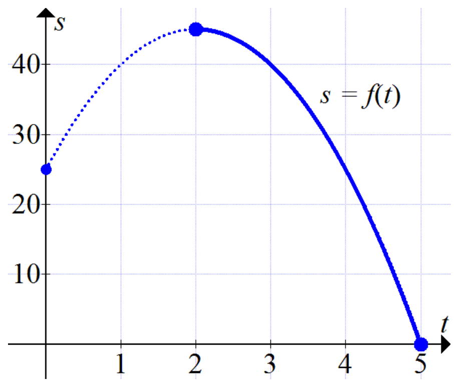
Inverse Functions

What is a function? Like other mathematical concepts, the idea of a *function* has four levels of meaning: physical (word problems), geometric (graphs), numerical (spreadsheets), and algebraic (formulas). We illustrate these with the following example.

- *Physical.* We throw a stone upward off a 25 m building, with an initial vertical speed of 20 m/sec; and we assume gravitational acceleration of 10 m/sec^2 downward. Let s be the height of the stone in meters, t sec after the throw. Thus, the function $s = f(t)$ is defined by looking to see the height at a given time. We limit our experiment to $t \in [0, 5]$.

- *Geometric.* The stone's height is plotted approximately by:



- *Numerical.* We collect the following observations for height each second:

t	0	1	2	3	4	5
$s = f(t)$	25	40	45	40	25	0

  This is only a rough model: we could take a smaller increment for the time inputs, or indeed imagine an infinite table to capture all inputs and outputs of $f(t)$.

- *Algebraic:* $s = f(t) = 25 + 20t - 5t^2$. This agrees with the physical version because the initial height is $f(0) = 25$; the initial velocity is $f'(0) = [20 - 10t]_{t=0} = 20$; and the constant acceleration is $f''(t) = -10$, negative meaning downward. This formula is only valid for $t \in [0, 5]$, since the stone stops at the ground, and like all physical models, it is slightly thrown off by subtle factors such as air resistance.

What is an inverse function? Continuing our example, we examine the inverse of the function $s = f(t)$ above. This means *reversing the roles of input and output* (the independent and dependent variables), so that time t becomes a function of height s: in symbols $t = f^{-1}(s)$. That is, f^{-1} is the rule that tells us at what time t the stone reaches a given height s.

Here we encounter a problem: since the stone rises and then falls, it reaches a given height s at *two* different times: for example, $f(1) = f(3) = 40$. Thus $f^{-1}(40) = 1$ and/or 3, with no single output, so f^{-1} is not a function. This happens because the original function f is not *one-to-one*: instead of taking different inputs to different outputs, it takes two different inputs $t = 1, 3$ to the same output $s = 40$. In graphical terms,

Notes by Peter Magyar `magyar@math.msu.edu`

$s = f(t)$ fails the *horizontal line test*, meaning a horizontal line like $s = 40$ intersects the graph more than once.

To fix this problem, we must restrict the domain of our function: we will look at the stone only for $t \in [2, 5]$ (solid part of the graph). Technically, this defines a new function:

$$f : [2, 5] \to [0, 45],$$

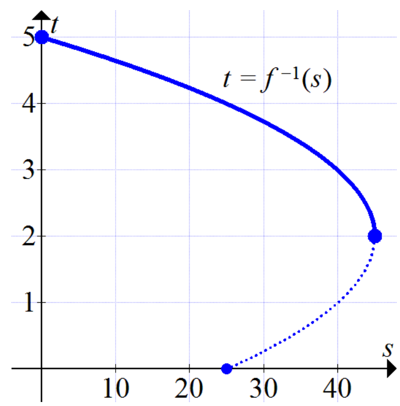
meaning the only allowed inputs are $t \in [2, 5]$, which produce outputs covering the interval $s \in [0, 45]$. This restricted function is one-to-one, satisfying the horizontal line test, so we can get an inverse function:

$$f^{-1} : [0, 45] \to [2, 5],$$

which again has four meanings:

- *Physical.* The inverse function $t = f^{-1}(s)$ gives the unique time $t \in [2, 5]$ for which the stone is at height s.

- *Geometric.* The graph $t = f^{-1}(s)$ is the original graph $s = f(t)$ flipped diagonally, so as to switch the vertical and horizontal axes.



