equals one, and its minimum value occurs at $x = \frac{3\pi}{2}$, where its value is -1 . These minima and maxima recur at every integer multiple of 2π of these values as well. Its amplitude is one half of the difference between the maximum and minimum, so that is $\frac{1}{2}(1 - (-1)) = \frac{1}{2} \cdot 2 = 1$. The zeros of the graph occur at $x = 0$, $x = \pi$, and at $x = 2\pi$ as well as at every positive and negative multiple of 2π of those. We can express this by saying that the zeros are at $x = 2\pi k$ and $x = \pi + 2\pi k$ for $k \in \mathbb{Z}$. Since in general a periodic function like sin will have infinitely-many zeros, most problems involving a sin function will limit the domain to an interval like $[0, 2\pi]$ to focus your attention on a small, finite set of zeros.

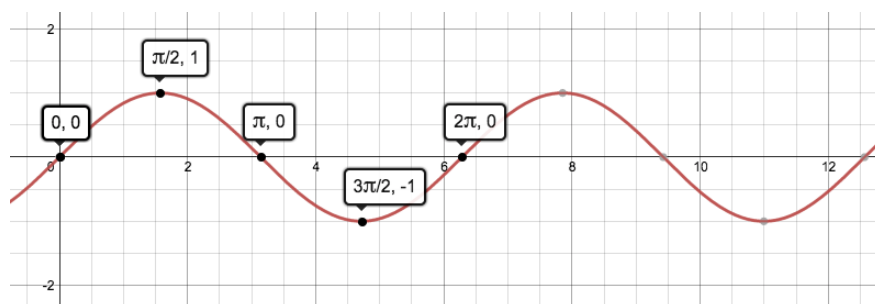


Figure 63: Portion of the Graph of $\sin x$

3.15.2 The Graph of $\cos x$

Figure 63 shows a portion of the graph of the $\cos x$ function. The domain of the function is all real numbers (\mathbb{R}), and its range is $\{x \in \mathbb{R} \mid -1 \leq x \leq 1\}$. It cycles smoothly between its maximum value at 1 and its minimum value at -1 over a period of 2π . That is, every positive and negative multiple of 2π the graph repeats itself. Its maximum value occurs at $x = 0$, where it equals one, and its minimum value occurs at $x = \pi$, where its value is -1 . These minima and maxima recur at every integer multiple of 2π of these values as well. Its amplitude is one half of the difference between the maximum and minimum, so that is $\frac{1}{2}(1 - (-1)) = \frac{1}{2} \cdot 2 = 1$. The zeros of the graph occur at $x = \frac{\pi}{2}$ and $x = \frac{3\pi}{2}$, as well as at every positive and negative multiple of 2π of those. We can express this by saying that the zeros are at $x = \frac{\pi}{2} + 2\pi k$ and $x = \frac{3\pi}{2} + 2\pi k$ for $k \in \mathbb{Z}$. Since in general a periodic function like cos will have infinitely-many zeros, most problems involving a cos function will limit the domain to an interval like $[0, 2\pi]$ to focus your attention on a small, finite set of zeros.

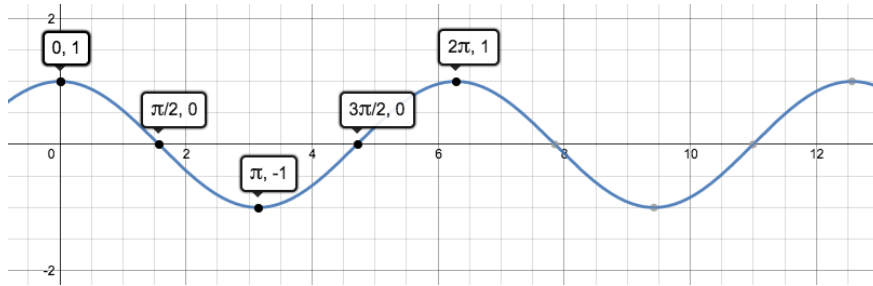


Figure 64: Portion of the Graph of $\cos x$

3.15.3 The Graph of $\tan x$

Figure 63 shows a portion of the graph of the $\tan x$ function. The domain of the function is all real numbers except integer multiples of $\frac{\pi}{2}$, where it has vertical asymptotes. Its range is all real numbers (\mathbb{R}). It cycles between $-\infty$ and ∞ repeating every π , so its period is π . As it has no minimum or maximum value, it does not make sense to talk about the amplitude. The zeros of the graph occur at $x = 0$ as well as at every positive and negative multiple of π . We can express this by saying that the zeros are at $x = \pi k$ for $k \in \mathbb{Z}$. Since in general a periodic function like \tan will have infinitely-many zeros, most problems involving a \tan function will limit the domain to an interval like $(-\frac{\pi}{2}, \frac{\pi}{2})$ to focus your attention on a small, finite set of zeros.

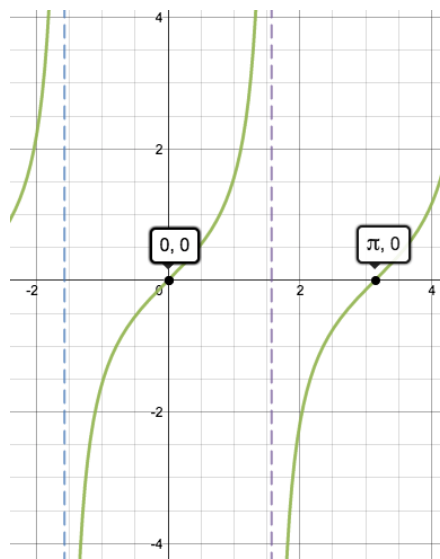


Figure 65: Portion of the Graph of $\tan x$

3.15.4 Transformations of Circular Functions

As described more generally in Section 2.15, “Transformations of Graphs”, the graphs of \sin , \cos , and \tan can be transformed in all the same ways. In its most general form, the graph $f(x)$ of a transformed sine function can be written with four coefficients that control the transformations of the original $\sin x$ graph:

$$f(x) = a \sin(b(x - c)) + d$$

From Section 2.15, “Transformations of Graphs”, you recognize that:

- a is a vertical stretch: it changes the amplitude from 1 to a
- b is a horizontal stretch: it changes the period from 2π to $\frac{2\pi}{b}$
- c is a horizontal translation: it shifts the graph c units to the right
- d is a vertical translation: it shifts the graph d units up

Of course if $a < 0$ then you can think of that as a reflection in the x -axis, followed by a vertical stretch of $|a|$. Similarly, if $b < 0$, then you can think of that as a reflection in the y -axis, followed by a horizontal stretch of $|b|$.

Notice that when written in this form, any coefficient in front of the x inside the sin function is factored out so that what is left inside the parenthesis has the x with a coefficient of 1. For example, given a function like $6 \sin(3x - 2)$, we cannot immediately “read off” the horizontal translation. We must first factor out the coefficient in front of x to get: $6 \sin\left(3\left(x - \frac{2}{3}\right)\right)$. Only when the x has a coefficient of 1 in front of it can we read off the horizontal translation. In this example, it is a shift to the right by $\frac{2}{3}$ (and not by 2!).

3.15.5 Using Transformation to Highlight Additional Identities

Figure 66 shows the graphs of both $\sin x$ and $\cos x$. You can see that the graphs’ shapes are identical, but that they are “out of phase” by $\frac{\pi}{2}$. It is visually clear that if we...

1. Shift $\cos x$ to the *right* by $\frac{\pi}{2}$, we get $\sin x$.
2. Shift $\cos x$ to the *left* by $\frac{\pi}{2}$ and reflect in the x -axis, we also get $\sin x$.
3. Shift the graph of $\sin x$ to the *left* by $\frac{\pi}{2}$ we get $\cos x$.
4. Shift $\sin x$ to the *right* by $\frac{\pi}{2}$ and reflect in the x -axis, we also get $\cos x$.

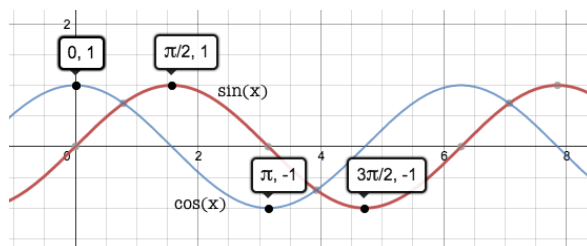


Figure 66: Portion of the Graphs of $\sin x$ and $\cos x$

These observations above explain the four handy identities below (for x in radians):

1. $\cos\left(x - \frac{\pi}{2}\right) = \sin x$
2. $-\cos\left(x + \frac{\pi}{2}\right) = \sin x$
3. $\sin\left(x + \frac{\pi}{2}\right) = \cos x$
4. $-\sin\left(x - \frac{\pi}{2}\right) = \cos x$

3.15.6 Determining Period from Minimum and Maximum

Sometimes a problem will tell you the minimum and maximum points on a sin or cos curve and ask you to find the period. We know that on both of these functions in one complete cycle, the minimum occurs only once and the maximum occurs only once. In the simplest version of the sin function with period 2π , the maximum occurs at $\frac{\pi}{2}$ and the minimum occurs at $\frac{3\pi}{2}$. This is $\frac{1}{4}$ the way around the circle and $\frac{3}{4}$ the way around the circle, respectively. Said another way, if you go from the minimum to the maximum you have gone $\frac{3}{4} - \frac{1}{4} = \frac{1}{2}$ of the way around, or half of the period.

Therefore, if a problem tells you the minimum and the maximum points, subtract their x -coordinates and that distance will be one half of the period. Double that value to get the period. To find out the value of the constant b that's needed to produce a transformed sin function that has the computed period value, use the equation:

$$\frac{2\pi}{b} = \text{period}$$

So, for example, if the minimum occurs at point $(-2,0)$ and the maximum occurs at point $(16,4)$, then the distance between their x -coordinates is $16 - (-2) = 18$ and that represents half of the period. The period is therefore $2 \times 18 = 36$, and the value of b that is needed in the general $a \sin(b(x - c)) + d$ form of the circular function would be:

$$\begin{aligned}\frac{2\pi}{b} &= 36 \\ \frac{2\pi}{36} &= b \\ b &= \frac{\pi}{18}\end{aligned}$$

3.16 Applications of the sin Function: Tide Example

During high tide, the water depth in a harbor is 22 m, and during low tide it is 10 m. Assuming a 12 h cycle:

- Find an expression for water depth t hours after low tide
- draw a graph of the function for a 48 h period
- State the times at which the water level is at a maximum, a minimum, and a mean sea level.

We need to find the appropriate values of a , b , c , and d to “adjust” the general form of the sin function ($a \sin(b(x - c)) + d$) to fit our data. We start by noting that the period is 12 hours. This means that we need $\frac{2\pi}{b} = 12$. Solving this simple equation gives $b = \frac{\pi}{6}$.

Next, we determine the amplitude of the function. This is one half of the difference between the maximum water depth of 22 m and the minimum water depth of 10 m: $a = \frac{1}{2}(22 - 10) = \frac{1}{2}(12) = 6$. After determining the period and the amplitude, we have a graph that looks like what you see in Figure 67. We can see that the period is correct because the graph completes one cycle exactly after twelve hours. We can see that the amplitude is right because it goes up to 6 and down to -6 .

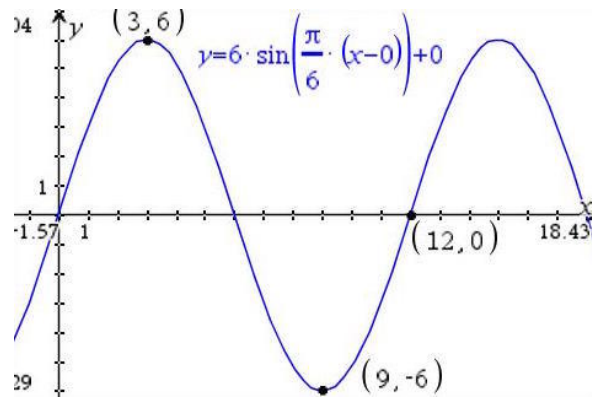


Figure 67: Initial sin Graph with Period and Amplitude Correct

The last steps are translating the graph horizontally and vertically to finish the job. Currently the graph has a maximum of 6, but in our problem statement the maximum is 22 m, so the value of d that we need to shift the graph up vertically is $d = 22 - 6 = 16$. After doing the vertical translation, we have the graph in Figure 68.

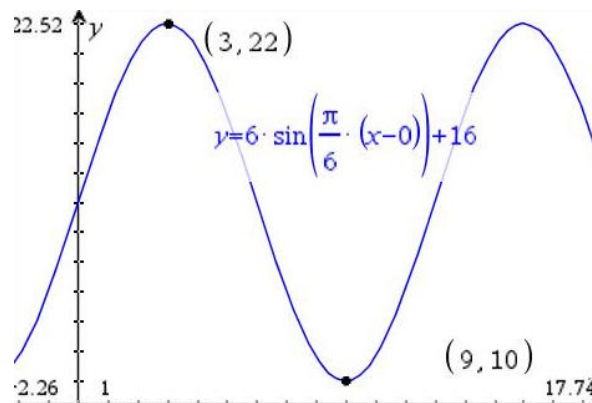


Figure 68: Initial sin Graph Attempt After Vertical Translation

Finally, the problem asks that the graph start at time $t = 0$ at the minimum value. At the moment the minimum value occurs after nine hours, so we need to shift the graph 9 units to the left to translate that point to be on the y -axis where $t = 0$. This means that our value for $c = -9$. This gives a final function of $6 \sin\left(\frac{\pi}{6}(x+9)\right) + 16$, whose graph appears in Figure 69.

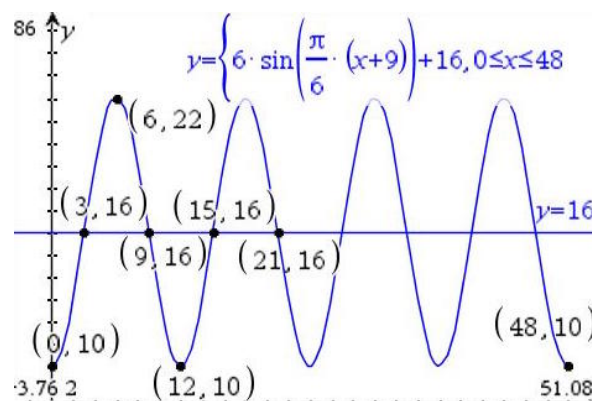


Figure 69: Final Tidal Graph Involving sin Graphed on Interval $[0, 48]$ Hours

So, the expression for the water depth t hours after low tide is $6 \sin\left(\frac{\pi}{6}(t+9)\right) + 16$. The graph of the tide depth over 48 hours appears in Figure 69. By also graphing the line $y = 16$, which is the horizontal “mid-line” halfway between the maximum and minimum values, and using the calculator to analyze the graph to find intersections, we can find all the points the problem asks for:

- The minimum tide depths occur at hours 0, 12, 24, 36, 48
- The maximum tide depths occur at hours 6, 18, 30, 42
- The mean tide depths occur at hours 3, 9, 15, 21, 27, 33, 39, 45

3.17 Applications of the \cos Function: Ferris Wheel Example

A ferris wheel of diameter 122 m rotates clockwise at constant speed. The wheel completes 2.4 rotations every hour. The bottom of the wheel is 13 m above the ground. A seat starts at the bottom of the wheel. The height above the ground of the seat after t minutes is $h(t) = 74 + a \cos(bt)$.

- Find the maximum height above the ground of the seat
- Show the period is 25 minutes
- Find the exact value of a and b
- Sketch the graph of h for $0 < t < 50$

At the lowest point on the bottom of the circle, we know the seat is 13 m above the ground. At the highest point, it will be an entire diameter higher than that, which is 122 m higher. So the maximum height above ground of the seat is $13 + 122 = 135$ m.

We’re told that in 60 minutes, the seat does 2.4 rotations. By dividing 60 by 2.4, we can find how many minutes it takes to perform a single rotation. $\frac{60}{2.4} = 25$, so a single rotation takes 25 minutes.

To find the value of b , we use the equation $\frac{2\pi}{b} = 25$. Solving for b gives $b = \frac{2\pi}{25}$.

To find the value of a , we find the difference between the maximum and minimum value of h , and then divide by two since the amplitude is the height above the horizontal midline of symmetry of the \cos function. To compute this, we calculate $\frac{135-13}{2} = \frac{122}{2} = 61$.

If we use these values of $a = 61$ and $b = \frac{2\pi}{25}$ in the equation for $h(t)$, then we get a graph like Figure 70. However, immediately we notice something is not right. The problem says that “A seat starts at the bottom of the wheel,” which means that the height of the seat at time zero $h(0)$ should equal 13. However, from the graph we see that $h(t) = 135$ and that it does take on the value of 135 until $h(12.5)$. So, things are backwards. We need the graph to start out “low” instead of starting “high” where it is now.

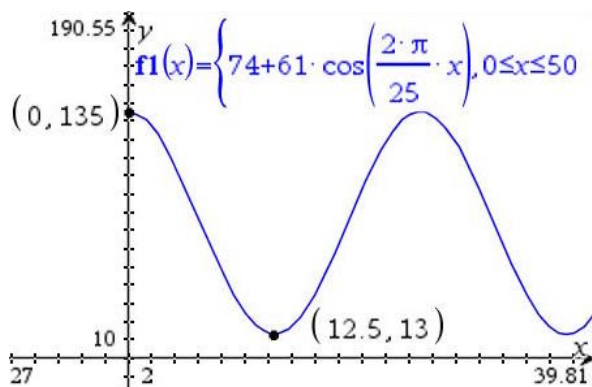


Figure 70: Initial Attempt at Graph of Ferris Wheel Seat Height

In order to correct this mistake, we simply need to reflect the graph vertically, reflecting it around its horizontal midline of symmetry. Recall from Section 2.15.2, “Vertical Reflection” that the way to reflect a graph vertically is to multiply the function \cos by -1 . This means that instead of using $a = 61$ we need to instead use $a = (-1) \cdot 61 = -61$. This keeps the amplitude of the \cos graph at the required value of 61, but flip-flops the graph so what was “up” becomes “down” instead. Figure 71 shows the graph of the resulting $h(t) = 74 - 61 \cos\left(\frac{2\pi}{25}t\right)$ function using this new, correct value for a , and restricts the graph as requested only to the values of t between 0 and 50 minutes. Note that the calculator requires the function variable to be x instead of allow us to use t , but the result is the same. Notice that now, as demanded by the problem statement, the height of the ferris wheel seat at time zero $h(0)$ equals 13 as it should, and one half of the way through the period of 25 minutes, the height at time 12.5 minutes $h(12.5) = 135$ which occurs when the seat is at its maximum height. So, the moral of the story is that sometimes after determining the amplitude, it might be necessary to use the negative value of that amplitude as the multiplier in front of the original function if the problem requires the graph to be flipped vertically. Therefore, $a = -61$.

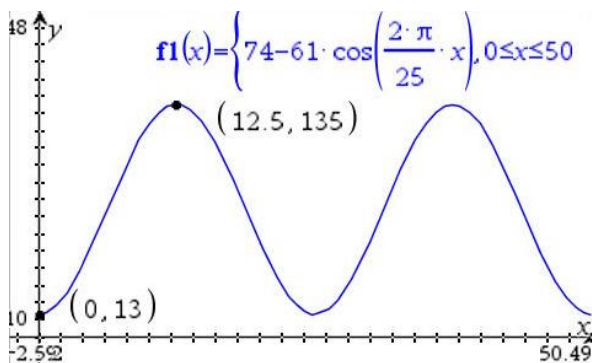


Figure 71: Correct Ferris Wheel Graph with Vertically Reflection

NOTE: Remember from Section 2.14.2, “Restricting the Domain of a Graph” that when entering a function to graph on the TI-NSpire, it’s possible to type the vertical bar “such that” symbol and provide a range of values that constrains the function domain. In this example, I typed the function in as $f1(x) = 74 - 61 \cos\left(\frac{2\pi}{25}x\right) \mid 0 \leq x \leq 50$

3.18 Solving Trigonometric Equations in a Finite Interval

When you are given an equation involving a trigonometric function that does **not** involve any squared functions (e.g. $\sin^2 x$, $\cos^2 x$, or $\tan^2 x$), then you can isolate the trig function on one side of the equation and then use your calculator to take the appropriate inverse function of both sides to solve the equation. The problem will always provide an interval in which to seek the solution so that the number of answers is a small, finite number.

3.19 Solving Quadratic Equations in \sin , \cos , and \tan

When you are given an equation involving squared trigonometric functions like $\sin^2 x$, $\cos^2 x$, or $\tan^2 x$, then to solve them, follow these steps:

1. If the equation involves different trigonometric functions (e.g. both $\sin x$ and $\cos x$), then use an appropriate identity to rewrite the equation to use only a single trigonometric function (e.g. $\sin^2 x + \cos^2 x = 1$).
2. Substitute a variable u for the trigonometric function to end up with a simple quadratic equation in u^2 and u .
3. Use the quadratic equation or factoring to find the two solutions for u .
4. Replace u by its trigonometric function again and solve the equations using the unit circle or inverse trig functions.

Consider the example equation $6 \cos^2 x + 7 \sin x = 8$.

Since it involves both cosine and sine functions, use an identity to replace the $\cos^2 x$ so that the equation ends up using only $\sin x$. The most common identity for this purpose is one described in Section 3.10, "The Pythagorean Identity". Since we know from this identity that $\sin^2 x + \cos^2 x = 1$, rearranging this equation we get $\cos^2 x = 1 - \sin^2 x$. Substituting the right hand side of this identity equation into the original problem, we get:

$$\begin{aligned}6 \cos^2 x + 7 \sin x &= 8 \\6(1 - \sin^2 x) + 7 \sin x - 8 &= 0 \\6 - 6 \sin^2 x + 7 \sin x - 8 &= 0 \\-6 \sin^2 x + 7 \sin x - 2 &= 0 \\6 \sin^2 x - 7 \sin x + 2 &= 0\end{aligned}$$

Next substitute $u = \sin x$ to get a simpler-looking quadratic equation to solve: $6u^2 - 7u + 2 = 0$. Using the quadratic equation with $a = 6$, $b = -7$, and $c = 2$, we solve for u :

$$u = \frac{-(-7) \pm \sqrt{(-7)^2 - 4 \cdot 6 \cdot 2}}{2 \cdot 6} = \frac{7 \pm \sqrt{49 - 48}}{12} = \frac{7 \pm 1}{12} = \frac{8}{12} \text{ or } \frac{6}{12} = \frac{2}{3} \text{ or } \frac{1}{2}$$

Now, substitute back the original expression for u so that we have:

$$\begin{array}{ll}u = \frac{2}{3} & \text{or} & u = \frac{1}{2} \\ \sin x = \frac{2}{3} & \text{or} & \sin x = \frac{1}{2} \\ \sin^{-1}(\sin x) = \sin^{-1} \frac{2}{3} & \text{or} & \sin^{-1}(\sin x) = \sin^{-1} \frac{1}{2} \\ x = 0.72973 & \text{or} & x = \frac{\pi}{6}\end{array}$$

Notice that one of the values of $\sin x$ was easily recognizable from the “key angles” of the unit circle, so in that case we provide the exact answer for x of $\frac{\pi}{6}$. In the other case, we needed to use the calculator to compute the inverse sine of $\frac{2}{3}$. To three significant figures, $x \in \{0.730, \frac{\pi}{6}\}$.

Before giving the final answer, we need to look again at the interval stipulated by the original problem. They ask us to find the solutions over the interval $[0, 4\pi]$. Given this, we need to expand the solution set to include any multiples of 2π that would still fall inside this interval. As a decimal, 4π is approximately 12.57 and 2π is approximately 6.28. So, for the decimal solution $x = 0.730$ we can add 6.28 once and still be inside the 12.57 upper bound. This gives us one additional decimal solution of $0.730 + 6.28 = 7.01$. Adding 2π to $\frac{\pi}{6}$ we can get another additional solution of $\frac{13\pi}{6}$ that is still inside the upper bound. Therefore, the final solution to the problem is: $x \in \left\{0.730, 7.01, \frac{\pi}{6}, \frac{13\pi}{6}\right\}$.

3.20 Solutions of Right Triangles

In addition to the Pythagorean Theorem, trigonometric functions give you a new way to solve for missing lengths in a right triangle when you know the measure of one of the angles. Figure 72 shows two different kinds of angles that can occur in problems like this, so-called angles of depression and elevation. For a particular problem, in practice, the angles are the same. One man’s angle of depression is another man’s angle of elevation as shown in the figure.

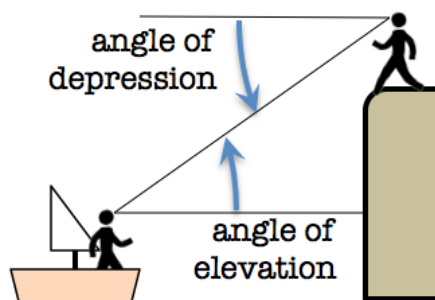


Figure 72: Angles of Depression and Elevation are Equal

Consider the example showing in Figure 73, where a man stands 100m away from the front door of the Chrysler Building, and measures an angle of elevation to the top of the building’s spire of 73° . How tall is the building?

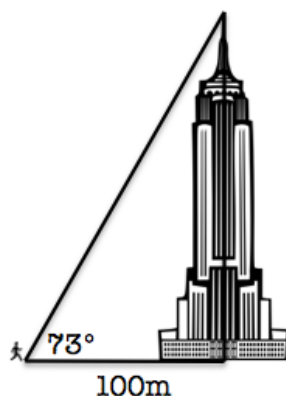


Figure 73: Using the tan Function to Calculate a Building’s Height

If we call the unknown height of the building x , then we can use the tangent function to represent an equation that relates the angle of elevation to the ratio of the opposite over adjacent sides of the right triangle formed by the little man, the front door of the building, and the top of the spire:

$$\tan 73^\circ = \frac{\textit{opposite}}{\textit{adjacent}} = \frac{x}{100}$$

Normally the tan function expects an angle measure in radians as its argument. On your calculator, you can enter the degree symbol after the 73 and it will automatically convert the degrees to radians and then compute the tangent. Be careful to add the degree symbol, or else the calculator will compute the tangent of 73 radians! The calculator gives 3.27 as the value of the $\tan 73^\circ$, so we have:

$$\begin{aligned}\tan 73^\circ = 3.27 &= \frac{x}{100} \\ 100 \cdot 3.27 &= x \\ x &= 327\end{aligned}$$

So, the building is 327 m tall.

3.21 The Cosine Rule

The Cosine Rule is a more generalized form of the Pythagorean Theorem. While the latter only is applicable to right triangles, this rule works for any triangle. The Cosine Rule is useful when you are given the lengths of two sides and the angle between them. It allows you to solve for the length of the missing side.

$$\begin{aligned}a^2 &= b^2 + c^2 - 2bc \cos A \\ b^2 &= a^2 + c^2 - 2ac \cos B \\ c^2 &= a^2 + b^2 - 2ab \cos C\end{aligned}$$

3.22 The Sine Rule

The Sine Rule says that in any triangle, the ratio of the length of a side to the sin of the angle opposite that side stays the same. This rule is useful when you know any angle in a triangle as well as the length of the side opposite. Then, depending on whether you know another angle or another side, you can use the Sine Rule to compute the missing quantity.

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

Consider the triangle in Figure 74. Suppose you are given the length of sides a and b , as well as angle A , and are asked to find angle B that you are told (and can see) is obtuse.

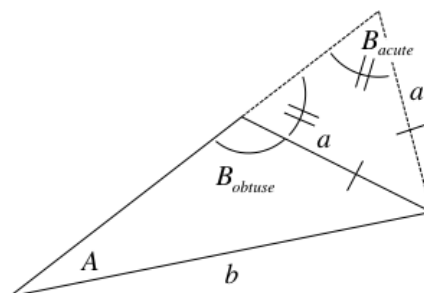


Figure 74: Sine Rule Always Finds the Acute Angle

Using the Sine Rule, you will therefore solve this equation for B :

$$\begin{aligned}\frac{a}{\sin A} &= \frac{b}{\sin B} \\ \sin B &= \frac{b \cdot \sin A}{a} \\ \sin^{-1}(\sin B) &= \sin^{-1}\left(\frac{b \cdot \sin A}{a}\right) \\ B &= \sin^{-1}\left(\frac{b \cdot \sin A}{a}\right)\end{aligned}$$

Since the range of the \sin^{-1} function is from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$, the value that you compute for B will be an acute angle. You can see in the figure above that the dotted line with the same length as side a can also form another triangle with side length b and angle A that has an acute angle labeled B_{acute} . That is the angle whose measure you find using the computation above. In the isosceles triangle with equal side-lengths a , we know that the two base angles are equal. So the value of the obtuse angle you are looking for B_{obtuse} is the complement of B_{acute} , which means that $B_{\text{obtuse}} = \pi - B_{\text{acute}}$ (or $180^\circ - B_{\text{acute}}$ if in degrees).

3.23 Area of a Triangle

If you know the length of two sides of a triangle, and the angle between them, then you can compute the area of the triangle with the formula:

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

4 Vectors

4.1 Vectors as Displacements in the Plane

A vector represents the shortest path in (two-dimensional) space starting at some initial point A and terminating at another point B . Unlike a segment, a vector is represented as an arrow, whose length indicates the distance required to move in a straight line from A to B , and whose arrow head indicates the direction of the movement. Two vectors are equal if they have the same magnitude and the same direction, irrespective of their initial point. As shown in Figure 75, the vector \vec{v} indicates the directed straight path from A to B , but it is the same vector if it is translated to have its initial point coincide with the origin of the Cartesian plane. In either position, the vector \vec{v} points in the same direction and has the same length (also known as *magnitude*).

