Calculus I, II, III

F22-W23

Louis Meunier

notes.louismeunier.net

Contents

Part I

Single-Variable Calculus

1 Limits

1.1 Tricks for solving limits

1.1.1 $\frac{x}{x}$

In some situations, you can multiply a limit by some factor of $\frac{x}{x}$ to simplify an otherwise difficult to evaluate limit

eg:

$$\begin{aligned}\lim_{x\to\infty}\frac{x+3}{\sqrt{9x^2-5x}} &= \lim_{x\to\infty}\frac{x+3}{\sqrt{9x^2-5x}}\cdot\frac{\frac{1}{x}}{\frac{1}{x}}\\ &= \lim_{x\to\infty}\frac{\frac{x}{x}+\frac{3}{x}}{\sqrt{(\frac{1}{x^2})(9x^2-5x)}}\\ &= \lim_{x\to\infty}\frac{1+\frac{3}{x}}{\sqrt{9-\frac{5}{x}}}\end{aligned} \tag{1}$$

The terms $\frac{3}{x}$ and $\frac{5}{x}$ both tend to 0 as $x\to\infty$, so the final limit simplifies:

$$\begin{aligned}\lim_{x\to\infty}\frac{1}{\sqrt{9}} &= \frac{1}{\sqrt{9}}\\ &= \frac{1}{3}\end{aligned} \tag{2}$$

1.1.2 Intuition with ∞

In limits involving fractions where both numerator and denominator are polynomials, you can intuitively calculate the limit using the coefficients/powers.

$$\lim_{x\to\infty}\frac{ax^m}{bx^n}$$

1. $\mathbf{m} > \mathbf{n} : \infty$
2. $\mathbf{m} < \mathbf{n} : 0$
3. $\mathbf{m} = \mathbf{n} : \frac{m}{n}$ *(since both of the powers are the same "strength" ie ∞, you can think of this as the infinities "canceling out"). To prove this idea actually works, use the method above.*

1.1.3 L'Hôpital's Rule

Theorem (L'Hôpital's Rule)**.** *In short, if:*

$$\frac{f(c)}{g(c)} = \frac{0}{0}, \frac{\infty}{\infty}, 0^0, \infty - \infty, ...(\textbf{\textit{indeterminate form}})$$

then:

$$\lim_{x \to c} \frac{f(x)}{g(x)} = \lim_{x \to c} \frac{f'(x)}{g'(x)}$$

This can also be repeated for multiple derivatives of f, g, as long as the resulting form is still indeterminate. Oftentimes, you will need to manipulate the limit to express it in an indeterminate form such that L'Hôpital's is usable.

While solving limits this way is very powerful (and, often, very easy), one should be careful to not overuse it when it is not applicable.

1.1.4 Solving for a given limit

When given a function with variables in place of constants and a final limit to solve for, there are some additional considerations to take into account.

eg:

$$\lim_{x \to 0} \frac{\sqrt{ax + b} - 2}{x} = 1$$

In order for this limit to even exist, $\lim_{x \to 0} \sqrt{ax + b} - 2$ MUST equal 0, following the quotient rule of limits, which states:

$$\lim_{x \to c} \frac{f(x)}{g(x)} = \frac{\lim_{x \to c} f(x)}{\lim_{x \to c} g(x)}$$

So:

$$\begin{aligned} \lim_{x \to 0} \sqrt{ax + b} - 2 &= 0 \\ \sqrt{a * 0 + b} - 2 &= 0 \\ \sqrt{b} &= 2 \\ b &= 4 \end{aligned} \tag{3}$$

To finish solving this limit, see:

1.1.5 "Radicalizing" fraction

Similar to the first point, but a little more complex/situation-specific.

When a limit is in the form:

$$\lim_{x\to 0} \frac{\sqrt[n]{(ax+b)}+c}{x}$$

The issue here is that we have an x alone on the bottom, which results in something being divided by 0. To solve this, we can radicalize the fraction; this results in an x "alone", canceling with that on the bottom, making the equation solvable.

For the example above:

$$\begin{aligned}
\lim_{x\to 0} \frac{\sqrt{ax+4}-2}{x} &= 1 \\
\lim_{x\to 0} \frac{\sqrt{ax+4}-2}{x} . \frac{\sqrt{ax+4}+2}{\sqrt{ax+4}+2} &= 1 \\
\lim_{x\to 0} \frac{ax}{x(\sqrt{ax+4}+2)} &= 1 \\
\lim_{x\to 0} \frac{a}{\sqrt{ax+4}+2} &= 1 \\
&\dots \\
a &= 4
\end{aligned} \tag{4}$$

1.1.6 Squeeze Theorem

Theorem (Squeeze Theorem)**.** *If:*

- $f(x) \leq g(x) \leq h(x)$
- $\lim_{x\to c} f(x) = \lim_{x\to c} h(x)$

then

$$\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = \lim_{x\to c} h(x)$$

This makes a lot more sense in practice:

$$\begin{aligned} \lim_{x\to 0} x^4 \sin\left(\frac{x}{2}\right) &= ? \\ -1 \leq \sin\left(\frac{x}{2}\right) &\leq 1 \\ -x^4 \leq x^4 \sin\left(\frac{x}{2}\right) &\leq x^4 \\ \lim_{x\to 0} -x^4 = \lim_{x\to 0} x^4 = 0 \therefore \lim_{x\to 0} x^4 \sin\left(\frac{x}{2}\right) &= 0 \end{aligned} \tag{5}$$

See Figure 1 to see this visually.

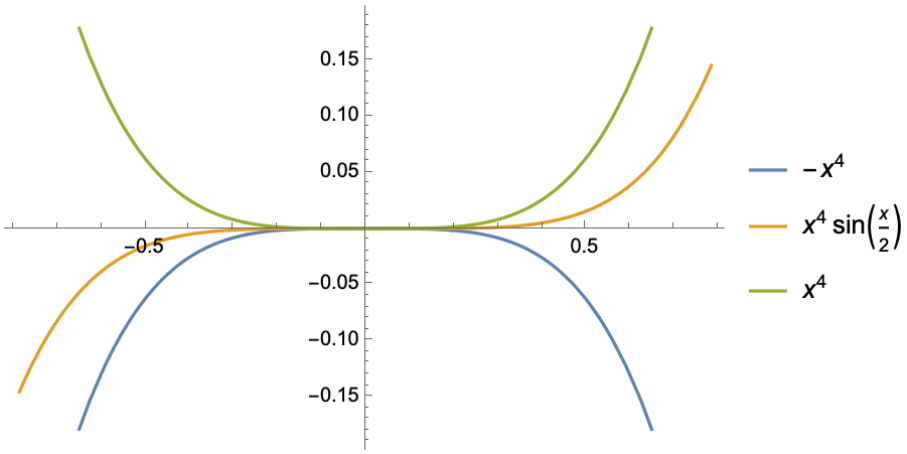


Figure 1: The squeeze theorem visualized

1.1.7 Using the definition of a derivative

A derivative, $f'(x)$, is defined as followed:

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

Alternatively, given $h = x - a$:

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

If you are given a limit in the form (or near) the above forms, you can use these definitions to solve the limit.

eg:

$$\lim_{x\to\pi}\frac{e^{\sin(x)}-1}{x-\pi}=\lim_{x\to\pi}\frac{e^{\sin(x)}-e^{\sin(\pi)}}{x-\pi}$$

Let f(x) = $e^{\sin(x)}$

$$=\lim_{x\to\pi}\frac{f(x)-f(\pi)}{x-\pi} \tag{6}$$

This is simply the definition of $f'(x)$ *at* $a=\pi$

$$=f'(\pi)=\cos(\pi)e^{\sin(\pi)}=-1$$

1.2 Intermediate Value Theorem

Theorem (Intermediate Value Theorem)**.** *If*

- f *is continuous on* $[a,b]$
- $f(a)\neq f(b)$
- $f(a)<=N<=f(b)$

there exists some c *in* (a,b) *such that* $N=f(c)$.

See Figure 2.

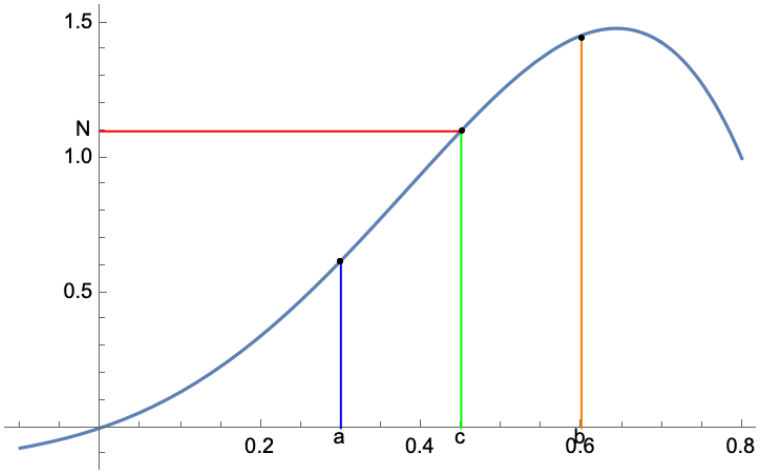


Figure 2: The intermediate value theorem visualized

2 Exponential and Logarithmic Functions

2.1 Exponential Functions

2.1.1 Defining e

An exponential function is any $f(x)$ of the form $f(x)=a^x$. To find the derivative, $f'(x)$:

$$\begin{aligned} f'(x) &= \lim_{h\to 0}[\frac{a^{x+h}-a^x}{h}] \\ &= a^x \lim_{h\to 0}[\frac{a^h-1}{h}] \end{aligned} \tag{7}$$

