Entry #1   Inverses in General

The term "inverse" is used in very different contexts in mathematics. For example, the multiplicative inverse of a number, the inverse of a function, and the inverse relationship between differentiation and antidifferentiation are common mathematical topics. Many but not all of these notions can be unified by noticing the operation relative to which the inverse is defined.

Operations

For a collection of objects (such as some number system, or some collection of functions), there might be several different binary operations of interest. For example, for a number system the operations ($*$) are likely to include addition ($a * b = a + b$) or multiplication ($a * b = a \cdot b$). For a collection of functions with a specified domain, taking values in a number system, the operation might be pointwise multiplication ($f * g = f \cdot g$). For a collection of functions with a specified domain, taking values in that domain, the operation could be composition ($f * g = f \circ g$). In this more general context for inverses, there is also an element that serves as the identity relative to that operation (an element 1 for which $a * 1 = a$ and $1 * a = a$ for all a). For example, the additive identity for numbers is 0 ($a + 0 = a$ for all a) and the multiplicative identity is 1 ($a \cdot 1 = a$ for all a). The identity function for pointwise multiplication of functions is the function 1 for which $1(x) = 1$ for all x because $f \cdot 1 = f$ and $1 \cdot f = f$. The identity for composition of functions is the identity function i for which $i(x) = x$ for all x because $f \circ i = f$ and $i \circ f = f$. Note that inverse functions also occur for functions that don't take values in their domain, as is described in Entry #3.

Inverses

In the general context, if the operation is denoted $*$ and the identity is denoted 1, then the inverse of an element a is the element a^{-1} that satisfies $a * a^{-1} = 1$ and $a^{-1} * a = 1$. This then yields the additive inverse for numbers ($^{-}a + a = 0$) and the multiplicative inverse for numbers ($a \cdot \dfrac{1}{a} = 1$). The statement that subtraction is the inverse process to addition or that division is the inverse process to multiplication can be seen to follow from this. When the operation for functions is composition, the inverse is the usual inverse function (for which

$f \circ f^{-1} = i$ and $f^{-1} \circ f = i$) and when the operation is pointwise multiplication, the inverse is the function $x \mapsto \dfrac{1}{f(x)}$.

The table below summarizes these relationships:

Objects	Operation	Identity	Inverse Object
numbers, a	addition, $+$	0	^{-}a
numbers, a	multiplication, $\cdot$	1	$a^{-1} = \dfrac{1}{a}$
functions, f	composition, $\circ$	$i(x) = x$ for all x	f^{-1}
functions, f	pointwise multiplication, $\cdot$	$1(x) = 1$ for all x	$\dfrac{1}{f}$
n by n matrices, M	matrix multiplication	identity matrix	inverse matrix M^{-1}

Notation

Exponential notation is commonly used to denote multiplicative inverses (e.g., $7^{-1} = \dfrac{1}{7}$) as it is a part of the general exponential notation (e.g., $7^{-3} = \left(7^{-1}\right)^{3} = \dfrac{1}{7^{3}}$). In similar fashion, exponential notation is used to express powers of functions (e.g., $(\ln x)^{-3} = \dfrac{1}{(\ln x)^{3}}$). For functions, however, exponential notation is also used to describe inverse relative to composition, as the inverse function to f is commonly denoted f^{-1}. This dual usage of exponential notation leads to two interpretations of f^{-1}, as the inverse function and as the reciprocal, but for functions this is always used to denote the inverse function, not the reciprocal. This confusion is most typically encountered by students when working with $\sin^{-1} x$. In this case, misunderstanding the notation leads students to interpret $\sin^{-1} x$ as $\csc x$. Some authors prefer to use Arcsin x to avoid this confusion.

