

Differentiation and Derivatives

So far in this course, you've learned how to evaluate limits for points on a line. Now you're going to build on that knowledge and look at a calculus technique called *differentiation*. In differentiation, we use our knowledge of limits to calculate the *derivative* of a function in order to determine the rate of change at an individual point on a line.

Let's remind ourselves of the problem we're trying to solve, here's a function:

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

We can visualize part of the line that this function defines using the following Python code:

```
In [1]: %matplotlib inline

# Here's the function
def f(x):
    return x**2 + x

from matplotlib import pyplot as plt

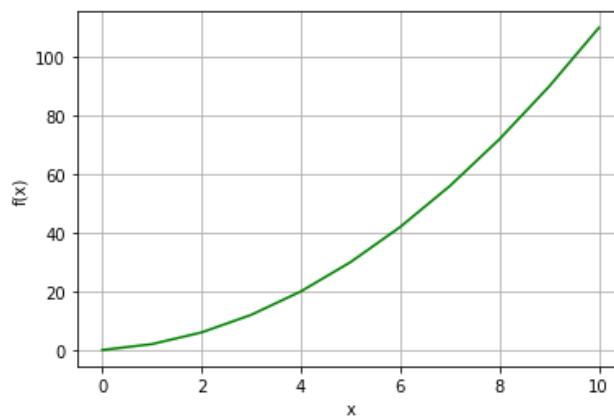
# Create an array of x values from 0 to 10 to plot
x = list(range(0, 11))

# Use the function to get the y values
y = [f(i) for i in x]

# Set up the graph
plt.xlabel('x')
plt.ylabel('f(x)')
plt.grid()

# Plot the function
plt.plot(x,y, color='green')

plt.show()
```



Now, we know that we can calculate the average rate of change for a given interval on the line by calculating the slope for a secant line that connects two points on the line. For example, we can calculate the average change for the interval between $x=4$ and $x=6$ by dividing the change (or *delta*, indicated as Δ) in the value of $f(x)$ by the change in the value of x :

$$m = \frac{\Delta f(x)}{\Delta x}$$

The delta for $f(x)$ is calculated by subtracting the $f(x)$ values of our points, and the delta for x is calculated by subtracting the x values of our points; like this:

$$m = \frac{f(x)_2 - f(x)_1}{x_2 - x_1}$$

So for the interval between $x=4$ and $x=6$, that's:

$$m = \frac{f(6) - f(4)}{6 - 4}$$

We can calculate and plot this using the following Python:

```

In [2]: %matplotlib inline

def f(x):
    return x**2 + x

from matplotlib import pyplot as plt

# Create an array of x values from 0 to 10 to plot
x = list(range(0, 11))

# Use the function to get the y values
y = [f(i) for i in x]

# Set the a values
x1 = 4
x2 = 6

# Get the corresponding f(x) values
y1 = f(x1)
y2 = f(x2)

# Calculate the slope by dividing the deltas
a = (y2 - y1)/(x2 - x1)

# Create an array of x values for the secant line
sx = [x1,x2]

# Use the function to get the y values
sy = [f(i) for i in sx]

# Set up the graph
plt.xlabel('x')
plt.ylabel('f(x)')
plt.grid()

# Plot the function
plt.plot(x,y, color='green')

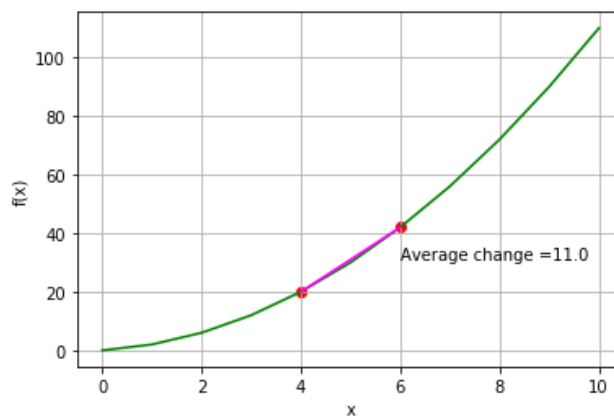
# Plot the interval points
plt.scatter([x1,x2],[y1,y2], c='red')

# Plot the secant line
plt.plot(sx,sy, color='magenta')

# Display the calculated average rate of change
plt.annotate('Average change = ' + str(a),(x2, (y2+y1)/2))

plt.show()