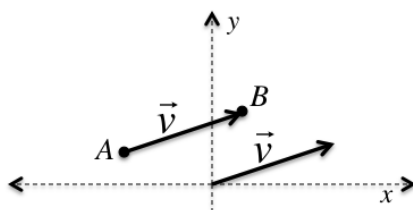


Figure 75: Vectors Represent Distance and Direction, Irrespective of Origin Point

We typically distinguish the name of a vector from other kinds of variable names with an arrow written above and a bolder type. If the initial and terminal points of a vector are labeled with letters like A and B , then an equivalent way of naming the vector uses these point labels listed in the order the arrow is pointing. So, for example, Figure 75 vector \vec{v} can also be written \overrightarrow{AB} .

4.2 Vectors as Displacements in Three Dimensions

A vector has exactly the same meaning in three dimensions as it does in the plane. It represents the shortest path through space starting at some initial point A and terminating at another point B . Figure 76 shows such a vector with initial point along the x -axis and terminating at a point above the y -axis in the yz -plane. As was the case in two dimensions, we can always translate a vector so that its initial point is the origin of the x -, y -, and z -axes without changing the vector. In other words, both vectors labeled \vec{v} in Figure 76 are the same vector: same magnitude or length, same direction.

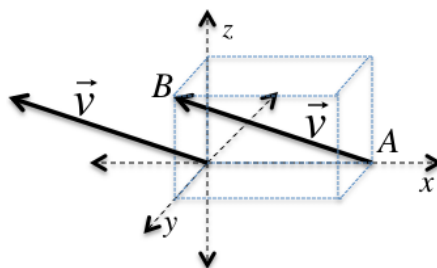


Figure 76: Vectors Also Represent Distance and Direction in 3 Dimensions

4.3 Terminology: Tip and Tail

Since vectors are drawn as arrows, the terminal point with the arrowhead is sometimes called the *tip* of the vector, while the initial point is called the *tail*. When one vector starts at exactly the place where another vector terminates, we say that they are placed “tip to tail”.

4.4 Representation of Vectors

Since vectors having the same length and direction are equivalent, we always describe a vector as if its initial point were at the origin. By doing this, we can use the (x, y) coordinates of its terminal point to describe it. Figure 77 shows two vectors \vec{v} and \vec{w} . The x - and y -coordinates of the point $(4, 3)$ for vector \vec{v} correspond to the fact that to get from the initial point at the origin to that terminal point, you need to “walk” 4 units to the right and 3 units up. Similarly, for \vec{w} , its terminal point $(-2, 2)$ means that to travel from the origin to there, you need to go 2 units left, and 2 units up. Even though the length of the vector is measured along the straight path from the initial point to the terminal point, the vector with initial point at the origin is uniquely identified by these x and y coordinates that represent the distance of the legs of a right triangle having the vector as the hypotenuse.

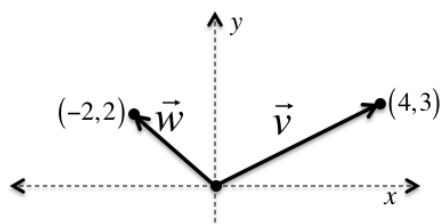


Figure 77: Coordinates of Terminal Point Uniquely Identify a Vector

To make it clear we are talking about a vector instead of just a point in the plane, we not only write the vector variable with the bold font and arrow above, we also write the coordinates of the vector’s terminal point in a column like this:

$$\vec{v} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$
$$\vec{w} = \begin{pmatrix} -2 \\ 2 \end{pmatrix}$$

In general, a vector in the plane is written in a column with its x -coordinate at the top, and y -coordinate at the bottom. A vector in three dimensions extends this to list the z -coordinate at the bottom, like this:

$$\vec{a} = \begin{pmatrix} 4 \\ 3 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{Two-dimensional vector}$$
$$\vec{b} = \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \text{Three-dimensional vector}$$

4.5 Magnitude of a Vector

Given a vector \vec{v} , its length (known as *magnitude*) is written $|\vec{v}|$. As shown in Figure 78, we calculate the magnitude using the Pythagorean Theorem with the x - and y -coordinates as the

legs of the right triangle for which the vector is the hypotenuse. This gives:

$$|\vec{v}| = \sqrt{4^2 + 3^2} = \sqrt{25} = 5$$

Therefore, in general we have:

$$\left| \begin{pmatrix} a \\ b \end{pmatrix} \right| = \sqrt{a^2 + b^2}$$

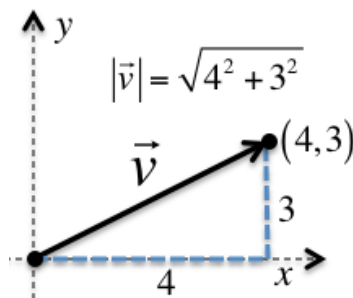


Figure 78: Using the Pythagorean Theorem to Find (2D) Vector Magnitude

As shown in Figure 79, for a three-dimensional vector we also calculate the magnitude using the Pythagorean Theorem. First we use the theorem to find the length of the diagonal in the xy -plane, which is $\sqrt{(-4)^2 + (-2)^2}$. Then we use the theorem again with the right triangle having this first diagonal as one leg and the segment parallel to the z -axis as the other leg. This gives:

$$\begin{aligned} |\vec{v}| &= \sqrt{\left(\sqrt{(-4)^2 + (-2)^2}\right)^2 + 3^2} \\ &= \sqrt{(-4)^2 + (-2)^2 + 3^2} = \sqrt{16 + 4 + 9} = \sqrt{29} \end{aligned}$$

Therefore, in general we have:

$$\left| \begin{pmatrix} a \\ b \\ c \end{pmatrix} \right| = \sqrt{a^2 + b^2 + c^2}$$

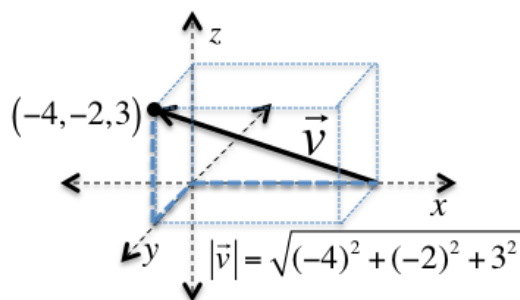


Figure 79: Using the Pythagorean Theorem to Find (3D) Vector Magnitude

4.6 Multiplication of a Vector by a Scalar

To multiply a vector by a scalar k , simply multiply each coordinate by k :

$$k \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} ka \\ kb \\ kc \end{pmatrix}$$

As shown in Figure 80, when you multiply a vector by a scalar number like 2, you stretch it 2 times in the same direction it's already pointing. A consequence is that the magnitude of the new vector is 2 times the magnitude of the original vector.

$$|k\vec{a}| = k|\vec{a}|$$

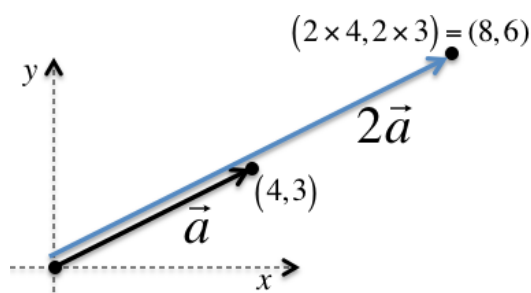


Figure 80: Multiplying a Vector by a Scalar Stretches It By That Factor

4.7 Negating a Vector

Given a vector \vec{a} , you can negate the vector by putting a minus sign in front of it: $-\vec{a}$. Just as with simple numbers, putting a minus sign in front is the same as multiplying the vector by

the scalar -1 . So if $\vec{a} = \begin{pmatrix} a \\ b \end{pmatrix}$, we have:

$$-\vec{a} = - \begin{pmatrix} a \\ b \end{pmatrix} = -1 \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -a \\ -b \end{pmatrix}$$

Note that the magnitude is not changed by this operation:

$$|-\vec{a}| = |\vec{a}|$$

Figure 81 shows the visual interpretation of what negating a vector does. It flips the vector around so that it's pointing in the opposite direction, without changing its length. As with any vector, we can then translate the negated vector so that its initial point is at the origin without changing it.

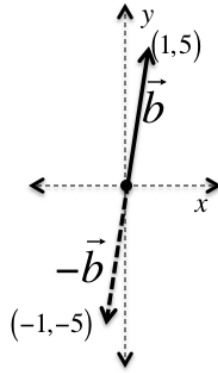


Figure 81: Negating a Vector Makes It Point in the Opposite Direction

4.8 Sum of Vectors

To add vectors, simply add their coordinates:

$$\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} d \\ e \end{pmatrix} = \begin{pmatrix} a + d \\ b + e \end{pmatrix}$$

Figure 82 shows the visual interpretation of adding vectors $\vec{a} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$ and $\vec{b} = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$ together.

Since we can move vector \vec{b} 's initial point around without changing it, we move it to coincide with the terminal point of \vec{a} as shown in the figure. The vector $\vec{a} + \vec{b}$ is the result of first traveling in the direction of \vec{a} for the distance $|\vec{a}|$, then traveling in the direction of \vec{b} for the distance $|\vec{b}|$. This is exactly the same destination you would get to by starting at the origin and traveling the distance of the diagonal of the parallelogram formed by the two vectors. For this reason vector addition is sometimes called the “parallelogram rule”.

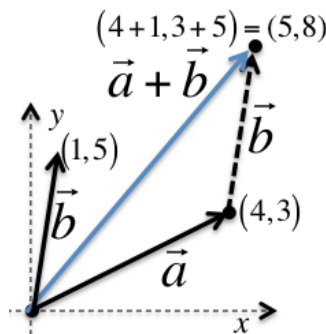


Figure 82: Adding Vectors Produces Their Parallelogram Diagonal

4.9 Difference of Vectors

To subtract vectors, simply subtract their coordinates in the correct order:

$$\begin{pmatrix} a \\ b \end{pmatrix} - \begin{pmatrix} d \\ e \end{pmatrix} = \begin{pmatrix} a - d \\ b - e \end{pmatrix}$$

Figure 83 shows the visual interpretation of $\vec{a} - \vec{b}$. Recall that $\vec{a} - \vec{b} = \vec{a} + (-\vec{b})$, so we can start by negating \vec{b} to flip its direction around, translate the resulting $-\vec{b}$ to start at the origin, and then finally use the “parallelogram rule” above for adding vectors \vec{a} and $-\vec{b}$. If we translate the resulting vector $\vec{a} - \vec{b}$ to start at the terminal point of \vec{b} , then we can say that $\vec{a} - \vec{b}$ is the vector that connects the tip of \vec{b} to the tip of \vec{a} .

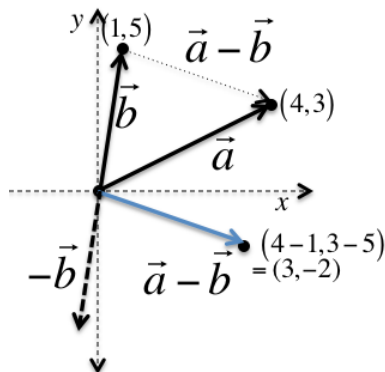


Figure 83: Vector $\vec{a} - \vec{b}$ Connects Tip of \vec{b} to Tip of \vec{a}

Just as subtraction of numbers is not commutative (e.g. $5 - 3 \neq 3 - 5$), the same is true of vector subtraction. Figure 84 shows the visual interpretation of $\vec{b} - \vec{a}$. As we did above, we start by negating \vec{a} and translating its initial point to the origin. Then, we add the vector \vec{b} with this $-\vec{a}$ using the parallelogram rule. If we translate the resulting vector $\vec{b} - \vec{a}$ to start at the terminal point of \vec{a} , then we can say that $\vec{b} - \vec{a}$ is the vector that connects the tip of \vec{a} to the tip of \vec{b} . If you compare the result to Figure 83, you’ll see that it’s the same vector but with the opposite direction. In other words, we have:

$$\vec{a} - \vec{b} = -(\vec{b} - \vec{a})$$

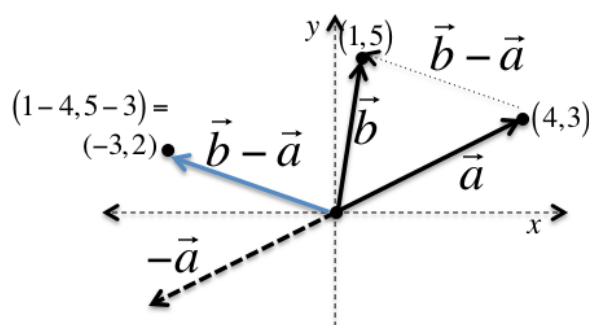


Figure 84: Vector $\vec{b} - \vec{a}$ Connects Tip of \vec{a} to Tip of \vec{b}

4.10 Unit Vectors

A unit vector is one whose length is equal to one unit. For example, the following three vectors all are unit vectors because their length is one:

$$\begin{aligned}\left| \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right| &= \sqrt{1^2 + 0^2} = \sqrt{1} = 1 \\ \left| \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right| &= \sqrt{0^2 + 1^2} = \sqrt{1} = 1 \\ \left| \begin{pmatrix} \frac{4}{5} \\ \frac{3}{5} \end{pmatrix} \right| &= \sqrt{\frac{4^2}{5^2} + \frac{3^2}{5^2}} = \sqrt{\frac{4^2 + 3^2}{5^2}} = \sqrt{\frac{16 + 9}{25}} = \sqrt{\frac{25}{25}} = \sqrt{1} = 1\end{aligned}$$

4.11 Scaling Any Vector to Produce a Parallel Unit Vector

You can make any vector \vec{v} into a unit vector by following these two steps:

- If $|\vec{v}|=1$, then you're done. It's already a unit vector
- Otherwise, multiply \vec{v} by the reciprocal of $|\vec{v}|$. This will scale \vec{v} to stretch or shrink it to have exactly the length 1 without changing its direction.

For example, consider the vector $\begin{pmatrix} 4 \\ 5 \end{pmatrix}$. We start by computing:

$$\left| \begin{pmatrix} 4 \\ 5 \end{pmatrix} \right| = \sqrt{4^2 + 5^2} = \sqrt{16 + 25} = \sqrt{41}$$

So the vector's length is not equal to 1, therefore it is not a unit vector. To turn it into a unit vector, multiply it by the reciprocal of its magnitude:

$$\frac{1}{\sqrt{41}} \begin{pmatrix} 4 \\ 5 \end{pmatrix} = \begin{pmatrix} \frac{4}{\sqrt{41}} \\ \frac{5}{\sqrt{41}} \end{pmatrix}$$

This vector should be a unit vector now. We can verify it by computing its magnitude:

$$\left| \begin{pmatrix} \frac{4}{\sqrt{41}} \\ \frac{5}{\sqrt{41}} \end{pmatrix} \right| = \sqrt{\left(\frac{4}{\sqrt{41}}\right)^2 + \left(\frac{5}{\sqrt{41}}\right)^2} = \sqrt{\frac{4^2 + 5^2}{41}} = \sqrt{\frac{41}{41}} = \sqrt{1} = 1$$

4.12 Position Vectors

Given a point A , the *position vector* of A (typically written \vec{a}) is the vector with initial point at the origin and terminal point at A . Given another point B , it also has an associated position vector \vec{b} that points from the origin to B . As shown in Figure 85, the vector \vec{AB} is therefore $\vec{b} - \vec{a}$.

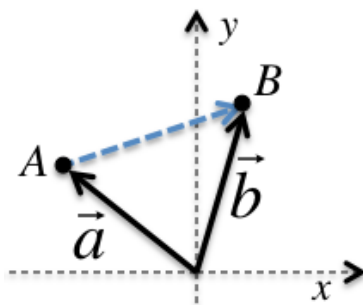


Figure 85: Vector \vec{AB} Connects Tip of Position Vector \vec{a} to Tip of Position Vector \vec{b}

4.13 Determining Whether Vectors are Parallel

Two vectors \vec{a} and \vec{b} are parallel if one is a scalar multiple of the other. Recall that you can change the initial point of \vec{a} and \vec{b} by translating them to the origin (without changing either one's direction!). They are parallel, then, if $\vec{a} = k\vec{b}$ for some scalar k . In a visual sense, it means that \vec{a} and \vec{b} are pointing in the same direction and you can stretch (or shrink) \vec{a} by some amount to get \vec{b} .

Given two vectors $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ and $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$, to determine if they are parallel compute

the three quotients: $\frac{b_1}{a_1}$, $\frac{b_2}{a_2}$, and $\frac{b_3}{a_3}$.

If any two of these quotients are not equal, then you can conclude that \vec{a} and \vec{b} are **NOT** parallel. If all three of the quotients are equal, then that value is the scalar multiple k by which \vec{a} was multiplied to get \vec{b} .

For example, consider the vectors $\vec{a} = \begin{pmatrix} 7 \\ -3 \\ 5 \end{pmatrix}$ and $\vec{b} = \begin{pmatrix} -91 \\ 39 \\ -65 \end{pmatrix}$. Computing the three quotients

we get:

$$\frac{-91}{7} = -13 \qquad \frac{39}{-3} = -13 \qquad \frac{-65}{5} = -13$$

Since all three quotients equal -13 , then we know that $\vec{b} = -13\vec{a}$ and this tells us that they are parallel, since one is a scalar multiple of the other.

4.14 Finding Parallel Vector with Certain Fixed Length

Given a vector \vec{v} of a given length $|\vec{v}|$, you may be asked to find a parallel vector \vec{w} with a specific *different* length L . The procedure is similar to that described in Section 4.11, “Scaling Any Vector to Produce a Parallel Unit Vector”, with one additional twist. Instead of scaling the original vector by multiplying by $\frac{1}{|\vec{v}|}$, scale it by that factor multiplied by the desired length L :

$$\vec{w} = \frac{L}{|\vec{v}|} \vec{v}$$

4.15 Scalar (or “Dot”) Product of Two Vectors

The scalar product (also known as the *dot product*) of vectors \vec{a} and \vec{b} is written $\vec{a} \cdot \vec{b}$ and is defined as follows:

$$\text{If } \vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \text{ and } \vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \text{ then } \vec{a} \cdot \vec{b} = a_1b_1 + a_2b_2$$

The scalar product is called that because the result is a scalar number, not a vector. It’s also called the dot product because the operation is written with a dot between the two vectors. If the vectors are in three dimensions, the scalar product also includes the product of their z -coordinate values:

$$\text{If } \vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \text{ and } \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \text{ then } \vec{a} \cdot \vec{b} = a_1b_1 + a_2b_2 + a_3b_3$$

4.16 Perpendicular Vectors

Two vectors \vec{a} and \vec{b} are perpendicular if and only if their scalar product is zero:

$$\vec{a} \cdot \vec{b} = 0 \iff \vec{a} \perp \vec{b}$$

4.17 Base Vectors for Two Dimensions

In two dimensions, base vectors are a set of two perpendicular unit vectors. Any vector in two dimensions can be expressed as the sum of appropriate scalar multiples of these two base vectors. The canonical base vectors comprise the set $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$. This is the unit vector

from the origin pointing a distance of one unit along the x -axis and the unit vector from the origin pointing a distance of one unit up the y -axis. By convention, the one pointing along the x -axis is called \vec{i} and the one pointing along the y -axis is called \vec{j} . Figure 86 shows the point

$(3, 4)$ in the plane, and its position vector. In column notation, this vector is written $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$.

However, if we multiply the \vec{i} unit vector by the scalar value 3 to get $3\vec{i}$ and we multiply the \vec{j} unit vector by the scalar value 4 to get $4\vec{j}$, then we can use the parallelogram rule from

Section 4.8, “Sum of Vectors” to add $3\vec{i} + 4\vec{j}$ to get the position vector for the point $(3, 4)$.

When we write a vector $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$ in the form $3\vec{i} + 4\vec{j}$, we say that it is in “base vector form”.

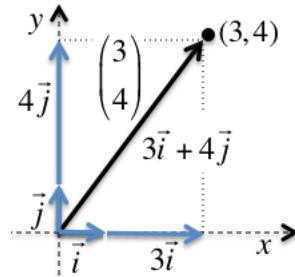


Figure 86: Any Vector Can Be Expressed as Sum of Scalar Multiples of Base Vectors

4.18 Base Vectors for Three Dimensions

In three dimensions, base vectors are a set of three perpendicular unit vectors. Any vector in three dimensions can be expressed as the sum of appropriate scalar multiples of these three base

vectors. The canonical base vectors comprise the set $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$. This is the unit

vector from the origin pointing a distance of one unit along the x -axis, the unit vector from the origin pointing a distance of one unit up the y -axis, and the unit vector from the origin pointing a distance of one unit up the z -axis. By convention, these are called \vec{i} , \vec{j} , and \vec{k} respectively. In an analogous way to what we illustrated in Section 4.17, “Base Vectors for Two Dimensions”, in three dimensions any vector can be written in base vector form. For example:

$$\begin{pmatrix} 3 \\ 17 \\ -2 \end{pmatrix} = 3\vec{i} + 17\vec{j} - 2\vec{k}$$

4.19 The Angle Between Two Vectors

Given the two vectors \vec{a} and \vec{b} in Figure 87, recall that the cosine of the angle θ between them is:

$$\cos \theta = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}| |\vec{b}|} \text{ therefore } \theta = \cos^{-1} \left(\frac{\vec{a} \cdot \vec{b}}{|\vec{a}| |\vec{b}|} \right)$$

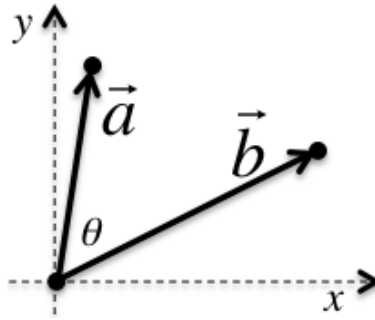


Figure 87: Angle θ Between \mathbf{a} and \mathbf{b} is $\cos^{-1}\left(\frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}\right)$

If vectors $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ are perpendicular, then the angle between them is $\frac{\pi}{2}$. We know from the unit circle “key angles” that the $\cos \frac{\pi}{2} = 0$. This means that $\frac{\vec{\mathbf{a}} \cdot \vec{\mathbf{b}}}{|\vec{\mathbf{a}}||\vec{\mathbf{b}}|} = \cos \frac{\pi}{2} = 0$, so for that to happen, it must be the case that $\vec{\mathbf{a}} \cdot \vec{\mathbf{b}} = 0$. This is the condition we learned about in Section 4.16, “Perpendicular Vectors”.

4.20 Vector Equation of a Line in Two and Three Dimensions

A line is a straight, one-dimensional figure that we normally describe using the equation $y = mx + b$ in slope-intercept form, or $y - y_1 = m(x - x_1)$ in point-slope form. In these two ways to express the set of points (x, y) that belong to a straight line, the slope m is the value that determines the “direction” the line is pointing, and the y-intercept $(0, b)$ or arbitrary point (x_1, y_1) defines a point that the line passes through. In both cases, once we’ve indicated a direction for the line and one point that it passes through, we have uniquely defined the line.

Using these same concepts, we can use a vector equation as an alternative way to describe a line. Figure 88 shows the technique we’ll adopt. Given a line L , we need the same two “ingredients” as in the two examples above: a *direction vector* $\vec{\mathbf{d}}$ that indicates the direction the line is pointing, and a *position vector* $\vec{\mathbf{p}}$ for some point that the line passes through.

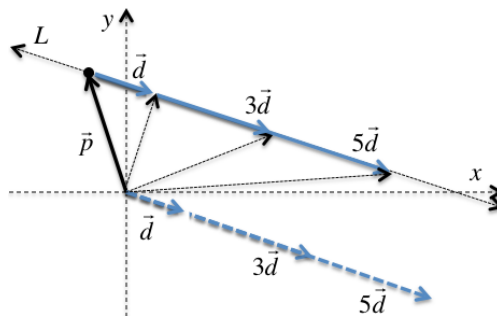


Figure 88: Vector Line $\mathbf{r} = \mathbf{p} + k\mathbf{d}$ Has Position Vector \mathbf{p} and Direction Vector \mathbf{d}

As shown in Figure 88, if we use the “parallelogram rule” to add $\vec{\mathbf{p}}$ and $\vec{\mathbf{d}}$ (after translating it to the origin as well), we get a position vector for some other point on the line L . If we repeat the process of adding $\vec{\mathbf{p}}$ to another multiple of the direction vector like $3\vec{\mathbf{d}}$, again we get another position vector for different point on the line L . Trying a third time, if we add $\vec{\mathbf{p}}$ to yet another multiple of the direction vector like $5\vec{\mathbf{d}}$, we get a third position vector for yet another point on the line L . If instead of adding the position vector $\vec{\mathbf{p}}$ only to a few different scalar multiples of the direction vector $\vec{\mathbf{d}}$, we added to $\vec{\mathbf{p}}$ every possible positive or negative real number multiple k of the direction vector, then we could end up producing position vectors

for every point on the line L .

Since the line L is made up of infinitely many position vectors, one that connects the origin to each point on the line, we can describe the set of these infinitely many position vectors using the vector equation:

$$\vec{r} = \vec{p} + k\vec{d} \quad (k \in \mathbb{R})$$

In other words, the line L is the set of all vectors \vec{r} that result by taking a position vector \vec{p} for some point on the line and adding to it every possible real number multiple of the direction vector \vec{d} that determines the line's direction.

Keep in mind that any position vector on the line is fine to use for the position vector of the vector line equation. If you happen to know more than one position vector on the line, just pick one of them. It doesn't matter which one you choose. Once you add all the scalar multiples of the direction vector to it, you will still end up producing the same line L .

4.21 Vector Equation of Line Passing Through Two Points

Given two points, for example $(5, 6)$ and $(1, 3)$, to determine the vector equation of the line that passes through both of them, do the following:

- Write the points in vector form and subtract one from the other to get the direction vector.
- Pick one of the two points (doesn't matter which) to use as the position vector

So, $\begin{pmatrix} 5 \\ 6 \end{pmatrix} - \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$. This is the direction vector. Subtracting them in the other order

would give $\begin{pmatrix} -4 \\ -3 \end{pmatrix} = -\begin{pmatrix} 4 \\ 3 \end{pmatrix}$, so it's just pointing in the opposite direction. It still ends up

generating the same line when we consider adding all real number multiples of the direction vector to the position vector.

For the position vector, pick either one. For example, we can use $\begin{pmatrix} 5 \\ 6 \end{pmatrix}$. So, writing the equation in vector line form we get:

$$\vec{r} = \begin{pmatrix} 5 \\ 6 \end{pmatrix} + k \begin{pmatrix} 4 \\ 3 \end{pmatrix} \quad (k \in \mathbb{R})$$

4.22 Finding the Cartesian Equation from a Vector Line

If you're given a line in vector line form $\vec{r} = \begin{pmatrix} 5 \\ 6 \end{pmatrix} + k \begin{pmatrix} 4 \\ 3 \end{pmatrix}$, you can easily convert it to

Cartesian form by following these steps:

- Write the vector \vec{r} as $\begin{pmatrix} x \\ y \end{pmatrix}$
- Write the system of two equations involving x and k as well as involving y and k
- Choose one of the two equations and solve for k
- Substitute that value of k into the other equation to eliminate k and end up with a Cartesian equation in only x and y .

In the example above, start by replacing \vec{r} with $\begin{pmatrix} x \\ y \end{pmatrix}$:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix} + k \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

So the system of equations is:

$$\begin{cases} x = 5 + 4k \\ y = 6 + 3k \end{cases}$$

Using the first equation to solve for k we get:

$$\begin{aligned} x &= 5 + 4k \\ x - 5 &= 4k \\ \frac{1}{4}x - \frac{5}{4} &= k \end{aligned}$$

Substituting this value of k into the other equation in the system, we can eliminate k and end up with a Cartesian equation for the line in only x and y :

$$\begin{aligned} y &= 6 + 3k \\ y &= 6 + 3\left(\frac{1}{4}x - \frac{5}{4}\right) \\ y &= 6 + \frac{3}{4}x - \frac{15}{4} \\ y &= \frac{3}{4}x + \frac{24}{4} - \frac{15}{4} \\ y &= \frac{3}{4}x + \frac{9}{4} \end{aligned}$$

4.23 The Angle Between Two Vector Lines

To compute the angle between two vector lines, find the angle between their respective direction vectors using the formula in Section 4.19, "The Angle Between Two Vectors".

4.24 Distinguishing Between Coincident and Parallel Lines

Two vector lines are parallel if their respective direction vectors are parallel (equal or one is a scalar multiple of the other). See Section 4.13, “Determining Whether Vectors are Parallel” for tip on determining if two vectors are parallel.

Two vector lines are coincident if they specify the same line. This means that they must be parallel, and you can find some point that exists on both lines.

4.25 Finding the Point of Intersection of Two Lines

To find the point of intersection of two vector lines, set the two vectors equal and solve the resulting system of simultaneous equations. Make sure that the two different equations use two distinct variables for the real number multiple of the direction vector. If the lines are in two dimensions, then they will either be parallel or they will intersect. In three dimensions the lines might be parallel, intersect, or be skew (no intersection but not parallel).

For example, given the two lines $\vec{r}_1 = \begin{pmatrix} 4 \\ 7 \\ 2 \end{pmatrix} + t \begin{pmatrix} 5 \\ 0 \\ -4 \end{pmatrix}$ and $\vec{r}_2 = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} + s \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix}$ find the point B where they intersect.