Entry #2 Inverses of Numbers

The number 0 is the unique additive identity for the set of numbers since $x + 0 = 0 + x = x$ for every number x. Every number x has a unique additive inverse, denoted by $-x$, with the property that

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

The number 1 is the unique multiplicative identity for the set of numbers since $x \times 1 = 1 \times x = x$ for every number x. Every nonzero number x has a unique multiplicative inverse, usually denoted by $1/x$, with the property that

$$x \times (1/x) = (1/x) \times x = 1.$$

Multiplicative and additive inverses of numbers can be connected to the idea of the inverse of a function (see Entry #3) by defining a function that accomplishes the multiplication or addition. For example, for each fixed real number r, define "adding r" as the function $A_r(x) = x + r$.

Similarly, define "subtracting r" as the function $S_r(x) = x - r$. Since $A_r(x)$ and $S_r(x)$ are inverse functions of each other, they provide a function context for the statement that "addition and subtraction are inverse operations." Notice that $A_{-r}(x) = S_r(x)$, which is another way of noticing that subtracting a number r is accomplished by adding $(-r)$. The number q is the additive inverse of r if and only if $A_q(x)$ and $A_r(x)$ are inverse functions. This serves as a natural connection between an "inverse operation" and an "inverse function."

This works for multiplication as well. For each non-zero number r, define "multiplying by r" as the function $M_r(x) = r \cdot x$. Similarly, define "dividing by r" as the function $D_r(x) = x/r$. Since $M_r(x)$ and $D_r(x)$ are inverse functions of each other, they provide a function context for the statement that "multiplication and division are inverse operations." Notice that $M_{1/r}(x) = D_r(x)$, which is another way of noticing that dividing by a nonzero number r is accomplished by multiplying by $(1/r)$. The number q is the multiplicative inverse of r if and only if $M_q(x)$ and

$D_r(x)$ are inverse functions. Again, this serves as a natural connection between an "inverse operation" and an "inverse function."

Entry #3 Inverses of Functions

A function $f : X \to Y$ having domain X and range Y is said to be *one-to-one* if $f(x_1) = f(x_2)$ implies $x_1 = x_2$. The function f is said to be *onto* if for each y in Y there is an x in X such that $f(x) = y$. If $f : X \to Y$ is both one-to-one and onto, then, for every b in the range Y of f, there exists an element a in the domain X of f such that $f(a) = b$. Because f is one-to-one, for each element b that element a is unique. This association of a in X with the given b in Y defines a function $g : Y \to X$, called the *inverse* of f . Thus, the inverse of f is defined to be that function g that reverses f:

$$g(b) = a \ \text{ if and only if } \ f(a) = b$$

for all a in the domain of f and all b in the range of f. More formally, the inverse g can be defined in terms of composition and the identity function,

$$g \circ f(a) = g(f(a)) = g(b) = a$$

for all a in the domain of f, and so $g \circ f = i$. Likewise,

$$f \circ g(b) = f(g(b)) = f(a) = b$$

for all b in the range of f and so $f \circ g = i$. The definition of the inverse function above,

$$g(b) = a \ \text{ if and only if } \ f(a) = b,$$

ensures that the inverse of a function is unique.   It also follows that if g is the inverse of f, then f is the inverse of g. The domain of g is the range of f and the range of g is the domain of f.

If g is the inverse for f, then this is typically written $g = f^{-1}$. With this notation, the statement above:

if g is the inverse of f, then f is the inverse of g

says that $g = f^{-1}$ implies $f = g^{-1}$. This also means that $(f^{-1})^{-1} = f$.

Not every function has an inverse. Specifically, a function that is not one-to-one does not have an inverse.

Horizontal Line Test

Some textbooks describe the *horizontal line test* as providing a necessary and sufficient graphical condition that a function be one-to-one. The statement that a function is one-to-one, which is necessary for it to have an inverse function, can be considered graphically. For a function to be one-to-one means that no horizontal line can intersect the graph of the function more than once. This is the horizontal line test. Although it does provide some visual insight into the idea of one-to-one, students should not conclude that the horizontal line test is a critical tool in determining

whether or not a function has an inverse. After all, the horizontal line test makes sense only in the context of examining a function's graph, and that graph is always meaningful or readily accessible.