```



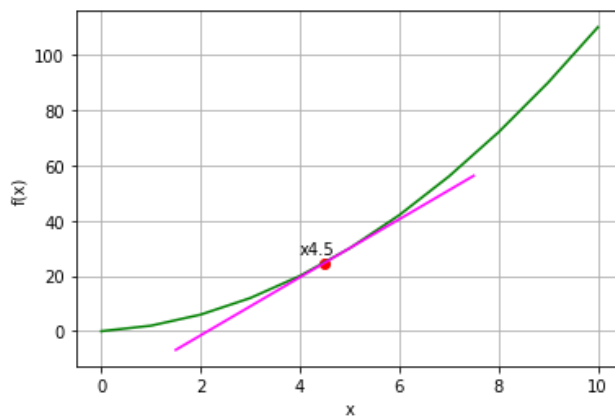
The average rate of change for the interval between $x=4$ and $x=6$ is $\frac{11}{1}$ (or simply 11), meaning that for every 1 added to x , $f(x)$ increases by 11. Put another way, if x represents time in seconds and $f(x)$ represents distance in meters, the average rate of change for distance over time (in other words, *velocity*) for the 4 to 6 second interval is 11 meters-per-second.

So far, this is just basic algebra; but what if instead of the average rate of change over an interval, we want to calculate the rate of change at a single point, say, where $x = 4.5$?

One approach we could take is to create a secant line between the point at which we want the slope and another point on the function line that is infinitesimally close to it. So close in fact that the secant line is actually a tangent that goes through both points. We can then calculate the slope for the secant line as before. This would look something like the graph produced by the following code:

In [3]: %matplotlib inline

```
def f(x):  
    return x**2 + x  
  
from matplotlib import pyplot as plt  
  
# Create an array of x values from 0 to 10 to plot  
x = list(range(0, 11))  
  
# Use the function to get the y values  
y = [f(i) for i in x]  
  
# Set the x1 point, arbitrarily 5  
x1 = 4.5  
y1 = f(x1)  
  
# Set the x2 point, very close to x1  
x2 = 5.000000001  
y2 = f(x2)  
  
# Set up the graph  
plt.xlabel('x')  
plt.ylabel('f(x)')  
plt.grid()  
  
# Plot the function  
plt.plot(x,y, color='green')  
  
# Plot the point  
plt.scatter(x1,y1, c='red')  
plt.annotate('x' + str(x1), (x1,y1), xytext=(x1-0.5, y1+3))  
  
# Approximate the tangent slope and plot it  
m = (y2-y1)/(x2-x1)  
xMin = x1 - 3  
yMin = y1 - (3*m)  
xMax = x1 + 3  
yMax = y1 + (3*m)  
plt.plot([xMin,xMax],[yMin,yMax], color='magenta')  
  
plt.show()
```



Calculating a Derivative

In the Python code above, we created the (almost) tangential secant line by specifying a point that is very close to the point at which we want to calculate the rate of change. This is adequate to show the line conceptually in the graph, but it's not a particularly generalizable (or accurate) way to actually calculate the line so that we can get the rate of change at any given point.

If only we knew of a way to calculate a point on the line that is as close as possible to point with a given x value.