We can rewrite this for the derivative of f at $x = 0$:

$$\begin{aligned} f'(0) &= a^0 \lim_{h\to 0}[\frac{a^h-1}{h}] \\ &= \lim_{h\to 0}[\frac{a^h-1}{h}] \end{aligned} \tag{8}$$

We can use this to find for what value of a that $f'(0) = f(0)$. $f(0) = a^0 = 1$, therefore $f'(0) = f(0) = 1 = \lim_{h\to 0}[\frac{a^h-1}{h}]$.

$$\begin{aligned} 1 &= \lim_{h\to 0}[\frac{a^h-1}{h}] \\ \mathbf{a = 2}: \lim_{h\to 0}[\frac{2^h-1}{h}] &= 0.693... \\ \mathbf{a = 3}: \lim_{h\to 0}[\frac{2^h-1}{h}] &= 1.099... \end{aligned} \tag{9}$$

By the Intermediate Value Theorem, we know that $2 < a < 3$, which can be defined by e.

$$1 = e^0 = \lim_{h\to 0}[\frac{e^h-1}{h}] \tag{10}$$

Extending this further for e;

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}e^x &= \lim_{h\to 0}[\frac{e^{x+h}-e^x}{h}] \\ &= e^x \lim_{h\to 0}[\frac{e^h-1}{h}] \\ \textit{Substituting from eqn. 10} \\ &= \mathbf{e^x} \end{aligned} \tag{11}$$

2.1.2 Generalizing $\frac{d}{dx}a^x$

From eqn. 8:

$$\frac{\mathrm{d}}{\mathrm{d}x}a^x = a^x \lim_{h\to 0}\frac{a^h-1}{h} = f'(0)a^x$$

We can rewrite using the rules of logs:

$$\frac{\mathrm{d}}{\mathrm{d}x}a^x = \frac{\mathrm{d}}{\mathrm{d}x}e^{\ln a^x} = \frac{\mathrm{d}}{\mathrm{d}x}e^{x\ln a}$$

We can let $u = x\ln a$, and thus $a^x = e^u$. We can then use implicit differentiation as follows:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}a^x &= \frac{\mathrm{d}}{\mathrm{d}x}e^u \\
&= \frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}}{\mathrm{d}u}[e^u] \\
&\textit{From eqn. 11} \\
&= e^u\frac{\mathrm{d}u}{\mathrm{d}x} \\
&= e^u\frac{\mathrm{d}(x\ln a)}{\mathrm{d}x}\textit{Differentiating in terms of x} \\
&= e^u\ln a \\
&= e^{\ln a^x}\ln a \\
&= a^x\ln a
\end{aligned} \tag{12}$$

Thus, $\frac{\mathrm{d}}{\mathrm{d}x}a^x = a^x\ln a$.

2.2 Logarithmic Functions

Let $f(x) = y = \log_a x$, so $x = a^y$. To find the derivative of $f(x)$:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}x &= \frac{\mathrm{d}}{\mathrm{d}x}a^y \\
1 &= [\frac{\mathrm{d}}{\mathrm{d}y}a^y]\frac{\mathrm{d}y}{\mathrm{d}x} \\
&\textit{Using eqn. 12 :} \\
1 &= a^y \ln a \frac{\mathrm{d}y}{\mathrm{d}x} \\
\frac{1}{a^y \ln a} &= \frac{\mathrm{d}y}{\mathrm{d}x} \\
&\textit{Subbing in for } y : \\
\frac{1}{a^{log_a x} \ln a} &= \frac{\mathrm{d}}{\mathrm{d}x} log_a x \\
\frac{1}{x \ln a} &= \frac{\mathrm{d}}{\mathrm{d}x} \log_a x
\end{aligned}
\tag{13}
$$

Therefore, $f'(x) = \frac{\mathrm{d}}{\mathrm{d}x} log_a x = \frac{1}{x \ln a}$.

When $a = e$:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x} log_e x &= \frac{\mathrm{d}}{\mathrm{d}x} \ln x \\
&= \frac{1}{x \ln e} \\
&= \frac{1}{x}
\end{aligned}
\tag{14}
$$

2.3 Obtaining e as a limit

Let $f(x) = \ln x, \therefore f(1) = \ln 1 = 0$. In addition, $f'(x) = \frac{1}{x}, \therefore f'(1) = 1/1 = 1$. We can use the definition of a limit to say the following:

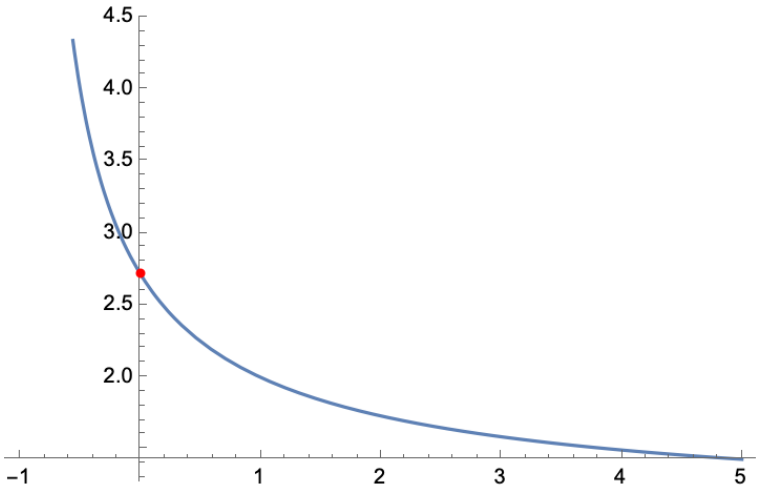


Figure 3: Eqn. 15 visualized

$$
\begin{aligned}
1 &= \lim_{h\to 0}[\frac{f(1+h)-f(1)}{h}] \\
&= \lim_{h\to 0}\frac{1}{h}[f(1+h)-f(1)] \\
&= \lim_{h\to 0}\frac{1}{h}[\ln(1+h)-\ln 1] \\
&= \lim_{h\to 0}\frac{1}{h}\ln(1+h) \\
&= \lim_{h\to 0}\ln[(1+h)^{\frac{1}{h}}] \\
&= \ln[\lim_{h\to 0}[(1+h)^{\frac{1}{h}}]] \\
&\therefore \lim_{h\to 0}[(1+h)^{\frac{1}{h}}] = e
\end{aligned}
\tag{15}
$$

Note that the penultimate step of this work relies on the following (very helpful) theorem:

Theorem. *If f is continuous at $\lim_{x\to a} g(x)$, then,*

$$\lim_{x\to a} f(g(x)) = f(\lim_{x\to a} g(x))$$

A proof can be found for this on page A39 of the Stewart's book.

Alternatively, we can write e as a limit going to ∞; Let $y = \frac{1}{h}$, and, substituting into eqn. 15:

$$\lim_{y\to\infty}[1+\frac{1}{y}]^y = e \tag{16}$$

3 Linearization

3.1 Linear Approximation

We can approximate a function by using the equation for its tangent at a particular point.

For example; For a function $y = f(x)$, there exists a point $P = (a, f(a))$. Let $L(x)$ be the line tangent to $f(x)$ at P. The slope of L would be defined as:

$$m_L = \frac{\Delta y}{\Delta x} = \frac{L(x) - f(a)}{x - a} = f'(a)$$

Solving for $L(x)$:

$$\begin{aligned} L(x) - f(a) &= (x - a)f'(a) \\ L(x) &= f(a) + (x - a)f'(a) \end{aligned} \tag{17}$$

This function is a fairly good approximation of $f(x)$ near $x = a$, at least for most continuous functions. It's also very important to note that, by definition $f(a) = L(a)$, AND $f'(a) = L'(a)$. This will be expanded upon later.

For example; take $f(x) = \sin(x)$, @ $a = 0$:

$$\begin{aligned} f(x) \approx L(x) &= f(0) + (x - 0)f'(0) \\ &= f(0) + x\cos(0) \\ &= 0 + x * 1 \\ &= x = L(x) \end{aligned} \tag{18}$$

Clearly, in the vicinity of a, this is a pretty good approximation, but it quickly becomes inaccurate as we move away.

3.2 Quadratic Approximation

From the section earlier, we defined $L(x)$, a linear approximation of $f(x)$, where $f(a) = L(a)$ AND $f'(a) = L'(a)$. Now, what if we wanted to continue this trend, and define a new function, $P_2(x)$, where $f(a) = P_2(a)$, $f'(a) = P_2'(a)$, and $f''(a) = P_2''(a)$? Rather than a "linear" approximation, this would, logically, become a "quadratic" approximation. We can say:

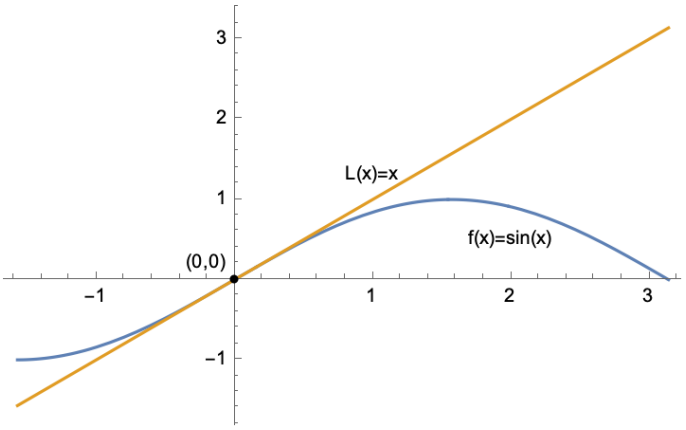


Figure 4: Linearization of $\sin(x)$ at $x = 0$

$$
\begin{aligned}
P_2(x) &= f(a) + (x-a)f'(a) + k(x-a)^2 \\
P_2(x) &= L(a) + k(x-a)^2 \\
P_2'(x) &= L'(a) + 2k(x-a) \\
P_2''(a) &= 2k \\
k &= \frac{P_2''(a)}{2} = \frac{f''(a)}{2}
\end{aligned}
\tag{19}
$$

Note that $L''(a) = 0$; since $L(x)$ is a line, its first derivative is a constant, and its second is 0.