Inverse functions undo each other

If we think of a function as a rule that assigns to each number in its domain a unique value, then the inverse can be thought of as the rule that undoes that assignment. For example, the function $f(x) = 3x + 4$ can be thought of as the rule defined by the following sequence of two steps: Multiply by 3 and then add 4. If these steps are undone, we get the following: Subtract 4 and then divide by 3. Undoing the steps works only if the given function can be expressed as the composition of two or more functions and is based on the identity $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$. In words, this identity says that the function $g^{-1} \circ f^{-1}$ undoes the assignment made by the function $(f \circ g)$. To verify this, observe that composition of functions is associative, so that, if $y = (f \circ g)(x)$, then $y = f(g(x))$ and

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

The formula for $g(x) = \dfrac{x-7}{x-9}$ is not written as the composition of two or more functions.

However, it can be rewritten as

$$g(x) = \frac{x-9+2}{x-9} = 1 + \frac{2}{x-9},$$

which is the composition of four functions: $f_1(x) = x - 9$, $f_2(x) = \dfrac{1}{x}$, $f_3(x) = 2x$,

$f_4(x) = 1 + x$, and $g(x) = f_4 \circ f_3 \circ f_2 \circ f_1(x)$. One can then recognize the inverse of g as a composition of inverses: $g^{-1}(x) = f_1^{-1} \circ f_2^{-1} \circ f_3^{-1} \circ f_4^{-1}(x)$, in the reverse order.

The statement that f and g are inverse functions can be expressed in the equations

$$f(g(x)) = x \text{ and } g(f(x)) = x.$$

The first of these can be interpreted as saying that f undoes g and the second as saying that g undoes f.

The idea that inverse functions undo each other can appear early in the development of algebraic skills. Consider the problem:

If I start with a number, multiply it by 3, and then add 4, I get 19.

With what number did I start?

Students solve this in their heads, reasoning that before you added 4 to get 19 you must have had 15, and then that before you multiplied by three to get 15 you must have had 5. Note that the students are using the inverse processes, in the reverse order. That is, they start with 19, subtract 4 to get 15, then divide by 3 to get 5. These steps describe the inverse function. Note that they also describe the steps involved in solving the algebraic equation to solve for x:

$$3x + 4 = 19$$

In this way, the idea of an inverse function appears early in consideration of linear functions, when solving what are called "two-step equations," and indeed in pre-algebraic contexts such as the number problem above.

Graphs of Inverse Functions

If f and g are inverse functions, then the point (b, a) is on the graph of g if and only if the point (a, b) is on the graph of f. So the collection of points that form the graph of g is exactly the collection of points on the graph of f, but with their coordinates interchanged. Interchanging the coordinates of a point is equivalent to interchanging the roles of x and y, which graphically amounts to interchanging the x-axis and y-axis (while retaining direction). This can be accomplished by reflecting about the line $y = x$. In this way, the graph of the inverse function to f is the reflection of the graph of f about the line $y = x$.
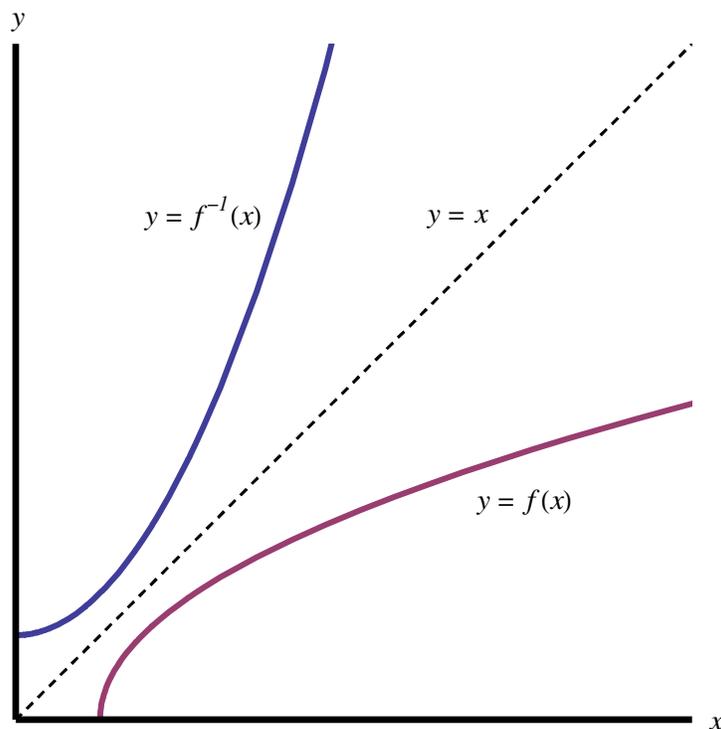
However, this graphical interpretation carries with it some risks of misuse. If the elements of the domain of f and the elements of the range of f are measured in different units, then placing the graphs of both f and its inverse together would render the unit attributes of both axes meaningless.