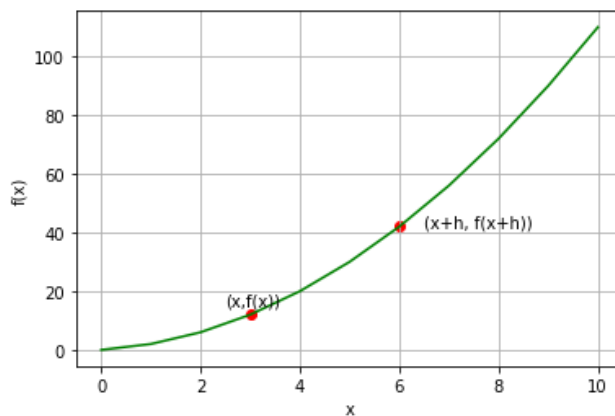
Oh wait, we do! It's a *limit*.

So how do we apply a limit in this scenario? Well, let's start by examining our general approach to calculating slope in a little more detail. Our tried and tested approach is to plot a secant line between two points at different values of x , so let's plot an arbitrary (x, y) point, and then add an arbitrary amount to x , which we'll call h . Then we know that we can plot a secant line between $(x, f(x))$ and $(x+h, f(x+h))$ and find its slope.

Run the cell below to see these points:

In [4]: %matplotlib inline

```
def f(x):  
    return x**2 + x  
  
from matplotlib import pyplot as plt  
  
# Create an array of x values from 0 to 10 to plot  
x = list(range(0, 11))  
  
# Use the function to get the y values  
y = [f(i) for i in x]  
  
# Set the x point  
x1 = 3  
y1 = f(x1)  
  
# set the increment  
h = 3  
  
# set the x+h point  
x2 = x1+h  
y2 = f(x2)  
  
# Set up the graph  
plt.xlabel('x')  
plt.ylabel('f(x)')  
plt.grid()  
  
# Plot the function  
plt.plot(x,y, color='green')  
  
# Plot the x point  
plt.scatter(x1,y1, c='red')  
plt.annotate('(x,f(x))',(x1,y1), xytext=(x1-0.5, y1+3))  
  
# Plot the x+h point  
plt.scatter(x2,y2, c='red')  
plt.annotate('(x+h, f(x+h))',(x2,y2), xytext=(x2+0.5, y2))  
  
plt.show()
```



As we saw previously, our formula to calculate slope is:

$$m = \frac{\Delta f(x)}{\Delta x}$$

The delta for $f(x)$ is calculated by subtracting the $f(x+h)$ and $f(x)$ values of our points, and the delta for x is just the difference between x and $x+h$; in other words, h :

$$m = \frac{f(x+h) - f(x)}{h}$$

What we actually need is the slope at the shortest possible distance between x and $x+h$, so we're looking for the smallest possible value of h . In other words, we need the limit as h approaches 0.

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

This equation is generalizable, and we can use it as the definition of a function to help us find the slope at any given value of x on the line, and it's what we call the *derivative* of our original function (which in this case is called f). This is generally indicated in *Lagrange* notation like this:

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

You'll also sometimes see derivatives written in *Leibniz's* notation like this:

$$\frac{d}{dx} f(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

Note: Some textbooks use h to symbolize the difference between x_0 and x_1 , while others use Δx . It makes no difference which symbolic value you use.

Alternate Form for a Derivative

The formula above shows the generalized form for a derivative. You can use the derivative function to get the slope at any given point, for example to get the slope at point a you could just plug the value for a into the generalized derivative function:

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}$$

Or you could use the alternate form, which is specific to point a :

$$f'(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$$

These are mathematically equivalent.

Finding the Derivative for a Specific Point

It's easier to understand differentiation by seeing it in action, so let's use it to find the derivative for a specific point in the function f .