If the two lines intersect, then there will be some values for s and t that produce the same position vector on both lines. So, we start by setting $\vec{r}_1 = \vec{r}_2$:

$$\begin{pmatrix} 4 \\ 7 \\ 2 \end{pmatrix} + t \begin{pmatrix} 5 \\ 0 \\ -4 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} + s \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix}$$

This implies the system of linear equations:

$$\begin{aligned} 4 + 5t &= 1 - s \\ 7 + 0t &= 3 + 2s \\ 2 - 4t &= -2 + 4s \end{aligned}$$

Given that the second equation has $0 \times t$, that one will only have variable s left.

$$\begin{aligned} 7 + 0t &= 3 + 2s \\ 7 &= 3 + 2s \\ 2 &= s \end{aligned}$$

So $s = 2$. Substituting that back into one of the other two equations in the system (let's pick the first one), we get:

$$\begin{aligned} 4 + 5t &= 1 - s \\ 4 + 5t &= 1 - (2) \\ 4 + 5t &= -1 \\ t &= -1 \end{aligned}$$

Now we check our work to ensure that using the values $s = 2$ and $t = -1$ produce the same vector:

$$\begin{pmatrix} 4 \\ 7 \\ 2 \end{pmatrix} + (-1) \begin{pmatrix} 5 \\ 0 \\ -4 \end{pmatrix} = \begin{pmatrix} 4 - 5 \\ 7 - 0 \\ 2 - (-4) \end{pmatrix} = \begin{pmatrix} -1 \\ 7 \\ 6 \end{pmatrix}$$
$$\begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} + (2) \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 - 2 \\ 3 + 4 \\ -2 + 8 \end{pmatrix} = \begin{pmatrix} -1 \\ 7 \\ 6 \end{pmatrix}$$

So, the point B where the two lines intersect is $(-1, 7, 6)$.

NOTE: If the problem appears on a calculator test, remember from Section 1.25, “Solving Systems of Three Linear Equations Using Technology” that you can enter the system of three linear equations in three variables into the calculator and it will give you the answer in a single step.

4.25.1 Finding Intersection Between a Line and an Axis

If you are given the equation of one vector line L_1 and asked to find where it intersects an axis like the x -axis, use a vector equation for the desired axis as your L_2 and then solve the problem as described in Section 4.25, “Finding the Point of Intersection of Two Lines”. Since the origin

$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ is a point on all three axes, it is the simplest vector to use for the position vector of the

axis line equations. Then you use the base vector $(\vec{i}, \vec{j}, \text{ or } \vec{k})$ from Section 4.18, “Base Vectors for Three Dimensions” (or the two-dimensional versions if you are working in two dimensions).

So, the vector equations for the three axes in three dimensions are:

$$\bullet \text{ } x\text{-axis: } \mathbf{r} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\bullet \text{ } y\text{-axis: } \mathbf{r} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\bullet \text{ } z\text{-axis: } \mathbf{r} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

The analogous vector equations for the axes in two dimensions are:

$$\bullet \text{ } x\text{-axis: } \mathbf{r} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\bullet \text{ } y\text{-axis: } \mathbf{r} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

5 Statistics

5.1 Concepts

5.1.1 Population versus Sample

Statistics are drawn from an overall set called a *population*. When you draw a subset of examples from the population, you are selecting a *sample*. Since it may be impractical (or impossible) to measure statistics about an entire population, we use the statistics for a sample to estimate population statistics. A *random sample* is a subset of the population that is chosen randomly, to increase the chances that the sample statistics fairly represent the whole population.

5.1.2 Discrete Data versus Continuous Data

Discrete data is counted, while *continuous* data is measured.

Discrete data typically have natural number values and have a distinct set of possibilities you can identify (and count) ahead of time. Examples include:

- dress sizes
- outcomes from rolling dice
- candies in a box

In contrast, *continuous* data occur somewhere in a real number interval of values, and it's not possible to count the number of distinct values ahead of time. Examples include:

- height or weight of students in a class
- times of a swimming race
- length of leaves fallen from a tree

5.2 Presentation of Data

Numerical data is typically arranged in a table, which has one row for each group or class of data. For discrete data you usually have one row per distinct data value, whereas for continuous data each row typically represents a subinterval of the overall range of possible values. Be careful when labeling the range groups to partition the data uniquely into groups. If the first group were labeled as $0 \leq v \leq 5$ and the second group as $5 \leq v \leq 10$, then the number 5 belongs to both groups. They should instead be labeled $0 \leq v < 5$ and the second group as $5 \leq v < 10$ so that any value (like 5 or 10) falls in exactly one subinterval.

5.2.1 Frequency Distribution Tables

The *frequency* is the number of times a particular value occurs in the sample. A frequency distribution table lists each possible value, or range of values, in increasing order, along with the frequency with which that value occurs. Table 4 shows an example of a frequency distribution table for family sizes of students in a classroom. The sum of the “Students” column gives the total number of values in the data set.

Family Size	Students
2	3
3	6
4	6
5	2
6	1
TOTAL	18

Table 4: Frequency Table of Family Size

5.2.2 Frequency Histograms with Equal Class Intervals

A frequency histogram that uses equal class intervals (or discrete values) is called a bar chart. It's the simplest way to visualize grouped data. Figure 89 shows an example.

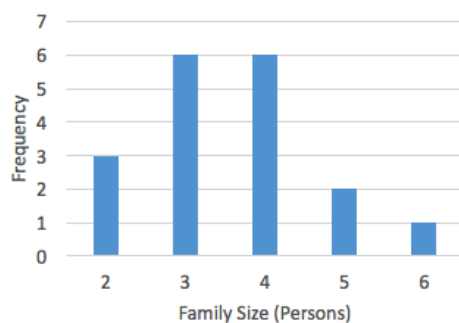


Figure 89: Bar Chart Showing Family Size Frequency Table Data from Table 4

5.3 Mean

The *mean* (also known as the average) is one measure of the central tendency of a list of data values. To calculate the *mean* of list of numbers, add up the numbers and divide by the total number of values. The symbol for the mean is \bar{x} , so the formula for the mean of a list of n numbers is:

$$\bar{x} = \frac{\sum x_i}{n}$$

Multiplying both sides of this equation by n gives a useful corollary of this formula: the sum of the n numbers in the list equals n times the mean:

$$n\bar{x} = \sum x_i$$

If the data is in a frequency table, then you calculate the total value by multiplying each value by its frequency and summing the results. The sum of the “Students” column gives the total number of values in the data set, while the sum of the “Value \times Frequency” column gives the sum of all values. The mean of the data set in Table 5 is $\frac{64}{18} = 3.56$.

Family Size	Students	Value \times Frequency
2	3	6
3	6	18
4	6	24
5	2	10
6	1	6
TOTAL	18	64

Table 5: Frequency Table of Family Size with Product

If the data is in a *grouped* frequency table, you use the *mid-interval value* of each group to represent the group for the purposes of multiplying by the frequency. The mean in this case is only an estimate since the actual data values in each group are unknown. Based on the data in Table 6, the estimated mean is $\frac{218.5}{13} = 16.8$

Race Time (min)	Mid Interval	Athletes	Value \times Frequency
$10 \leq t < 13$	11.5	2	23.0
$13 \leq t < 16$	14.5	3	43.5
$16 \leq t < 19$	17.5	5	87.5
$19 \leq t < 22$	20.5	2	41.0
$22 \leq t < 25$	23.5	1	23.5
TOTAL		13	218.5

Table 6: Frequency Table with Grouped Data Uses Mid-Interval Values

5.4 Median

The *median* of a set of numbers is the figure that sits in the middle of the list when its values are ordered from lowest to highest. After ordering the list, if it has an odd number of items, then the median is the middle number. If it has an even number of items, then the median is the average of the two central numbers. In contrast to the mean, the median value is not affected by very large or very small values in the list, only by the number of items in the list and their relative ordering.

If you have a frequency table like Table 4 in Section 5.2.1, “Frequency Distribution Tables”, then after ensuring that the data values are listed in increasing order, you can calculate the median as follows: start by dividing the total number of students by 2. Then calculate a running total of the “Frequency” column until you get to a data value whose frequency causes the running total to meet or exceed the middle number. If there are an even number of items, you might have to take the average of the data value from two contiguous values if their frequency count falls exactly between the two middle numbers. In this example, there are 18 students and the 9th value has family size 3 while the 10th student has family size 4. So here, the median is the average of 3 and 4 which is $\frac{3+4}{2} = 3.5$.

If you have a frequency table of grouped data like in Table 6 in Section 5.3, “Mean”, then since there are 13 athletes total the median value will be the 7th race time value. It falls in the third range of values from 16-19. Since there are 5 values in the previous two categories, the 7th data value is $\frac{2}{5}$ of the way across the interval (since it’s the 2nd out of 5 items in that range). So the interpolated median is $16 + \frac{2}{5} \cdot 3 = 17.2$. We multiply the $\frac{2}{5}$ times 3 because the range is 3 wide.

5.5 Mode

The *mode* is the value that occurs the most frequently in the data set. In a frequency table, it is the data value with the largest frequency. If the frequency table is reporting on grouped data ranges, then the best you can say is which group has the highest frequency. It is possible for a data set to have more than one mode or modal group.

5.6 Cumulative Frequency

5.6.1 Cumulative Frequency Table

It is a bit easier to compute the median from a frequency table if it is a *cumulative frequency table*. Table 7 shows an alternative presentation of the data from Table 4 using *cumulative frequencies*. Rather than listing the frequency of each data value, we list the running total of how many students have the current family size or lower.

Family Size	Cumulative Students
≤ 2	3
≤ 3	9
≤ 4	15
≤ 5	17
≤ 6	18

Table 7: Cumulative Frequency Table of Family Size

5.6.2 Cumulative Frequency Graphs

Once you have a cumulative frequency table, you can graph that data to produce a *cumulative frequency graph*. You use the values as the *x-axis* data and the cumulative frequency for the *y-axis* data. For discrete values like in Table 7, you plot the points representing each data value and its corresponding cumulative frequency. Then you connect the points with a straight line or smooth curve. Figure 90 shows the result of doing this for the cumulative frequency data in Table 7.

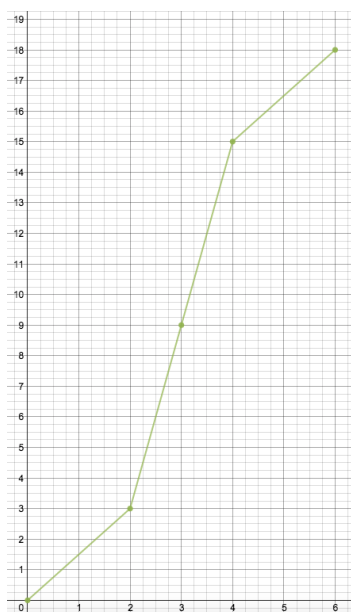


Figure 90: Cumulative Frequency Graph of Data in Table 7

Consider another data set with ranges of test scores shown in Table 8. The table lists the individual frequency for each grade band, and contains an additional “C.F.” column where the running total of the cumulative frequency appears.

Test Score	Frequency	C.F.
0-10	7	7
10-30	11	18
30-50	14	32
50-60	16	48
60-80	20	68
80-100	9	77
100-120	3	80

Table 8: Cumulative Frequency Table of Test Scores

Figure 91 shows the related cumulative frequency diagram. Notice that we use the **upper bound** of each range as the x -coordinate value to plot (e.g. 10, 30, 50, 60, 80, 100, 120). The figure also shows the lines for the first quartile (Q_1), median (which is technically the 2nd quartile), the third quartile (Q_3). These quartile lines are obtained by taking the maximum frequency value (80 in this example) and dividing it evenly into four segments. The maximum frequency of 80 divided by 4 equals 20, so the first quartile occurs at a frequency of 20, second quartile (or median) at frequency of 40, the third quartile at the frequency of 60, and the fourth quartile at the frequency of 80. By drawing a horizontal line at the quartile boundaries on the frequency axis until it intersects with the Cumulative Frequency Graph, we can then draw the vertical line down from each intersection to find an estimate among the values on the “data value” axis for the first quartile, median, and third quartile. Once you find the first quartile value and the third quartile value, you can subtract the former from the latter to calculate the interquartile range.

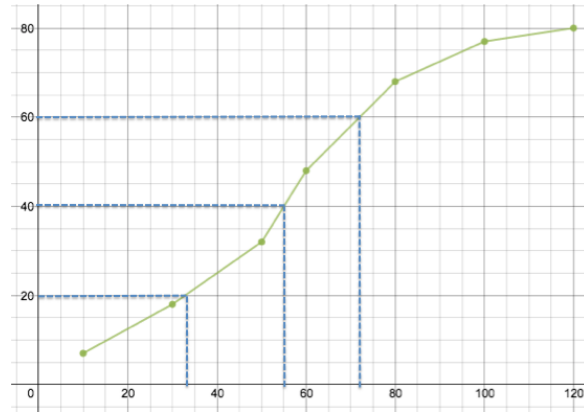


Figure 91: Cumulative Frequency Graph of Data in Table 8

5.7 Dispersion

A data set of one thousand values whose mean is 85 can all be clumped near the mean, or can be more spread out. The range, interquartile range, variance, and standard deviations are calculations that help us understand how spread out (or “dispersed”) the data values are.

5.7.1 Range

The *range* is the simplest calculation to compute. Just subtract the smallest value from the largest value.

5.7.2 Quartiles

In addition to the median, which is the data value right in the middle of the ordered list of data, it can be useful to understand the values at the one quarter and three quarters boundaries as well. The value that corresponds to the one quarter boundary is called the *lower quartile*, while the value that corresponds to the three quarters boundary is the *upper quartile*. To find the lower quartile, use the same technique as finding the median but do it on the ordered set of data values that come before the median. To find the upper quartile, use the same technique as finding the median, but do it on the ordered set of data values that come after the median.

5.7.3 Interquartile Range

To calculate the interquartile range, subtract the lower quartile value from the upper quartile value.

5.7.4 Box and Whiskers Plots

A box and whiskers plot is a quick way of visualizing how spread out the data set is. The diagram in Figure 92 is typically drawn above a number line with a thin line whose left and right edges mark the overall minimum and maximum data values. The left and right edges of a box in the middle of the extremity line indicate the lower quartile and upper quartile values. A further vertical line inside the box indicates the median.

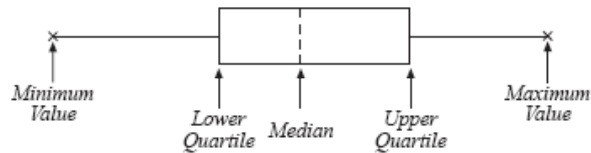


Figure 92: Box and Whiskers¹

5.7.5 Percentiles

To calculate median and quartiles, we divide the cumulative frequency range into four equal intervals with a midpoint and the midpoints of the lower and upper halves. If instead you divide the cumulative frequency range into ten equal intervals, you can use that division to find the data values in the 10 *percentile* ranges. Figure 93 shows the cumulative frequency graph for the data in Table 8 in Section 5.6.2, “Cumulative Frequency Graphs”. The figure also shows the lines for the ten percentiles boundaries on the frequency axis. These percentile lines are obtained by taking the maximum frequency value (80 in this example) and dividing it evenly into ten segments. The maximum frequency of 80 divided by 10 equals 8, so the 10 percentile occurs at a frequency of 8, the 20 percentile occurs at a frequency of 16, the 30 percentile at a frequency of 24, and so on until the 100 percentile that occurs at the frequency value 80. By drawing a horizontal line at each percentile boundary on the frequency axis until it intersects with the Cumulative Frequency Graph, we can then draw the vertical line down from each intersection to find an estimate among the values on the “data value” axis for the 10 percentile, 20 percentile, 30 percentile, and so on.

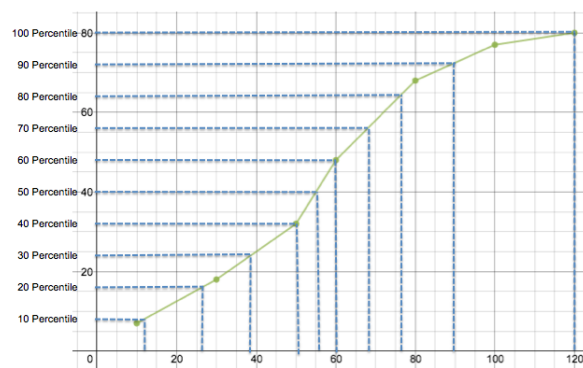


Figure 93: Cumulative Frequency Graph of Data in Table 8

5.7.6 Variance

The variance is a measure of how spread out the data is. It is an average of the squared difference between each data value and the mean. It is the square of the standard deviation. While the formula to calculate the variance is shown below, in general you will always use your calculator to compute it.

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

¹Source: www.cimt.plymouth.ac.uk/projects/mepres/book9/bk9i16/bk9_16i4.htm

5.7.7 Standard Deviation

The standard deviation is another measure of how spread out the data is. It is simply the square root of the variance. Since the variance is computed by taking the average of the *square* of the difference between each data value and the mean, by taking the square root of the variance we end up with a number that is more easily comparable to the original data in the list. While the formula to calculate the standard deviation is shown below, in general you will always use your calculator to compute it.

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

5.8 Use Cumulative Frequency Graph to Find Median and Quartiles

You can use a cumulative frequency diagram to estimate the median, upper quartile, and lower quartile values, as well as calculate the interquartile range. To find the median, draw a horizontal line at the point on the y -axis at exactly one half of the range of cumulative frequency. In the example in Figure 91 the cumulative frequency range is 80, so we draw the horizontal line at $y = 40$. At the point where this horizontal line intersects the cumulative frequency graph, draw a vertical line down to the x -axis. The point where it intersects the x -axis is an estimate of the median. In this example, it is 55.

To estimate the upper quartile value, repeat the same process but start the horizontal line at one half of the *upper half* of the cumulative frequency range. Half way between 40 and 80 is 60, so drawing a horizontal line across at $y = 60$ and then down from where it intersects the cumulative frequency graph gives an upper quartile value of about 72.

To estimate the lower quartile value, repeat the same process but start the horizontal line at one half of the *lower half* of the cumulative frequency range. Half way between 0 and 40 is 20, so drawing a horizontal line across at $y = 20$ and then down from where it intersects the cumulative frequency graph gives an upper quartile value of about 33.

To find the interquartile range, subtract the lower quartile value from the upper quartile value. In this example, it is $72 - 33 = 39$.

To find percentage values, you would use a similar technique but rather than drawing a horizontal line at one quarter, one half and three quarters points in the cumulative frequency range on the y -axis, you would instead divide that range into ten equal intervals and draw the horizontal lines over to the graph and then vertical lines down to the x -axis from each of the 10 interval boundary points.

5.9 Computing Statistics Using Your Calculator

In a few clicks, your graphing calculator will compute all of the relevant statistics we've mentioned in these notes. With the TI-Nspire calculator, you start by entering a data list. The procedure is slightly different depending on whether you are entering a "raw" list of data that includes duplicates or else using a list with distinct data values and a frequency count.

5.9.1 Calculating Statistics for a Single List

In your calculator, start by inserting a new “Lists and Spreadsheets” tab into your document. In the first row of the table (colored grey and identified by an icon of a tag/label at the far left) enter a name for your column of data in the cell labeled **A**. I like to use the word **data**. Skip the second row of the table (also grey, with an equals sign at the left) and enter your data values in the cells of column **A** starting with the row labeled 1. After you have entered every data value into this “data” column, you can compute the statistics by following these steps:

- With your cursor in a cell in the first column, press [menu] and choose **Statistics** → **Stat Calculations** → **One-Variable Statistics...**
- In the small *One-Variable Statistics* dialog, enter 1 for **Num of Lists** and click [OK]
- In the larger *One-Variable Statistics* dialog, choose 'data' from the **X1 List** dropdown list
- Click [OK]

Figure 94 shows the before and after screenshots after following these instructions. The statistics are all calculated and placed in the spreadsheet column you indicated in the dialog. By default, it will have chosen the first empty column in your spreadsheet. If you scroll through the results, you'll find the minimum (**MinX**), maximum (**MaxX**), median (**MedianX**), lower quartile (**Q₁X**), upper quartile (**Q₃X**), number of items (**n**), the mean (\bar{x}), the sum of all the items ($\sum x$), and the standard deviation (σx). **NOTE:** Do not use the statistic labeled (**sx**). You need to use σx always on the IB Exams.

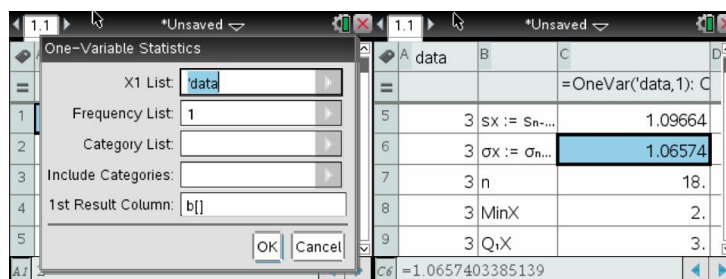


Figure 94: One-Variable Statistics on a Simple List of Data

5.9.2 Calculating Statistics for a List with Frequency

If your data includes values and frequency information, you can enter the data in that way. The steps are similar to those in Section 5.9.1, “Calculating Statistics for a Single List”. Add a “Lists and Spreadsheets” type of tab to your document. Label column **A** with the name **data** and column **B** with the label **freq**. Enter your data values in the “data” column and the corresponding frequency values in the “freq” column. Then do these steps:

- With your cursor in a cell in the first column, press [menu] and choose **Statistics** → **Stat Calculations** → **One-Variable Statistics...**
- In the small *One-Variable Statistics* dialog, enter 1 for **Num of Lists** and click [OK]
- In the larger *One-Variable Statistics* dialog, choose 'data' from the **X1 List** dropdown list and choose 'freq' from the **Frequency List** dropdown list
- Click [OK]

Figure 95 shows the before and after screenshots after following these instructions. The statistics are all calculated and placed in the spreadsheet column you indicated in the dialog. By default,

it will have chosen the first empty column in your spreadsheet. If you scroll through the results, you'll find the minimum (MinX), maximum (MaxX), median (MedianX), lower quartile (Q_1X), upper quartile (Q_3X), number of items (n), the mean (\bar{x}), the sum of all the items ($\sum x$), and the standard deviation (σx). **NOTE:** Do not use the statistic labeled (sx). You need to use σx always on the IB Exams.

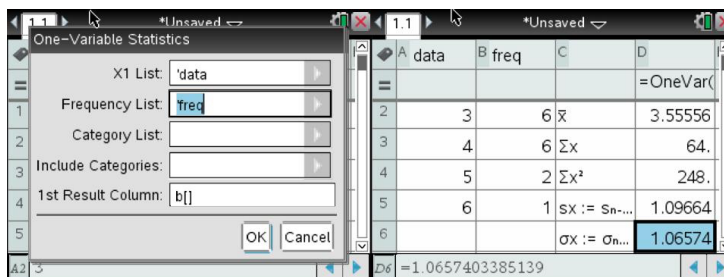


Figure 95: One-Variable Statistics on a List of Data with Frequencies

5.10 Linear Correlation of Bivariate Data

5.10.1 Scatter Diagrams

You can graph the relationship between any two variables using a scatter diagram. If there is some correlation between the variables, then the points you plot should show some kind of trend. For example, suppose you graph the number of hours each student in a class studied for an exam on one axis against the grade each respective student attained on an exam on the other axis. You would expect the scatter plot to exhibit a trend that studying for more hours produced a higher grade. This would illustrate a correlation between the two variables (*hours of study* and *grade*).

Student	Hours of Study	Test Score
Joe	3	82
June	1	35
Judy	6	85
Jason	4.5	77
James	5	80
Julia	8	99
Jupiter	2	45

Table 9: Hours of Study versus Test Score

To graph this data on the TI-Nspire, add a “Lists and Spreadsheets” tab to your document and label the two columns in the top row with **hours** and **score**. Proceed to enter your data from the table into these two columns. Then, to create the scatter plot do the following:

- Press $[Ctrl]+[Doc]$ to insert a page, and choose **Add Graphs**
- In the graph tab, press $[menu]$ and choose **Graph Entry/Edit** → **Scatter Plot**
- In the popup window that appears with the cursor next to the $x \leftarrow$, press $[var]$ and choose the **hours** data column.

- Then tab to place the cursor next to the $y \leftarrow$, press $[var]$ and choose the `score` data column.
- Press $[enter]$, then use the graph zoom functions to view all the data

You should end up with a scatter plot that looks like Figure 96.

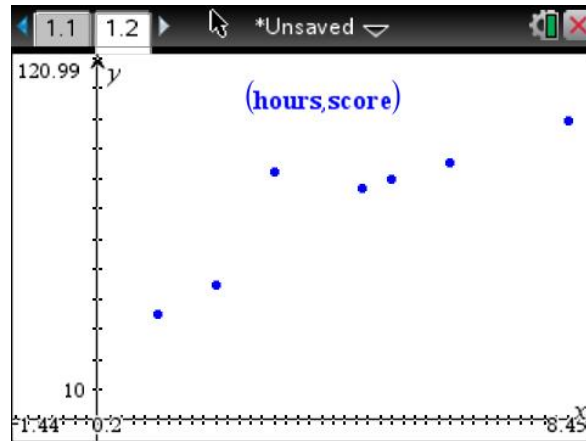


Figure 96: Scatter Plot of Hours Studied Versus Test Score

5.10.2 Pearson's Product-Moment Correlation Coefficient r

The Pearson's Product-Moment Correlation Coefficient is a number between -1 and 1 that your calculator can compute to tell you how strong of a relationship there is between the two data series you have scatter-plotted. The number r is one of the statistics that your calculator will compute when you ask it to calculate the line of best fit (also known as linear regression). Figure 97 shows some typical values of r and how that relates to the strength of correlation.

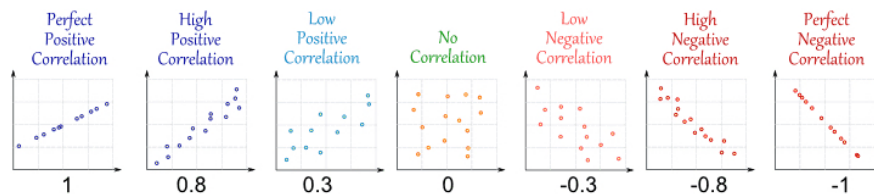


Figure 97: Examples of Different Levels of Correlation²

5.10.3 Lines of Best Fit

The line of best fit is a straight line that minimizes the squared difference between the data points and the points on the line. It's also called the "least squares regression line", and your calculator will compute it for you. In the process of determining the line of best fit, it will compute the m (slope) and b (y -intercept) of that line's equation ($y = mx + b$). It will also calculate the correlation coefficient, which is shown as the value r value in Figure 98.

5.10.4 Equation of the Regression Line

To compute the equation of the regression line on the TI-Nspire, go back to the tab containing your labeled data columns and do the following:

²Source: mathsisfun.com/data/correlation.html

- Press [Menu] and choose **Statistics** → **Stat Calculations** → **Linear Regression (mx+b)**
- In the *Linear Regression* window that appears, choose **hours** for the *X List* and choose **score** for the *Y List* data column.
- Note the name of the equation in the *Save RegEqn to* field.
- Click [OK]

This will produce the additional column of data as shown in Figure 98. The highlighted box is the correlation coefficient, labeled r .

A	B	C	D
hours	score		=LinRegM
6	85	m	8.58078
4.5	77	b	35.6953
5	80	r^2	0.807503
8	99	r	0.898612
2	45	Resid	{20.5623...

Figure 98: Slope, y -Intercept, and Correlation Coefficient of Linear Regression

To add the regression line to the scatter plot graph, click back to activate the graph tab and do the following:

- Press [Menu] and choose **Graph Entry/Edit** → **Function**
- Arrow up to see the $f1(x)$ function has already been filled in with the m slope value and b value of the y -intercept.
- Then press [Enter] to see the line on the graph

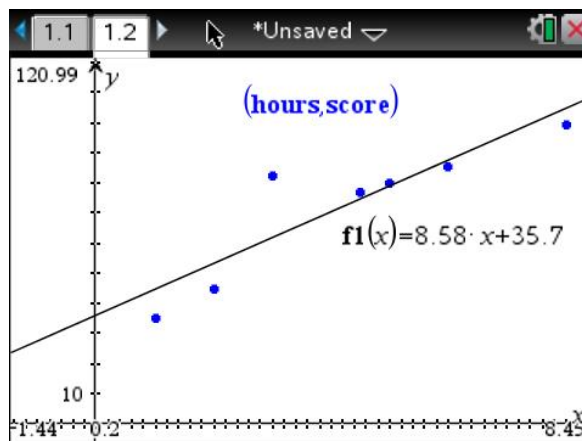


Figure 99: Linear Regression Line of Best Fit for Hours/Grade Data

6 Probability

6.1 Probability Concepts

Probability measures the chance that some event will occur. If you purchase a single ticket for a raffle, and you know that 250 raffle tickets were sold in total, then you know that your chances of winning are one in 250. Expressed as a fraction, your chances of winning are $\frac{1}{250}$. Expressed as a decimal, your chances are 0.004. Expressed as a percentage, your chances are 0.4%.

A probability is always a number between 0 and 1. A probability of zero (0) means that it is impossible for something to occur, while a probability of one (1) means that it will always happen. As with most things in life, the true probability of something's occurring lies somewhere in between impossible (0) and certain (1).

For a given situation, the *sample space* is the universe U of all possible things that can occur. An *event* A is one of the things that can occur. An uppercase letter P represents the probability (or chance) that a given event will occur. Instead of always writing out the words “the probability that event A will happen”, we write $P(A)$.