The purpose of this work was to find some k that we can multiply by a quadratic factor $(x-a)^2$ to maintain the previously stated desired properties of P_2. As such, we can rewrite:

$$P_2(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2}f''(a)$$

3.2.1 Taylor Polynomials

We can, logically, continue this trend for cubic, quartic, etc. approximations. In general, you can approximate any function $f(x)$ to the Nth power;

$$f(x) \approx P_N(x) = \sum_{n=0}^{N} \frac{(x-a)^n f^{(n)}(a)}{n!}$$

This approximation:

- passes through the point $(a, f(a))$
- has the same slope, concavity, ..., ie $P_N^{N-1}(a) = f_N^{N-1}(a)$

This Nth degree polynomial is called a **Taylor Polynomial**, and, intuitively, better approximates $f(x)$ as

$N \to \infty$.

When $a = 0$, this polynomial becomes called a **Maclaurin Polynomial**:

$$P_n(x) = \sum_{n=0}^{N} \frac{x^n f^n(0)}{n!}$$

3.2.2 Relationship between $f(x)$ and $P_N(x)$

While, intuitively, it makes sense that $P_N(x)$ becomes a better approximation of $f(x)$ as $N \to \infty$, this isn't necessarily true.

Theorem (Taylor's Theorem)**.**

$$f(x) = P_N(x) + R_N(x);$$

where $P_N(x)$ is an approximation (via Taylor Polynomial) of $f(x)$, and $R_N(x)$ is the error in said approximation, defined as:

$$R_N(x) = \frac{(x-a)^{N+1}}{(N+1)!} f^{(N+1)}(c), c \in \mathbb{R} \text{ and between } x, a$$

.

We can then rewrite Taylor's Theorem, as N goes to ∞:

$$f(x) = \lim_{N\to\infty} [\sum_{n=0}^{N} \frac{(x-a)^n f^N(a)}{n!} + \frac{(x-a)^{N+1}}{(N+1)!} f^{(N+1)}(c)]$$

Therefore, $f(x) = P_N(x)$ if, and only if, $\lim_{N\to\infty} R_N(x) = 0$. Logically, this should make sense; if there is no remainder, then the approximation is equal to the original function.

3.3 Taylor Polynomials, Applied Example

Using Taylor Polynomials, we can explore a number of interesting properties of functions, and even derive some unique expressions. For instance (assuming $R_N(x) = 0$):

$$\begin{aligned} e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + ... \\ e^{ix} &= 1 + ix + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + ... \\ e^{ix} &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} ... + i[x - \frac{x^3}{3!} + \frac{x^5}{5!} ...] \end{aligned} \tag{20}$$

This may look like it just made things more complicated. However, take a look at the Taylor Polynomial representations of $\sin(x)$:

$$\begin{aligned} f(x) = \sin(x) &= \sum_{n=0}^{\infty} \frac{(x-a)^n f^n(a)}{n!} \\ f'(x) &= \cos(x) \\ f''(x) &= -\sin(x) \\ f'''(x) &= -\cos(x) \\ f^{IV}(x) &= \sin(x) \\ \therefore f(x) &= 0 + x + 0 + \frac{-x^3}{3!} + 0 + \frac{x^5}{5!} + ... \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} \end{aligned} \tag{21}$$

and, very similarly, for $\cos x$:

$$\begin{aligned} f(x) &= \cos(x) \\ &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + ... \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} \end{aligned} \tag{22}$$

Thus, we can rewrite equation 20 by substituting in $\sin$ and $\cos$, for the odd and even factors respectively:

$$e^{ix} = \cos(x) + i\sin(x) \tag{23}$$

This is a very helpful equation, known as **Euler's Formula**. Some applications are present in Section 5.

From here, we can solve for $x = \pi$, to obtain:

$$\begin{aligned} e^{i\pi} &= \cos(\pi) + i\sin(\pi) \\ &= -1 + i * 0 \\ e^{i\pi} + 1 &= 0 \end{aligned} \tag{24}$$

This ("the most beautiful formula in mathematics") is known as **Euler's Identity**.

To take this formula further we can solve for $\sin(x)$ and $\cos(x)$ in the imaginary plane.

$$\begin{aligned}
e^{ix} = \cos(x) + i\sin(x) &\implies \cos(x) = e^{ix} - i\sin(x) \\
e^{-ix} = \cos(x) - i\sin(x) &\implies \cos(x) = e^{-ix} + i\sin(x) \\
e^{ix} - i\sin(x) &= e^{-ix} + i\sin(x) \\
e^{ix} - e^{-ix} &= 2i\sin(x) \\
\sin(x) &= \frac{e^{ix} - e^{-ix}}{2i} \\
\cos(x) &= e^{ix} - i\sin(x) \\
\cos(x) &= e^{ix} - i(\frac{e^{ix} - e^{-ix}}{2i}) \\
\cos(x) &= \frac{e^{ix} + e^{-ix}}{2}
\end{aligned} \tag{25}$$

From here, we can solve for $\cos(ix)$ and $\sin(ix)$:

$$\begin{aligned}
\cos(ix) &= \frac{e^{i^2x} + e^{-i(ix)}}{2} \\
&= \frac{e^{x} + e^{-x}}{2} = \cosh(x) \\
\sin(ix) &= i(\frac{e^{i^2x} - e^{-i(ix)}}{2}) \\
&= i(\frac{e^{x} - e^{-x}}{2}) = i\sinh(x)
\end{aligned} \tag{26}$$

And thus, we can define the hyperbolic trigonometric functions using the complex plane, Taylor Series, and Euler's Identity.

4 Trigonometry

4.1 Inverse Trig Functions

As with all functions, we can define the inverse of the basic trig functions, as the functions that undo their respective trig functions. The following shows each of the three basic trig functions, their inverses, and how to derive their respective derivatives.

- $y = \arcsin x = sin^{-1}x$

$$\begin{aligned} x &= \sin y \\ \frac{\mathrm{d}}{\mathrm{d}x}x &= \frac{\mathrm{d}}{\mathrm{d}x}\sin y \\ 1 &= \frac{\mathrm{d}}{\mathrm{d}y}[\sin y]\frac{\mathrm{d}y}{\mathrm{d}x} \\ 1 &= \cos y\frac{\mathrm{d}y}{\mathrm{d}x} \\ \frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{1}{\cos y} \\ &= \frac{1}{\sqrt{1-\sin^2(y)}} = \frac{1}{\sqrt{1-x^2}} \end{aligned} \tag{27}$$

- $y = \arccos x = cos^{-1}x$

$$\begin{aligned} x &= \cos y \\ \frac{\mathrm{d}}{\mathrm{d}x}x &= \frac{\mathrm{d}}{\mathrm{d}x}\cos y \\ 1 &= \frac{\mathrm{d}}{\mathrm{d}y}[\cos y]\frac{\mathrm{d}y}{\mathrm{d}x} \\ 1 &= -\sin y\frac{\mathrm{d}y}{\mathrm{d}x} \\ \frac{\mathrm{d}y}{\mathrm{d}x} &= -\frac{1}{\sin y} \\ &= -\frac{1}{\sqrt{1-\cos^2(y)}} = -\frac{1}{\sqrt{1-x^2}} \end{aligned} \tag{28}$$

- $y = \arctan x = \tan^{-1} x$

$$\begin{aligned} x &= \tan y \\ \frac{\mathrm{d}}{\mathrm{d}x}x &= \frac{\mathrm{d}}{\mathrm{d}x}\tan y \\ 1 &= \frac{\mathrm{d}}{\mathrm{d}y}\tan y\frac{\mathrm{d}y}{\mathrm{d}x} \\ 1 &= \sec^2 y\frac{\mathrm{d}y}{\mathrm{d}x} \\ \frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{1}{\sec^2 y} \\ &= \frac{1}{1+\tan^2 y} = \frac{1}{1+x^2} \end{aligned} \tag{29}$$

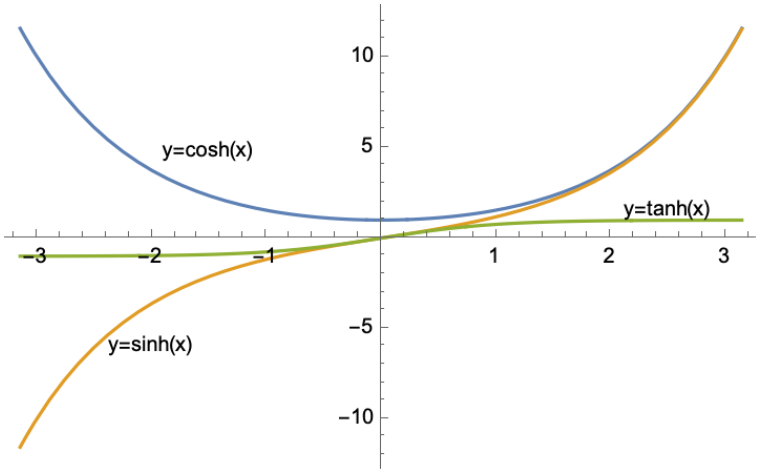


Figure 5: $\cosh(x), \sinh(x), \tanh(x)$

4.2 Hyperbolic Functions

4.2.1 Definitions, Identities, Derivatives

Def:

$$\begin{aligned} \sinh x &= \frac{e^x - e^{-x}}{2} \\ \cosh x &= \frac{e^x + e^{-x}}{2} \end{aligned} \tag{30}$$

Hyperbolic trigonometric functions are the "hyperbolic equivalent" of the basic trig functions. This is an admittedly abstract definition, but there is a more concrete definition in section 3.3, using Taylor Polynomials and complex numbers.