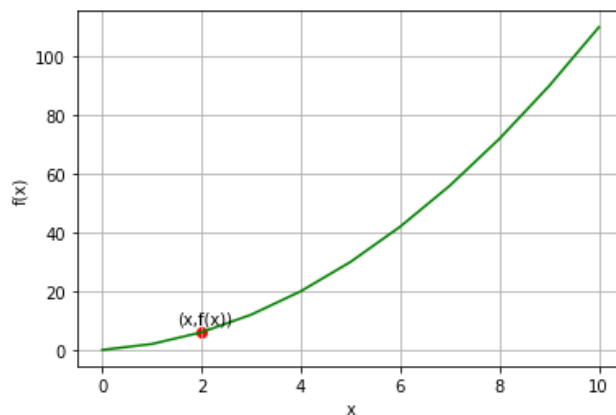
Here's the definition of function f :

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

Let's say we want to find $f'(2)$ (the derivative for f when x is 2); so we're trying to find the slope at the point shown by the following code:


```
In [5]: %matplotlib inline
```

```
def f(x):  
    return x**2 + x  
  
from matplotlib import pyplot as plt  
  
# Create an array of x values from 0 to 10 to plot  
x = list(range(0, 11))  
  
# Use the function to get the y values  
y = [f(i) for i in x]  
  
# Set the point  
x1 = 2  
y1 = f(x1)  
  
# Set up the graph  
plt.xlabel('x')  
plt.ylabel('f(x)')  
plt.grid()  
  
# Plot the function  
plt.plot(x,y, color='green')  
  
# Plot the point  
plt.scatter(x1,y1, c='red')  
plt.annotate('(x,f(x))', (x1,y1), xytext=(x1-0.5, y1+3))  
  
plt.show()
```



Here's our generalized formula for finding a derivative at a specific point (a):

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}$$

So let's just start by plugging our a value in:

$$f'(2) = \lim_{h \rightarrow 0} \frac{f(2+h) - f(2)}{h}$$

We know that $f(x)$ encapsulates the equation $x^2 + x$, so we can rewrite our derivative equation as:

$$f'(2) = \lim_{h \rightarrow 0} \frac{((2+h)^2 + 2+h) - (2^2 + 2)}{h}$$

We can apply the distribution property to $(2+h)^2$ using the rule that $(a+b)^2 = a^2 + b^2 + 2ab$:

$$f'(2) = \lim_{h \rightarrow 0} \frac{(4 + h^2 + 4h + 2 + h) - (2^2 + 2)}{h}$$

Then we can simplify $2^2 + 2$ (2^2 is 4, plus 2 gives is 6):

$$f'(2) = \lim_{h \rightarrow 0} \frac{(4 + h^2 + 4h + 2 + h) - 6}{h}$$

We can combine like terms on the left side of the numerator to make things a little clearer:

$$f'(2) = \lim_{h \rightarrow 0} \frac{(h^2 + 5h + 6) - 6}{h}$$

Which combines even further to get rid of the 6:

$$f'(2) = \lim_{h \rightarrow 0} \frac{h^2 + 5h}{h}$$

And finally, we can simplify the fraction:

$$f'(2) = \lim_{h \rightarrow 0} h + 5$$

To get the limit when h is approaching 0, we can use direct substitution for h :

$$f'(2) = 0 + 5$$

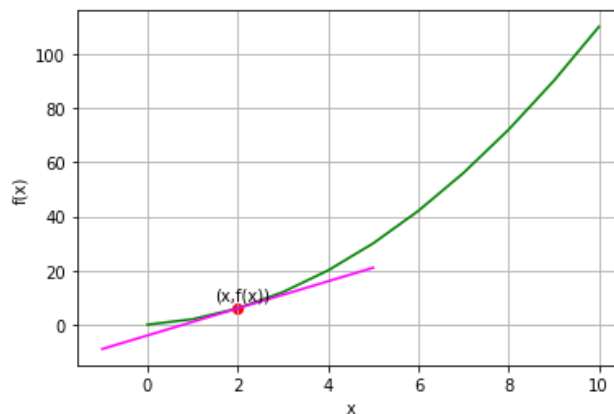
so:

$$f'(2) = 5$$

Let's draw a tangent line with that slope on our graph to see if it looks right:

In [6]: %matplotlib inline

```
def f(x):  
    return x**2 + x  
  
from matplotlib import pyplot as plt  
  
# Create an array of x values from 0 to 10 to plot  
x = list(range(0, 11))  
  
# Use the function to get the y values  
y = [f(i) for i in x]  
  
# Set the point  
x1 = 2  
y1 = f(x1)  
  
# Specify the derivative we calculated above  
m = 5  
  
# Set up the graph  
plt.xlabel('x')  
plt.ylabel('f(x)')  
plt.grid()  
  
# Plot the function  
plt.plot(x,y, color='green')  
  
# Plot the point  
plt.scatter(x1,y1, c='red')  
plt.annotate('(x,f(x))',(x1,y1), xytext=(x1-0.5, y1+3))  
  
# Plot the tangent line using the derivative we calculated  
xMin = x1 - 3  
yMin = y1 - (3*m)  
xMax = x1 + 3  
yMax = y1 + (3*m)  
plt.plot([xMin,xMax],[yMin,yMax], color='magenta')  
  
plt.show()
```



Finding a Derivative for Any Point

Now let's put it all together and define a function that we can use to find the derivative for any point in the f function:

Here's our general derivative function again:

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

We know that $f(x)$ encapsulates the equation $x^2 + x$, so we can rewrite our derivative equation as:

$$f'(x) = \lim_{h \rightarrow 0} \frac{((x+h)^2 + x+h) - (x^2 + x)}{h}$$

We can apply the distribution property to $(x+h)^2$ using the rule that $(a+b)^2 = a^2 + b^2 + 2ab$:

$$f'(x) = \lim_{h \rightarrow 0} \frac{(x^2 + h^2 + 2xh + x+h) - (x^2 + x)}{h}$$

Then we can use the distributive property to expand $-(x^2 + x)$, which is the same thing as $-1(x^2 + x)$, to $-x^2 - x$:

$$f'(x) = \lim_{h \rightarrow 0} \frac{x^2 + h^2 + 2xh + x+h - x^2 - x}{h}$$

We can combine like terms on the numerator to make things a little clearer:

$$f'(x) = \lim_{h \rightarrow 0} \frac{h^2 + 2xh + h}{h}$$

And finally, we can simplify the fraction:

$$f'(x) = \lim_{h \rightarrow 0} 2x + h + 1$$

To get the limit when h is approaching 0, we can use direct substitution for h :

$$f'(x) = 2x + 0 + 1$$

so:

$$f'(x) = 2x + 1$$

Now we have a function for the derivative of f , which we can apply to any x value to find the slope of the function at $f(x)$.

For example, let's find the derivative of f with an x value of 5:

$$f'(5) = 2 \cdot 5 + 1 = 10 + 1 = 11$$

Let's use Python to define the $f(x)$ and $f'(x)$ functions, plot $f(5)$ and show the tangent line for $f'(5)$:

In [7]: %matplotlib inline

```
# Create function f
def f(x):
    return x**2 + x

# Create derivative function for f
def fd(x):
    return (2 * x) + 1

from matplotlib import pyplot as plt

# Create an array of x values from 0 to 10 to plot
x = list(range(0, 11))

# Use the function to get the y values
y = [f(i) for i in x]

# Set the point
x1 = 5
y1 = f(x1)

# Calculate the derivative using the derivative function
m = fd(x1)

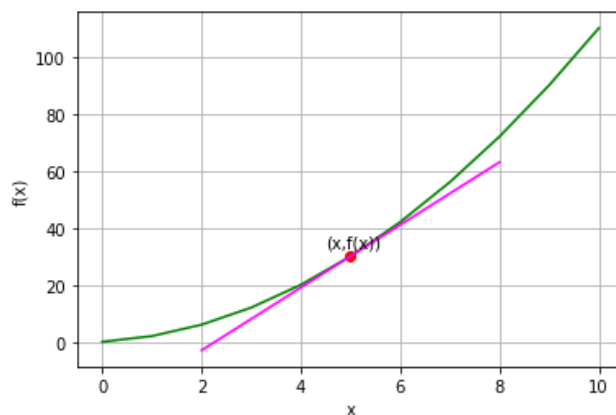
# Set up the graph
plt.xlabel('x')
plt.ylabel('f(x)')
plt.grid()

# Plot the function
plt.plot(x,y, color='green')

# Plot the point
plt.scatter(x1,y1, c='red')
plt.annotate('(x,f(x))', (x1,y1), xytext=(x1-0.5, y1+3))

# Plot the tangent line using the derivative we calculated
xMin = x1 - 3
yMin = y1 - (3*m)
xMax = x1 + 3
yMax = y1 + (3*m)
plt.plot([xMin,xMax],[yMin,yMax], color='magenta')

plt.show()
```