6.2 Probability of an Event

If we use $n(A)$ to mean the number of ways that event A can occur, then the probability of an event A is:

$$P(A) = \frac{\text{Number of ways } A \text{ can occur}}{\text{Number of all possible events in the sample space}} = \frac{n(A)}{n(U)}$$

6.3 Probability of an Event's Not Occurring

If A is some event whose probability we're interested in, $P(A)$ represents the probability that A will occur. If we visualize the universe of all things that can occur as a diagram, then as shown in Figure 100 on the left, we can represent $P(A)$ as a shaded part of the entire set of possibilities U . All of the things in the universe *other than* A are called the “complement” of A and we use the “prime” symbol from set theory to represent it. As shown in the diagram on the right in Figure 100, this A' represents the probability that anything *other than* A happens. In other words, it represents the probability that A does not happen. Since it is certain that either A will happen or something *other than* A will happen, it makes intuitive sense that $P(A) + P(A') = 1$.

If we rearrange the equality above, we see that $P(A') = 1 - P(A)$. In other words, if $P(A)$ is the probability that A will happen, $1 - P(A)$ is the probability that A will not occur.

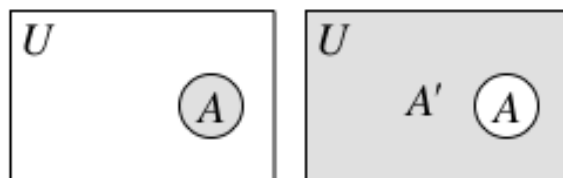


Figure 100: Visualizing the Probability of Event A and A'

6.4 Independent, Dependent, and Mutually Exclusive Events

Two events A and B are *independent* if the chance that A occurs is not affected by whether B occurs or not. For example, if you roll a die three times and it comes up 4 each time, what

is the chance that the next roll will also be 4? The chance is simply $\frac{1}{6} \approx 16.7\%$, just like any other die roll. The previous rolls did not affect the chances on the next roll.

However, if you are playing blackjack with a single deck of cards, and the dealer just drew a Queen of Hearts, since there is now one fewer card in the deck, the probability of drawing a Queen or of drawing a card with the suit of Hearts is now different. In this case, we say the events are *dependent* since the outcome of one can affect the probability of the other.

Two events A and B are *mutually exclusive* if the two events cannot happen at the same time.

6.5 Probability of A and B for Independent Events

When events A and B are independent, to calculate the probability that both event A and event B will occur, you multiply their probabilities together: $P(A \cap B) = P(A) \times P(B)$.

NOTE: $P(A \cap B) = P(A) \times P(B)$ only works when A and B are independent! See Section 6.10, “Conditional Probability of Dependent Events” for how to work with probability of dependent events.

Consider a deck of playing cards, and the events H and K about an arbitrary card:

H = The suit of the card is *Hearts*

K = The card is a *King*

Then, to compute the probability that a card is *both* a King and has the suit Hearts, we multiply the probabilities:

$$P(H \cap K) = P(H) \times P(K)$$

Since a probability is a fraction between 0 and 1, multiplying two positive fractions that are both less than one produces a number that is smaller than either of these two original quantities. In the example with playing cards shown in Figure 101 we would have $P(H) = \frac{13}{52} = \frac{1}{4}$ and $P(K) = \frac{4}{52} = \frac{1}{13}$. Therefore, the probability of drawing a card that is both a King and has suit of Hearts, is:

$$P(H \cap K) = P(H) \times P(K) = \frac{1}{4} \times \frac{1}{13} = \frac{1}{52}$$

We can get the same answer by counting the number of ways that these events overlap (or intersect). In Figure 101, we see that $P(H \cap K)$ is represented by the intersection of the circles representing $P(H)$ for Hearts and $P(K)$ for Kings. Since there is only one card out of 52 that is both a King and a Heart, we again get a probability of $\frac{1}{52}$.

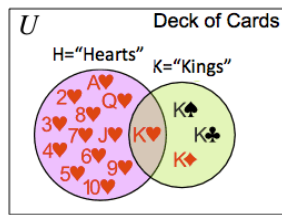


Figure 101: Visualizing Overlapping Events ³

6.6 Probability of A or B for Mutually Exclusive Events

To calculate the probability that either event A or event B will occur, you add their probabilities together: $P(A \cup B) = P(A) + P(B)$, however as shown in Figure 102, this only works if the two events are *mutually exclusive*. This means that it would never be possible that both A and B could happen at the same time. Consider a set of playing cards, and let the statements A and K about an arbitrary card be the following:

A = The card is an *Ace*

K = The card is a *King*

Then we see that the probability that a card is either an Ace or a King is just the sum of the separate probabilities $P(A \cup K) = P(A) + P(K)$.

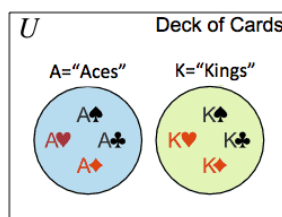


Figure 102: Visualizing Mutually Exclusive Events⁴

6.7 Probability of Mutually Exclusive Events

If events A and B are mutually exclusive, that means they cannot happen at the same time. Therefore, the probability of having both A and B occur is zero:

$$P(A \cap B) = 0$$

This works both ways, if you can demonstrate that $P(A)$ and $P(B)$ have no overlap (or intersection), then you can say that A and B are mutually exclusive.

6.8 Probability of A or B for Non-Mutually Exclusive Events

If we consider the events H and K from Section 6.5, “Probability of A and B for Independent Events”, we see in Figure 101 that simply adding together $P(H)$ and $P(K)$, we would end up counting the King of Hearts twice and get the wrong answer. So, in general when computing the probability of combined events, we need to add the probabilities of the individual events, and then subtract the probability of their *both* occurring to avoid double-counting any events in the intersection.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

³Source: mathsisfun.com/data/probability-events-mutually-exclusive.html

⁴Source: mathsisfun.com/data/probability-events-mutually-exclusive.html

6.9 Lists and Tables of Outcomes

If the number of events is small, create a list of all the possible outcomes, then count the number of outcomes that meet the criteria you're looking for to compute:

$$\frac{\text{Number of outcomes that meet the criteria}}{\text{All possible outcomes}}$$

6.10 Conditional Probability of Dependent Events

If an event A depends on an event B then we talk about the *conditional probability* of A given B and write $P(A | B)$. This means that assuming B has already happened, what is the probability A will happen. The formula to calculate this conditional probability is:

$$P(A | B) = \frac{\text{Probability of } A \text{ and } B \text{ happening}}{\text{Probability of } B \text{ that already occurred}} = \frac{P(A \cap B)}{P(B)}$$

The order of the A and B is important. That is, in general $P(A | B) \neq P(B | A)$. To calculate the probability of B given A , the formula is:

$$P(B | A) = \frac{\text{Probability of } A \text{ and } B \text{ happening}}{\text{Probability of } A \text{ that already occurred}} = \frac{P(A \cap B)}{P(A)}$$

When you are calculating a conditional probability $P(A | B)$ using a list or table of results, first narrow down the set of all possible outcomes to only those in the sample space where event B is true. Then, out of this subset, compute how many of the remaining possibilities would meet the criteria to make event A true. The conditional probability will then be the fraction:

$$P(A | B) = \frac{\text{Number of outcomes where } A \text{ is true (In subset of outcomes where } B \text{ is true)}}{\text{Number of outcomes in the sample space where } B \text{ is true}}$$

6.11 Testing for Independent Events

If A and B are independent events, then the probability that A will occur will be the same whether or not B occurred, so you have:

$$P(A) = P(A | B) = P(A | B')$$

However, if you are asked to test that two events are independent, it is enough to show that:

$$P(A \cap B) = P(A) \times P(B)$$

6.12 Venn Diagrams

When there are multiple events, draw a Venn diagram to count the outcomes that qualify for each event. When you draw your diagram, draw one circle for each event and make the circles overlap in case there are some cases that fall in the intersection. The space inside the outer rectangle but not inside any of the circles represents the sample space outcomes that don't qualify for any of the events. If any of the quantities is unknown, call one of them x and then express the others in terms of the known information and this x variable. The problem will give you some of the probabilities or number of qualifying members to fill in, leaving some for you to determine using the facts provided. If nothing else, you know that all of the probabilities must add up to 1 (or if you're counting set members, must add up to the total number of things).

6.13 Tree Diagrams

Tree diagrams are good for determining the probability of a sequence of events and help you ensure that you have accounted for all the different possibilities of one event occurring (or not occurring) followed by another event happening (or not). Each time the tree branches into multiple outcomes, the probability of those outcomes adds up to one (1) because they account for all choices. The probabilities at the leaf nodes on the right of the diagram represent the conditional probability of following the particular set of paths to get from the root of the tree to that particular leaf node. Each of these leaf node probabilities is calculated by multiplying the probabilities along each path line to get to the destination node in question. Then, if there are multiple paths in the tree that satisfy the conditions of the problem, you add together the qualifying leaf node probabilities.

Example Consider an example of a bag containing 4 white balls and 3 black balls. If two balls are picked in succession without replacement, what is the probability of picking one white ball and one black ball?

Solution We draw the tree diagram in Figure 103. For the first event “Pick 1” we draw branches for the two possible outcomes: one for drawing a black ball, and one for drawing a white ball. We write the probability of drawing a black ball ($\frac{3}{7}$) along that branch and write the probability of drawing a white ball ($\frac{4}{7}$) along its branch. For the second event “Pick 2” the number of white and black balls remaining is different depending on what happened on “Pick 1”. On the upper branch of the tree, corresponding to having first picked a black ball, we draw two new branches representing the possible outcomes of the second pick. Along this branch, during pick 1 a black ball was drawn so there are 4 white balls and 2 black balls left. We write the probability of drawing a black ($\frac{2}{6} = \frac{1}{3}$) or white ($\frac{4}{6} = \frac{2}{3}$) ball along the respective branches. On the lower branch of the tree, corresponding to having first picked a white ball, we draw two new branches representing the possible outcomes of the second pick. Along this branch, during pick 1 a white ball was drawn so there are 3 white balls and 3 black balls left. We write the probability of drawing a black ($\frac{3}{6} = \frac{1}{2}$) or white ($\frac{3}{6} = \frac{1}{2}$) ball along the respective branches. The four final leaf nodes of the tree represent the conditional probability of following that respective set of branches. The probabilities along the paths taken to reach each final leaf node are multiplied together. Then, to get the final answer to the question, we add together the different leaf node probabilities that represent drawing one white ball and one black ball. This includes:

$$P(\text{one white and one black ball}) = P(B_2 | W_1) + P(W_2 | B_1) = \frac{2}{7} + \frac{2}{7} = \frac{4}{7}$$

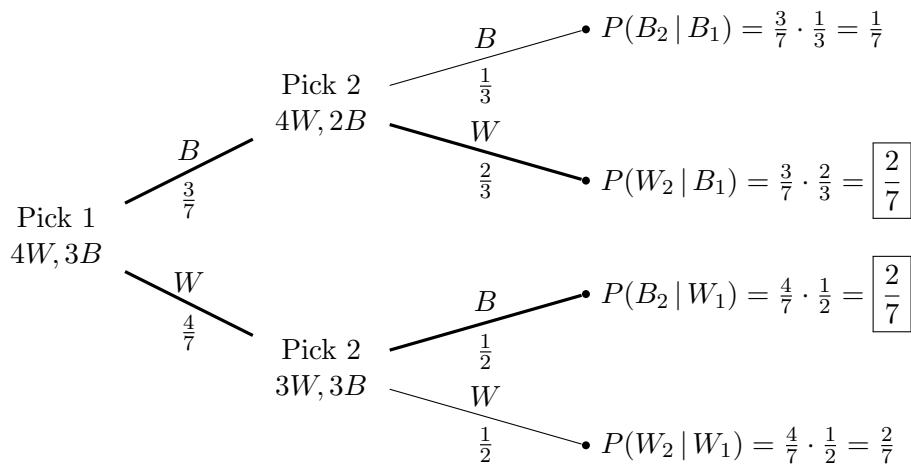


Figure 103: Tree Diagram for Successive Events

6.14 Probabilities With and Without Replacement

For repeated trials, understand whether the problem is asking for the probability of a series of events *with replacement* or *without replacement*. If your experiment involves replacing the selected element after each trial, then the probability of each trial is independent of the other trials that came before or which will come after. Since each trial is an independent event, we know from Section 6.5, “Probability of A and B for Independent Events” that:

$$P(E_1 \cap E_2 \cap \dots \cap E_n) = P(E_1) \times P(E_2) \times \dots \times P(E_n)$$

On the other hand, if the problem asks for the probability involving a series of events without replacement, then you need to use a tree diagram to work out the conditional probabilities for all the possible branches involved in the multiple trials. This is because when you are working *without replacement* the outcome of a previous trial affects the probabilities of the next trial as we saw in the example in Section 6.13, “Tree Diagrams”.

6.15 Discrete Random Variables and Their Probability Distributions

When you calculate the probability of an random event, if the number of possible outcomes is discrete (i.e. countable) then you can use a probability distribution table to document the probability of each.

6.15.1 Explicitly Listed Probabilities

Table 10 shows the probability distribution table for rolling a fair die. The values in the top row of the table, labeled with a lower case x , represent the distinct discrete values that a fair die roll can produce. The values in the bottom row of the table, labeled with $P(X = x)$, represent the probability of each possible outcome. The uppercase letter X in the notation $P(X = x)$ represents the discrete random variable. On each trial it will randomly take on one of the possible values listed in the top row of the table. So, the notation $P(X = x)$ can be read “The probability that the random variable X will take on the value x ” for each possible value x . Since the probability distribution table exhaustively lists every possible outcome for a given trial, then all of the probabilities in the $P(X = x)$ row must add up to one (1).

x	1	2	3	4	5	6
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Table 10: Probability Distribution of Rolling a Die

If all of the probabilities in the $P(X = x)$ row have the same value, then we say that the distribution is *uniform* or that the event has all *equally likely outcomes*.

6.15.2 Probability Distribution Given by a Function

The probability distribution can also be given by a function like this:

$$P(X = x) = \begin{cases} kx(x - 1) & x = 2, 3, 4, 5, 6 \\ 0 & \text{Otherwise} \end{cases}$$

This means that there are five possible values for x (2,3,4,5,6), and that for each value x the respective probability is given by $kx(x - 1)$. Table 11 shows the probabilities obtained by using this formula, substituting in the correct value of x in each column.

x	2	3	4	5	6
$P(X = x)$	$k \cdot 2 \cdot 1$	$k \cdot 3 \cdot 2$	$k \cdot 4 \cdot 3$	$k \cdot 5 \cdot 4$	$k \cdot 6 \cdot 5$

Table 11: Probability Distribution Based on a Function (with k to be determined)

Since the probabilities across the second row of the table must add up to one, we can set up an equation and solve for k :

$$\begin{aligned} k \cdot 2 \cdot 1 + k \cdot 3 \cdot 2 + k \cdot 4 \cdot 3 + k \cdot 5 \cdot 4 + k \cdot 6 \cdot 5 &= 1 \\ 2k + 6k + 12k + 20k + 30k &= 1 \\ 70k &= 1 \\ k &= \frac{1}{70} \end{aligned}$$

So, substituting the value k back into each probability in Table 11 and reducing the fraction to lowest terms, we get the probability distribution that corresponds to the function definition in Table 12.

x	2	3	4	5	6
$P(X = x)$	$\frac{1}{35}$	$\frac{3}{35}$	$\frac{6}{35}$	$\frac{2}{7}$	$\frac{3}{7}$

Table 12: Probability Distribution Based on a Function

6.15.3 Explicit Probability Distribution Involving an Unknown

Instead of a function as in the example in the previous section, you might be given explicit probabilities in the distribution involving an unknown variable k as shown in Table 13.

x	-1	0	2	3
$P(X = x)$	0.2	$10k^2$	0.4	$3k$

Table 13: Probability Distribution Based on an Unknown

Since the probabilities across the second row of the table must add up to one, we can setup an equation and solve for k :

$$\begin{aligned}
 10k^2 + 3k + 0.2 + 0.4 &= 1 \\
 10k^2 + 3k - 0.4 &= 0 \\
 100k^2 + 30k - 4 &= 0 \\
 (20k - 2)(5k + 2) &= 0 \\
 k &= 0.1 \text{ or } -0.4
 \end{aligned}$$

Since the probability $3k$ in the table cannot be negative, we discard the solution $k = -0.4$ and use only $k = 0.1$. Substituting the value k back into each probability in Table 13, we get the probability distribution in Table 14.

x	-1	0	2	3
$P(X = x)$	0.2	0.1	0.4	0.3

Table 14: Probability Distribution Based on an Unknown

6.16 Expected Value (Mean), $E(X)$ for Discrete Data

The expected value is the mean of the data that you would get by randomly picking a large number of trials from the data values having the given probability distribution. The expected value formula allows us to calculate the mean based only on the values that occur and the probability that each value occurs. In particular, we don't need to know the number of trials to predict what the mean will be. To find the *expected value* $E(X)$ of a discrete probability distribution, multiply each value x by its respective probability and add the results together. So the definition is:

$$E(X) = \sum x \cdot P(X = x)$$

As we'll see, the concept is logically equivalent to the technique explained in Section 5.3, "Mean" for calculating the mean. In that section we computed the mean of the data in Table 5 by adding up the *Value* \times *Frequency* products (for each distinct value) and then dividing by the total number of items:

$$\bar{x} = \frac{2 \cdot 3 + 3 \cdot 6 + 4 \cdot 6 + 5 \cdot 2 + 6 \cdot 1}{18}$$

Notice that we can divide each product in the sum in the numerator by the total number of items in the denominator without changing the value of the mean.

$$\begin{aligned}
 \bar{x} &= \frac{2 \cdot 3 + 3 \cdot 6 + 4 \cdot 6 + 5 \cdot 2 + 6 \cdot 1}{18} && \text{Sum of (value} \times \text{frequency) over total} \\
 &= 2 \cdot \frac{3}{18} + 3 \cdot \frac{6}{18} + 4 \cdot \frac{6}{18} + 5 \cdot \frac{2}{18} + 6 \cdot \frac{1}{18} && \text{Sum of (value} \times \text{frequency percentage)}
 \end{aligned}$$

What do these individual fractions represent? They are the fraction of total student family size values that each family size's *frequency* represents. For example, 3 students had the value 2 for family size, so in the data set, the value 2 occurs $\frac{3}{18}$ of the time. Expressed as a percentage, it means that the value 2 occurs 16.7% of the time in this data set. It is the *frequency percentage* that represents the probability of randomly picking a value 2 from the values in the data set. You have 3 chances in 18 to do so.

$$\begin{aligned} \bar{x} &= 2 \cdot \frac{3}{18} + 3 \cdot \frac{6}{18} + 4 \cdot \frac{6}{18} + 5 \cdot \frac{2}{18} + 6 \cdot \frac{1}{18} && \text{Sum of (value} \times \text{frequency percentage)} \\ &= 2 \cdot P(X = 2) + 3 \cdot P(X = 3) && \text{Expected value of frequency distribution} \\ &\quad + 4 \cdot P(X = 4) + 5 \cdot P(X = 5) \\ &\quad + 6 \cdot P(X = 6) \end{aligned}$$

Table 15 shows the family size data from Section 5.3, “Mean”, but calculates the Frequency Percentage instead of the “Value \times Frequency”.

Family Size	Students	Frequency Percentage $P(X = x)$
2	3	$\frac{3}{18} = 0.167$
3	6	$\frac{6}{18} = 0.333$
4	6	$\frac{6}{18} = 0.333$
5	2	$\frac{2}{18} = 0.111$
6	1	$\frac{1}{18} = 0.056$
TOTAL	18	1

Table 15: Frequency Table of Family Size with Product

We can write this same information in a probability distribution table as shown in Table 16.

x	2	3	4	5	6
$P(X = x)$	0.167	0.333	0.333	0.111	0.056

Table 16: Probability Distribution of Student Family Size Data Set

Using the formula for the expected value, we can calculate the mean. The result is the same value we obtained in Section 5.3, “Mean” for the same dataset, but done using the probability distribution instead.

$$\begin{aligned} E(x) &= 2 \cdot P(X = 2) + 3 \cdot P(X = 3) + 4 \cdot P(X = 4) + 5 \cdot P(X = 5) + 6 \cdot P(X = 6) \\ &= 2 \cdot 0.167 + 3 \cdot 0.333 + 4 \cdot 0.333 + 5 \cdot 0.111 + 6 \cdot 0.056 \\ &= 3.56 \end{aligned}$$

6.16.1 Expected Value for a “Fair” Game

A game is said to be a “fair” if the expected value for points (or winnings) is zero.

6.17 Binomial Distribution

The binomial probability distribution is a special kind of discrete distribution. You use it when you know:

- The number of trials is fixed
- There are only two possible outcomes
 - Success or Failure
 - “A happens” or “A does not happen” [A or A’]
- Each trial is independent
- Probability of success remains the same

Assume you have a binomial distribution where the probability of success is p and the total number of trials is n . In general, the probability $P(X = k)$ of having k successes out of those n total trials is given by the formula:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

When you perform n trials and succeed k times, then $n - k$ times you must fail as shown in Figure 104.

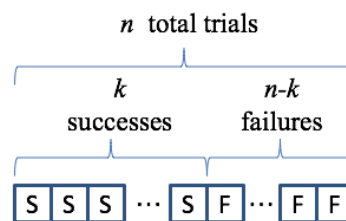


Figure 104: Visualizing Successful and Failed Trials

The term p^k in the formula above is the probability of getting k successes. And we know that if p is the probability of success, then $1 - p$ is the probability of failure. Therefore, the term $(1 - p)^{n-k}$ in the formula represents the probability of getting $n - k$ failures. However, you don’t necessarily have k successes in a row, followed by $n - k$ failures in a row. Your k successes and $n - k$ failures might occur in many different sequences. This is easier to understand if we look at a specific example. Imagine doing five trials and getting 3 successes and 2 failures. In other words, we’ve let $n = 5$ and $k = 3$, so that the number of failures ($n - k$) must be $5 - 3 = 2$. Figure 105 shows the ten ways you can get three successes in five trials.

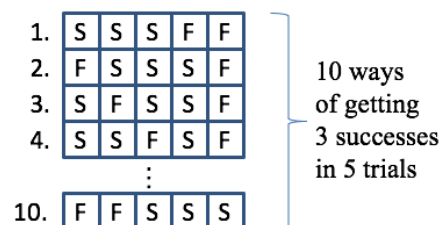


Figure 105: Visualizing Ten Ways to Get Three Successes in Five Trials

The $\binom{n}{k}$ coefficient in the formula represents the number of different ways out of n trials that you can have k successes. By multiplying the probabilities of getting k successes and $n - k$ failures by the number of different ways this outcome can occur, we get the probability of all the different possible combinations in which you might succeed k times out of n trials.

Example Consider the example of the mailman delivering mail. Let's say that the probability that he delivers mail to my mailbox is 2 days in 7. I want to calculate the probability of receiving mail 5 days out of 14.

The number of trials is fixed at 14. Every day, either the mailman brings mail or he does not bring mail, so there are only two possibilities. Whether he brings mail yesterday does not affect whether he will bring mail today, so each trial is independent. The probability of success (i.e. his bringing me mail) remains the same, which is $\frac{2}{7} = 0.286$. Therefore we say that the distribution is $X \sim B(14, 0.286)$.

There are three parts to the calculation of the probability:

$$\underbrace{\binom{14}{5}}_{\substack{\text{Number of combinations of} \\ \text{5 mail-getting days in 14}}} \times \underbrace{(0.286)^5}_{\substack{\text{Probability of} \\ \text{getting mail} \\ \text{5 times}}} \times \underbrace{(1 - 0.286)^9}_{\substack{\text{Probability of} \\ \text{not getting mail} \\ \text{9 times}}} = 0.1847$$

So the probability that he brings me mail on 5 days out of 14 is 18.5%.

If the problem asks for the probability that the mailman brings mail on all 14 days, that is a simpler calculation of just $(0.286)^{14}$. Similarly, the probability that the mailman does not bring mail on any of the 14 days is $(1 - 0.286)^{14}$. However, the more interesting problems involving success on some subset of the total number of trials involve the more complex calculation. Luckily, your calculator makes it easy.

6.17.1 Using the Calculator for Binomial Distribution Problems

Once you have identified the problem as a binomial distribution, your calculator can quickly get you the answer. There are several kinds of problems you will encounter:

- Probability of success k times out of n trials
- Probability of success k or fewer times out of n trials
- Probability of success k or more times out of n trials
- Probability of success from k to m times out of n trials

All of these problems can be handled using the “Binomial Cumulative Distribution Function” (`binomCdf`) function on your calculator. To access that function from a calculator tab, do the following:

- With your cursor in a calculator tab, press `[menu]` and choose **Statistics** → **Distributions** → **Binomial Cdf...**
- In the *Binomial Cdf* dialog, enter the number of trials and the probability of success, the lower bound of the number of successes you want to find and the upper bound on the number of successes
- Click `[OK]`

In order to find the probability of success...

- Exactly k times out of n trials, enter the value k for both the lower bound and upper bound.
- k or fewer times out of n trials, enter the value zero (0) for the lower bound and k for the upper bound.

- k or more times out of n trials, enter the value k for the lower bound and n for the upper bound.
- from k to m times out of n trials, enter the value k for the lower bound and m for the upper bound.

So, for example, using the TI-Nspire CX calculator, we could calculate the probability that my mailman brings mail on 5, 6, or 7 days out of 14, by entering the values shown in Figure 106 to get the answer of 0.346.

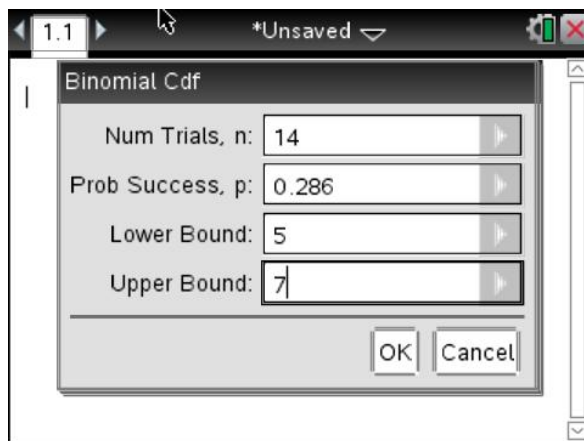


Figure 106: Calculating Cumulative Probability Using the Binomial Distribution

6.17.2 Tip for Calculators Whose `binomCdf` Does Not Have Lower Bound

If your calculator's `binomCdf` function does not allow you to input a lower bound, you need to do problems that ask for “ k or more times” and “between k and m ” times a slightly different way. These calculator's `binomCdf` only allows specifying an upper bound, and calculates the cumulative probability from zero times up through the upper bound number of times.

Doing “ k or More Successes” on TI-84 If a binomial distribution problem has 10 trials, a success probability of 0.18, and asks you to calculate the probability of success on 7 or more times, start by remembering that the probability of “7 or more successes” on a problem with 10 trials means $P(7) + P(8) + P(9) + P(10)$. Also remember that the sum of the probabilities for all trials must equal one (1):

$$P(0) + P(1) + P(2) + P(3) + P(4) + P(5) + P(6) + P(7) + P(8) + P(9) + P(10) = 1$$

The sum of $P(0) + P(1) + P(2) + P(3) + P(4) + P(5) + P(6)$ can be computed using your calculator's `binomCdf(10,0.18,6)` function that accumulates probabilities always starting at zero successes up to your upper bound number of successes, so we can substitute that into the above expression above and solve for the quantity we're being asked to calculate like this:

$$\begin{aligned} P(0) + P(1) + P(2) + P(3) + P(4) + P(5) + P(6) \\ + P(7) + P(8) + P(9) + P(10) &= 1 \\ \text{binomCdf}(10,0.18,6) + P(7) + P(8) + P(9) + P(10) &= 1 \\ P(7) + P(8) + P(9) + P(10) &= 1 - \text{binomCdf}(10,0.18,6) \end{aligned}$$

Therefore, to calculate the probability of “ k or more successes”, use your calculator to find $1 - \text{binomCdf}(n,p,k-1)$.

Doing “Between k and m Successes” on TI-84 If a binomial distribution problem has 10 trials, a success probability of 0.18, and asks you to calculate the probability of between 4 and 6 success, start by remembering that `binomCdf(10,0.18,6)` represents the sum of the probabilities of success for zero up to 6 times, and then rearranging the equation to get your desired cumulative probability on one side:

$$\begin{aligned}\text{binomCdf}(10,0.18,6) &= P(0) + P(1) + P(2) + P(3) \\ &\quad + P(4) + P(5) + P(6) \\ \text{binomCdf}(10,0.18,6) - P(0) - P(1) - P(2) - P(3) &= P(4) + P(5) + P(6) \\ \text{binomCdf}(10,0.18,6) - (P(0) + P(1) + P(2) + P(3)) &= P(4) + P(5) + P(6)\end{aligned}$$

The sum of $P(0)+P(1)+P(2)+P(3)$ can be computed using your calculator’s `binomCdf(10,0.18,3)` function that accumulates probabilities always starting at zero successes up to your upper bound number of successes, so we can substitute that into the above expression above and solve for the quantity we’re being asked to calculate like this:

$$\begin{aligned}\text{binomCdf}(10,0.18,6) - (P(0) + P(1) + P(2) + P(3)) &= P(4) + P(5) + P(6) \\ \text{binomCdf}(10,0.18,6) - \text{binomCdf}(10,0.18,3) &= P(4) + P(5) + P(6)\end{aligned}$$

Therefore, to calculate the probability of “Between k and m successes”, use your calculator to find `binomCdf(n,p,m) - binomCdf(n,p,k-1)`.