Finding the inverse function

To find an expression for the inverse function for a one-to-one function f, we seek a function g for which

$$g(b) = a \text{ if and only if } f(a) = b$$

for all a in the domain of f and all b in the range of f. Thus we seek to solve the equation $f(a) = b$ for a. If we start with $f(x) = y$, then we will be solving for x to find $x = g(y)$. Because of the tradition of using x for the variable in functions, once we have the equation $x = g(y)$, it is standard practice to interchange the roles of x and y in this equation to give the general form for the function g: $g(x) = y$. This general process for finding the inverse of a function, solving for x and then interchanging x and y, is often reversed (first interchange x and y and then solve for y) in algebra texts. Note that the units and the possible values of x in the expression $f(x)$ in general will differ from the units and the possible values of x in the expression $g(x)$. Accordingly, it makes no sense to superimpose the graph of the inverse on that of the original function. For

example, if we have population expressed as a function of time, and if we displayed both the function and its inverse on the same set of coordinate axes, then, for example, the horizontal axis would represent both time and population; likewise for the vertical axis.

Inverse of a Point

There are two ways in which the word "inverse" might be used in association with a function f that is one-to-one and onto. First, there is the inverse function g of f, described above, having the property that $g(b) = a$ if and only if $f(a) = b$ for all a in the domain of f and all b in the range of f. Second, for each b in the range of f, the point a in the domain of f for which $f(a) = b$ can be referred to as the inverse image of the point b.

Entry #4   Some Common Inverse Functions

Exponents and Logarithms

For a fixed positive number b, raising b to the power x is a function of x for any real number x:

$$E_b(x) = b^x.$$

This is the exponential function with base b. Its domain is the set of all real numbers and its range is the set of all positive numbers. Also, for a fixed positive number b, taking the logarithm to the base b of a positive real number x is also a function of x:

$$L_b(x) = \log_b(x).$$

This is the logarithmic function with base b. Its domain is the set of all positive numbers and its range is the set of all real numbers. The exponential function with base b and the logarithmic function with base b are inverses of one another. The definition of inverse functions can be viewed to say that $L_b\left(E_b\left(x\right)\right) = x$ for all real numbers x and $E_b\left(L_b\left(x\right)\right) = x$ for all positive numbers x. That is, $\log_b\left(b^x\right) = x$ for all real numbers x and $b^{\log_b(x)} = x$ for all positive numbers x. In the case where $b = e$, $\ln\left(e^x\right) = x$ for all real numbers x and $e^{\ln(x)} = x$ for all positive numbers x.

Powers and Roots

For a fixed odd positive integer n, raising any positive real number x to the power n is a one-to-one function of x: $P_n(x) = x^n$. Similarly, for an odd positive integer n, taking the nth root of a positive real number x is also a one-to-one function of x: $R_n(x) = x^{1/n}$. These functions are inverses of one another. Thus, raising to an odd power and taking the corresponding root are inverse functions.

In the case of raising to a positive even power n, the function $P_n(x) = x^n$ is not one-to-one as a function of all real numbers, and so it does not have an inverse function. For example, the function $f(x) = x^2$ and the function $g(x) = x^{1/2}$ are not true inverse functions because, for example, $g\left(f\left(x\right)\right) = \sqrt{x^2} = |x|$. If these were inverse functions we would have found $g(f(x)) = x$. This situation is explained next.