The derivatives are as follows; for $\sinh x$:

$$\frac{\mathrm{d}}{\mathrm{d}x} \sinh x = \frac{\mathrm{d}}{\mathrm{d}x}[\frac{e^x - e^{-x}}{2}] = \frac{e^x + e^{-x}}{2} = \cosh x;$$

and for $\cosh x$:

$$\frac{\mathrm{d}}{\mathrm{d}x} \cosh x = \frac{\mathrm{d}}{\mathrm{d}x}[\frac{e^x + e^{-x}}{2}] = \frac{e^x - e^{-x}}{2} = \sinh x.$$

Note that this is not quite the same relationship between the derivatives of $\sin x$ and $\cos x$.

Similarly, there exist several properties of hyperbolic functions that, while similar in appearance to their equivalent trig functions, have some key differences.

$$\cosh^2 x - \sinh^2 x = 1$$
$$\sinh(x+y) = \sinh x \cosh y + \cosh x \sinh y \tag{31}$$
$$\cosh(x+y) = \sinh x \sinh y + \cosh x \cosh y$$

We can also define $\tanh x$:

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{\frac{e^x - e^{-x}}{2}}{\frac{e^x + e^{-x}}{2}} = \frac{e^x - e^{-x}}{e^x + e^{-x}};$$

and its derivative:

$$\frac{\mathrm{d}}{\mathrm{d}x}\tanh x = \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\sinh x}{\cosh x}\right] = \frac{\sinh x \sinh x - \cosh x \cosh x}{\cosh^2} = \frac{\sinh^2 x - \cosh^2 x}{\cosh^2 x} = \frac{1}{\cosh^2 x} = \operatorname{sech}^2 x.$$

4.2.2 Inverse Hyperbolic Functions and their Derivatives

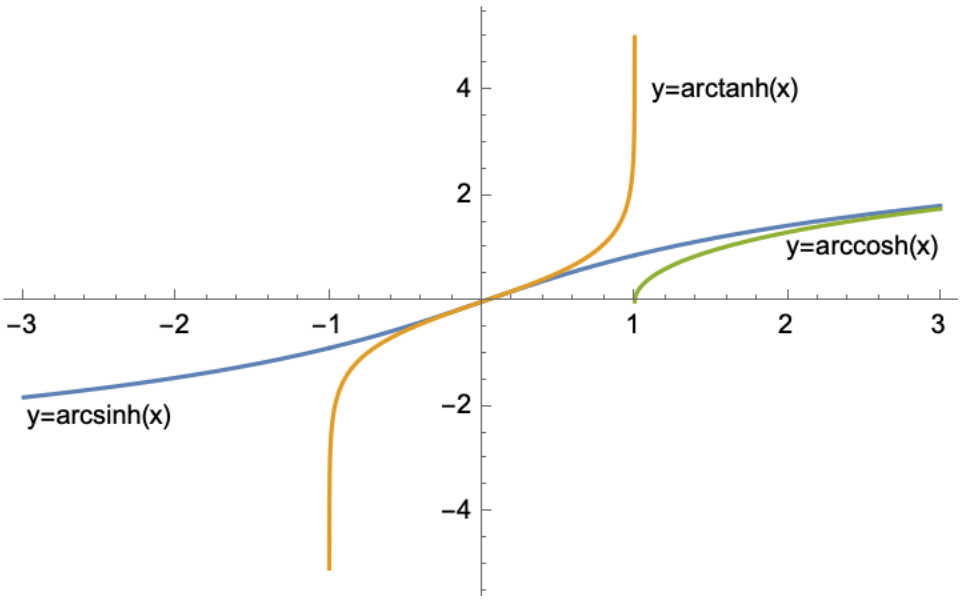


Figure 6: Inverse Hyperbolic Functions

$$
\begin{aligned}
y = \sinh^{-1}(x) \Rightarrow x &= \sinh(y) \\
\frac{\mathrm{d}}{\mathrm{d}x}x &= \frac{\mathrm{d}}{\mathrm{d}x}\sinh(y) \\
1 &= \frac{\mathrm{d}y}{\mathrm{d}x}\cosh(y) \\
\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{1}{\cosh(y)} = \frac{1}{\sqrt{1+\sinh^2(y)}} = \frac{1}{\sqrt{1+x^2}} \\
y = \cosh^{-1}(x) \Rightarrow x &= \cosh(y) \\
\frac{\mathrm{d}}{\mathrm{d}x}x &= \frac{\mathrm{d}}{\mathrm{d}x}\cosh(y) \\
1 &= \frac{\mathrm{d}y}{\mathrm{d}x}\sinh y \\
\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{1}{\sinh y} = \frac{1}{\sqrt{\cosh^2(y)-1}} = \frac{1}{\sqrt{x^2-1}} \\
y = \tanh^{-1}(x) \Rightarrow x &= \tanh(y) \\
\frac{\mathrm{d}}{\mathrm{d}x}x &= \frac{\mathrm{d}}{\mathrm{d}x}\tanh(y) \\
1 &= \frac{\mathrm{d}}{\mathrm{d}y}\tanh(y)\frac{\mathrm{d}y}{\mathrm{d}x} \\
1 &= \operatorname{sech}^2(y)\frac{\mathrm{d}y}{\mathrm{d}x} \\
\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{1}{\operatorname{sech}^2(y)} = \frac{1}{1-\tanh^2(y)} = \frac{1}{1-x^2}
\end{aligned}
\tag{32}
$$

Note that for $\cosh^{-1}(x)$, the domain of $\cosh(x)$ must be restricted to $x >= 0$ so that $\cosh^{-1}(x)$ is a valid one-to-one function. You can see this graphically in figure 6.

4.2.3 Explicitly Expressing Inverse Hyperbolic Functions

Using the rules of natural logarithms, we can re-express each inverse hyperbolic function explicitly in terms of x.

$$
\begin{aligned}
y = \sinh^{-1}(x) \Rightarrow x = \sinh y &= \frac{e^y - e^{-y}}{2} \\
2x &= e^y - e^{-y} \\
2xe^y &= (e^y - e^{-y})e^y \\
2xe^y &= e^{2y} - 1 \\
0 &= e^{2y} - 2xe^y - 1
\end{aligned}
\tag{33}
$$

This is just a quadratic formula in terms of e^y, and can be solved accordingly:

$$\begin{aligned} e^y &= \frac{2x \pm \sqrt{4x^2+4}}{2} \\ e^y &= x + \sqrt{x^2+1} \end{aligned} \tag{34}$$

We don't take the negative, as e^y is always > 0.

$$\begin{aligned} \ln(e^y) &= \ln\Big(x + \sqrt{x^2+1}\Big) \\ y = \sinh^{-1}(x) &= \ln\Big(x + \sqrt{x^2+1}\Big) \end{aligned} \tag{35}$$

A very similar process follows for $\cosh$ and $\tanh$, so it won't be shown, but for reference:

$$\begin{aligned} y = \cosh^{-1} x &= \ln\Big(x + \sqrt{x^2-1}\Big) \\ y = \tanh^{-1} x &= \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right) \end{aligned} \tag{36}$$

5 Complex Numbers

5.1 Complex Quadratic Equations

To introduce the concept of complex numbers, take an arbitrary quadratic equation of the form

$$\alpha x^2 + \beta x + \gamma = 0, \alpha, \beta, \gamma \in \mathbb{R}$$

.

We can solve this for x as follows:

$$x = \frac{-\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}$$

.

If the **discriminant**, $\beta^2 - 4\alpha\gamma$, is < 0, then the equation has no real roots. However, it does have two imaginary roots, of the form

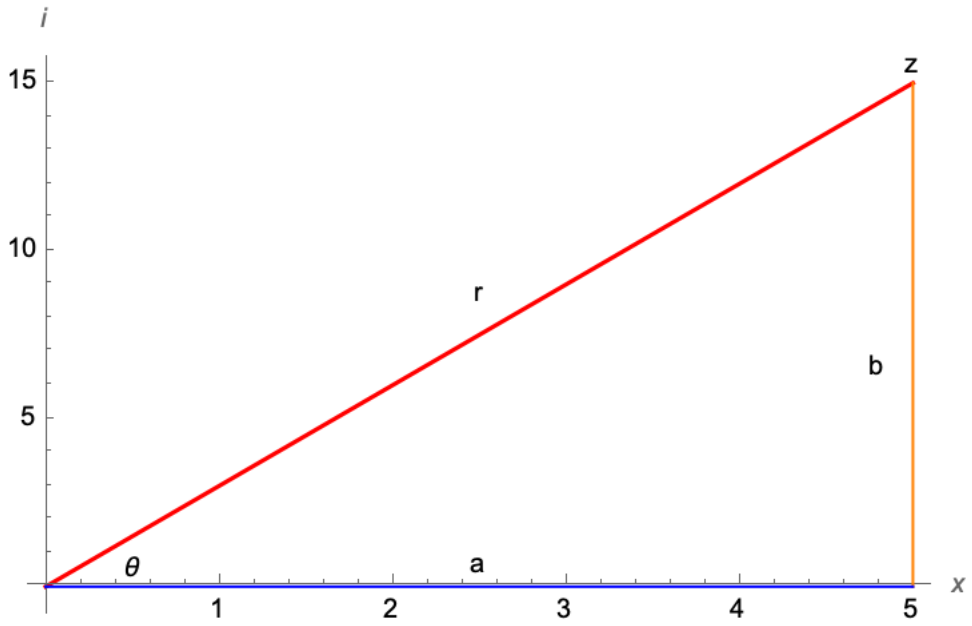


Figure 7: The complex plane

$$\begin{aligned} x &= \frac{-\beta \pm \sqrt{(-4\alpha)(-\beta^2)}}{2\alpha} \\ &= \frac{-\beta}{2\alpha} \pm i\frac{\sqrt{4\alpha\gamma - \beta^2}}{2\alpha} \\ &= a \pm ib; a, b \in \mathbb{R}. \end{aligned} \tag{37}$$

This final, simplified form, $a \pm ib$, is also known as "complex conjugates", which has one part (a) that is real, and one part (ib) that is imaginary. Generally, this is written as $z = a + ib$

5.2 De Moivre's Formula and Related

Take the equation from the previous section, $z = a + ib$. Figure 7 shows this formula in the complex plane, with r representing the distance between z and the origin $(0, 0)$.