- *Numerical.* We restrict the $s = f(t)$ table to $t \in [2, 5]$, and switch the two rows so that s is the input row, t is the output row.

s	45	40	25	0
$t = f^{-1}(s)$	2	3	4	5

  We can tidy this, rearranging and supplementing the data columns:

s	0	5	10	15	20	25	30	35	40	45
$t = f^{-1}(s)$	5.0	4.8	4.6	4.4	4.2	4.0	3.7	3.4	3.0	2.0

- *Algebraic:* We must go from $s = 25 + 20t - 5t^2$ to $t =$ formula in s. This just means to *solve* the original equation for t in terms of s. The Quadratic Formula solves $ax^2+bx+c = 0$ as $x = \frac{-b\pm\sqrt{b^2-4ac}}{2a}$, so:

$$s = 25 + 20t - 5t^2 \quad \Longleftrightarrow \quad (-5)t^2 + 20t + (25-s) = 0$$

$$\iff \quad t \;=\; \frac{-20 \pm \sqrt{20^2 - 4(-5)(25{-}s)}}{2(-5)} \;=\; 2 \pm \sqrt{9 - \tfrac{1}{5}s}.$$

The $\pm$ sign gives a choice between the two values t_1, t_2 in the original domain $t \in [0, 5]$ which correspond to s. We want the larger choice, namely the $+$ sign, to get $t \in [2, 5]$ as required:

$$t = 2 + \sqrt{9 - \tfrac{1}{5}s}.$$

Note: In the relationships $s = f(t)$ and $t = f^{-1}(s)$, the variables t, s are merely suggestive, recalling the physical meaning of these functions. On the algebraic level, we don't really care what the variables mean, and we sometimes change letters to make x the input variable for every function:

$$f(x) = 25 + 20x - 5x^2, \qquad f^{-1}(x) = 2 + \sqrt{9 - \tfrac{1}{5}x}\ .$$

Formal definitions.

Definition: Consider a function $f : A \to B$ with inputs in the set A and outputs covering the set B. Suppose f is one-to-one, meaning if $x_1 \neq x_2$, then $f(x_1) \neq f(x_2)$. (The graph $y = f(x)$ satisfies the horizontal line test.)

Then we define the *inverse function* $f^{-1} : B \to A$ as $f^{-1}(b) = a$, where $a \in A$ is the unique value with $f(a) = b$.

Inverse Theorem: The function f and its inverse f^{-1} satisfy:

$$f^{-1}(f(a)) = a \quad \text{and} \quad f(f^{-1}(b)) = b$$

for all $a \in A$, $b \in B$. That is, f and f^{-1} undo each other under composition.

The proof is just applying the definitions: if $f(a) = b$, then $f^{-1}(b) = a$, which means $f^{-1}(f(a)) = f^{-1}(b) = a$; and similarly for the other equation.

In our previous example, this means:

$$f^{-1}(f(t)) \;=\; 2 + \sqrt{9 - \tfrac{1}{5}(25{+}20t{-}5t^2)} \;=\; t,$$

$$f(f^{-1}(s)) \;=\; 25 + 20\left(2{+}\sqrt{9 - \tfrac{1}{5}s}\right) - 5\left(2{+}\sqrt{9 - \tfrac{1}{5}s}\right)^2 \;=\; s.$$

These equations are not obvious, but they must hold after simplification.

Another Example. For the function: $y = f(x) = \frac{x+1}{x+2}$, we can get its inverse by solving for x in terms of y:

$$\begin{aligned} y = \tfrac{x+1}{x+2} \quad &\iff \quad x{+}1 = (x{+}2)y \\ &\iff \quad x{+}1 = xy + 2y \\ &\iff \quad x - xy = 2y - 1 \\ &\iff \quad x(1{-}y) = 2y - 1 \\ &\iff \quad x = f^{-1}(y) = \tfrac{2y-1}{1-y}\ . \end{aligned}$$

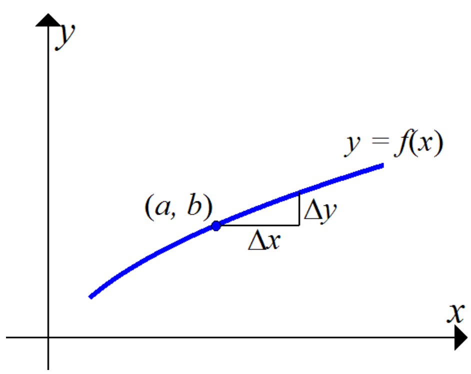
Changing the input variable to x, we get: $f^{-1}(x) = \frac{2x-1}{1-x}$.