Inverse of a Function that is Not One-to-One

If a function f is not one-to-one, then there are two distinct values a_1 and a_2 in its domain for which $f(a_1) = f(a_2)$. If the domain of f were restricted so as not to include two such values, then the function would be one-to-one, and an inverse for f could be found. This occurs frequently in mathematics, but the understanding that it is only an inverse when the function is viewed with the restricted domain is a frequent source of confusion. If the inverse function created in this fashion is g, then we have, for a in the restricted domain, $f(a) = b$ if and only if $g(b) = a$. This can be restated $f(g(b)) = b$ and $g(f(a)) = a$, provided a is in the restricted domain.

Even Powers

If the function f is an even power function, such as $f(x) = x^2$, then f is not one-to-one because for any value a, $f(a) = f(-a)$. If we restrict the domain to be just the non-negative real numbers, then the function is one-to-one and we find the inverse function $g(x) = x^{1/2} = \sqrt{x}$. Hence, we have

$f(g(x)) = x$ for all x in the domain of g (i.e., $\left(\sqrt{x}\right)^2 = x$ for $x \geq 0$) and $g(f(x)) = x$ for x in the

restricted domain of f (i.e., $\sqrt{x^2} = |x| = x$ for $x \geq 0$). The same is true for $f(x) = x^n$, where n is any

positive even number. The function $g(x) = x^{1/n}$ will be an inverse function, in the sense that

$f(g(x)) = x$ for all x in the domain of g and $g(f(x)) = x$ for x in the restricted domain of f.

Inverse Trigonometric Functions

The trigonometric functions are periodic, and thus are not one-to-one. For example, the graph of the sine function attains each value in the interval $[-1,1]$ for infinitely many values of x (see figure below).

$$y = \sin x$$
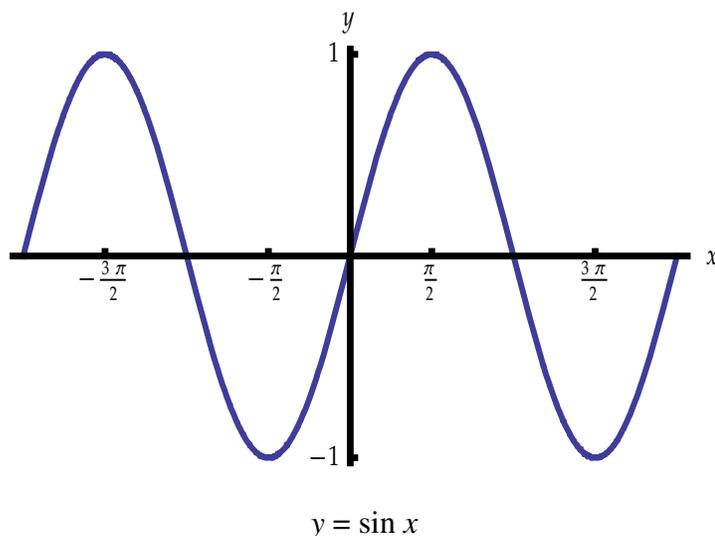
If the domain is restricted appropriately, an inverse can be found. The agreement for $f(x) = \sin x$ is to restrict the domain to $\left[-\dfrac{\pi}{2}, \dfrac{\pi}{2} \right]$, and then the function $g(x) = \text{Arcsin}(x)$ (also denoted

$\sin^{-1}(x)$) becomes an inverse over this restricted domain. The domain of $g(x) = \text{Arcsin}(x)$ is $[-1,1]$ and so $f(g(x)) = x$ (i.e., $\sin(\text{Arcsin}(x)) = x$) for all x in $[-1,1]$. Composing in the other order we

have $g(f(x)) = x$ for in the restricted domain for $\sin(x)$, that is $\text{Arcsin}(\sin(x)) = x$ for x in $\left[-\dfrac{\pi}{2}, \dfrac{\pi}{2} \right]$.