Using this visualization, we can rewrite the formula in the polar form

$$z = r\cos\theta + ir\sin\theta$$

.

We can use **Euler's formula** to simplify this;

$$\begin{aligned} z &= r(\cos\theta + i\sin\theta) \\ z &= re^{i\theta} \end{aligned} \tag{38}$$

From this definition, we can define several other properties to aid in solving equations involving complex numbers.

- **Products:** $z_1 z_2 = r_1 e^{i\theta_1} r_2 e^{i\theta_2} = r_1 r_2 e^{i(\theta_1+\theta_2)}$
- **Quotients:** $\frac{z_1}{z_2} = \frac{r_1}{r_2} e^{i(\theta_1-\theta_2)}$
- **Powers:** $z^n = r^n e^{in\theta}, n \in \mathbb{Z}$

Example: calculate $(-1-i)^{20}$. See figure 8 for a visual representation.

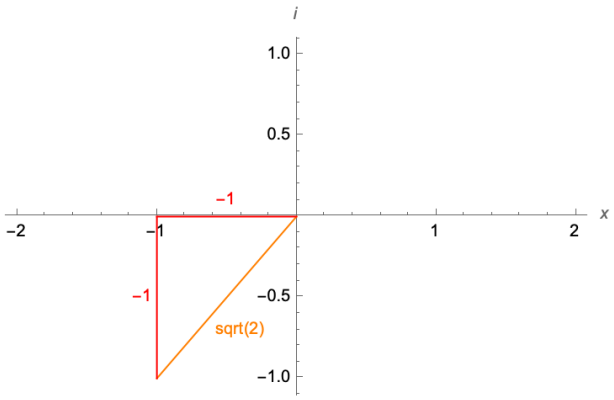


Figure 8: $(-1-i)$ represented on the complex plane

$$\begin{aligned} (-1-i)^{20} = z^n &= r^n e^{in\theta} \\ &= \sqrt{(2)}^{20} e^{i20*\frac{5\pi}{4}} \\ &= 2^{10} e^{25\pi i} \\ &= 1024 e^{24\pi i} e^{\pi i} \\ &= 1024 * (1)(-1) \\ &= -1024 \end{aligned} \tag{39}$$

Note that the second to last step is done using Euler's Identity; specifically:

$$\begin{aligned} e^{i\pi} &= -1 \\ e^{ni\pi} &= e^{(i\pi)^n} = -1^n, n \in \mathbb{Z} \\ n &= \{0, 2, 4, ...\}, e^{ni\pi} = 1 \\ n &= \{1, 3, 5, ...\}, e^{ni\pi} = -1 \end{aligned} \tag{40}$$

5.2.1 Roots

We can use some of the identities above to find a formula to find the real AND complex roots of number:

$$\begin{aligned} z &= re^{i\theta} = re^{i(\theta+2k\pi)}, k \in \mathbb{Z} \geq 0 \\ z^{\frac{1}{n}} &= z^{\frac{1}{n}} e^{\frac{i(\theta+2k\pi)}{n}}, k = 0, 1, \ldots, (n-1) \end{aligned} \tag{41}$$

For example, to find all of the cube roots of 8:

$$\begin{aligned} z &= 8 = 8e^{2k\pi i} \\ z^{\frac{1}{3}} &= 8^{\frac{1}{3}} = 8^{\frac{1}{3}} e^{\frac{2k\pi i}{3}}, k = 0, 1, 2 \\ z_{k=0} &= 8^{\frac{1}{3}} e^{0} = 2 \\ z_{k=1} &= 8^{\frac{1}{3}} e^{\frac{2\pi i}{3}} = 2e^{\frac{2\pi}{3}} = 2[cos\frac{2\pi}{3} + i \sin\frac{2\pi}{3}] = 2[-\frac{1}{2} + i\frac{\sqrt{(3)}}{2}] = -1 + i\sqrt{3} \\ z_{k=2} &= 8^{\frac{1}{3}} e^{\frac{4\pi i}{3}} = ... = -1 - i\sqrt{3} \end{aligned} \tag{42}$$

Thus, the cube roots of 8 are $2, -1 + i\sqrt{3}, -1 - i\sqrt{3}$.

Another method of solving for the roots of a complex number (which can also be extended to other problems involving complex numbers, but is most obvious when calculating roots) involves a little more intuition rather than just using the polar form immediately.

Say you are solving for the square roots of $-8 - 15i$, ie $(-8 - 15i)^{\frac{1}{2}}$. We can then say:

$$\begin{aligned} (-8 - 15i)^{\frac{1}{2}} &= a + bi, a, b \in \mathbb{R} \\ -8 - 15i &= (a + bi)^2 = a^2 + 2abi - b^2 = a^2 - b^2 + 2abi \\ &\Rightarrow -8 = a^2 - b^2 \\ &\Rightarrow -15 = 2ab \end{aligned} \tag{43}$$

Logically, the roots of a complex number involve some real part (a) and some imaginary part (bi). Using this knowledge, we can use the above steps to find a and b using more basic algebra. Doing the work out fully will result in a quadratic formula for a and b, which can finally be solved for the values of the roots (the $\pm$ in the quadratic formula allows for the creation of two roots).

5.2.2 Logarithms

We can use some of the identities above to also find a log involving complex numbers:

$$\begin{aligned} z &= re^{i\theta} \\ \ln(z) &= \ln\left(re^{i\theta}\right) \\ &= \ln(r) + \ln\left(e^{i\theta}\right) \\ &= \ln(r) + i\theta \end{aligned} \tag{44}$$

For example, to find $\ln(-1)$:

$$\begin{aligned} \ln(-1) &= \ln(1) + i(\pi + 2n\pi) \\ &= i(\pi + 2n\pi), n = 0, \pm 1, \pm 2, ... \end{aligned} \tag{45}$$

6 Applications of Differentiation

6.1 Related Rates

"Related rates" problems are very common, and often follow very similar patterns that can be exploited to make individual problems very straightforward. Generally:

1. Draw a picture to represent the scenario
2. Write a formula involving the variable whose change is given, in relation to the variable whose change you're solving for
   (a) Many times, said formula will have to be simplified/rewritten in terms of a single variable. This often involves similar triangles in geometric questions.
3. Differentiate, then solve accordingly
   (a) This pretty much always involves implicit differentiation

The best way to approach these questions is to simply practice. The patterns between questions will quickly become obvious.

6.2 Rolle's Theorem

Theorem (Rolle's Theorem)**.** *If a function f*

- *is continuous on* $[a, b]$
- *is differentiable on* (a, b)
- $f(a) = f(b)$

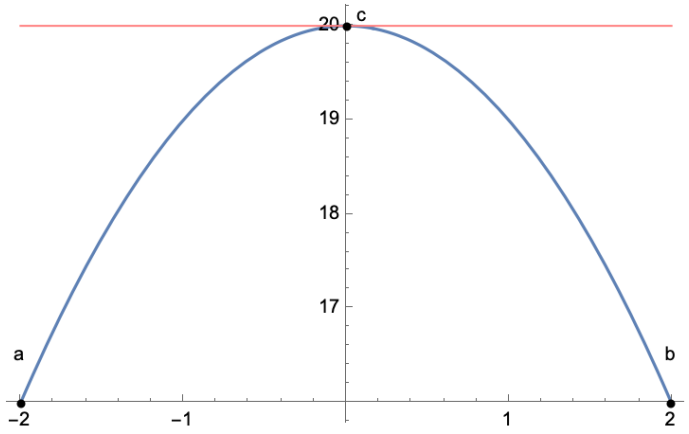


Figure 9: Rolle's Theorem visualized

then there is a number c *in* (a, b) *such that* $f'(c) = 0$.

The proof for this theorem is fairly straightforward (and intuitive), and also makes a lot of sense graphically (see figure 9).

6.3 Mean Value Theorem

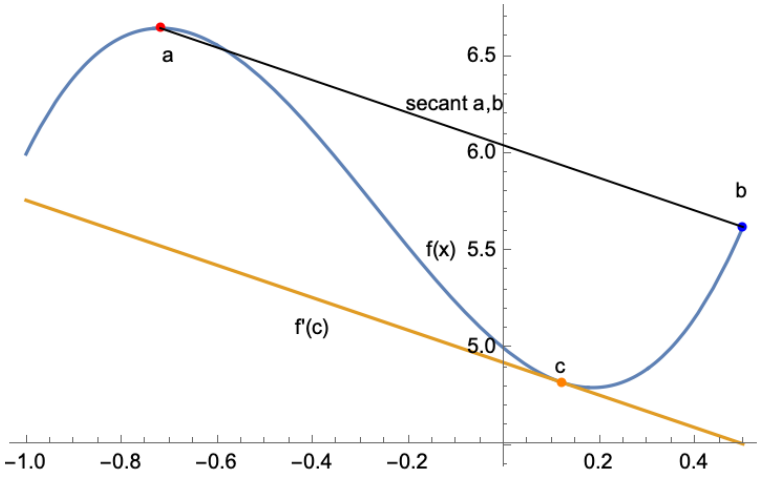


Figure 10: The Mean Value Theorem visualized

Theorem (Mean Value Theorem)**.** *If a function* f *is*

- *continuous on* $[a, b]$
- *differentiable on* (a, b)

then there exists a c *in* (a, b) *such that* $\frac{f(b)-f(a)}{b-a} = f'(c)$. *See figure 10.*

This theorem is essentially an application of Rolle's Theorem, but for any a, b that are not necessarily equal. To prove it, using Rolle's Theorem:

Proof. Take a function $f(x)$ that is continuous on the closed interval $[a, b]$, and differentiable on the open interval (a, b). We can define $g(x)$ as the secant line between the two points $(a, f(a))$ and $(b, f(b))$ as $g(x) =$

$(\frac{f(b)-f(a)}{b-a})(x-a)+f(a)$. We can then define a function $h(x)$ as the difference between the secant line and the original function, $h(x)=g(x)-f(x)$.