Differentiability

It's important to realize that a function may not be *differentiable* at every point; that is, you might not be able to calculate the derivative for every point on the function line.

To be differentiable at a given point:

- The function must be *continuous* at that point.
- The tangent line at that point cannot be vertical
- The line must be *smooth* at that point (that is, it cannot take on a sudden change of direction at the point)

For example, consider the following (somewhat bizarre) function:

$$q(x) = \begin{cases} \frac{40,000}{x^2}, & \text{if } x < -4, \\ (x^2 - 2) \cdot (x - 1), & \text{if } x \neq 0 \text{ and } x \geq -4 \text{ and } x < 8, \\ (x^2 - 2), & \text{if } x \neq 0 \text{ and } x \geq 8 \end{cases}$$

In [8]: %matplotlib inline

```
# Define function q
def q(x):
    if x != 0:
        if x < -4:
            return 40000 / (x**2)
        elif x < 8:
            return (x**2 - 2) * x - 1
        else:
            return (x**2 - 2)

# Plot output from function q
from matplotlib import pyplot as plt

# Create an array of x values
x = list(range(-10, -5))
x.append(-4.01)
x2 = list(range(-4,8))
x2.append(7.9999)
x2 = x2 + list(range(8,11))

# Get the corresponding y values from the function
y = [q(i) for i in x]
y2 = [q(i) for i in x2]

# Set up the graph
plt.xlabel('x')
plt.ylabel('q(x)')
plt.grid()

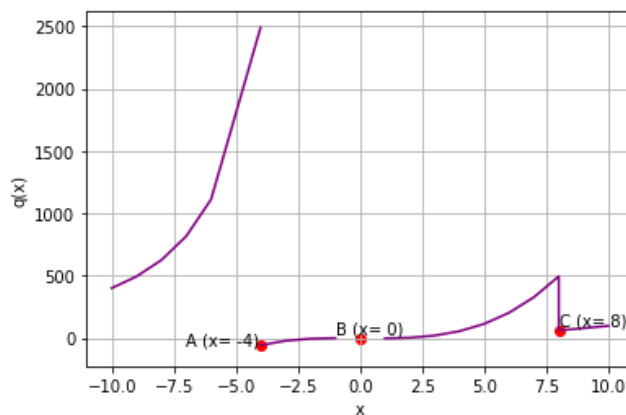
# Plot x against q(x)
plt.plot(x,y, color='purple')
plt.plot(x2,y2, color='purple')

plt.scatter(-4,q(-4), c='red')
plt.annotate('A (x= -4)',(-5,q(-3.9)), xytext=(-7, q(-3.9)))

plt.scatter(0,0, c='red')
plt.annotate('B (x= 0)',(0,0), xytext=(-1, 40))

plt.scatter(8,q(8), c='red')
plt.annotate('C (x= 8)',(8,q(8)), xytext=(8, 100))

plt.show()
```



The points marked on this graph are non-differentiable:

- Point **A** is non-continuous - the limit from the negative side is infinity, but the limit from the positive side ≈ -57
- Point **B** is also non-continuous - the function is not defined at $x = 0$.
- Point **C** is defined and continuous, but the sharp change in direction makes it non-differentiable.