One can, of course, compute $\text{Arcsin}(\sin(x))$ for a value of x outside of this domain, in which case $\text{Arcsin}(\sin(x)) \neq x$. The other inverse trigonometric functions are similarly defined by first restricting their domains. For $\text{Arccos}(x)$ the restricted domain for $f(x)=\cos(x)$ is $[0, \pi]$ and for

$\text{Arctan}(x)$ the restricted domain for $f(x) = \tan(x)$ is $\left(-\dfrac{\pi}{2}, \dfrac{\pi}{2} \right)$.

The quadratic formula and the inverse of each branch of a quadratic function

With its maximal domain of the real line, a quadratic function is not one-to-one and therefore does not have an inverse. By restricting the domain appropriately, a quadratic function is one-to-one and has an inverse. If $a > 0$ and $f(x) = ax^2 + bx + c$ for $x \geq -b/2a$, then f has the inverse

$$g(x) = \frac{-b + \sqrt{4ax + b^2 - 4ac}}{2a} \quad \text{for } x \geq \frac{4ac - b^2}{4a};$$

if $a > 0$ and $f(x) = ax^2 + bx + c$ for $x \leq -b/2a$, then f has the inverse

$$g(x) = \frac{-b - \sqrt{4ax + b^2 - 4ac}}{2a} \quad \text{for } x \leq \frac{4ac - b^2}{4a} \; ;$$

Together, these two formulas used with x equal to zero give us the quadratic formula.

Inverse Relations

A *relation* is a collection of ordered pairs and a *function* is a special kind of relation, namely, one for which $(x_1, y_1) = (x_1, y_2)$ implies that $y_1 = y_2$. The inverse of a relation is defined to be that relation obtained by reversing the coordinates of each ordered pair. In this context, every function has an inverse. However, the inverse may not itself be a function. For example, consider the squaring function $s(x) = x^2$. As a relation, its inverse is the set of all order pairs of the form (x^2, x). This inverse, of course, is not a function since, for example, the inverse contains both the order pair (4, 2) as well as the ordered pair (4, –2).

In trigonometry we sometimes see inverses of the functions introduced without the usual restriction on the domain. For example, arcsin(x) is defined to be any y for which sin(y) = x. So, for example, arcsin(0) would mean any integer multiple of π. This inverse of the sine function is a relation but, of course, is not a function.

Entry #5   Differentiation and Integration

The operations of differentiation and indefinite integration (or *antidifferentiation*) that are studied in calculus can be thought of as inverses of each other in a rather loose sense. Indeed, the very definition of the indefinite integral of a function *f* carries with it the notion of inverse: The indefinite integral (or antiderivative) of *f* is a function whose derivative is *f*. Although viewing these two operations as inverses of each other can help in understanding what indefinite integration is, there is a difficulty regarding uniqueness. It is a fundamental result in calculus that each integrable function has infinitely many antiderivatives. Specifically, if *G* is an antiderivative of *f*, then, for every constant *C*, the function *H* defined by $H(x) = G(x) + C$ also is an antiderivative of *f*. This observation is based on two properties established in calculus: (1) The derivative of the sum of two differentiable functions is the sum of the individual derivatives and (2) the derivative of a constant function is zero.

So, if we view indefinite integration as the inverse of differentiation, what we mean is that the operator ϕ that associates with a function its antiderivative has as its domain the set of all integrable functions and has as its range the set of all equivalence classes of differentiable functions. Thus, for a given *f*, $\phi(f) = \int f$ is not a single function, but, rather, a family of functions. Two differentiable functions are equivalent if they differ by a constant. The inverse of ϕ associates with each such equivalence class the unique function that is the derivative of every function in that equivalence class.

Another view of the operator ϕ that associates with a function its antiderivative is that of observing that although there is no inverse operator for ϕ in the strict "inverse function" sense, there is an inverse for ϕ that is a relation. The inverse of ϕ associates with each function *f* in its domain and of the functions *G* for which the derivative of *G* is *f*.