We can then compute $h(a)$ and $h(b)$ respectively:

$$\begin{aligned}
h(a) &= g(a)-f(a)\\
&= (\frac{f(b)-f(a)}{b-a})(a-a)+f(a)-f(a)\\
&= 0\\
h(b) &= g(b)-f(b)\\
&= (\frac{f(b)-f(a)}{b-a})(b-a)+f(a)-f(b)\\
&= (\frac{f(b)-f(a)}{b-a})(b-a)\\
&= 0
\end{aligned}$$

Since $h(a)=h(b)=0$, and $h(x)$ is simply the difference between a linear function and an arbitrary continuous function on $[a,b]$ and is thus continuous, we can say that, by Rolle's Theorem, there exists a $c\in(a,b)$ such that $h'(c)=0$. We can thus say that:

$$\begin{aligned}
0=h'(c) &= g'(c)-f'(c)\\
&= \frac{f(b)-f(a)}{b-a}-f'(c)\\
f'(c) &= \frac{f(b)-f(a)}{b-a}
\end{aligned}$$

□

This, and Rolle's, are useful in questions involving proving the number of roots in an equation, as well as proving other theorems.

6.4 Optimization

Optimization questions are (*at least to me*) very similar to related rates questions. In both, you have to find some sort of formula representing a situation, then differentiating. However, for optimization questions, the intention is to solve for a min/max value, by finding when the derivative ("rate of change") is 0. Logically this should make a lot of sense: when something is not changing ($\frac{\mathrm{d}x}{\mathrm{d}y}=0$), then it has to be either increasing until that point and then begin to decrease, or vice versa. From here, checking the second derivative will reveal whether it was a min or max.

In many situations, you will have to ensure that your min/max fits within some endpoints, which depend on the situation at hand; eg, if you have a geometric question, then your value logically can't be negative.

A summary of general steps:

- Draw a diagram, assign relevant variables, create a formula
- Differentiate the formula, set it equal to 0, and solve for the unknown(s) that satisfy the question
- Check the sign of the second derivative of the formula to ensure that your value really was a min/max
- Check the "endpoints" of the problem (are either logical, or imposed by the wording in the question), and ensure that your min/max falls within that range
  - If it doesn't, this usually means your ACTUAL min/max is one of the endpoints, so be careful for this

6.5 Error and Uncertainty

6.5.1 Uncertainty

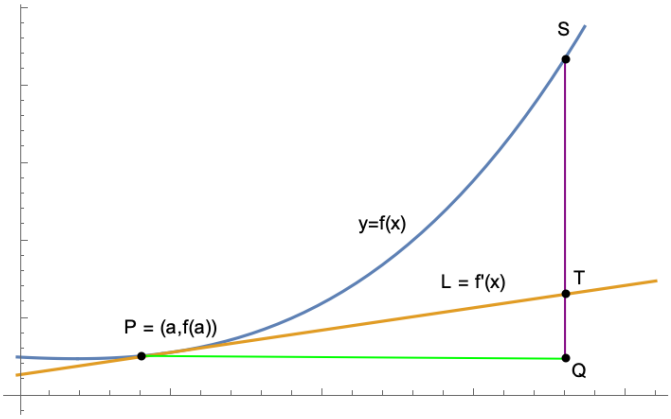


Figure 11: Visual representation of using a differential to approximate error

One very helpful use of derivatives is in estimating how some variable changes when given the relative change in another. This same idea is very similar to the rationale behind using differentials to estimate an error in a measurement, but this will be further explained a little later.

To show how a derivative of a function at a point can be used to estimate its change over a particular range, consider a function $y = f(x)$, as shown in figure 11.

As discussed in 3, we can define L as the linear approximation of $f(x)$ at P as $L = f(a) + (x - a)f'(a)$

As shown in the graph, $PQ = \Delta x = dx$; note that this is the change in x of both $f(x)$ AND L. Conversely, $QS = \Delta y$, the change in $f(x)$, and $QT = dy$, the change in L. In summary:

$$\begin{aligned} L &= f(a) + (x-a)f'(a) \\ PQ &= \Delta x = dx \\ QS &= \Delta(f(x)) = \Delta y \\ QT &= \Delta(L(x)) = dy \end{aligned} \tag{46}$$

From here, we can define θ as the angle between PQ and PT. As such:

$$\begin{aligned} \frac{QT}{PQ} &= \tan\theta = f'(a) \\ \therefore QT &= f'(a)PQ \\ dy &= f'(a)\Delta x = f'(a)dx \\ \therefore QS &\approx QT \therefore \Delta y \approx dy = f'(a)\Delta x = f'(a)dx \end{aligned} \tag{47}$$

As shown, dy is a good approximation for Δy when Δx is relatively small. This reasoning can also be applied to the logic behind why a linear approximation $L(x)$ of a function $f(x)$ at a is most accurate when $x \approx a$.

For example: say you were asked to find $\cos\left(\frac{\pi}{3} + 0.05\right)$. In this case, we can say $f(x) = \cos(x)$, $a = \frac{\pi}{3}$, and $\Delta x = dx = 0.05$. From here:

$$\begin{aligned} f'(x) &= -\sin(x) \\ f'(a) &= -\sin\left(\frac{\pi}{3}\right) = -\frac{\sqrt{3}}{2} \\ dy &\approx f'(a)dx = \frac{-\sqrt{3}}{2} * 0.05 \approx -0.043 \end{aligned} \tag{48}$$

In reality, $\cos\left(\frac{\pi}{3} + 0.05\right) = -0.044$, demonstrating how accurate (and easy-to-compute) this method is.

6.5.2 Error

Using very similar reasoning to above, we can use differentials to find the range of error in a calculation. If given the error (change) of one variable in a formula, you can then use implicit differentiation, and solve for the error (change) in the desired variable.

Oftentimes, "relative change" is also asked for. This is simple $\frac{dx}{x}$.

For example: given a circle with radius $r = 24$, with a max error in measurement of 0.2. To find the max error in the circle's area:

$$\begin{aligned} A &= \pi r^2 = f(r) \\ \Delta A \approx \mathrm{d}A &= f'(r)dr \\ &= 2\pi r dr \\ &= 2\pi(24)(0.2) = 9.6\pi \approx 30.16 \\ \textit{rel. error:} \frac{\Delta A}{A} &\approx \frac{\mathrm{d}A}{A} = \frac{1}{60} \end{aligned} \tag{49}$$

6.6 Newton's Method

Newton's Method is an application of derivatives that can be used to estimate the solutions/roots of a function. The steps to the method are as follows, for a function $f(x)$:

1. Pick some x_1, then find the point $P = (x_1, f(x_1))$
2. Find the tangent line of $f(x)$ at the point P, defined by $y - f(x_1) = f'(x_1)(x - x_1)$
3. Find the point x_2 where the tangent line intersects the x-axis, by setting $y = 0$:

$$\begin{aligned} 0 &= f(x_1) + f'(x_1)(x_2 - x_1) \\ x_2 &= x_1 - \frac{f(x_1)}{f'(x_1)} \end{aligned} \tag{50}$$

4. Repeat steps 1-3 for increasing x_n and x_{n+1}. As n increases, the x_n will approach a root of $f(x)$. The general formula follows:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

See figure 12 to see this process visually.

6.7 Derivative-Related Theorems

Following are a number of theorems related to derivatives, and limits. While many of them are already described above, some are not.

These theorems have various applications in determining limits and derivatives and can aid in graphing functions.

1. **Intermediate Value Theorem**

   If f is continuous on $[a, b]$, let N be any number between $f(a)$ and $f(b)$, where $f(a) \neq f(b)$, there exists a c in (a, b) such that $f(c) = N$.

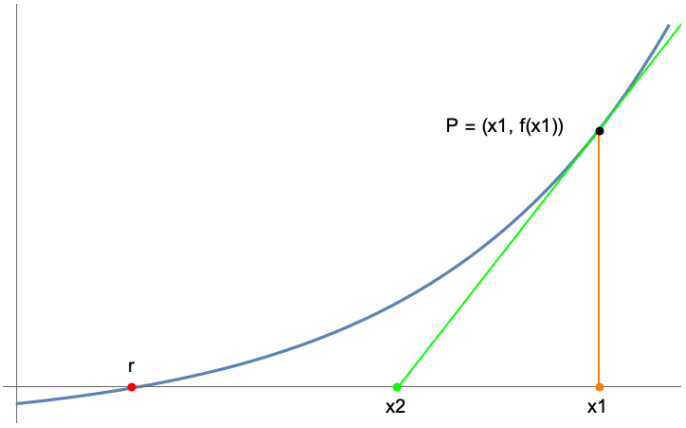


Figure 12: Newton's Method Visualized

2. **Extreme Value Theorem**

   If f is continuous on $[a, b]$, then f attains a minimum and maximum value (extrema) on that interval. *Not that these extrema may occur at a or b; make sure to check this case when answering questions using this theorem.*

3. **Fermat's Theorem**

   If f has a local extrema at c, and $f'(c)$ exists, then $f'(c) = 0$.