Derivatives of Equations

We've been talking about derivatives of *functions*, but it's important to remember that functions are just named operations that return a value. We can apply what we know about calculating derivatives to any equation, for example:

$$\frac{d}{dx}(2x + 6)$$

Note that we generally switch to *Leibniz's* notation when finding derivatives of equations that are not encapsulated as functions; but the approach for solving this example is exactly the same as if we had a hypothetical function with the definition $2x + 6$:

$$\frac{d}{dx}(2x + 6) = \lim_{h \rightarrow 0} \frac{(2(x + h) + 6) - (2x + 6)}{h}$$

After factoring out the $2(x+h)$ on the left and the $-(2x - 6)$ on the right, this is:

$$\frac{d}{dx}(2x + 6) = \lim_{h \rightarrow 0} \frac{2x + 2h + 6 - 2x - 6}{h}$$

We can simplify this to:

$$\frac{d}{dx}(2x + 6) = \lim_{h \rightarrow 0} \frac{2h}{h}$$

Now we can factor h out entirely, so at any point:

$$\frac{d}{dx}(2x + 6) = 2$$

If you run the Python code below to plot the line created by the equation, you'll see that it does indeed have a constant slope of 2:


```
In [9]: %matplotlib inline
from matplotlib import pyplot as plt

# Create an array of x values from 0 to 10 to plot
x = list(range(1, 11))

# Use the function to get the y values
y = [(2*i) + 6 for i in x]

# Set up the graph
plt.xlabel('x')
plt.xticks(range(1,11, 1))
plt.ylabel('y')
plt.yticks(range(8,27, 1))
plt.grid()

# Plot the function
plt.plot(x,y, color='purple')

plt.show()
```



Derivative Rules and Operations

When working with derivatives, there are some rules, or shortcuts, that you can apply to make your life easier.

Basic Derivative Rules

Let's start with some basic rules that it's useful to know.

- If $f(x) = C$ (where C is a constant), then $f'(x) = 0$

This makes sense if you think about it for a second. No matter what value you use for x , the function returns the same constant value; so the graph of the function will be a horizontal line. There's no rate of change in a horizontal line, so its slope is 0 at all points. This is true of any constant, including symbolic constants like π (pi).

So, for example:

$$f(x) = 6 \quad \therefore \quad f'(x) = 0$$

Or:

$$f(x) = \pi \quad \therefore \quad f'(x) = 0$$

- If $f(x) = Cg(x)$, then $f'(x) = Cg'(x)$

This rule tells us that if a function is equal to a second function multiplied by a constant, then the derivative of the first function will be equal to the derivative of the second function multiplied by the same constant. For example:

$$f(x) = 2g(x) \quad \therefore \quad f'(x) = 2g'(x)$$

- If $f(x) = g(x) + h(x)$, then $f'(x) = g'(x) + h'(x)$

In other words, if a function is the sum of two other functions, then the derivative of the first function is the sum of the derivatives of the other two functions. For example:

$$f(x) = g(x) + h(x) \quad \therefore \quad f'(x) = g'(x) + h'(x)$$

Of course, this also applies to subtraction:

$$f(x) = k(x) - l(x) \quad \therefore \quad f'(x) = k'(x) - l'(x)$$

As discussed previously, functions are just equations encapsulated as a named entity that return a value; and the rules can be applied to any equation. For example:

$$\frac{d}{dx}(2x + 6) = \frac{d}{dx}2x + \frac{d}{dx}6$$

So we can take advantage of the rules to make the calculation a little easier:

$$\frac{d}{dx}(2x) = \lim_{h \rightarrow 0} \frac{2(x+h) - 2x}{h}$$

After factoring out the $2(x+h)$ on the left, this is:

$$\frac{d}{dx}(2x) = \lim_{h \rightarrow 0} \frac{2x + 2h - 2x}{h}$$

We can simplify this to:

$$\frac{d}{dx}(2x) = \lim_{h \rightarrow 0} \frac{2h}{h}$$

The Quotient Rule

The *quotient rule* applies to functions that are defined as a quotient of one expression divided by another; for example:

$$r(x) = \frac{s(x)}{t(x)}$$

In this situation, you can apply the following quotient rule to find the derivative of $r(x)$:

$$r'(x) = \frac{s'(x)t(x) - s(x)t'(x)}{[t(x)]^2}$$

Here are our definitions for $s(x)$ and $t(x)$:

$$s(x) = 3x^2$$

$$t(x) = 2x$$

Let's start with $s'(x)$:

$$s'(x) = \lim_{h \rightarrow 0} \frac{3(x+h)^2 - 3x^2}{h}$$

This factors out to:

$$s'(x) = \lim_{h \rightarrow 0} \frac{3x^2 + 3h^2 + 6xh - 3x^2}{h}$$

Which when we cancel out the $3x^2$ and $-3x^2$ is:

$$s'(x) = \lim_{h \rightarrow 0} \frac{3h^2 + 6xh}{h}$$

Which simplifies to:

$$s'(x) = \lim_{h \rightarrow 0} 3h + 6x$$

With h approaching 0, this is:

$$s'(x) = 6x$$

Now let's look at $t'(x)$:

$$t'(x) = \lim_{h \rightarrow 0} \frac{2(x+h) - 2x}{h}$$

We can just factor out the $2(x+h)$ on the left:

$$t'(x) = \lim_{h \rightarrow 0} \frac{2x + 2h - 2x}{h}$$

Which can be cleaned up to:

$$t'(x) = \lim_{h \rightarrow 0} \frac{2h}{h}$$

Enabling us to factor h out completely to give a constant derivative of 2:

$$t'(x) = 2$$

So now we can calculate the derivative for the quotient of these functions by plugging the function definitions and the derivatives we've calculated for them into the quotient rule equation:

The Chain Rule

The *chain rule* takes advantage of the fact that equations can be encapsulated as functions, and since functions contain equations, it's possible to nest one function within another.

For example, consider the following function:

$$u(x) = 2x^2$$

We could view the definition of $u(x)$ as a composite of two functions,; an *inner* function that calculates x^2 , and an *outer* function that multiplies the result of the inner function by 2.

$$u(x) = \widehat{2(x^2)}$$

To make things simpler, we can name these inner and outer functions:

$$i(x) = x^2$$

$$o(x) = 2x$$

Note that x indicates the input for each function. Function i takes its input and squares it, and function o takes its input and multiplies it by 2. When we use these as a composite function, the x value input into the outer function will be the output from the inner function.

Let's take a look at how we can apply these functions to get back to our original u function:

$$u(x) = o(i(x))$$

So first we need to find the output of the inner i function so we can use it as the input value for the outer o function. Well, that's easy, we know the definition of i (square the input), so we can just plug it in:

$$u(x) = o(x^2)$$

We also know the definition for the outer o function (multiply the input by 2), so we can just apply that to the input:

$$u(x) = 2x^2$$

OK, so now we know how to form a composite function. The *chain rule* can be stated like this:

$$\frac{d}{dx}[o(i(x))] = o'(i(x)) \cdot i'(x)$$

Alright, let's start by plugging the output of the inner $i(x)$ function in:

$$\frac{d}{dx}[o(i(x))] = o'(x^2) \cdot i'(x)$$

Now let's use that to calculate the derivative of o , replacing each x in the equation with the output from the i function (x^2):

$$o'(x) = \lim_{h \rightarrow 0} \frac{2(x^2 + h) - 2x^2}{h}$$

This factors out to:

$$o'(x) = \lim_{h \rightarrow 0} \frac{2x^2 + 2h - 2x^2}{h}$$

Which when we cancel out the $2x^2$ and $-2x^2$ is:

$$o'(x) = \lim_{h \rightarrow 0} \frac{2h}{h}$$