6.18 Mean and Variance of the Binomial Distribution

The expected mean of the binomial distribution is just the number of trials times the probability of success. That is:

$$\text{If } X \sim B(n,p), \text{ then } E(X) = np$$

The expected variance of the binomial distribution is the number of trials times the probability of success times the probability of failure. That is:

$$\text{If } X \sim B(n,p), \text{ then } \text{Var}(X) = np(1-p) \text{ and therefore } \text{StdDev}(X) = \sqrt{np(1-p)}$$

6.18.1 Example of Mean and Variance of Binomial Distribution

The probability of obtaining heads on a biased coin is 0.4. The coin is tossed 600 times. Find the mean number of heads and the standard deviation of the number of heads. Then, find the probability that the number of heads is less than one standard deviation away from the mean.

The coin toss comes up heads or not heads, so there are only two possibilities, and the probability of coming up heads is the same on every trial. So, this is a binomial distribution with 600 trials and a probability of success (i.e. coming up heads) of 0.4. The mean number of heads for the binomial distribution is just the number of trials times the probability of success, so $600 \times 0.4 = 240$. Therefore, in 600 tosses we expect a mean of 240 heads. The standard deviation is the square root of the number of trials times the probability of success times that of failure, so $\sqrt{600 \times 0.4 \times 0.6} = 12$.

Given a mean of 240, values that are **less than or equal to** one standard deviation away from the mean are those between $240 - 12 = 228$ and $240 + 12 = 252$. However, the problem here only asks for those values **less than** one standard deviation away, so those are values between $240 - 11 = 229$ and $240 + 11 = 251$. So, using the calculator we can compute the cumulative probability of this range of successes using `binomCdf(600,0.4,229,251)` which equals 0.66210447239411 or to three significant figures 0.662.

6.19 Normal Distribution and Curves

The Normal Distribution models many frequently occurring distributions of continuous data like the heights or weights of students in a class, the times of a swimming race, or the lengths of leaves fallen from a tree. Its probability density function is a continuous function in x that has a “bell” shape that depends on the mean (μ) and standard deviation (σ) of the data set being considered. The curve has the following properties:

- It is symmetric about the mean μ
- The median and the mode are the same as the mean
- The area under the curve between $x = a$ and $x = b$ represents the probability of randomly choosing data values in a given range $a \leq x \leq b$
- The total area under the curve is one (1) (i.e. 100%)
- 68% of the values fall within one standard deviation from the mean
- 95% of the values fall within two standard deviations from the mean
- 99% of the values fall within three standard deviations from the mean

6.20 Graphing the Normal Distribution and Computing Bounded Area

Your calculator provides the built-in function `normPdf(x, μ , σ)` that represents the Normal Distribution’s continuous probability density function for the data set with mean μ and standard deviation σ . To visualize any data set’s probability density function, graph `normPdf(x, μ , σ)`, supplying appropriate values for μ and σ . Then, to compute the probability that a random variable X will take on a value between any two values a and b , use the calculator’s **Analyze Graph** \rightarrow **Bounded Area** capability to compute the area under the curve between these two values.

For example, consider a data set of the heights of Ghanaian women, where the population mean $\mu = 159$ cm and the standard deviation $\sigma = 4.9$ cm. Start by graphing the function `normPdf(x, 159, 4.9)` and adjusting the graph window to be centered around the mean value of 159. From the previous section, we know that 99% area under the graph should fall within three standard deviations of the mean. So, you can use this fact to size the window to show the x -axis from $159 - 3 \cdot 4.9 = 144.3$ to $159 + 3 \cdot 4.9 = 173.3$.

Next, suppose you want to compute the probability that a random variable X ’s value falls in the range from 154.1 to 163.9. One way to do this is to compute the area under the probability density function we have graphed between these two values. To do that follow these steps:

- While on a graph tab showing the appropriate `normPdf(x, μ , σ)` function, press [*menu*]
- Choose **Analyze Graph** \rightarrow **Bounded Area**
- Click on the `normPdf(x, μ , σ)` function to identify the first curve
- Click on the x -axis to identify the second curve
- Instead of using the pointer, just type in 154.1 and press [*Enter*] to set the lower bound
- Type in 163.9 and press [*Enter*] to set the upper bound

After following these steps, as shown in Figure 107 the bounded area under the curve appears shaded in grey and the value of the bounded area appears (in this example 0.683) appears on the graph.

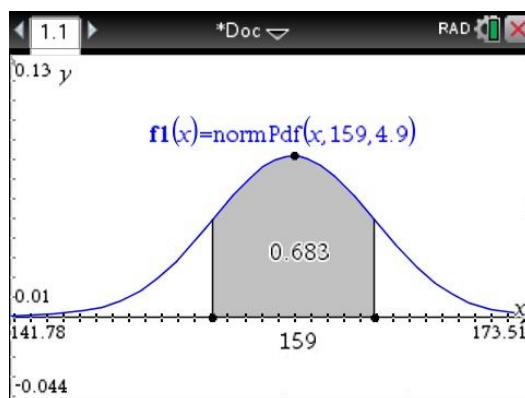


Figure 107: Area Under Normal PDF for $\mu = 159$ and $\sigma = 4.9$ Between 154.1 and 163.9

Notice that the two boundaries selected were exactly one standard deviation away from the mean value of 159. That is, $154.1 = 159 - 1 \cdot 4.9$ and $163.9 = 159 + 1 \cdot 4.9$. As we would expect from the rules of thumb outlined in Section 6.19, “Normal Distribution and Curves”, 68% of the values fall within one standard deviation of the mean in a normally distributed data set. So, the computed area of 0.683 under the curve from 154.1 to 163.9 agrees with this expectation as this is 68.3%.

In general, we can remember that the probability that a random variable X takes on a value:

- **between a and b** , equals the area under the normal curve bounded by a and b .
- **greater than a** , equals the area under the normal curve bounded by a and ∞ .
- **less than a** , equals the area under the normal curve bounded by $-\infty$ and a .
- **exactly a** , equals 0 (there is no area under the normal curve between a and a).

For the bounded area graph cases involving $-\infty$ and ∞ , to avoid rescaling the graph, use a number like $\mu + 4\sigma$ instead of ∞ and a number $\mu - 4\sigma$ instead of $-\infty$.

It’s good to visualize and understand the normal distribution probability calculation using the area under the curve. However, in Section 6.22, “Using the Calculator for Normal Distribution Problems” we’ll learn an even simpler way to get the answer using the `normCdf()` function without using a graph.

6.21 Standardizing Normal Variables to Get z -values (z -scores)

In order to work with a single curve for all different possible data sets having a normal distribution, we standardize data values so that we can use the Standard Normal Distribution for all calculations of probability. The standard normal distribution has a mean $\mu = 0$ and standard deviation $\sigma = 1$. We can convert back and forth between the actual data values in a given data set and so-called z -value for the standard normal distribution using the formula below. If you know the mean μ and the standard deviation σ of the values in the actual data set, then for a given data value x , you compute its corresponding z -value by subtracting the mean and dividing by the standard deviation:

$$z = \frac{x - \mu}{\sigma}$$

Of course, this formula can be used in reverse by rearranging it to solve for x . If you know the mean μ and the standard deviation σ of the values in the actual data set, then for a given

z -value, you compute its corresponding actual data value x by multiplying by the standard deviation and adding the mean:

$$x = \sigma z + \mu$$



Figure 108: Standardizing Data to Work with Standard Normal Distribution⁵

6.22 Using the Calculator for Normal Distribution Problems

Once you have identified a normal distribution problem, your calculator can quickly get you the answer. There are several kinds of problems you will encounter:

- Probability of getting a value less than or equal to k
- Probability of getting a value between k and m
- Probability of getting a value greater than or equal to k .

All of these problems can be handled using the “Normal Cumulative Distribution Function” (`normalCdf`) function on your calculator. This function computes the area under the normal distribution’s probability density function between the two values you supply. As explained in Section 6.20, “Graphing the Normal Distribution and Computing Bounded Area”, this area is the probability that a random variable X ’s value will fall in the range specified by those two values. To access the `normalCdf` function from a calculator tab, do the following:

- With your cursor in a calculator tab, press `[menu]` and choose **Statistics** → **Distributions** → **Normal Cdf...**
- In the *Normal Cdf* dialog, enter the lower bound (or `-9E999` for $-\infty$), the upper bound (or `9E999` for $+\infty$), the mean μ , and the standard deviation σ .
- Click `[OK]`

In order to find the probability of getting...

- a value less than or equal to k , then use `normalCdf` and enter the value `-9E999` for the lower bound and k for the upper bound.
- a value between k and m , then use `normalCdf` and enter the value k for the lower bound and m for the upper bound.
- a value greater than or equal to k , then use `normalCdf` and enter the value k for the lower bound and `9E999` for the upper bound.

If the value you are entering is a standardized z -value, then enter $\mu = 0$ and $\sigma = 1$. Otherwise, enter the correct μ and σ values for your dataset. Sometimes you may need to use the z -value standardization formula first to determine a missing μ or σ value, given other provided information in the problem.

⁵Source: mathsisfun.com/data/standard-normal-distribution.html

6.22.1 Tip for Calculators Whose `normalCdf` Does Not Have Lower Bound

If your calculator's `normalCdf` function does not allow you to input a lower bound, you need to do problems that ask for $P(X \geq k)$, and $P(k \leq X \leq m)$ a slightly different way.

Doing $P(X \geq k)$ on TI-84 To calculate $P(X \geq k)$, compute $1 - \text{normalCdf}(k, \mu, \sigma)$. That is, calculate $P(X < k)$ and subtract it from 1.

Doing $P(k \leq X \leq m)$ on TI-84 To calculate $P(k \leq X \leq m)$, compute $\text{normalCdf}(m, \mu, \sigma) - \text{normalCdf}(k, \mu, \sigma)$. That is, calculate $P(X < k)$ and subtract it from $P(X < m)$.

6.23 Using Inverse Normal Cumulative Density Function

If you know the probability (V) that a random variable X 's value is less than some unknown number k , use the Inverse Normal function to find the value k that makes $P(X < k) = V$.

- With your cursor in a calculator tab, press [menu] and choose **Statistics** → **Distributions** → **Inverse Normal...**
- In the *Inverse Normal* dialog, enter the probability value V for the *Area*, the mean μ , and the standard deviation σ .
- Click [OK]

If you are told that the probability V represents...

- $P(x \leq k)$, then use `invNorm(V, μ , σ)`
- $P(x \geq k)$, then use `invNorm(1-V, μ , σ)`

If the value you are entering is a standardized z -value, then enter $\mu = 0$ and $\sigma = 1$. Otherwise, enter the correct μ and σ values for your dataset. Sometimes you may need to use the z -value standardization formula first to determine a missing μ or σ value, given other provided information in the problem.

6.24 Determining z Value and σ from the Probability

Sometimes, a problem will give you the mean μ of your data but not provide the value of the standard deviation σ . In this case, you must use additional information provided in the problem (or determined in previous steps) in order to calculate the value of σ .

If you know the probability (i.e. area) that corresponds to one particular value in the data, and you also know the mean μ , then you can figure out the σ from these values using the approach described here. Consider the following example:

The weights of fish in a lake are normally distributed with a mean of 760 g and standard deviation σ . It is known that 78.7% of the fish have weights between 705 g and 815 g. Find the probability that a fish weighs less than 815 g and write down the standardized (z) value for 815 g. Figure 109 shows the information we've been given.

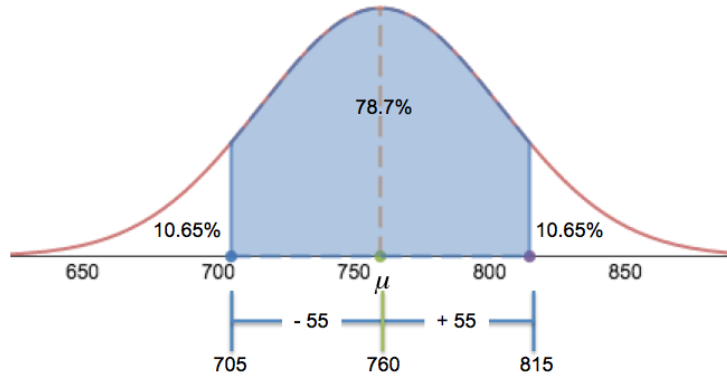


Figure 109: Visualizing Normalized Data for the Fish Problem

We notice that the range of fish weights we're given 705 g and 815 g are symmetric about the mean 760 g. One is 55 g less than the mean and the other is 55 g greater than the mean. Since the normal distribution is symmetric about the mean, this allows us to conclude that the area remaining under the curve below 705 g must be equal to the area remaining under the curve above 815 g. That combined area remaining is $100\% - 78.7\% = 21.3\%$, and since the two “tails” are symmetric, we know that half of this remaining area must be in the left tail and half in the right tail. So $\frac{21.3\%}{2} = 10.65\%$ in each tail. Therefore, the probability (i.e. area) that a fish weighs less than 815 g is the sum of the left tail (10.65%) and the central area given in the problem between 705 g and 815 g. This equals $78.7 + 10.65 = 89.35\%$.

Now, we can use this probability (area) to find the standard deviation. We can use the *standard* normal distribution, whose mean and standard deviation we always know are $\mu_{\text{std}} = 0$ and $\sigma_{\text{std}} = 1$ to find the z -value corresponding to an area of 89.35%. Using our calculator's `invNorm(Area, μ_{std} , σ_{std})` function. The value of `invNorm(0.8935, 0, 1)` = 1.24536. So this is the z -value corresponding to the area of 89.35%.

Finally, we can use this z -value to compute the standard deviation of the original data set of fish weights. To do that we use the formula from Section 6.21, “Standardizing Normal Variables to Get z -values (z -scores)” and fill in the values we know for z , μ , and x .

$$z = \frac{x - \mu}{\sigma}$$

$$1.24536 = \frac{815 - 760}{\sigma}$$

$$\sigma = \frac{55}{1.24536}$$

$$\sigma = 44.1639$$

7 Calculus

Calculus is the study of rates of change and areas under a curve and how these two concepts are closely related. Isaac Newton (in England) and Gottfried Leibniz (in Germany) invented this area of mathematics independently around 1690 while each was trying to find a more effective way to study the relationships between the distance traveled by the planets, the velocity at which they travel, and their acceleration with respect to time.

7.1 Overview of Concepts for Differential Calculus

Newton and Leibniz used calculus to study rates of change. They did this by studying how the position of a moving body differs from one moment to the next. The techniques they invented became known as Differential Calculus, the study of instantaneous rates of change.

7.1.1 Rate of Change in Distance

Figure 110 shows one of the simplest examples of a rate of change from the world around you. Every time you drive, or ride in a car, you see an odometer and a speedometer. The odometer measures the current distance your car has driven (in kilometers), while the speedometer measures your car's current speed (in kilometers per hour). Your car's current speed is the rate at which your distance traveled is changing per unit of time (in this case, hours).



Figure 110: Speedometer Shows Current Speed While Odometer Shows Current Distance⁶

These two measuring devices are intimately connected. Glancing at the speedometer's current reading of 154 km/h, you can predict the value the odometer will have one hour from now if you continue at the current rate of speed: $28888 = 28734 + 154$. This is true because right now the odometer is showing 28734 and (at the current rate of speed) you will travel 154 kilometers in the next one hour. After a second hour at this constant rate of 154 km/h, the speedometer will read 29042 as shown in the distance/time graph in Figure 111. In contrast, reading the odometer before leaving and again after arriving, you can compute your average speed (in kilometers per hour) by subtracting the initial distance reading from the final distance measurement and then dividing by the total travel time.

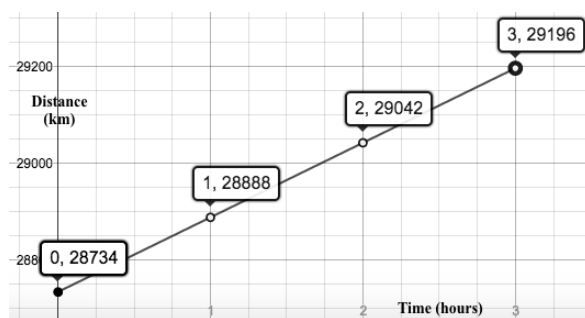


Figure 111: Distance versus Time Graph

⁶Source: 123rf.com/photo_14194592_vector-illustration-of-a-speedometer.html

Assuming a constant speed of 154 km/h , Figure 112 shows a graph of the speed versus time. Since the speed is constant, the graph is a flat line that stays at the value 154 for three hours.

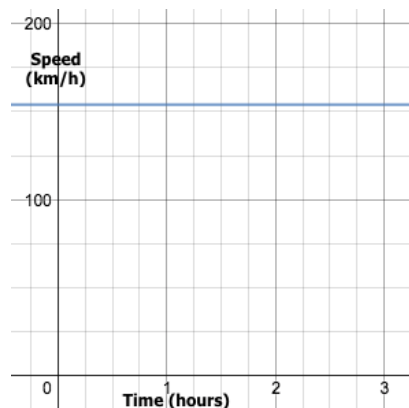


Figure 112: Graph of Constant Speed of 154 km/h over Time

In reality, your car's speed is never constant. You speed up, slow down, and even can drive in reverse (backwards). When going faster, you cover more distance in a given unit of time, and when going slower, less distance. When driving backwards your distance traveled becomes smaller than it was before.

Of course the concepts are not limited to a car. Imagine you take a walk to a park six kilometers away from home with your new Apple Watch. It tells you your walking speed (in km/h) and the distance in kilometers that you are away from home. Figure 113 shows a graph of your distance from home versus the time you've walked. Notice that you got to the park after 1.48 hours and started walking back home, but after 4.147 hours you realize you left something at the park so you turn around and walk back to the park again.

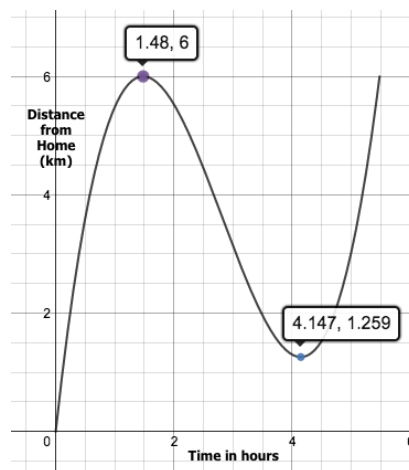


Figure 113: Distance versus Time Graph of Walk to Park with Apple Watch

If you would have walked at a constant speed (as in our 154 km/h car example graphed in Figure 111 above) then the distance time graph would have been a straight line sloping upwards until you got to the park then a straight line sloping downwards for your walk home. Instead, this distance graph has a shape that is not straight. In the first half hour you cover almost 3.75 kilometers so you are walking quite fast, but then in the second half hour you slow down your pace and only cover two kilometers. During the third half hour you must have been tired and walking even more slowly since you only do a half a kilometer during that time.

7.1.2 Derivative Gives Instantaneous Rate of Change Using Tangent Lines

Newton and Leibniz were interested in measuring the rate of change at any given instant of time, when observing a body in motion. Their clever idea, when studying a graph of the distance versus time, was to use the tangent line to the curve at a given point $(t, f(t))$ as the measure of the rate of change of the distance at that very instant in time t . Figure 114 shows the tangent lines at a number of points along the distance graph from Figure 113. The slope (or “gradient”) of the tangent line represents the rate of change of your distance at that moment in time. Tangent lines that are sloping upwards (going left to right) represent intervals of time where the value of the function is increasing. The steeper the up-sloping line, the faster the value is increasing at that moment. Tangent lines sloping downward (going left to right) represent intervals of time where the value of the function is decreasing. The steeper the down-sloping line, the faster the value is decreasing at that moment. Places where the slope of the tangent is perfectly horizontal (slope = zero) represent a point at which the function might be switching from increasing to decreasing (a local maximum value) or switching from decreasing to increasing (a local minimum value).

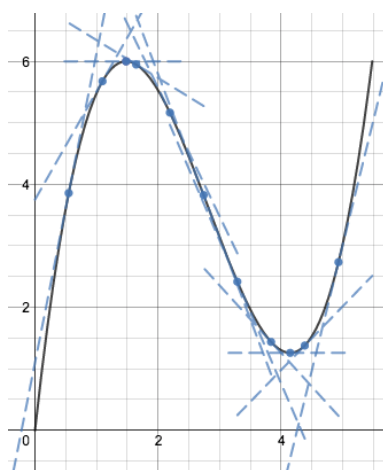


Figure 114: Example Function with Tangent Lines

Notice in Figure 114 that the maximum distance away from home (when you got the park and turned around) has a slope of zero (a horizontal line) at that point on the graph. Similarly when you realized later you had to turn around and head back to the park, you were temporarily at a minimum distance away from home compared to the time just before and just after that point.

The slope of the tangent at a given point in time is just a scalar number: positive numbers are sloping upwards, zero is horizontal, and negative numbers are sloping downwards. If we plot these different slope values as points $(x, \text{slope-of-}f(x))$ on the same axes for the same values of x and “connect the dots”, we end up with Figure 115.

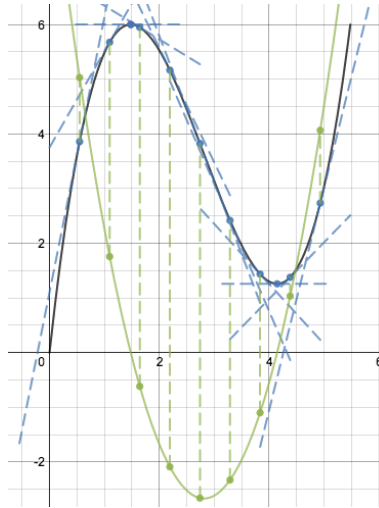


Figure 115: Graphing the Derivative

We see that there is a pattern to the values that these slopes take on. If we call the original distance versus time function $f(t)$, then there appears to be a *different* smooth function that passes through all of the dots that represent the values of the slopes of the tangents of the original function $f(x)$. We call this new function the “derivative of $f(t)$ ” and we write it using the notation $f'(t)$. This derivative function $f'(t)$ provides a formula to compute the slope of the tangent to $f(t)$ at any value of t that we care to know it.

7.1.3 Relationships Between Function and Derivative

There are several interesting things to notice about the relationship between the function $f(t)$ and its derivative $f'(t)$. Perhaps the first thing we notice is shown in Figure 116, that the maximum and minimum points of the original function $f(t)$ correspond exactly to the “zeros” of the derivative $f'(t)$. That is, where the derivative $f'(t) = 0$. The derivative function $f'(t)$ gives the slope of the tangent to $f(t)$ at the value t , and at the values of t where $f(t)$ has minimums and maximums, the slope is zero, so this makes intuitive sense.

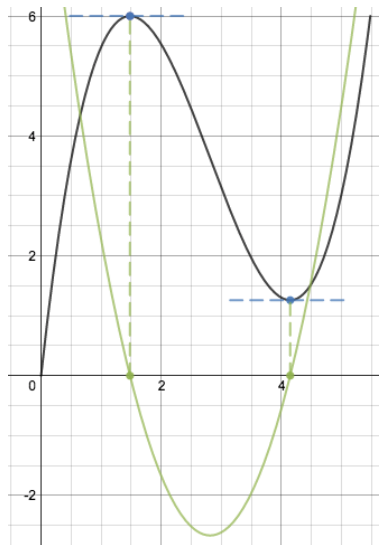


Figure 116: Minimum and Maximum of Original Function Occur Where Derivative Has Zeros

Another interesting observation shown in Figure 117 is that the intervals where the derivative $f'(t)$'s value is positive – that is, where its graph is above the x -axis – correspond exactly to

the intervals where the original function $f(t)$ is *increasing*. Likewise, we see that the intervals where the derivative $f'(t)$'s value is negative – that is, where its graph is below the x -axis – correspond exactly to the intervals where the original function $f(t)$ is *decreasing*.

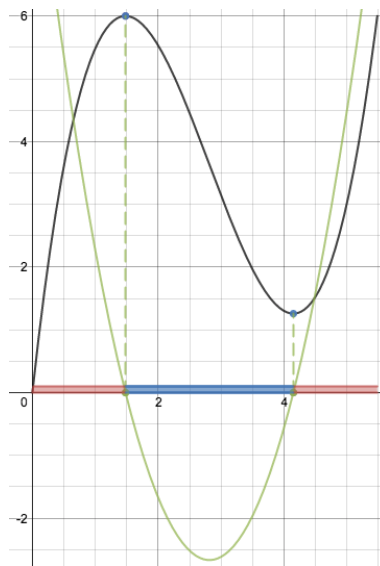


Figure 117: Original Function Increases Where Derivative is Positive, Decreases Where Negative

If we shift our attention to the $f'(t)$ derivative function and consider the slopes of *its* tangent lines in a number of places, we can plot the values of *these* slopes as we've done in Figure 118 to see that they also follow a pattern.

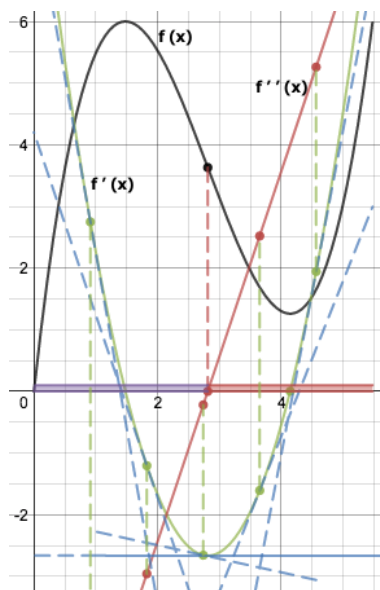


Figure 118: f'' Reveals Min/Max of f' and Concavity of f

Connecting the dots reveals that they lie along a straight line in this example. Since this is the derivative function of $f'(t)$, we might rightly call this function $(f')'(t)$, however we simplify the notation and just write $f''(t)$ instead. As we'd expect from above, we can observe that where the line representing $f''(t)$ falls below the x -axis (meaning its values are negative), the $f'(t)$ function is decreasing and where this line falls above the x -axis (meaning its values are positive) the $f'(t)$ function is increasing. However, notice something else. The interval where the $f''(t)$ is negative, also corresponds to where the original function $f(t)$ is concave down. Similarly, the interval

where the $f''(t)$ is positive, also corresponds to where the original function $f(t)$ is concave up. What's more, point where the line $f''(t)$ cross the x -axis (i.e. where $f''(t) = 0$) corresponds both to where $f'(t)$ has a minimum as well as to where the original function f switches from being concave down to being concave up. This is known as a point of inflection of f .

7.1.4 Summary of Derivative Concepts

Given a function $f(x)$, its (first) derivative $f'(x)$...

- Is a related function that gives the slope of the tangent to the original $f(x)$ at any point
- Has a negative value wherever $f(x)$ is decreasing
- Has a positive value wherever $f(x)$ is increasing
- Equals zero at values of x where $f(x)$ has “critical points” like maxima (where sign of slope switches from + to -) and minima (where sign of slope switches from - to +)

and its second derivative $f''(x)$...

- Is a related function that gives the concavity of $f(x)$ at any point
- Has a negative value wherever $f(x)$ is concave down
- Has a positive value wherever $f(x)$ is concave up
- Equals zero at values of x where $f(x)$ has potential “points of inflection” where concavity changes from up to down or vice versa.

In particular, $f''(x)$ is also the first derivative of $f'(x)$, so at the same time it...

- Gives slope of the tangent to the first derivative $f'(x)$ at any point
- Has a negative value wherever $f'(x)$ is decreasing
- Has a positive value wherever $f'(x)$ is increasing
- Equals zero at values of x where $f'(x)$ has “critical points” like maxima (where sign of slope switches from + to -) and minima (where sign of slope switches from - to +)

7.2 Equations of Tangents and Normals

The tangent to $f(x)$ at a point $x = p$ is a line that goes through the point $(p, f(p))$ and which has a slope equal to the derivative of $f(x)$ at $x = p$. In other words, the slope of the tangent line is $f'(p)$. Using the “point-slope” form of the equation of a line, this means that the tangent has the equation:

$$y - f(p) = f'(p)(x - p)$$

By multiplying out the $f'(p)$ into the binomial in parenthesis on the right hand side of the equation, and adding $f(p)$ to both sides, you can rewrite this tangent line in “slope-intercept” form instead if the problem asks for that.

The normal to $f(x)$ at a point $x = p$ is the line that is perpendicular to the tangent to $f(x)$ at $x = p$. It also goes through the point $(p, f(p))$, but its slope is the negative reciprocal of the slope of the tangent line. Therefore, in point-slope form, the equation of the normal line is:

$$y - f(p) = -\frac{1}{f'(p)}(x - p)$$

Again, you can multiply out the quantities on the right hand side and add $f(p)$ to both sides to rewrite this equation in slope-intercept form if the problem asks for that instead. Figure 119 shows both tangent and normal on the graph of $f(x)$ at $x = p$.