   Note that the inverse of this theorem is not necessarily true: if $f'(b) = 0$, then b is not necessarily a local extrema.

4. **Rolle's Theorem**

   If f is continuous on $[a, b]$, f' is continuous on (a, b), and $f(a) = f(b)$, then there exists a c in (a, b) such that $f'(c) = 0$.

6.8 Graph Sketching

Using calculus (and other properties of functions), we can sketch curves with fairly high accuracy, without needing calculators. To do so, there is a "shopping list" to follow, to graph a function $f(x)$:

1. **Domain**

   Find all x for which $f(x)$ is defined. If $f(x)$ is undefined for some x, you should also find how f behaves as it approaches x (is it a hole? An asymptote? Or does the function simply only exist on a closed interval?)

2. **Intercepts**

   Find both the x and y intercepts, i.e. the ordered pair $(x, f(x))$ at $y = 0$ and $x = 0$ respectively.

3. **Symmetry (even/odd)**

   - $f(x) = f(-x)$: $f(x)$ is **even**

- $f(-x) = -f(x)$: $f(x)$ is **odd**

Graphically, an even function is symmetrical over the y axis, and an odd function is symmetrical over $y = x$.

4. **Asymptotes**

   Determine $\lim_{x\to\infty} f(x)$ and $\lim_{x\to-\infty}$, to determine whether the function has any **horizontal asymptotes**. In addition, find the limit of $f(x)$ to any holes in the domain, from both right and left, to determine if the function has any **vertical asymptotes**.

5. **Intervals of Increase/Decrease & Local Min/Max**

   Find $f'(x)$, and find when $f'(x)$ is greater/less than 0 to find when $f(x)$ is increasing/decreasing respectively. The points where $f'(x) = 0$ are also the place where the function has a local min/max, AS LONG AS $f'(x)$ is not an inflection point...

6. **Concavity & Points of Inflection**

   Find $f''(x)$, and find when $f''(x)$ is greater/less than 0 to find when $f(x)$ is concave up/down respectively. The points where $f''(x) = 0$ are also the inflection points of $f(x)$.

7 Integration

7.1 Defining Antiderivatives

For a function $f(x)$, we can say that:

$$\int f(x)\,\mathrm{d}x = F(x) + k$$

,

where $\frac{\mathrm{d}}{\mathrm{d}x}[F(x) + k] = f(x)$, and k is a constant.

For example: $\int x^2\,\mathrm{d}x = \frac{x^3}{3} + k$.

You can (and should) think of $\int\,\mathrm{d}x$ and $\frac{\mathrm{d}}{\mathrm{d}x}$ as *inverse operators* of one another, hence the term "antiderivative". As such, numerous derivative rules have corresponding rules for integrals, which are discussed later. Also note that this type of integration is called *indefinite* because it defines the integral of a function over an *indefinite range*, ie, for all x. *Definite integrals* will be discussed later and have their own very particular applications, and define an integral $\int_m^n$ over a range $[m, n]$.

7.2 Techniques of Integration

7.2.1 Known Formulas of Derivatives

eg) $\int \frac{1}{1+x^2}\,\mathrm{d}x = \arctan x + c$

"Simply" recognize that the integral to solve for is the derivative of a common function. Most other techniques of integration involve simplifying/modifying a given integral into this type of easy-to-solve form.

7.2.2 By Substitution

For many integrals, we can substitute part of a complex expression for a single variable (typically u; this technique is often shorthanded "u-sub"), and, following some manipulation, find an integral much more easily. The best way to understand this method is through an example:

$$\int \sqrt{x^3+2}x^2\,\mathrm{d}x$$

We can let $u = x^3 + 2$, and differentiating u in terms of x, we get

$$\begin{aligned} \frac{\mathrm{d}u}{\mathrm{d}x} &= 3x^2 \\ du &= 3x^2\mathrm{d}x \\ x^2\mathrm{d}x &= \frac{\mathrm{d}u}{3} \end{aligned} \tag{51}$$

From here, we can substitute our expressions back into the original integral, and solve:

$$\begin{aligned} \int x^2\sqrt{x^3+2}\,\mathrm{d}x &= \int u^{\frac{1}{2}}\frac{\mathrm{d}u}{3} \\ &= \frac{1}{3}\frac{u^{\frac{3}{2}}}{\frac{3}{2}} + k \\ &= \frac{2(x^3+2)^{\frac{3}{2}}}{9} + k \end{aligned} \tag{52}$$

Note that solving this integral involved using the inverse of the "power rule" of derivatives, which should hopefully be intuitive.

7.2.3 By Parts

Given two functions u and v, we can write:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}(uv) &= \frac{\mathrm{d}u}{\mathrm{d}x}v + u\frac{\mathrm{d}v}{\mathrm{d}x} \\
\int (\frac{\mathrm{d}}{\mathrm{d}x}(uv) &= \frac{\mathrm{d}u}{\mathrm{d}x}v + u\frac{\mathrm{d}v}{\mathrm{d}x})\mathrm{d}x \\
\int \frac{\mathrm{d}}{\mathrm{d}x}(uv)\,\mathrm{d}x &= \int \left[v\frac{\mathrm{d}u}{\mathrm{d}x}\right]\,\mathrm{d}x + \int \left[u\frac{\mathrm{d}v}{\mathrm{d}x}\right]\,\mathrm{d}x \\
uv &= \int v\mathrm{d}u + \int u\,\mathrm{d}v \\
\int u\mathrm{d}v &= uv - \int v\,\mathrm{d}u
\end{aligned}
\tag{53}
$$

For example, to calculate the integral of xe^{ax}:

$$
\begin{aligned}
&\int xe^{ax}\,\mathrm{d}x \\
&\qquad x = u, \mathrm{d}u = \mathrm{d}x \\
&\qquad e^{ax}\mathrm{d}x = \mathrm{d}v, v = \frac{e^{ax}}{a} \\
&\qquad\qquad \int xe^{ax}\,\mathrm{d}x = x\frac{e^{ax}}{a} - \int \frac{e^{ax}}{a}\,\mathrm{d}x \\
&\qquad\qquad = \frac{xe^{ax}}{a} - \frac{e^{ax}}{a^2} + C
\end{aligned}
\tag{54}
$$

Generally, you want to find some u such that du becomes "simpler", and a dv such that v does not get "too much more complicated".

Another integral that is commonly computed using integration by parts is as follows:

$$
\begin{aligned}
&\int \ln x\,\mathrm{d}x \\
&\qquad\qquad = x\ln x - \int x\frac{\mathrm{d}x}{x} \\
&\qquad\qquad = x\ln x - x + C
\end{aligned}
\tag{55}
$$

7.2.4 Trigonometric Integrals

Finding the integral of a combination of trigonometric functions is largely a question of using trig. identities in such a way that the integral becomes easier to solve using more "basic" methods.

While a lot of this becomes intuitive over time, some general guidelines are as follows:

- $\int \sin^m(x)\cos^n(x)\,\mathrm{d}x$

If n **is odd (**$n = 2k + 1$**)**, rewrite all $\cos$, except one, with $\cos^2(x) = 1 - \sin^2(x)$.

$$\begin{aligned} \int \sin^m(x) \cos^{2k+1}(x) \, \mathrm{d}x &= \int \sin^m(x)(\cos^2 x)^k \cos(x) \, \mathrm{d}x \\ &= \int \sin^m(x)(1 - \sin^2(x))^k \cos(x) \, \mathrm{d}x \end{aligned} \tag{56}$$

From here, substitute $u = \sin(x)$ and solve accordingly.

Very similarly, **if** m **is odd (**$m = 2k + 1$**)**:

$$\begin{aligned} \int \sin^{2k+1}(x) \cos^n(x) \, \mathrm{d}x &= \int (\sin^2(x))^k \cos^n(x) \sin(x) \, \mathrm{d}x \\ &= \int (1 - \cos^2 x)^k \cos^n(x) \sin^n(x) \, \mathrm{d}x \end{aligned} \tag{57}$$

Let $u = \cos(x)$, etc..

If both m **and** n **are even**, you can use the half-angle theorems.

- $\int \tan^m(x) \sec^n(x) \, \mathrm{d}x$

  Very similar to the strategy for $\sin$ and $\cos$: if n is even, substitute $\sec^2(x) = 1 + \tan^2(x)$ for all $\sec$ except one, and the opposite for $\tan^2(x)$ if m is even.

7.2.5 Trigonometric Substitution

Using a strategy similar to "u-substitution", we can carefully replace variables in an integral with trigonometric functions to make the integral easier to solve.

In general, we can say:

$$\int f(x) \, \mathrm{d}x = \int f(g(t))g'(t) \, \mathrm{d}t$$

There are three main forms to look out for in integrals that can be used to effectively substitute trigonometric functions. Notice that, in each case, a form similar to a common trigonometric identity is present.