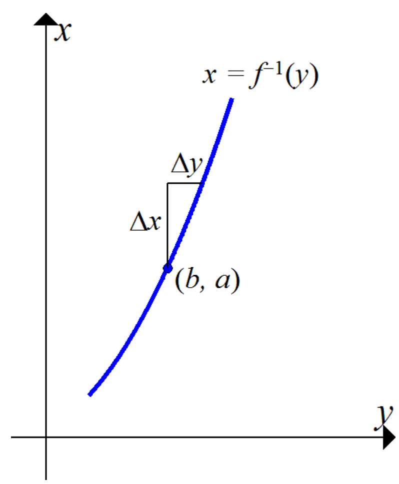
Note that the natural domain of $f(x)$, the set of inputs for which the formula $\frac{x+1}{x+2}$ makes sense, is all $x \neq -2$. Also, we can write $f(x) = \frac{x+2-1}{x+2} = 1 - \frac{1}{x+2}$, so the range of $f(x)$, the set of all outputs, is all $y \neq 1$. This is reversed for the inverse function: $f^{-1}(y)$ has domain $y \neq 1$ and range $x \neq 2$.

Derivatives of inverse functions.

> *Inverse Derivative Theorem.* Suppose $f : A \to B$ has inverse $f^{-1} : B \to A$, that $f(a) = b$ for some values a, b, and that $f(x)$ is differentiable at $x = a$. Then $f^{-1}(y)$ is differentiable at $y = b$:
>
> $$(f^{-1})'(b) \;=\; \frac{1}{f'(a)} \;=\; \frac{1}{f'(f^{-1}(b))}.$$
>
> In Leibnitz notation with $y = f(x)$ and $x = f^{-1}(y)$, we have:
>
> $$\left.\frac{dx}{dy}\right|_{y=b} = \frac{1}{\left.\frac{dy}{dx}\right|_{x=a}}.$$

Proof: This is most clear geometrically, considering that in the graph $y = f(x)$, we have $f'(a) \approx \frac{\Delta y}{\Delta x}$, the rise-over-run for a small interval near $x = a$. In the inverse graph $x = f^{-1}(y)$, the vertical and horizontal increments (rise and run) are switched, so that $(f^{-1})'(b) \approx \frac{\Delta x}{\Delta y} \approx \frac{1}{f'(a)}$. Taking $\Delta x, \Delta y \to 0$ turns the approximations into exact equalities in the limit.





A different, algebraic proof comes from the equations in the Inverse Theorem. Taking x as the input variable, we have $f(f^{-1}(x)) = x$. Applying the Chain Rule with $f^{-1}(x)$ as the inside function gives:

$$[f(f^{-1}(x))]' \;=\; (x)' \qquad \Longrightarrow \qquad f'(f^{-1}(x)) \cdot (f^{-1})'(x) = 1$$

$$\Longrightarrow \qquad (f^{-1})'(x) \;=\; \frac{1}{f'(f^{-1}(x))},$$

which is the formula of the Theorem when $x = b$.

EXAMPLE: Let $y = f(x) = x^3 + x + 1$. Since $f'(x) = 3x^2 + 1 > 0$ for all x, we see that $f(x)$ is increasing everywhere, and different inputs $x_1 < x_2$ must go to different outputs $f(x_1) < f(x_2)$; thus $f(x)$ is one-to-one. Hence there is an inverse function $f^{-1}(x)$, even though there is no neat formula for it.

Nevertheless, we know $f(1) = 3$ and $f'(1) = 4$, so:

$$(f^{-1})'(3) \;=\; \frac{1}{f'(f^{-1}(3))} \;=\; \frac{1}{f'(1)} \;=\; \frac{1}{4}.$$