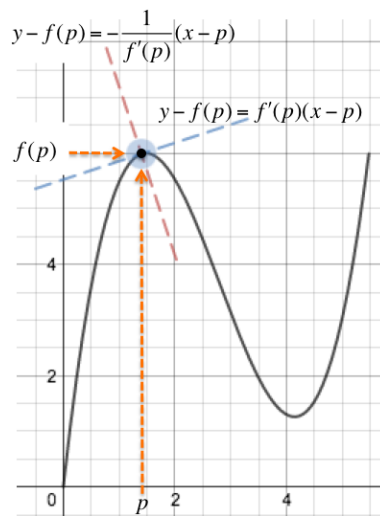


Figure 119: Equations of Tangent and Normal to $f(x)$ at $x = p$

7.3 Notation for Derivatives

If a function is named $f(x)$, as we've seen above its first derivative is written using the notation $f'(x)$. However, you will also see the notation $\frac{df}{dx}$, the derivative of the function f with respect to x . The second derivative is written either $f''(x)$ or $\frac{d^2 f}{dx^2}$. For higher order derivative, when you have taken the derivative of f a total of n times, you'll see $f^{(n)}(x)$ or $\frac{d^n f}{dx^n}$.

If a function is written as $y = f(x)$ then writing $\frac{dy}{dx}$ is the same as $f'(x)$.

7.4 Graphing Derivatives with Your Calculator

Suppose you are given a problem with the function $f(x) = \sqrt{6x - x^2}$ that forms the half-circle from 0 to 6 depicted in Figure 120. You are asked to consider a point x along the x -axis that defines a rectangle with two of its corners sitting on the half-circle. The area of the rectangle $A(x)$ is the length $(6 - 2x)$ times the height $\sqrt{6x - x^2}$. The problem asks to determine the values of x where the rate of change of the area function $A(x)$ is decreasing.

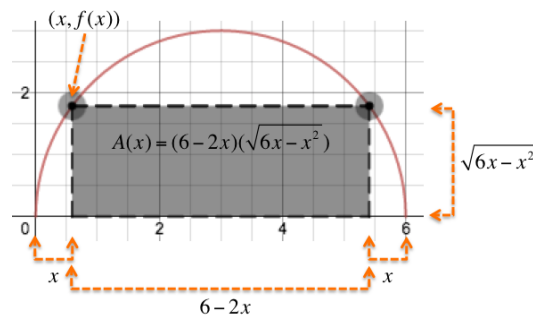


Figure 120: Formula for the Area of a Rectangle in a Half-Circle

Recall that the “rate of change of $A(x)$ ” is one way to refer to the derivative $A'(x)$ or $\frac{d}{dx}A(x)$. Without having to compute the derivative function $A'(x)$ yourself, you can easily graph the derivative of the function $A(x)$ by using your calculator’s built-in derivative function, then you can analyze the graph to see where it is positive, negative, and where it has zeros. The steps to do this depend on whether you’re using a TI-Nspire CX calculator or the TI-84 Silver Edition.

7.4.1 Graphing Derivatives on the TI-Nspire CX

On the TI-Nspire CX, there is a $\frac{d}{dx}$ operator in the palette of symbols that you can directly enter as part of a function to graph. By putting the definition of function $A(x)$ directly inside the $\frac{d}{dx}$ operator as shown in Figure 121, the graph produced is the graph of the derivative of $A(x)$, that is, of $A'(x)$ without having to compute the derivative yourself by hand. Then, using the normal “Analyze Graph” or “Trace” functions, you can study the graph of the derivative to find where it is zero. This means that solving any kind of minimum/maximum or other problem involving derivatives becomes very straightforward. Here we’ve used the *Analyze Graph* \rightarrow *Zeros* twice to find that that the derivative $A'(x)$ equals zero at $x = 0.879$ and $x = 5.12$. By studying the graph, we see that the values of the derivative are positive for $x < 0.879$ and for $x > 5.12$ while the values of the derivative are negative for $0.879 < x < 5.12$. Therefore, the answer to the question of where the function $A(x)$ is decreasing is precisely between $0.879 < x < 5.12$ where its derivative is negative.

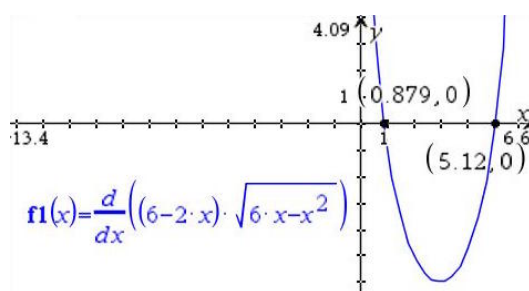


Figure 121: Graph of Derivative of $A(x)$ with Zeros Found Using Analyze Graph

If we were asked to find the maximum and minimum values of $A(x)$ we would need to identify which of the two zeros of $A'(x)$ represents the minimum and which represents the maximum. Looking at the graph, we see that the value of the derivative changes from positive (above the x -axis) to negative (below the x -axis) at $x = 0.879$ and that at $x = 5.12$ the value of the derivative changes from negative to positive. Therefore, the original function $A(x)$ representing the area of the rectangle must have a maximum at $x = 0.859$ and a minimum at $x = 5.12$. To find the maximum value of the area itself, we just compute $A(0.879) = 9$ and to find the minimum value of the area we compute $A(5.12) = -9$. In this particular problem, however, we don’t want to allow x to get bigger than 3 because then there is no rectangle left to compute the area of, so the only interesting point is the maximum area of 9 that occurs when $x = 0.879$.

7.4.2 Graphing Derivatives on the TI-84 Silver Edition

On the TI-84 Silver Edition, you use the `nDeriv` function to enter the derivative operator and then enter the variable x in the $\frac{d}{dx}$ symbol that appears, enter the equation whose derivative you want to graph, and then enter a value of $x = x$ in the specific point box at the end of the formula as shown in Figure 122. Once graphed, you can use the analyze graph features as you can with any other graph.

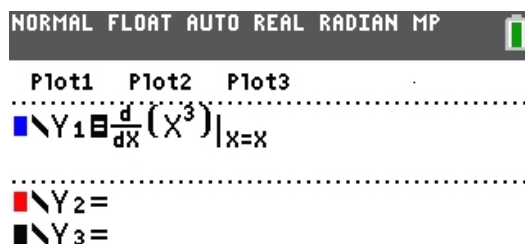


Figure 122: Graphing a Derivative using nDeriv on TI-84 Silver Edition

For more information, see <http://bit.ly/ib-sl-maths-graph-deriv-ti-84>

7.5 Computing Derivatives at a Point with Your Calculator

Suppose you are given a function like $f(x) = \frac{\sin x + x^2}{e^x + x^3}$ and you are asked to find the equation of the tangent line at $x = 0.5$. One step in finding the equation of the tangent is calculating the derivative $f'(0.5)$ which gives the slope of the tangent line at $x = 0.5$. The other step is finding the point $(0.5, f(0.5))$ through which the line passes where it is tangent to the curve $f(x)$ at that value of x . You can compute the value of the derivative of a function at a specific point using your calculator without having to mechanically determine the derivative function yourself. The steps to do this depend on whether you're using a TI-Nspire CX calculator or the TI-84 Silver Edition.

7.5.1 Computing Derivative at a Point on the TI-Nspire CX

On the TI-Nspire CX, in a calculator tab use the “Menu” button to pick the “Calculus → Numerical Derivative at a Point” option. In the dialog box that appears, enter the variable your equation will use in the *Variable* field (it defaults to x). In the *Value* field, enter the value of x for which you want to compute the derivative. For example, given the problem above you would enter the value 0.5 there. In the *Derivative* picklist, choose either the “1st Derivative” or the “2nd Derivative” depending on which you want to compute. Before clicking OK the dialog will look like what you see in Figure 123.

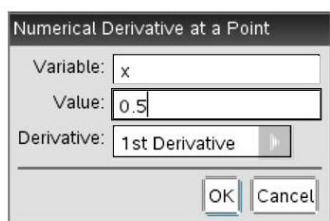


Figure 123: Computing a Derivative at a Specific Point (Step 1) on TI-Nspire CX

A template $\frac{d}{dx} () |_{x=0.5}$ appears and you proceed to enter in the equation whose derivative you want to calculate at $x = 0.5$ as shown in Figure 124. When you press “Enter” the calculator displays the result of 0.502408.

$$\frac{d}{dx} \left(\frac{\sin(x) + x^2}{e^x + x^3} \right) |_{x=0.5} \quad 0.502408$$

Figure 124: Computing a Derivative at a Specific Point (Step 2) on TI-Nspire CX

You then can use your calculator to compute $f(0.5)$ by evaluating $\frac{\sin 0.5 + 0.5^2}{e^{0.5} + 0.5^3} = 0.41124$. With this information in hand, we now know the slope of the tangent line at $x = 0.5$ is 0.502408 and the line goes through the point $(0.5, 0.41124)$, so using point-slope form, the equation of the tangent (rounding to three significant figures) is:

$$y - 0.411 = 0.502(x - 0.500)$$

7.5.2 Computing Derivative at a Point on the TI-84 Silver Edition

To compute a derivative at a point on the TI-84 Silver Edition, use the `nDeriv` function. For more information, see: <http://bit.ly/ib-sl-maths-graph-deriv-ti-84>

7.6 Rules for Computing Derivatives

It is amazing that there are simple rules to compute the derivative of all of the common functions you work with in calculus. The formulas are in the formula booklet but for completeness we recall them here.

7.6.1 Derivative of x^n

To compute the derivative of x raised to any rational number power $n \in \mathbb{Q}$, follow this rule:

$$f(x) = x^n \quad \rightarrow \quad f'(x) = n \cdot x^{n-1}$$

You multiply by the current exponent and change the power on x to one less than its current value. The following examples help clarify how this works in different situations:

Positive Whole Number Power Example

$$f(x) = x^7 \quad \rightarrow \quad f'(x) = 7 \cdot x^{7-1} = 7 \cdot x^6$$

Negative Whole Number Power Example

$$f(x) = \frac{1}{x^7} = x^{-7} \quad \rightarrow \quad f'(x) = -7 \cdot x^{-7-1} = -7 \cdot x^{-8} = \frac{7}{x^8}$$

Positive Fractional Power Examples

$$f(x) = \sqrt{x} = x^{\frac{1}{2}} \quad \rightarrow \quad f'(x) = \frac{1}{2} \cdot x^{\frac{1}{2}-1} = \frac{1}{2} \cdot x^{\frac{1}{2}-\frac{2}{2}} = \frac{1}{2} \cdot x^{-\frac{1}{2}} = \frac{1}{2} \cdot \frac{1}{x^{\frac{1}{2}}} = \frac{1}{2\sqrt{x}}$$

$$f(x) = \sqrt{x^3} = x^{\frac{3}{2}} \quad \rightarrow \quad f'(x) = \frac{3}{2} \cdot x^{\frac{3}{2}-1} = \frac{3}{2} \cdot x^{\frac{3}{2}-\frac{2}{2}} = \frac{3}{2} \cdot x^{\frac{1}{2}} = \frac{3}{2}\sqrt{x}$$

Negative Fractional Power Examples

$$f(x) = \frac{1}{\sqrt{x}} = x^{-\frac{1}{2}} \quad \rightarrow \quad f'(x) = -\frac{1}{2} \cdot x^{-\frac{1}{2}-1} = -\frac{1}{2} \cdot x^{-\frac{1}{2}-\frac{2}{2}} = -\frac{1}{2} \cdot x^{-\frac{3}{2}} = -\frac{1}{2} \cdot \frac{1}{x^{\frac{3}{2}}} = -\frac{1}{2\sqrt{x^3}}$$

$$f(x) = \frac{1}{\sqrt{x^3}} = x^{-\frac{3}{2}} \quad \rightarrow \quad f'(x) = -\frac{3}{2} \cdot x^{-\frac{3}{2}-1} = -\frac{3}{2} \cdot x^{-\frac{3}{2}-\frac{2}{2}} = -\frac{3}{2} \cdot x^{-\frac{5}{2}} = -\frac{3}{2} \cdot \frac{1}{x^{\frac{5}{2}}} = -\frac{3}{2\sqrt{x^5}}$$

7.6.2 Derivative of $\sin x$

To compute the derivative of $\sin x$, follow this rule:

$$f(x) = \sin(x) \quad \rightarrow \quad f'(x) = \cos x$$

7.6.3 Derivative of $\cos x$

To compute the derivative of $\cos x$, follow this rule:

$$f(x) = \cos(x) \quad \rightarrow \quad f'(x) = -\sin x$$

7.6.4 Derivative of $\tan x$

To compute the derivative of $\tan x$, follow this rule:

$$f(x) = \tan(x) \quad \rightarrow \quad f'(x) = \frac{1}{\cos^2 x}$$

7.6.5 Derivative of e^x

To compute the derivative of e^x , follow this rule:

$$f(x) = e^x \quad \rightarrow \quad f'(x) = e^x$$

7.6.6 Derivative of $\ln x$

To compute the derivative of $\ln x$, follow this rule:

$$f(x) = \ln x \quad \rightarrow \quad f'(x) = \frac{1}{x}$$

7.7 Differentiating a Scalar Multiple

When a function is a scalar multiple of another function, you compute its derivative by taking the same scalar multiple of the derivative of that function. Said another way, “The derivative of a scalar multiple of a function is the scalar multiple of the derivative.” The scalar constant just goes along for the ride as you take the derivative of the function with which it is multiplied.

$$f(x) = k \cdot u(x) \quad \rightarrow \quad f'(x) = k \cdot u'(x)$$

Example

$$f(x) = 16x^7 \quad \rightarrow \quad f'(x) = 16 \cdot 7 \cdot x^6 = 112x^6$$

7.8 Differentiating a Sum or Difference

When a function is comprised of multiple terms that are added or subtracted, you compute its derivative by taking the derivative of each term and combining the results. Said another way, “The derivative of a sum/difference is the sum/difference of the derivatives.”

$$\begin{aligned} f(x) &= u(x) + v(x) & \rightarrow & \quad f'(x) = u'(x) + v'(x) \\ f(x) &= u(x) - v(x) & \rightarrow & \quad f'(x) = u'(x) - v'(x) \end{aligned}$$

Example

$$f(x) = 16x^7 + 5 \sin x + 2 \ln x \quad \rightarrow \quad f'(x) = 16 \cdot 7 \cdot x^6 + 5 \cos x + 2 \cdot \frac{1}{x} = 112x^6 + 5 \cos x + \frac{2}{x}$$

7.9 Example Using Derivation Rules

Let $f(x) = px^3 + px^2 + qx$. Find $f'(x)$, and then given that $f'(x) \geq 0$ show that $p^2 \leq 3pq$. Using the rules from the preceding sections about x^n , scalar multiples, and sums, we have:

$$\begin{aligned} f(x) &= px^3 + px^2 + qx \\ f'(x) &= 3 \cdot px^2 + 2 \cdot px + q \\ &= 3px^2 + 2px + q \end{aligned}$$

We see that $f'(x)$ is a quadratic equation, and therefore its graph is a parabola. We're told that $f'(x) \geq 0$, so this parabola must be completely above the x -axis, or at best have its vertex touching the x -axis. Said another way, it means that the parabola might have no real roots (if it never touches the x -axis) or at most one real root (if it touches the x -axis with its vertex). Recall from Section 2.16.5, "Using the Discriminant to Find the Number of Zeros", that if a quadratic function has no roots, then its discriminant is less than zero, and that if it has a single root (also known as two equal roots), then its discriminant equals zero. Therefore, the fact that our quadratic equation $f'(x) \geq 0$ implies that its discriminant is less than or equal to zero. The discriminant of $f'(x)$ is $(2p)^2 - 4 \cdot (3p) \cdot q$, so we have:

$$\begin{aligned} (2p)^2 - 4 \cdot (3p) \cdot q &\leq 0 \\ 4p^2 - 12pq &\leq 0 \\ 4(p^2 - 3pq) &\leq 0 \\ p^2 - 3pq &\leq 0 \\ p^2 &\leq 3pq \end{aligned}$$

7.10 "Chain Rule": Differentiating Composed Functions

The rules for calculating derivatives are meant to be applied to simple functions of x . When you encounter more complicated functions, you can always think of them as compositions of two simpler functions. Recall from Section 2.9, "Composing One Function with Another" that a composed function "chains" together the output of one function to the input of the next. This rule is known as the Chain Rule because when computing the derivative of a composition, you must also "chain" together the derivative of the outer function with the derivative of the inner function using multiplication.

$$f(x) = u(v(x)) \quad \rightarrow \quad f'(x) = u'(v(x)) \cdot v'(x) \text{ or alternatively written } \frac{du}{dv} \cdot \frac{dv}{dx}$$

You cannot *accidentally* use the chain rule in a situation where it is not required. If you adopt the convention of always doing the chain rule, at the very worst the derivative of the inner function will be equal to one (1) so multiplying by that value of one will not "harm" anything. For example, if you have the function $f(x) = x^7$ if you do the derivative to get $7x^6$ and then multiply by the derivative of x (the thing being raised to the seventh power), you will still get the correct answer because the derivative of x is 1. Over time you will realize that you only need to use the chain rule when the inner function is something more "complicated" than a simple x , but in the meantime it doesn't hurt to always use it.

Example 1 Consider the function:

$$f(x) = 16(x^2 + 5x + 2)^7$$

This looks like a simpler function $u(x) = 16x^7$ composed with another function $v(x) = x^2 + 5x + 2$. If we make these substitutions, then we can rewrite $f(x)$ as:

$$f(x) = u(v(x))$$

When we take the derivative of a composite function like this, we first take the derivative of the outer function (keeping the inner function in place as it is), then multiply that result by the derivative of the inner function with which it is composed. To calculate the derivative of the outer function $u(x)$, with respect to the inner function $v(x)$ we follow the derivative rules for the simple functions, but keep the composed function $v(x)$ inside intact. Then we multiply this result by the derivative of the inner function with respect to x . For example:

$$\begin{aligned} f(x) &= 16(x^2 + 5x + 2)^7 = 16(v(x))^7 && \rightarrow \\ f'(x) &= 16 \cdot 7(v(x))^6 \cdot v'(x) \\ &= 112(v(x))^6 \cdot (2x + 5) \\ &= 112(x^2 + 5x + 2)^6 \cdot (2x + 5) \end{aligned}$$

Example 2 Consider the function:

$$f(x) = 6 \sin(2e^x + 3x^3)$$

If we let $u(x) = 6 \sin x$ and $v(x) = 2e^x + 3x^3$, then we can rewrite $f(x) = u(v(x))$. As before, we have:

$$\begin{aligned} f(x) &= 6 \sin(2e^x + 3x^3) = 6 \sin(v(x)) && \rightarrow \\ f'(x) &= 6 \cos(v(x)) \cdot v'(x) \\ &= 6 \cos(v(x)) \cdot (2e^x + 3 \cdot 3x^2) \\ &= 6 \cos(2e^x + 3x^3) \cdot (2e^x + 9x^2) \\ &= 6 \cos(2e^x + 3x^3) \cdot (2e^x + 9x^2) \end{aligned}$$

Example 3 Consider the function:

$$f(x) = 4e^{3x^2+2x}$$

If we let $u(x) = 4e^x$ and $v(x) = 3x^2 + 2x$, then we can rewrite $f(x) = u(v(x))$. As before, we have:

$$\begin{aligned} f(x) &= 4e^{3x^2+2x} = 4e^{v(x)} && \rightarrow \\ f'(x) &= 4e^{v(x)} \cdot v'(x) \\ &= 4e^{v(x)} \cdot (3 \cdot 2x + 2) \\ &= 4e^{3x^2+2x} \cdot (6x + 2) \end{aligned}$$

Example 4 Sometimes you may have to apply the Chain Rule more than one time if the inner function is itself a composition of other functions. Consider the function:

$$f(x) = \sqrt{2e^{3 \sin x}}$$

Start by letting $u(x) = \sqrt{x} = x^{\frac{1}{2}}$ and $v(x) = 2e^{3\sin x}$, then we can rewrite $f(x) = u(v(x))$. As before, we have:

$$f(x) = \sqrt{2e^{3\sin x}} = (2e^{3\sin x})^{\frac{1}{2}} \rightarrow$$

$$f'(x) = \frac{1}{2} (v(x))^{-\frac{1}{2}} \cdot v'(x)$$

At this point, when we need to find $v'(x)$, we realize that the $v(x)$ is really a composite of a simpler function $w(x) = 2e^x$ and $z(x) = 3\sin x$. So we can rewrite $v(x) = w(z(x))$ and find $v'(x)$:

$$v(x) = 2e^{3\sin x} = 2e^{z(x)} \rightarrow$$

$$v'(x) = 2e^{z(x)} \cdot z'(x)$$

$$= 2e^{z(x)} \cdot 3\cos x$$

$$= 2e^{3\sin x} \cdot 3\cos x = 6e^{3\sin x} \cos x$$

Substituting this answer for $v'(x)$ back into our partial result from before, we end up with the final answer:

$$f(x) = \sqrt{2e^{3\sin x}} = (2e^{3\sin x})^{\frac{1}{2}} \rightarrow$$

$$f'(x) = \frac{1}{2} (v(x))^{-\frac{1}{2}} \cdot v'(x)$$

$$= \frac{1}{2} (v(x))^{-\frac{1}{2}} \cdot 6e^{3\sin x} \cos x$$

$$= \frac{1}{2\sqrt{2e^{3\sin x}}} \cdot 6e^{3\sin x} \cos x$$

$$= \frac{3e^{3\sin x} \cos x}{\sqrt{2e^{3\sin x}}}$$

7.11 “Product Rule”: Differentiating Product of Functions

You use the product rule to compute the derivative of a function $f(x)$ that is the product of two other functions.

$$f(x) = u(x) \cdot v(x) \rightarrow f'(x) = u(x) \cdot v'(x) + v(x) \cdot u'(x)$$

$$= u(x) \cdot \frac{dv}{dx} + v(x) \cdot \frac{du}{dx}$$

In other words, the derivative of the product of two functions is “the first times the derivative of the second, plus the second times the derivative of the first”.

Given a function $f(x) = 5x^2$, if rather than recognizing this as a constant 5 times the function x^2 , suppose you accidentally employ the product rule by letting $u(x) = 5$ and $v(x) = x^2$. Rest easy, because even if you do this you won't get the wrong answer. Things still work out in the end because...

$$u(x) = 5 \qquad v(x) = x^2$$

$$\frac{du}{dx} = 0 \qquad \frac{dv}{dx} = 2x$$

Then when you combine these elements using the “recipe” given by the product rule, you get:

$$\begin{aligned} f'(x) &= u(x) \cdot \frac{dv}{dx} + v(x) \cdot \frac{du}{dx} \\ &= 5 \cdot 2x + x^2 \cdot 0 \\ &= 10x \end{aligned}$$

This is the same answer you would have gotten had you realized that the 5 was just a constant that went “along for the ride” as you took the derivative of x^2 . Therefore, the general rule is that you apply the product rule only when the two separate functions that you can identify are being multiplied each contains a variable x , rather than just being a constant.

When calculating the derivative of the first function or of the second function, if it is a composed function then you will also need to employ the chain rule as illustrated by the following examples.

Example 1 Consider the function:

$$f(x) = \ln(3x) e^{-0.2x^2}$$

Start by identifying the two functions being multiplied together, then calculate their derivatives on their own before combining them using the “recipe” indicated by the product rule:

$$\begin{aligned} u(x) &= \ln(3x) & v(x) &= e^{-0.2x^2} \\ \frac{du}{dx} &= \frac{1}{3x} \cdot 3 = \frac{1}{x} & \frac{dv}{dx} &= e^{-0.2x^2} \cdot (-0.2 \cdot 2 \cdot x) = -0.4xe^{-0.2x^2} \end{aligned}$$

Notice above that while computing the derivatives $\frac{du}{dx}$ and $\frac{dv}{dx}$ that we used the Chain Rule. Finally assemble the answer by substituting the partial answers above into the formula for the product rule:

$$\begin{aligned} f'(x) &= u(x) \cdot \frac{dv}{dx} + v(x) \cdot \frac{du}{dx} \\ &= \ln(3x) \cdot (-0.4xe^{-0.2x^2}) + e^{-0.2x^2} \frac{1}{x} \\ &= -0.4x \ln(3x) e^{-0.2x^2} + \frac{e^{-0.2x^2}}{x} \\ &= \frac{e^{-0.2x^2} (1 - 0.4x^2 \ln(3x))}{x} \end{aligned}$$

Example 2 Consider the function:

$$f(x) = \sin(2x) e^{3x}$$

Start by identifying the two functions being multiplied together, then calculate their derivatives on their own before combining them using the “recipe” indicated by the product rule:

$$\begin{aligned} u(x) &= \sin(2x) & v(x) &= e^{3x} \\ \frac{du}{dx} &= \cos(2x) \cdot 2 = 2 \cos(2x) & \frac{dv}{dx} &= e^{3x} \cdot 3 = 3e^{3x} \end{aligned}$$

Notice above that while computing the derivatives $\frac{du}{dx}$ and $\frac{dv}{dx}$ that we used the Chain Rule. Finally assemble the answer by substituting the partial answers above into the formula for the

product rule:

$$\begin{aligned} f'(x) &= u(x) \cdot \frac{dv}{dx} + v(x) \cdot \frac{du}{dx} \\ &= \sin(2x) \cdot 3e^{3x} + e^{3x} \cdot 2 \cos(2x) \\ &= e^{3x} (3 \sin(2x) + 2 \cos(2x)) \end{aligned}$$

Example 3 Consider the function:

$$f(x) = (8x^3 - 6x)^4 (2x + 1)^3$$

Start by identifying the two functions being multiplied together, then calculate their derivatives on their own before combining them using the “recipe” indicated by the product rule:

$$\begin{aligned} u(x) &= (8x^3 - 6x)^4 & v(x) &= (2x + 1)^3 \\ \frac{du}{dx} &= 4 \cdot (8x^3 - 6x)^3 \cdot (8 \cdot 3x^2 - 6) & \frac{dv}{dx} &= 3 \cdot (2x + 1)^2 \cdot (2) \\ &= 4 \cdot (8x^3 - 6x)^3 \cdot (24x^2 - 6) & &= 6(2x + 1)^2 \\ &= 4 \cdot (8x^3 - 6x)^3 \cdot 6 \cdot (4x^2 - 1) \\ &= 24(8x^3 - 6x)^3 (4x^2 - 1) \end{aligned}$$

Notice above that while computing the derivatives $\frac{du}{dx}$ and $\frac{dv}{dx}$ that we used the Chain Rule. Finally assemble the answer by substituting the partial answers above into the formula for the product rule:

$$\begin{aligned} f'(x) &= u(x) \cdot \frac{dv}{dx} + v(x) \cdot \frac{du}{dx} \\ &= (8x^3 - 6x)^4 \cdot 6(2x + 1)^2 + (2x + 1)^3 \cdot 24(8x^3 - 6x)^3 (4x^2 - 1) \\ &= (8x^3 - 6x)^3 (2x + 1)^2 (6(8x^3 - 6x) + 24(2x + 1)(4x^2 - 1)) \\ &= 6(8x^3 - 6x)^3 (2x + 1)^2 ((8x^3 - 6x) + 4(2x + 1)(4x^2 - 1)) \\ &= 6(8x^3 - 6x)^3 (2x + 1)^2 ((8x^3 - 6x) + 32x^3 - 8x + 16x^2 - 4) \\ &= 6(8x^3 - 6x)^3 (2x + 1)^2 (40x^3 + 16x^2 - 14x - 4) \\ &= 12(8x^3 - 6x)^3 (2x + 1)^2 (20x^3 + 8x^2 - 7x - 2) \end{aligned}$$

Example 4 Consider the function:

$$f(x) = e^{3x^2} \sin\left(\frac{1}{x^2}\right)$$

Start by identifying the two functions being multiplied together, then calculate their derivatives on their own before combining them using the “recipe” indicated by the product rule:

$$\begin{aligned} u(x) &= e^{3x^2} & v(x) &= \sin\left(\frac{1}{x^2}\right) \\ \frac{du}{dx} &= e^{3x^2} \cdot (6x) & \frac{dv}{dx} &= \cos\left(\frac{1}{x^2}\right) \cdot \left(\frac{-2}{x^3}\right) \end{aligned}$$

Notice above that while computing the derivatives $\frac{du}{dx}$ and $\frac{dv}{dx}$ that we used the Chain Rule. Finally assemble the answer by substituting the partial answers above into the formula for the

product rule:

$$\begin{aligned} f'(x) &= u(x) \cdot \frac{dv}{dx} + v(x) \cdot \frac{du}{dx} \\ &= e^{3x^2} \cdot \cos\left(\frac{1}{x^2}\right) \cdot \left(\frac{-2}{x^3}\right) + \sin\left(\frac{1}{x^2}\right) \cdot e^{3x^2} \cdot (6x) \\ &= e^{3x^2} \left(\frac{-2}{x^3} \cos\left(\frac{1}{x^2}\right) + 6x \sin\left(\frac{1}{x^2}\right) \right) \end{aligned}$$

7.12 “Quotient Rule”: Differentiating Quotient of Functions

You use the quotient rule to compute the derivative of a function $f(x)$ that is the quotient of two other functions.

$$\begin{aligned} f(x) = \frac{u(x)}{v(x)} &\quad \rightarrow \quad f'(x) = \frac{v(x) \cdot u'(x) - u(x) \cdot v'(x)}{(v(x))^2} \\ &= \frac{v(x) \cdot \frac{du}{dx} - u(x) \cdot \frac{dv}{dx}}{(v(x))^2} \end{aligned}$$

When I was in high school, we weren’t allowed to use a formula booklet that reminded us of the important, essential formulas. A math teacher in high school taught me the following silly way to remember the quotient rule that harkens back to the Disney movie *Snow White and the Seven Dwarves* and its song “Hi Ho, Hi Ho, It’s off work we go”. If you think of the quotient as having a “Hi” part (the high part, the numerator) and a lower part called “Ho”, then you can remember the quotient rule using the phrase “Ho-dee-hi minus Hi-dee-ho over Ho Ho”:

$$f(x) = \frac{hi}{ho} \quad \rightarrow \quad f'(x) = \frac{ho \, d(hi) - hi \, d(ho)}{ho \cdot ho}$$

Given $f(x) = \frac{5}{x^2}$, if rather than recognizing this as a constant 5 times the function $\frac{1}{x^2}$, suppose that you accidentally employ the quotient rule by letting $u(x) = 5$ and $v(x) = \frac{1}{x^2}$.

Rest easy, because even if you do this you won’t get the wrong answer. Things still work out in the end because...

$$\begin{aligned} u(x) &= 5 & v(x) &= x^2 \\ \frac{du}{dx} &= 0 & \frac{dv}{dx} &= 2x \end{aligned}$$

Then when you combine these elements using the “recipe” given by the quotient rule, you get:

$$\begin{aligned} f'(x) &= \frac{v(x) \cdot \frac{du}{dx} - u(x) \cdot \frac{dv}{dx}}{(v(x))^2} \\ &= \frac{x^2 \cdot 0 - 5 \cdot 2x}{(x^2)^2} \\ &= \frac{-10x}{x^4} = \frac{-10}{x^3} \end{aligned}$$

Which is the same answer you would have gotten had you realized that the 5 was just a constant that went “along for the ride” as you took the derivative of $5 \cdot \frac{1}{x^2}$.

Therefore, the general rule is that you apply the quotient rule only when the two separate

functions that you can identify are being divided each contains a variable x , rather than just being a constant.

When calculating the derivative of the first function or of the second function, if it is a composed function then you will also need to employ the chain rule as illustrated by the following examples.

Example 1 Consider the function:

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

Start by identifying the functions in the numerator and denominator, then calculate their derivatives on their own before combining them using the “recipe” indicated by the quotient rule:

$$\begin{aligned} u(x) &= e^{2x} & v(x) &= 2x \\ \frac{du}{dx} &= e^{2x} \cdot 2 = 2e^{2x} & \frac{dv}{dx} &= 2 \end{aligned}$$

Notice above that while computing one the derivatives $\frac{du}{dx}$ and $\frac{dv}{dx}$ that we used the Chain Rule. Finally assemble the answer by substituting the partial answers above into the formula for the product rule:

$$\begin{aligned} f'(x) &= \frac{v(x) \cdot \frac{du}{dx} - u(x) \cdot \frac{dv}{dx}}{(v(x))^2} \\ &= \frac{2x \cdot 2e^{2x} - e^{2x} \cdot 2}{(2x)^2} \\ &= \frac{2e^{2x}(2x - 1)}{4x^2} \\ &= \frac{e^{2x}(2x - 1)}{2x^2} \end{aligned}$$

Example 2 Consider the function:

$$f(x) = \frac{2x^3 + 1}{x^2 + 1}$$

Start by identifying the functions in the numerator and denominator, then calculate their derivatives on their own before combining them using the “recipe” indicated by the quotient rule:

$$\begin{aligned} u(x) &= 2x^3 + 1 & v(x) &= x^2 + 1 \\ \frac{du}{dx} &= 2 \cdot 3x^2 = 6x^2 & \frac{dv}{dx} &= 2x \end{aligned}$$

Notice above that while computing one the derivatives $\frac{du}{dx}$ and $\frac{dv}{dx}$ that we used the Chain Rule. Finally assemble the answer by substituting the partial answers above into the formula for the product rule:

$$\begin{aligned} f'(x) &= \frac{v(x) \cdot \frac{du}{dx} - u(x) \cdot \frac{dv}{dx}}{(v(x))^2} \\ &= \frac{(x^2 + 1) \cdot 6x^2 - (2x^3 + 1) \cdot 2x}{(x^2 + 1)^2} \\ &= \frac{6x^2(x^2 + 1) - 2x(2x^3 + 1)}{(x^2 + 1)^2} \end{aligned}$$

Example 3 Consider the function:

$$f(x) = \frac{e^{3x}}{\ln(x^2 - x)}$$

Start by identifying the functions in the numerator and denominator, then calculate their derivatives on their own before combining them using the “recipe” indicated by the quotient rule:

$$\begin{aligned} u(x) &= e^{3x} & v(x) &= \ln(x^2 - x) \\ \frac{du}{dx} &= e^{3x} \cdot 3 = 3e^{3x} & \frac{dv}{dx} &= \frac{1}{(x^2 - x)} \cdot (2x - 1) = \frac{2x - 1}{x^2 - x} \end{aligned}$$

Notice above that while computing the derivatives $\frac{du}{dx}$ and $\frac{dv}{dx}$ that we used the Chain Rule. Finally assemble the answer by substituting the partial answers above into the formula for the quotient rule:

$$\begin{aligned} f'(x) &= \frac{v(x) \cdot \frac{du}{dx} - u(x) \cdot \frac{dv}{dx}}{(v(x))^2} \\ &= \frac{\ln(x^2 - x) \cdot 3e^{3x} - e^{3x} \cdot \frac{2x-1}{x^2-x}}{(\ln(x^2 - x))^2} \\ &= \frac{e^{3x} (3(x^2 - x) \ln(x^2 - x) - (2x - 1))}{(\ln(x^2 - x))^2 (x^2 - x)} \end{aligned}$$

7.13 Using the First Derivative to Find Local Maxima and Minima

The first derivative of a function f provides information on its instantaneous rate of change at any point x as measured by the slope of the line tangent to the curve $f(x)$ at any point $(x, f(x))$. The function f is increasing wherever $f'(x)$ has a positive value and is decreasing wherever $f'(x)$ has a negative value. The function f has a critical point at $x = p$ if $f'(p) = 0$. You can use the sign of the derivative just to the left or right of $x = p$ to determine whether the point is a minimum, a maximum, or neither of these. If the sign of the first derivative changes from positive to negative at $x = p$, then the point $(p, f(p))$ is a maximum. If the sign of the first derivative changes from negative to positive at $x = p$, then the point $(p, f(p))$ is a minimum. If the sign does not change, then the point is neither a minimum nor a maximum. As mentioned in the previous section, the second derivative can also be used to test whether a critical point is a minimum or maximum.

If the point $(p, f(p))$ is a “local minimum” then its y -coordinate $f(p)$ is the smallest value of the function f when compared to those function values for points in the vicinity of $x = p$. It may not necessarily be the absolute minimum value of the function. Analogously, if the point $(p, f(p))$ is a “local maximum” then its y -coordinate $f(p)$ is the largest value of the function f when compared to those function values for points in the vicinity of $x = p$. It may not necessarily be the absolute maximum value of the function. When the domain of a function is limited by the problem, remember to also check the function values at the domain interval’s endpoints for possible values of local or absolute maxima or minima.

Your calculator can help you find local maxima and minima. If you have graphed the function f , see section Section 2.14.4, “Finding a Maximum Value in an Interval” for how to analyze the graph to maximum points and section Section 2.14.5, “Finding a Minimum Value Value in an Interval” for how to find minimum points.

For additional information, see Section 7.1.3, “Relationships Between Function and Derivative”.

7.14 Analyzing Zeros of Derivative Graph to Find Maxima and Minima

Consider the function $f(x) = x^2$ and the point $(2, 1)$. Let the point (x, x^2) be an arbitrary point on the curve $f(x)$, find the minimum distance between the point and the curve.

Start by drawing a figure like the one in Figure 125 to visualize the distance between the given point and the curve, and to derive a formula for the distance in terms of x . The distance between an arbitrary point (x, x^2) on the curve $f(x)$ and the point $(2, 1)$ is the length of the hypotenuse of a right triangle where the horizontal leg has length $x - 2$ and the vertical leg has length $x^2 - 1$. These "leg" lengths are just the difference between the x -coordinates of the points (x, x^2) and $(2, 1)$ as well as the difference between their y -coordinates.

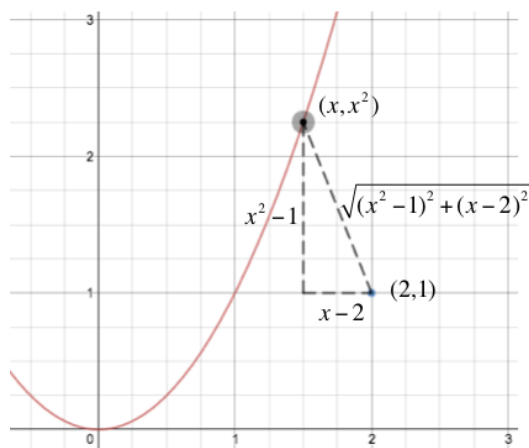


Figure 125: Sketching the Distance Between Point and Curve

Using the Pythagorean Theorem, we know that the length of the hypotenuse must be the square root of the sum of the squares of the leg distances, so this gives the expression for the distance between the point and the curve, for any value of x . We can call this distance-between-point-and-curve function $D(x)$:

$$D(x) = \sqrt{(x^2 - 1)^2 + (x - 2)^2}$$

We know that the minimum value of $D(x)$ must occur where its first derivative $\frac{d}{dx}D(x)$ equals zero. If we can use our calculator, the job is almost done. Figure 126 shows using the technique described in Section 7.4, "Graphing Derivatives with Your Calculator" to graph $\frac{d}{dx}D(x)$.

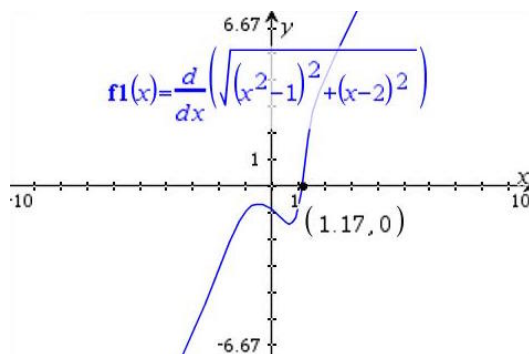


Figure 126: Finding Zeros of Derivative of $D(x)$ to Find Minimum of $D(x)$

Once we've graphed the derivative, we can use the "Analyze Graph" → "Zeros" feature of the calculator to find the zero shown in the figure at $x = 1.17$. We know the point is a minimum

because we can see in the graph that the derivative value is below the x -axis before this zero and above the x -axis after this zero. That means that the gradient values of the function $D(x)$ are changing at $x = 1.17$ from negative to positive.

If the problem had asked “At what value of x is the distance between the curve $f(x)$ and the point $(2, 1)$ at its minimum?” then $x = 1.17$ would have been the answer. However, pay attention that the problem has asked for the minimum value of the distance. Therefore, we need to compute the value of the distance function $D(x)$ at this point:

$$D(1.17) = \sqrt{(1.17^2 - 1)^2 + (1.17 - 2)^2} = 0.908$$

7.15 Using the Second Derivative to Determine Concavity and Points of Inflection

The second derivative of a function f provides information on its concavity. You compute a function’s second derivative by computing its first derivative f' , then repeating the process to compute the derivative of f' . The function f is concave down wherever $f''(x)$ has a negative value and is concave up wherever $f''(x)$ has a positive value. When evaluating a critical point at $x = p$ where the first derivative $f'(p) = 0$, you can evaluate the second derivative $f''(p)$ to determine whether the point $(p, f(p))$ is a minimum or a maximum. If $f''(p) < 0$ then f is concave down there, so $(p, f(p))$ must be a maximum. If $f''(p) > 0$ then f is concave up there, so $(p, f(p))$ must be a minimum.

A point of inflection of the function f is a point $(p, f(p))$ where the function f changes concavity. You can determine the points of inflection by first finding where $f''(x) = 0$ and then for each of these points finding the sign of values of $f''(x)$ for values of x just to the left and just to the right of the candidate point of inflection. If the sign of $f''(x)$ changes from positive to negative around the point $x = p$, or changes from negative to positive there, then the point $(p, f(p))$ is a point of inflection. If the sign remains positive or remains negative across a candidate point of inflection, then it does not qualify. Another even easier way to find a point of inflection is to graph the second derivative function, and analyze the graph to find its zeros. You can visually see where the values of the function are positive (above the x -axis) and negative (below the x -axis) to help you determine if the sign of the second derivative is changing at the points around where the second derivative equals zero (that is, crosses the x -axis). See Section 7.4, “Graphing Derivatives with Your Calculator” for more information.

Your calculator can compute the second derivative at a particular point $x = p$ for you. Just follow the steps outlined in Section 7.5, “Computing Derivatives at a Point with Your Calculator” but instead choose “2nd Derivative” instead of “1st Derivative”.

For additional information, see Section 7.1.3, “Relationships Between Function and Derivative”.

7.16 Overview of Concepts for Integral Calculus

Newton and Leibniz used calculus to study areas under curves. They did this by studying how to approximate the area of an arbitrary shape by dividing the area into smaller and smaller rectangles. The techniques they invented became known as Integral Calculus, the study of areas under curves. In the process, they proved an amazing link that they named the “Fundamental Theorem of Calculus” between the Differential Calculus described in the previous section and the area under a curve covered in this section.

7.16.1 Area Under Curve

To estimate the area under an arbitrary curve $f(x)$ between points $x = a$ and $x = b$, Newton and Leibniz started by dividing the interval (a, b) into a number of equal segments. At each segment boundary they drew a vertical segment up to intersect the curve $f(x)$ and used this as the height of a rectangle. Figure 127 shows the result when using a small number of segments. By adding together the areas of all the rectangles, they achieved a rough approximation of the area under the curve. As you can see in the figure, sometimes the rectangle undershoots the actual area, and in other places overshoots the real area.

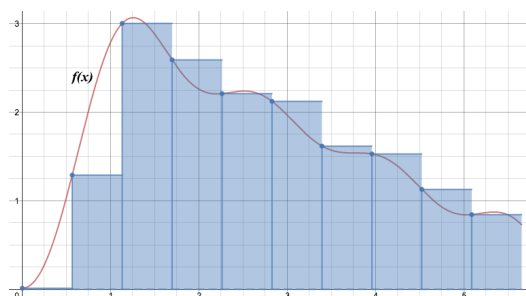


Figure 127: Estimating the Area Under the Curve Using Rectangles

Figure 128 shows how their approximation improved by dividing the interval into a larger number of equal segments, each having a smaller width.

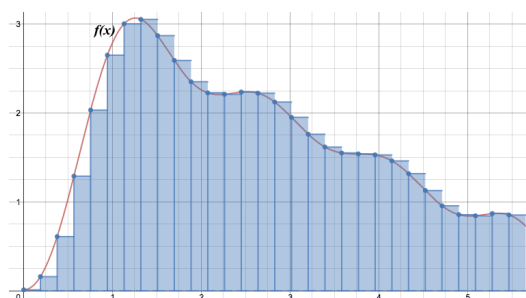


Figure 128: Estimating the Area Under the Curve Using Smaller Rectangles

Figure 129 shows how their approximation improved even further by allowing the number of segments to increase, thus shrinking even further the width of the equal intervals. They saw that adding together this larger number of ever-narrower rectangles would eventually converge on the actual area under the curve if they allowed the width of the rectangles to approach zero.

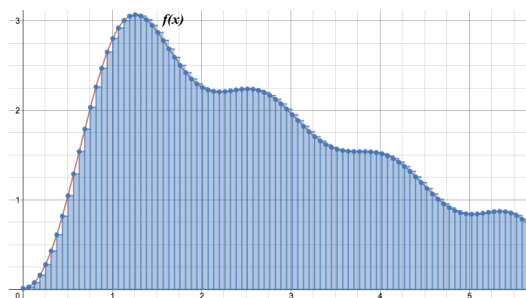


Figure 129: Estimating the Area Under the Curve Using Even Smaller Rectangles

The process of integration formalizes the calculation of the exact area under a curve. The notation Newton and Leibniz invented looks like a stylized letter “S” for sum, since the calculation

of the area involves summing up the areas of the infinitely many, infinitesimally-thin rectangles whose heights are given by the value of the function $f(x)$ being integrated. The area under the curve $f(x)$ from $x = a$ to $x = b$ is written using the notation below. As the value x ranges from $x = a$ to $x = b$, the height of the rectangle at any given point is given by $f(x)$ and the infinitesimally-thin width of the rectangle is represented by the dx , where the d stands for the “delta” or very small change in the value of x .

$$\int_{x=a}^{x=b} (\text{rectangle height}) \times (\text{rectangle width}) = \int_a^b f(x) dx$$

7.16.2 Relationship Between Derivation and Integration

While studying how the graph of an object’s “distance versus time” related to its graph of “speed versus time”, Newton and Liebniz proved a surprising connection between the two. The area under the speed graph between two moments in time t_{start} and t_{end} was equal to the distance traveled during that interval. Figure 130 shows this effect for a three-hour car journey at a constant speed of 154 km/h. The graph on the left is a distance versus time graph. It shows that we began the journey with our odometer showing 28734 kilometers. After three hours the odometer measuring our distance traveled showed 29196. Subtracting our starting distance from our ending distance yields a total distance traveled of $29196 - 28734 = 462$ kilometers. The graph on the right is a speed versus time graph. It is a straight, flat line because our speed – the rate of change in our distance per hour – was constant at 154 km/h for three hours. If we multiply the number of hours traveled by the speed, we get $154 \text{ km/h} \times 3 \text{ hours} = 462 \text{ km}$. The calculation is none other than the (in this case, rectangular) area under the graph of the speed for the duration of time we traveled.

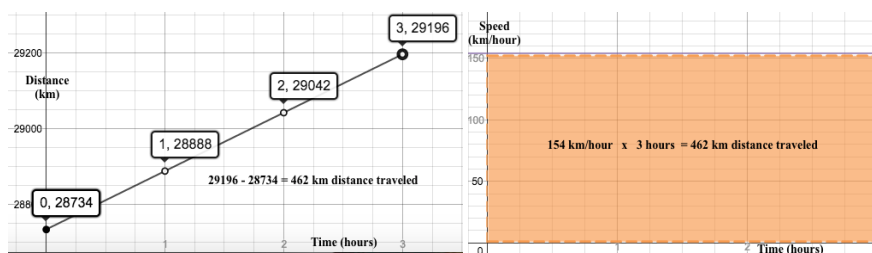


Figure 130: Area Under the “Speed versus Time” Graph Gives Distance Traveled

Since the speed is the rate of change in distance, and the rate of change function is the derivative, we can state this relationship in another way. Newton and Liebniz proved that the area under the graph of the derivative was equal to the difference in the original function when evaluated at the start and end of the interval they were observing. They named this important result the “Fundamental Theorem of Calculus”. Figure 131 shows a more general representation of this essential link between differential calculus (the derivative) and integral calculus (the area under a curve). The shaded area between the graph of the derivative function $f'(x)$ and the x -axis over the interval (a, b) is exactly equal to the (vertical) difference between $f(b) - f(a)$. This is true for any continuous function $f(x)$.

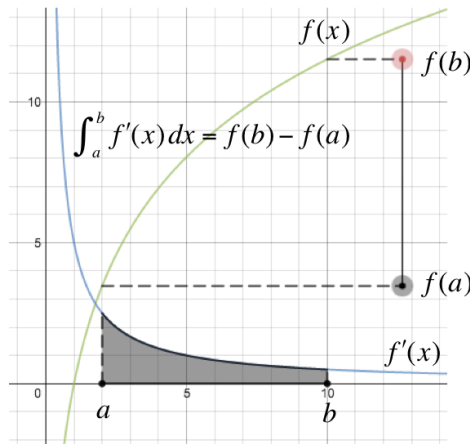


Figure 131: Area Under Derivative Curve Equals the Change in the Original Function

Figure 132 shows an alternative way to think about this important result. If we are given a function $f(x)$ and want to calculate the area under $f(x)$ from $x = a$ to $x = b$, all we have to do is find a new function $F(x)$ such that $F'(x)$ (its derivative) equals our original function $f(x)$. The Fundamental Theorem of Calculus tells us that $\int_a^b F'(x) dx = F(b) - F(a)$, that the area under the curve of $F'(x)$ from $x = a$ to $x = b$ is exactly equal to the difference between the values of the original function at $x = b$ and $x = a$. However, since we found the function $F(x)$ precisely so that $F'(x)$ is exactly our original function $f(x)$, we have $\int_a^b f(x) dx = F(b) - F(a)$. We call this function $F(x)$ the anti-derivative of $f(x)$.

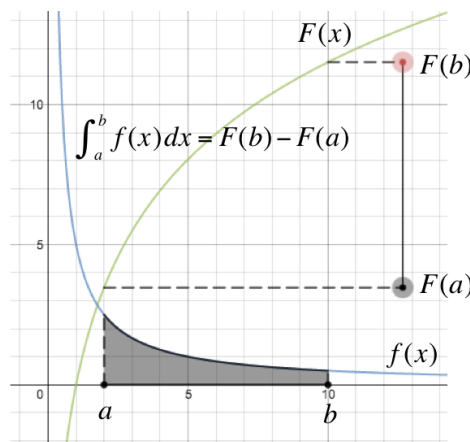


Figure 132: Area Under Curve Equals the Change in its Antiderivative Function

7.17 Anti-Differentiation to Compute Integrals

The process of finding a function $F(x)$ whose derivative is a given function $f(x)$ that you start with is called “anti-differentiation”. It’s like running the rules of differentiation in reverse. Given any function $f(x)$, if we can find its “anti-derivative”, then we can find the area under the curve of $f(x)$ between $x = a$ and $x = b$ simply by computing $F(b) - F(a)$ using this anti-derivative function. Therefore, finding the antiderivative is intimately related to computing an integral to find the area under a curve.

For example, consider the function $f(x) = x^2$. In order to compute the area under this curve from $x = 5$ to $x = 10$, we need to find a function $F(x)$ such that $F'(x) = x^2$, in other words,

a function $F(x)$ whose derivative equals our original $f(x)$. If we let $F(x) = \frac{x^3}{3}$ we can take the derivative $F'(x)$ and see that it comes out to $3 \cdot \frac{x^2}{3} = x^2$. So choosing $F(x) = \frac{x^3}{3}$ fits our purposes to be able to compute the area under the original $f(x)$ from $x = 5$ to $x = 10$:

$$\int_5^{10} x^2 dx = \left[F(x) \right]_5^{10} = F(10) - F(5) = \frac{10^3}{3} - \frac{5^3}{3} = \frac{1000 - 125}{3} = \frac{875}{3} \approx 291.67$$

7.17.1 Indefinite Integration as Anti-Differentiation

An integral with specific boundary conditions from $x = a$ to $x = b$ represents the area under the curve $f(x)$ from $x = a$ to $x = b$.

$$\int_a^b f(x) dx$$

However, if you omit the boundary conditions, then this represents the anti-derivative:

$$\int f(x) dx = F(x)$$

When computing the anti-derivative, there will always be infinitely-many possible answers. For example, consider again the function $f(x) = x^2$. What if we chose the function $F(x) = \frac{x^3}{3} + 1$ for the antiderivative? If we compute its derivative we get $F'(x) = x^2$ which is the original $f(x)$ function we started with. So, above we used $F(x) = \frac{x^3}{3}$ and here we see that another function $F(x) = \frac{x^3}{3} + 1$ also satisfies the conditions to be the antiderivative of $f(x)$. In fact, so does $F(x) = \frac{x^3}{3} + 2$, since it also has x^2 as its derivative. It should be clear that we could add any constant C to the function $\frac{x^3}{3}$ and when performing the derivative, that constant becomes zero and disappears.

This is why when writing the anti-derivative, we always write a solution that includes a “+ C ” at the end, which we call the “constant of integration”. For our function $f(x) = x^2$, we’ll write its anti-derivative as:

$$\int x^2 dx = \frac{x^3}{3} + C$$

Where that C represents any of infinitely many possible constant values that could be used. On any given problem, as we’ll see in section Section 7.22, “Using Additional Information to Determine Constant of Integration”, if some additional information is provided, then we can deduce the exact constant value for a particular set of problem conditions.

7.18 Rules for Computing the Indefinite Integral (Antiderivative)

Anti-derivation is running the rules of derivatives in reverse. Therefore, it makes sense that for a given function $f(x)$, you can follow a rule to find the new function $F(x)$ which has $f(x)$ as its derivative. When computing an integral involving $f(x)$, the first step is finding the antiderivative function $F(x)$ using these rules. Then, if boundary conditions are given, you evaluate the antiderivative function at the two boundaries of integration to compute their difference to find the area under the original curve. If no boundary conditions are provided, then you stop after determining the anti-derivative function, after making sure you added the “constant of integration” (C).

When boundary conditions are provided, we call that a *definite integral*. We *definitely* know the interval over which to compute the area. In contrast, when no boundary conditions are provided, it is an *indefinite integral*. We do not know the specific interval over which to compute the area, but in any event we have found an interesting piece of information: the anti-derivative function $F(x)$ determined can be used to find the area over any subsequent interval (a, b) we might be given by computing $F(b) - F(a)$.

7.18.1 Indefinite Integral of x^n ($n \in \mathbb{Q}$)

To compute the indefinite integral (anti-derivative) of x raised to any rational number power, follow this rule:

$$f(x) = x^n \quad \rightarrow \quad F(x) = \frac{x^{n+1}}{n+1} + C$$

You change the power on x to one greater than its current value, then divide by the value of this new exponent. The following examples help clarify how this works in different situations:

Positive Whole Number Power Example

$$f(x) = x^7 \quad \rightarrow \quad F(x) = \frac{x^8}{8} + C$$

Negative Whole Number Power Example

$$f(x) = \frac{1}{x^7} = x^{-7} \quad \rightarrow \quad F(x) = \frac{x^{-7+1}}{-7+1} = \frac{x^{-6}}{-6} = -\frac{1}{6x^6} + C$$

Positive Fractional Power Examples

$$f(x) = \sqrt{x} = x^{\frac{1}{2}} \quad \rightarrow \quad F(x) = \frac{x^{\frac{1}{2}+1}}{\frac{1}{2}+1} = \frac{x^{\frac{3}{2}}}{\frac{3}{2}} = \frac{2}{3}x^{\frac{3}{2}} = \frac{2}{3}\sqrt{x^3} + C$$

$$f(x) = \sqrt{x^3} = x^{\frac{3}{2}} \quad \rightarrow \quad F(x) = \frac{x^{\frac{3}{2}+1}}{\frac{3}{2}+1} = \frac{x^{\frac{5}{2}}}{\frac{5}{2}} = \frac{2}{5}x^{\frac{5}{2}} = \frac{2}{5}\sqrt{x^5} + C$$

Negative Fractional Power Examples

$$f(x) = \frac{1}{\sqrt{x}} = x^{-\frac{1}{2}} \quad \rightarrow \quad F(x) = \frac{x^{-\frac{1}{2}+1}}{-\frac{1}{2}+1} = \frac{x^{\frac{1}{2}}}{\frac{1}{2}} = 2\sqrt{x} + C$$

$$f(x) = \frac{1}{\sqrt{x^3}} = x^{-\frac{3}{2}} \quad \rightarrow \quad F(x) = \frac{x^{-\frac{3}{2}+1}}{-\frac{3}{2}+1} = \frac{x^{-\frac{1}{2}}}{-\frac{1}{2}} = -\frac{2}{\sqrt{x}} + C$$

7.18.2 Indefinite Integral of $\sin x$

To compute the indefinite integral (anti-derivative) of $\sin x$, follow this rule:

$$f(x) = \sin(x) \quad \rightarrow \quad F(x) = -\cos x + C$$

7.18.3 Indefinite Integral of $\cos x$

To compute the indefinite integral (anti-derivative) of $\cos x$, follow this rule:

$$f(x) = \cos(x) \quad \rightarrow \quad F(x) = \sin x + C$$

7.18.4 Indefinite Integral of $\frac{1}{x}$

To compute the indefinite integral (anti-derivative) of $\frac{1}{x}$, follow this rule:

$$f(x) = \frac{1}{x} \quad \rightarrow \quad F(x) = \ln x + C$$

7.18.5 Indefinite Integral of e^x

To compute the indefinite integral (anti-derivative) of e^x , follow this rule:

$$f(x) = e^x \quad \rightarrow \quad F(x) = e^x + C$$

7.19 Integrating Constant Multiple of a Function

The integral of a constant k times a function $f(x)$ is equal to k times the integral:

$$\int k \cdot f(x) dx = k \cdot \int f(x) dx$$

In simple terms, this rule means that you can “move a constant outside the integral” without changing the answer. This should make intuitive sense, because it’s like “factoring out” a common factor k from the height of each rectangle the integral is adding together.

Remember that integrating is adding up the areas of infinitely many, infinitesimally-thin rectangles whose height is given by the function $f(x)$. Imagine taking only five of these rectangles as an approximation with heights given by the function $f(x)$ at five points x_1, x_2, x_3, x_4 , and x_5 . The sum of the areas of these five rectangles would be the expression below, adding together the width \times height calculation of all five.

$$\int kf(x) dx \approx kf(x_1) dx + kf(x_2) dx + kf(x_3) dx + kf(x_4) dx + kf(x_5) dx$$

However, we see there is a constant k in every term that we can factor out, so we can write:

$$\begin{aligned} \int kf(x) dx &\approx kf(x_1) dx + kf(x_2) dx + kf(x_3) dx + kf(x_4) dx + kf(x_5) dx \\ &\approx k \left(f(x_1) dx + f(x_2) dx + f(x_3) dx + f(x_4) dx + f(x_5) dx \right) \end{aligned}$$

The remaining expression inside the parenthesis is the sum of the five rectangles without the constant k factor. That is just the integral of $f(x)$:

$$\begin{aligned} \int kf(x) dx &\approx k \left(f(x_1) dx + f(x_2) dx + f(x_3) dx + f(x_4) dx + f(x_5) dx \right) \\ &\approx k \int f(x) dx \end{aligned}$$

So, we can see that this would also make sense if there were 1000, or a million, or infinitely many rectangles. If the height of each rectangle has a common factor k , we can factor that outside the sum represented by the integral and multiply the factor k later after taking the sum of heights of the rectangles without that multiplier.

7.20 Integrating Sums and Differences of Functions

The integral of a sum or difference of functions is the sum or difference of the integrals:

$$\begin{aligned}\int (f(x) + g(x)) dx &= \int f(x) dx + \int g(x) dx \\ \int (f(x) - g(x)) dx &= \int f(x) dx - \int g(x) dx\end{aligned}$$

In simple terms, this rule means that when you compute the integral of functions added together, you can do so “term by term” and add the results together.

Remember again that integrating is adding up the areas of infinitely many, infinitesimally-thin rectangles whose height is given by the function $f(x)$. Imagine taking only three of these rectangles as an approximation with heights given by the function $f(x)$ at three points x_1 , x_2 , x_3 . The sum of the areas of these three rectangles would be the expression below, adding together the width \times height calculation of all three.

$$\int (f(x) + g(x)) dx \approx (f(x_1) + g(x_1)) dx + (f(x_1) + g(x_2)) dx + (f(x_1) + g(x_3)) dx$$

However, if we multiply the dx representing the infinitesimally-thin rectangle width we see we get:

$$\begin{aligned}\int (f(x) + g(x)) dx &\approx (f(x_1) + g(x_1)) dx + (f(x_1) + g(x_2)) dx + (f(x_1) + g(x_3)) dx \\ &\approx f(x_1) dx + g(x_1) dx + f(x_1) dx + g(x_2) dx + f(x_1) dx + g(x_3) dx\end{aligned}$$

Then we can regroup the terms with f and the terms with g , then factor out the common dx to get:

$$\begin{aligned}\int (f(x) + g(x)) dx &\approx (f(x_1) + g(x_1)) dx + (f(x_1) + g(x_2)) dx + (f(x_1) + g(x_3)) dx \\ &\approx f(x_1) dx + g(x_1) dx + f(x_1) dx + g(x_2) dx + f(x_1) dx + g(x_3) dx \\ &\approx \left(f(x_1) + f(x_2) + f(x_3) \right) dx + \left(g(x_1) + g(x_2) + g(x_3) \right) dx\end{aligned}$$

The expression inside the first parenthesis is the sum of the heights of the three rectangles given by the $f(x)$ function. The expression inside of the second parenthesis is the sum of the heights of the three rectangles given by the $g(x)$ function. But these are just the separate integrals over f and g like this:

$$\int (f(x) + g(x)) dx = \int f(x) dx + \int g(x) dx$$

We can see that this would also make sense if there were 1000, or a million, or infinitely many rectangles. If the height of each rectangle is composed of the sum of the functions f and g , then we can compute the sum of the areas of the rectangles whose heights are given by f and the sum of the areas of the rectangles whose heights are given by g and add those together. The same thing works if they were subtracted, of course.

7.21 Integration Using Substitution

As with the rules of derivatives, the rules for computing the anti-derivative work for functions having a “simple” form:

$$x^n, e^x, \frac{1}{x}, \cos x, \sin x$$

In practice, the integrals you'll need to compute will look a bit more involved, with some other quantity instead of just a simple x . For example:

$$(2x + 1)^n, e^{3x-1}, \frac{1}{4x+2}, \cos(5x-3), \sin(6x+4)$$

This does not present a problem, however. We will handle this by using substitution to make the integrals look “simple” again. Using this approach, we can then use the anti-differentiation rules given in Section 7.18, “Rules for Computing the Indefinite Integral (Antiderivative)”.

7.21.1 Simple Example Involving Linear Factor

Consider the integral:

$$\int \cos(5x-3) dx$$

We don't have a rule for computing the anti-derivative of $\cos(5x-3)$, only one for $\cos u$, so start by making the substitution:

$$u = 5x - 3$$

If we change the variable in the integral from x to u , we also have to substitute an appropriate “infinitesimal width” du to match the new variable u . The way we do this is to take the derivative of our new variable u to see what it “requires” to be present before we make an appropriate substitution for the existing dx . So, we compute the derivative of u :

$$\begin{aligned} u &= 5x - 3 \\ \frac{du}{dx} &= 5 \\ du &= 5 dx \end{aligned}$$

We see that in order to substitute a du , we need to have $5 dx$ to match the $u = 5x - 3$ substitution we have done. That's not a problem. We can multiply the original equation by one without changing its meaning, and in the process introduce the value 5 that we are missing to accompany the existing dx .

$$\int \cos(5x-3) dx = \int \cos(5x-3) \cdot \frac{1}{5} \cdot 5 dx$$

The $\frac{1}{5}$ is a constant, so we can move it outside the integral as explained in Section 7.19, “Integrating Constant Multiple of a Function”. Then, with all the right “ingredients” in place for the substitution, we can carry out our substitution to end up with the simpler expression in terms of u and du .

$$\int \cos(5x-3) dx = \int \cos(5x-3) \cdot \frac{1}{5} \cdot 5 dx = \frac{1}{5} \int \cos\left(\underbrace{5x-3}_u\right) \cdot \underbrace{5 dx}_{du} = \frac{1}{5} \int \cos(u) du$$

The rule for antidifferentiation of $\cos u$ tells us that it becomes $\sin u + C$, so we have:

$$\int \cos(5x-3) dx = \frac{1}{5} \int \cos(u) du = \frac{1}{5} \sin(u) + C$$

Then, finally, we replace the substitution variable u with the expression it represents to end up with an answer back in terms of x 's we started with:

$$\int \cos(5x - 3) dx = \frac{1}{5} \sin(u) + C = \frac{1}{5} \sin(5x - 3) + C$$

As a final step, we can check our work by taking the derivative of $\frac{1}{5} \sin(5x - 3) + C$ to make sure that we get back to the original function we integrated. Doing the derivative we get:

$$\frac{d}{dx} \left(\frac{1}{5} \sin(5x - 3) + C \right) = \frac{1}{5} \cdot \cos(5x - 3) \cdot 5 = \cos(5x - 3)$$

So we are sure the answer is correct.

7.21.2 Example Involving Additional Factor of x

Consider the integral:

$$\int \frac{3x}{6x^2 + 1} dx$$

When you have multiple expressions involving x , one expression will be your u substitution, and any remaining factors will be involved in the set of “ingredients” needed for your corresponding du substitution. For example, here we let:

$$\begin{aligned} u &= 6x^2 + 1 \\ \frac{du}{dx} &= 12x \\ du &= 12x dx \end{aligned}$$

So here, to make the substitution for du we need $12x dx$, however we only have a factor of $3x dx$ in the numerator. To get the 12 we need, we can add a factor of 4 inside the integral, and $\frac{1}{4}$ outside the integral.

$$\begin{aligned} \int \frac{3x}{6x^2 + 1} dx &= \frac{1}{4} \int \frac{1}{6x^2 + 1} \cdot 4 \cdot 3x dx \\ &= \frac{1}{4} \int \underbrace{\frac{1}{6x^2 + 1}}_u \cdot \underbrace{12x dx}_{du} \\ &= \frac{1}{4} \int \frac{1}{u} du \end{aligned}$$

Our rule for anti-differentiating $\frac{1}{u}$ says it turns into $\ln u + C$, so we have:

$$\int \frac{3x}{6x^2 + 1} dx = \frac{1}{4} \int \frac{1}{u} du = \frac{1}{4} \ln(u) + C$$

Then, finally, we replace the substitution variable u with the expression it represents to end up with an answer back in terms of x 's we started with:

$$\int \frac{3x}{6x^2 + 1} dx = \frac{1}{4} \ln(u) + C = \frac{1}{4} \ln(6x^2 + 1) + C$$

As a final step, we can check our work by taking the derivative of $\frac{1}{4} \ln(6x^2 + 1) + C$ to make sure that we get back to the original function we integrated. Doing the derivative we get:

$$\frac{d}{dx} \left(\frac{1}{4} \ln(6x^2 + 1) + C \right) = \frac{1}{4} \cdot \frac{1}{6x^2 + 1} \cdot 12x = \frac{3x}{6x^2 + 1}$$

So we are sure the answer is correct.

7.21.3 Example Involving Other Additional Factors

Consider the integral:

$$\int (\sin 3x) (\cos 3x) dx$$

If we let $u = \sin 3x$ then $du = 3 \cos 3x$. We can introduce the missing factor of 3 inside the integral, and the reciprocal $\frac{1}{3}$ outside the integral, we then end up with a simpler integral to perform:

$$\frac{1}{3} \int \underbrace{(\sin 3x)}_u \underbrace{3(\cos 3x)}_{du} dx = \frac{1}{3} \int u du = \frac{1}{3} \frac{u^2}{2} + C = \frac{1}{3} (\sin 3x)^2 + C = \frac{1}{3} \sin^2 3x + C$$

In this particular example, there was a second route we could have taken that would have also worked. If instead we let $u = \cos 3x$ then $du = -3 \sin 3x$. We can introduce the missing factor of -3 inside the integral, and the reciprocal $-\frac{1}{3}$ outside the integral, we then end up with a simpler integral to perform:

$$-\frac{1}{3} \int (\cos 3x) \underbrace{(-3)(\sin 3x)}_{du} dx = -\frac{1}{3} \int u du = -\frac{1}{3} \frac{u^2}{2} + C = -\frac{1}{3} (\cos 3x)^2 + C = -\frac{1}{3} \cos^2 3x + C$$

You might ask, why did taking two alternative approaches produce different answers? Since we know the identity that $\sin^2 x + \cos^2 x = 1$, we can rearrange that to get $\cos^2 x = 1 - \sin^2 x$ and use that to substitute $1 - \sin^2 3x$ for $\cos^2 3x$ above:

$$-\frac{1}{3} \cos^2 3x + C = -\frac{1}{3} (1 - \sin^2 3x) = \frac{1}{3} \sin^2 3x \underbrace{-\frac{1}{3} + C}_{\text{both constants}} = \frac{1}{3} \sin^2 3x + D$$

So the answers are the same, except for the constant of integration, which doesn't matter when computing the definite integral.

7.21.4 Summary of Strategy for Integration by Substitution

When computing integrals in IB Maths Standard Level, you will only be asked to integrate one of the five indefinite integrals in the formula booklet:

$$u^n, e^u, \frac{1}{u}, \cos(u), \sin(u)$$

Your challenge is to identify the correct substitution to perform in order to use one of the five anti-derivative rules corresponding to these functions. After performing the u -substitution, take the derivative du to find what other “ingredients” are necessary to be present in order to substitute the simpler du expression. If any other factors containing a variable x are necessary to be present to constitute the correct du expression, then those additional factors **must already be present in the problem**. At most, you can introduce a constant inside the integral and outside the integral.

7.22 Using Additional Information to Determine Constant of Integration

If you are given additional information, you can determine a specific value for C instead of leaving a generic constant of integration. For example, consider the problem:

$$s(t) = \int 6e^{2t} + t dt$$

If we are given the additional piece of information that $s(0) = 10$, we can determine the constant of integration. Start by doing the anti-derivative to find the indefinite integral. We can do the integral term by term, and factor out the constant 6 to get:

$$\begin{aligned} s(t) &= \int 6e^{2t} + t dt \\ &= 6 \int e^{2t} dt + \int t dt \end{aligned}$$

The second term does not require any substitution because it is already of the simple form x^n , so we can already follow the rules to compute the antiderivative of that term:

$$\begin{aligned} s(t) &= 6 \int e^{2t} dt + \int t dt \\ &= 6 \int e^{2t} dt + \frac{t^2}{2} + C \end{aligned}$$

For the remaining integral, substitute:

$$\begin{aligned} u &= 2t \\ du &= 2 dt \end{aligned}$$

We can introduce a factor of 2 inside the integral and a factor of $\frac{1}{2}$ outside the integral to get the missing “ingredient” we need to substitute the du , then do the integration following the anti-derivative rules for e^u , and finally substitute back in the expression for u involving the original variable t :

$$\begin{aligned} s(t) &= 6 \cdot \frac{1}{2} \int e^{2t} 2 dt + \frac{t^2}{2} + C \\ &= 3 \int e^u du + \frac{t^2}{2} + C \\ &= 3e^u + \frac{t^2}{2} + C \\ &= 3e^{2t} + \frac{t^2}{2} + C \end{aligned}$$

So we have found that $s(t)$ with a constant of integration C . Now we can use the additional piece of information provided that $s(0) = 10$ to find out the exact value C must have for this particular problem at hand:

$$\begin{aligned} s(0) = 10 &= 3e^{2 \cdot 0} + \frac{0^2}{2} + C \\ 10 &= 3 \cdot 1 + 0 + C \\ 10 &= 3 + C \\ 7 &= C \end{aligned}$$

So the exact answer to the problem is:

$$s(t) = 3e^{2t} + \frac{t^2}{2} + 7$$

7.23 Computing the Definite Integral

This section works the same example both analytically and using a calculator to show you the steps involved.

7.23.1 Computing the Definite Integral Analytically

To compute a definite integral $\int_a^b f(x) dx$ analytically, follow these steps:

1. Find the antiderivative $F(x)$ using substitution if necessary
2. Write the antiderivative inside of square brackets with the boundaries $\left[F(x) \right]_a^b$
3. Calculate $(F(b)) - (F(a))$ being careful about minus signs

For example, consider the integral $\int_0^{0.5} \left(x + \frac{\sin 3x}{(\cos 3x)^2} \right) dx$

Since the integrand has two terms, we can split up the integral to do it term by term:

$$\int_0^{0.5} x dx + \int_0^{0.5} \frac{\sin 3x}{(\cos 3x)^2} dx$$

The antiderivative of the first term is easy, so we can do that term first:

$$\begin{aligned} & \left[\frac{x^2}{2} \right]_0^{0.5} + \int_0^{0.5} \frac{\sin 3x}{(\cos 3x)^2} dx \\ & \left[\left(\frac{0.5^2}{2} \right) - \left(\frac{0^2}{2} \right) \right] + \int_0^{0.5} \frac{\sin 3x}{(\cos 3x)^2} dx \\ & 0.125 + \int_0^{0.5} \frac{\sin 3x}{(\cos 3x)^2} dx \end{aligned}$$

For integrating the second term, we substitute:

$$\begin{aligned} u &= \cos 3x \\ du &= -\sin(3x) \cdot 3 = -3 \sin 3x dx \end{aligned}$$

We introduce a factor of -3 inside the integral and $-\frac{1}{3}$ outside the integral to get the required $-3 \sin 3x dx$, then substitute in our u and du and do the antiderivative. We temporarily write the bounds with $x = a$ and $x = b$ to remind ourselves they were for the original variable x :

$$\begin{aligned} 0.125 + \left(-\frac{1}{3} \right) \int_{x=0}^{x=0.5} \underbrace{\frac{1}{(\cos 3x)^2}}_u \cdot \underbrace{(-3) \sin 3x dx}_{du} \\ 0.125 - \frac{1}{3} \int_{x=0}^{x=0.5} \frac{1}{u^2} du \\ 0.125 - \frac{1}{3} \int_{x=0}^{x=0.5} u^{-2} du \\ 0.125 - \frac{1}{3} \left[\frac{u^{-2+1}}{-2+1} \right]_{x=0}^{x=0.5} \\ 0.125 - \frac{1}{3} \left[\frac{u^{-1}}{-1} \right]_{x=0}^{x=0.5} \\ 0.125 + \frac{1}{3} \left[\frac{1}{u} \right]_{x=0}^{x=0.5} \end{aligned}$$

Now we can substitute back in the expression representing our u , drop the temporary reminder “ $x =$ ” notation from the bounds since the expression is back in terms of the original x variable, and finish the job:

$$\begin{aligned}
 & 0.125 + \frac{1}{3} \left[\frac{1}{u} \right]_{x=0}^{x=0.5} \\
 & 0.125 + \frac{1}{3} \left[\frac{1}{\cos 3x} \right]_0^{0.5} \\
 & 0.125 + \frac{1}{3} \left[\frac{1}{\cos(3 \cdot 0.5)} - \frac{1}{\cos(3 \cdot 0)} \right] \\
 & 0.125 + \frac{1}{3} \left[\frac{1}{0.070737} - \frac{1}{1} \right] \\
 & 0.125 + \frac{1}{3} [14.136833 - 1] \\
 & 0.125 + \frac{1}{3} [13.136833] \\
 & \approx 4.503944
 \end{aligned}$$

7.23.2 Computing the Definite Integral Using Technology

On the TI-Nspire CX, with a calculator tab open in your current document, you can choose the menu option *Calculus* \rightarrow *Numerical Integral* to enter the problem from the previous section. Figure 133 shows the result.

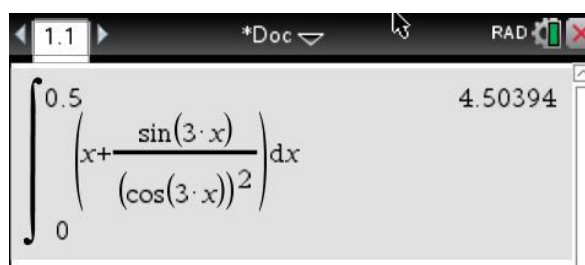


Figure 133: Computing a Definite Integral on the TI-Nspire CX

7.23.3 Computing Normal Distribution Probability Using a Definite Integral

We learned in Section 6.20, “Graphing the Normal Distribution and Computing Bounded Area” that for a normally distributed data set with mean μ and standard deviation σ , the probability that a random variable X will fall in a range $a \leq x \leq b$, written $P(a \leq X \leq b)$, is the area under the curve $\text{normPdf}(x, \mu, \sigma)$ between a and b . We also know that the definite integral of any function $f(x)$ between a and b represents the area under the curve $f(x)$ between a and b . Putting these two ideas together, we can write:

$$\begin{aligned}
 P(a \leq X \leq b) &= \text{Area under } \text{normPdf}(x, \mu, \sigma) \text{ between } a \text{ and } b \\
 &= \int_a^b \text{normPdf}(x, \mu, \sigma) \\
 &= \text{normCdf}(a, b, \mu, \sigma)
 \end{aligned}$$

Using this new knowledge, we can recompute the probability calculation from Section 6.20, “Graphing the Normal Distribution and Computing Bounded Area” using the following integral:

$$P(154.1 \leq X \leq 163.9) = \int_{154.1}^{163.9} \text{normPdf}(x, 159, 4.9) = \text{normCdf}(154.1, 163.9, 159, 4.9)$$

As shown in Figure 134, we see that the definite integral produces the same 0.683 result (to three significant figures) as the *Analyze Graph* → *Bounded Area* computation. It also confirms that the `normCdf()` function we learned about in Section 6.22, “Using the Calculator for Normal Distribution Problems” produces the identical result.

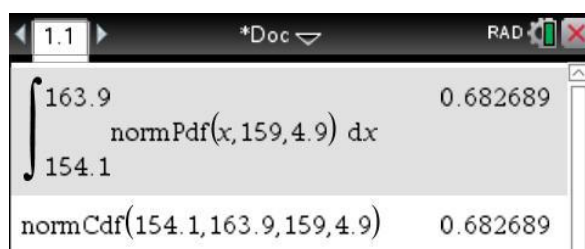


Figure 134: Computing Probability Using Integral of `normPDF(x)`

7.24 Displacement s , Velocity v , and Acceleration a

The displacement function $s(t)$ measures the distance a body has traveled away from its starting point. The velocity function $v(t)$ measures the rate of change in the distance, so $v(t) = s'(t)$. The acceleration function $a(t)$ measures the rate of change in the velocity (how quickly the body is speeding up or slowing down), so $a(t) = v'(t) = s''(t)$. This leads to these relationships illustrated in Table 17.

Concept	B	Rate of Change	Area Under
Displacement	$s(t)$		$\int v(t)$
Velocity	$v(t)$	$s'(t)$	$\int a(t)$
Acceleration	$a(t)$	$v'(t)$	

Table 17: Relationship Between Displacement, Velocity, and Acceleration

7.25 Determining Position Function from Acceleration

Newton and Leibniz used integration to study how acceleration due to gravity affected the path of an object in flight. Suppose you throw a ball upward at a height of 1.5 meters with initial velocity of 6 m/s. We can use the relationships between acceleration, velocity, and displacement mentioned in the previous section to determine the velocity the ball will land with.

Recalling that acceleration due to gravity is a constant -9.8 m/s^2 , so as a function of time, $a(t) = -9.8$. We integrate the acceleration function $a(t)$ to get the velocity function $v(t)$:

$$v(t) = \int a(t) dt = \int -9.8 dt = -9.8t + C$$

Use the initial velocity $v(0) = 6 \text{ m/s}$ to determine the value of the constant of integration C :

$$\begin{aligned} v(0) &= -9.8 \cdot 0 + C = 6 \\ C &= 6 \end{aligned}$$

Therefore, the velocity function is $v(t) = -9.8t + 6$.

Next, integrate the velocity function to get the position function $s(t)$:

$$s(t) = \int v(t) dt = \int -9.8t + 6 dt = \frac{-9.8t^2}{2} + 6t + C = -4.9t^2 + 6t + C$$

Use the initial position $s(0) = 1.5$ to determine the value of the constant of integration:

$$\begin{aligned} s(0) &= -4.9 \cdot 0^2 + 6 \cdot 0 + C = 1.5 \\ C &= 1.5 \end{aligned}$$

Therefore, the position function is $s(t) = -4.9t^2 + 6t + 1.5$.

Figure 135 shows the graph of the position function from time $t = 0$ until the position above ground (i.e. the height) is zero. Using a graphing calculator (or the quadratic equation), we find that the position function $s(t) = -4.9t^2 + 6t + 1.5 = 0$ when $t = 1.437$ seconds.

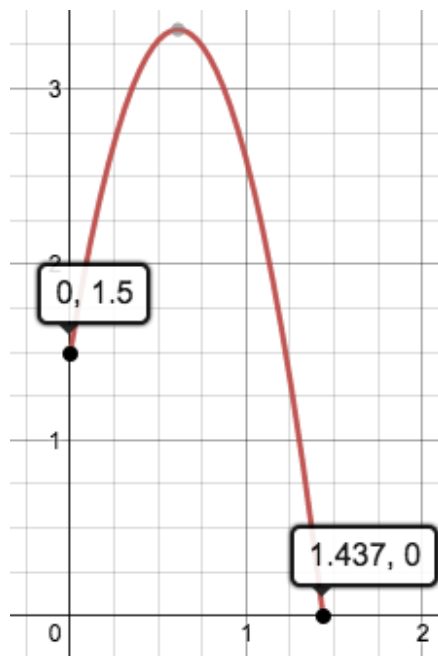


Figure 135: Graph of $s(t) = -4.9t^2 + 6t + 1.5$

The ball hits the ground (i.e. the height given by the position function is zero) when $t = 1.437$ seconds, so we can calculate the velocity of the ball by plugging $t = 1.437$ into the velocity function:

$$v(1.437) = -9.8 \cdot 1.437 + 6 = -8.0826 \approx -8.1 \text{ m/s}$$

7.26 Total Area Under a Curve

If you need to compute the total area under a function $f(x)$ over an interval where the function is both above and below the x -axis, you can use two approaches:

- Integrate the absolute value of the function $\int_a^b |f(x)| dx$
- Identify the zeros, use them to compute separate integrals for $f(x)$ over each interval separately, then add together the parts, taking the absolute value of the areas that are negative.

7.27 Area Between Curves

To find the area between two curves $f(x)$ and $g(x)$ between $x = a$ and $x = b$ compute the definite integral of the absolute value of the difference between the functions:

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

In practice, the procedure you follow to compute the answer depends on whether you are able to use your calculator.

7.27.1 Area Between Curves Using Calculator

Suppose you're given two curves $g(x) = \cos(x^2)$ and $f(x) = e^x$ as shown in Figure 136 and you are asked to find the area between the curves from -1.4 to 0.3 . Using your calculator's "Analyze Graph" feature, you can identify that over the interval in question, the two curves intersect at $x = -1.112$ and again at $x = 0$. On the intervals $(-1.4, -1.112)$ and $(0, 0.3)$ you see that $f(x) > g(x)$, while on the interval $(-1.112, 0)$ you have $g(x) > f(x)$.

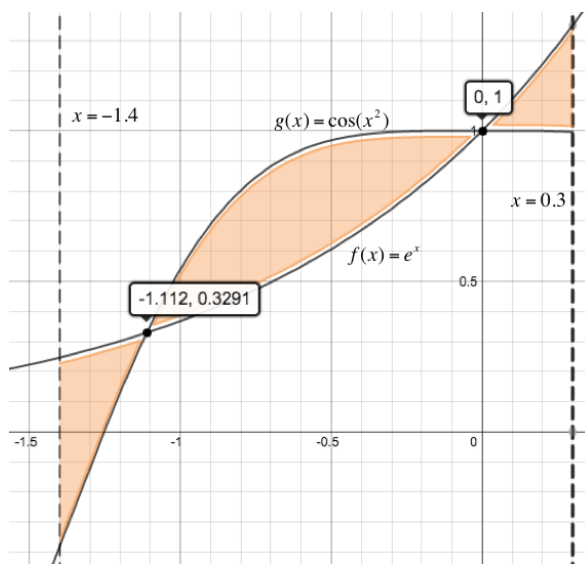


Figure 136: Computing the Area Between Two Curves

If you are using your calculator, as shown in Figure 137 you can simply bang in the absolute value of the difference of the functions and compute the answer without worrying about the points of intersection or the intervals where one function might flip-flop to being larger than the other.

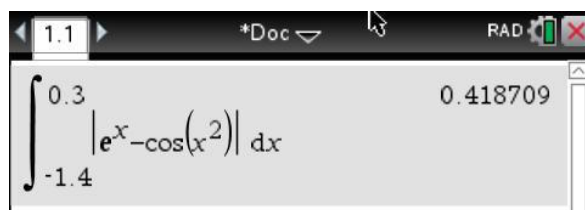


Figure 137: Integrating Absolute Value of Difference of Functions on TI-Nspire CX

7.27.2 Area Between Curves Analytically (Without the Calculator)

If you must calculate the area between curves $f(x)$ and $g(x)$ analytically, then you must pay attention to which function is greater than the other over which intervals. You start by identifying the places where the functions intersect by setting $f(x) = g(x)$ and solving for x . Then you need to perform the area calculation by splitting up the integral into separate parts, one per interval. In each part, choose the difference that will result in a positive quantity by putting the larger function first in the subtraction. For example, using the curves in Figure 136, you would have to compute:

$$\int_{-1.4}^{0.3} |f(x) - g(x)| dx = \int_{-1.4}^{-1.112} \underbrace{f(x) - g(x)}_{f(x) > g(x) \text{ here}} dx + \int_{-1.112}^0 \underbrace{g(x) - f(x)}_{g(x) > f(x) \text{ here}} dx + \int_0^{0.3} \underbrace{f(x) - g(x)}_{f(x) > g(x) \text{ here}} dx$$

Of course, in a non-calculator problem the points of intersection and the function values to compute would be something simpler than the functions and values in the previous section, but the concept of identifying the intersection points and performing the integral as separate parts added together (with the larger function written first in the subtraction) is still the key point to remember here.

7.28 Net Change in Displacement versus Total Distance Traveled

It's important to understand the difference between displacement and distance traveled so you know how to calculate each one when requested.

7.28.1 Difference Between Distance and Displacement

Figure 138 illustrates displacement. It is a vector quantity that tells how a body's current position related to its starting point.

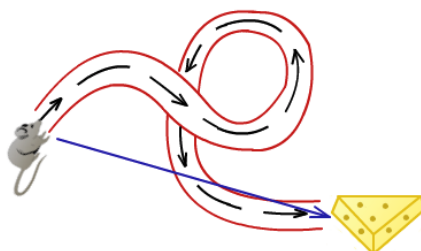


Figure 138: Displacement is Vector Measure from Initial Position to Final Position⁷

Figure 139 illustrates distance. It is a scalar quantity that measures how far a body has to travel along a path to get from its initial position to its current position.

⁷Source: tutorvista.com/physics/difference-between-distance-and-displacement

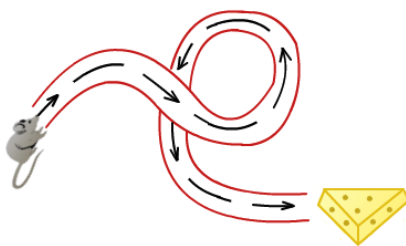


Figure 139: Distance is Scalar Measure of Actual Length of Path Traveled⁸

7.28.2 Computing Total Distance Traveled

The velocity function $v(t)$ is the derivative of the displacement function $s(t)$, giving the rate of change in distance at any time t . Vice versa, the displacement function $s(t)$ is the integral of the velocity function $v(t)$. The Fundamental Theorem of Calculus tells us that the area under the velocity curve equals the change in the displacement function. Since displacement has direction and can be backwards and forwards, when you compute the integral of the velocity function, you end up determining the **net change in displacement**. If some of the area under the velocity function is negative, it “cancels out” the area under the velocity function that might be positive, possibly resulting in a net change in displacement that is zero or maybe even negative.

If you want to calculate the total distance traveled (regardless of direction) you need to compute the area of the absolute value of the velocity function so that any negative values don't cancel out positive values in the distance calculation. Therefore, to compute the total distance traveled from time t_1 to t_2 , calculate the definite integral between t_1 and t_2 of the absolute value of the velocity function, with respect to time.

$$\text{Total Distance Traveled} = \int_{t_1}^{t_2} |v(t)| dt$$

⁸Source: tutorvista.com/physics/difference-between-distance-and-displacement