- $\sqrt{a^2 - x^2} \Rightarrow x = a \sin\theta$
- $\sqrt{a^2 + x^2} \Rightarrow x = a \tan\theta$
- $\sqrt{x^2 - a^2} \Rightarrow x = a \sec\theta$

Using this strategy makes a lot more sense when used in an example; say $\int \frac{\sqrt{9-x^2}}{x^2} \, \mathrm{d}x$. This is of the first form

listed above, so we can substitute $x = 3\sin\theta$, meaning $dx = 3\cos\theta d\theta$. Assume $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$

$$\begin{aligned}
\int \frac{\sqrt{9-x^2}}{x^2}\,\mathrm{d}x &= \int \frac{\sqrt{9-9\sin^2(\theta)}}{9\sin^2(\theta)} 3\cos(\theta)\,\mathrm{d}\theta \\
&= \int \frac{\sqrt{9\cos^2(\theta)}}{9\sin^2(\theta)} 3\cos(\theta)\,\mathrm{d}\theta \\
&= \int \frac{3\cos(\theta)}{9\sin^2(\theta)} 3\cos(\theta)\,\mathrm{d}\theta \\
&= \int \frac{\cos^2)(\theta)}{\sin^2(\theta)}\,\mathrm{d}\theta = \int \cot^2(\theta)\,\mathrm{d}\theta \\
&= \int (\csc^2(\theta) - 1)\,\mathrm{d}\theta = -\cot(\theta) - \theta + C
\end{aligned} \tag{58}$$

From here, the integral can be rewritten in terms of x using some trig rules (drawing a diagram of a triangle can help here):

$$-\cot(\theta) - \theta + C = -\frac{\sqrt{9-x^2}}{x} - \sin^{-1}(\frac{x}{3}) + C$$

7.3 The Fundamental Theorem of Calculus

Consider a function $f(x)$ that is continuous and non-negative over $a \leq x \leq b$. We can divided (a, b) into subintervals of width $\Delta x = \frac{b-a}{N}$, and consider rectangles of area $f(x_i^*)\Delta x$, where x_i^* is the midpoint of the ith subinterval. The area of all of these integrals can be written

$$\sum_{i=1}^{N} f(x_i^*)\Delta x.$$

Taking the limit of this sum ($\lim_{N\to\infty}$) represents subdividing (a, b) into infinitely many subintervals (and more accurately approximating the area under the curve). If this limit converges, we say

$$\lim_{N\to\infty} \sum_{i=1}^{N} f(x_i^*)\Delta x = \int_a^b f(x)\,\mathrm{d}x.$$

Specifically, $f(x)$ must be *piecewise* continuous over (a, b); it can have (finite) jump discontinuities (over a finite interval), but does not have to be continuous nor differentiable.

Consider the integral

$$\int_a^x f(t)\,\mathrm{d}t = g(x).$$

We can say that, for some small h, $g(x + h) - g(x) \approx f(x^*)h$, where x^* is the midpoint of $(x, x + h)$. Thus,

$f(x^*) \approx \frac{g(x+h)-g(x)}{h}$. Approaching h to 0, we have

$$\lim_{h\to 0} f(x^*) = \lim_{h\to 0}\left[\frac{g(x+h)-g(x)}{h}\right] = g'(x).$$

Thus,

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(t)\,\mathrm{d}t = f(x).$$

7.4 Applications of Integrals

7.4.1 Area Under a Curve

In short, the area under $f(x)$ from a to b is

$$\int_a^b f(x)\,\mathrm{d}x,$$

assuming $f(x)$ is piecewise continuous over (a,b).

7.4.2 Area Between Curves

The area between two curves $f(x)$ and $g(x)$ from a to b is

$$\int_a^b f(x)\,\mathrm{d}x - \int_a^b g(x)\,\mathrm{d}x = \int_a^b f(x) - g(x)\,\mathrm{d}x,$$

assuming $f(x) \geq g(x)$ over (a,b) (among the other conditions from FTC). If, on the other hand, f is less than g at some point in (a,b), we must do the opposite subtraction in this range. Say that $f > g$ on (a,c), and $f < g$ on (c,b), we can then write the are between the two curves as

$$\int_a^c f(x) - g(x)\,\mathrm{d}x + \int_c^b g(x) - f(x)\,\mathrm{d}x.$$

This can be rewritten any number of ways using properties of integrals (eg, $\int_a^b f(x)\,\mathrm{d}x = \int_b^a f(x)\,\mathrm{d}x$). It can also be generalized to any number of "switches" in the inequality.

7.4.3 Curve Length

Take a curve $f(x)$ on the interval (a,b), and two points $P_i = (x_i, f(x_i))$ and $P_{i-1} = (x_{i-1}, f(x_{i-1}))$ in this interval. The length of the section of curve between these two points is (very) approximately equal to the length of the secant line between them, with this approximation improving the closer the points are. It follows that we can take the sum of "all possible" values of these these points P_i between a and b to approximate the length of f on (a,b); thus, we can say

$$L \approx \lim_{N\to\infty} \sum_{i=1}^{N} |P_{i-1}P_i|.$$

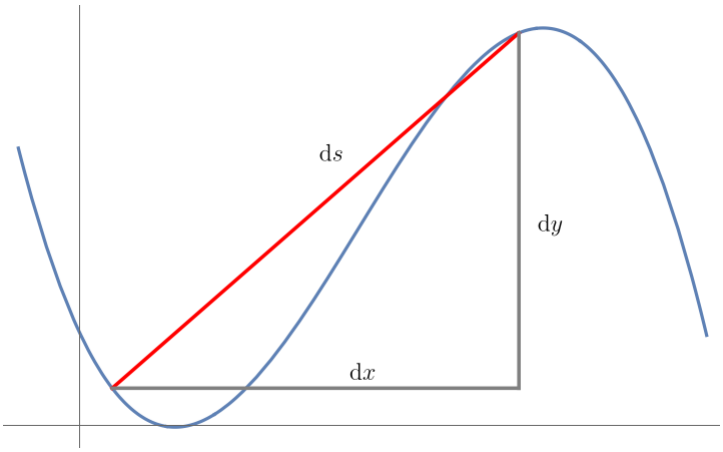


Figure 13: Visualization of finding the curve length.

The secant length between P_i and P_{i-1} can be rather simply rewritten using the standard distance formula;

$$\begin{aligned}|P_iP_{i-1}| &= \sqrt{(x_i - x_{i-1})^2 + (f(x_i) - f(x_{i-1}))^2} \\ &= \sqrt{1 + (\frac{f(x_i) - f(x_{x-1})}{x_i - x_{i-1}})^2}(x_i - x_{i-1}) \\ &= \sqrt{1 + f'(x^*)^2}(x_i - x_{i-1}) = \sqrt{1 + f'(x^*)^2}\,\Delta x\end{aligned}$$

Where $\Delta x = x_i - x_{i-1}$; notice that $f'(x^*)$ is simply the slope of the secant line, and x^* is the value of x in the interval (x_i, x_{i-1}) with the same slope of the secant (which must exist by Mean Value Theorem). Substituting this into the original limit;

$$\begin{aligned}L &= \lim_{N\to\infty} \sum_{i=1}^{N} \sqrt{1 + (f'(x^*))^2}\,\Delta x \\ &= \int_a^b \sqrt{1 + (f'(x))^2}\,\mathrm{d}x\end{aligned}$$

To try to understand this more intuitively and in terms of differentials, consider the length of the curve at a particular, infinitesimally small range to be $\mathrm{d}s$, ie, the "instantaneous change in length". At this small scale, it can be considered to be the hypotenuse of a right triangle made up of sides $\mathrm{d}x$ and $\mathrm{d}y$, the changes in x and y respectively. Thus, we can say $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 \implies \mathrm{d}s = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} = \sqrt{1 + \frac{(\mathrm{d}y)^2}{(\mathrm{d}x)^2}}\,\mathrm{d}x$. Summating this over an entire interval yields the same integral as above.

7.4.4 Volume of Revolution

Consider a function $f(x)$, which is rotated about the x-axis to form a solid. We can subdivide this solid into disks of area $A = \pi r^2 = \pi f(x_i)^2$. The width of each particular disk is a small interval $\Delta x = x_i - x_{i-1}$, and we

can thus estimate the volume of the entire solid (between two points a, b) as

$$V \approx \lim_{N\to\infty} \sum_{i=1}^{N} \pi(f(x_i))^2 \Delta x = \int_a^b \pi(f(x))^2 \,\mathrm{d}x$$

7.4.5 Surface Area of Revolution

Consider a function $f(x)$ rotated about the x-axis to form a solid. Subdividing this solid into disks, the surface area of each disk is $2\pi rh = 2\pi f(x_i)\Delta s$, where Δs is the change in the length of the curve between x_i and x_{i-1}. Using the definitions of curve length from earlier, we can write

$$\begin{aligned} SA &= \lim_{N\to\infty} \sum_{i=1}^{N} 2\pi f(x_i)\Delta s \\ &= \int_a^b 2\pi f(x_i)\,\mathrm{d}s \\ &= 2\pi \int_a^b f(x_i)\sqrt{1 + (\frac{\mathrm{d}y}{\mathrm{d}x})^2}\,\mathrm{d}x \end{aligned}$$

7.4.6 Gamma Function

The **Gamma Function** is defined

$$\Gamma(n) = \int_0^\infty e^{-x}x^{n-1}\,\mathrm{d}x,$$

and is a generalization of the factorial operator, i.e., $\Gamma(n+1) = n\Gamma(n)$. The Gamma function appears in many contexts (such as statistics), but one "helpful" (situationally) use case is the fact

$$\int_0^{\pi/2} \sin^{2m-1}\theta \cos^{2n-1}\theta \,\mathrm{d}\theta = \frac{\Gamma(n)\Gamma(m)}{2\Gamma(n+m)}; m, n > 0,$$

which is proven in 10.3.2.

8 Differential Equations

8.1 Definitions

A **differential equation** is an equation that contains an unknown function and some of its derivatives. We can more specifically describe differential equations by their **order**, which is the number of the highest derivative in the equation, ie, a first-order differential equation is of the form $F(x, y, y')$.

8.2 Variables Separable

While not all differential equations are solvable (or at least, easily solvable), one (of many) special cases is of the form:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)g(y) = \frac{f(x)}{h(y)}, g(y) = \frac{1}{h(y)}, h(y) \neq 0$$

This is the form of a function of x times a function of y, indicating the parts are *separable*. We can rewrite this:

$$\begin{aligned} \frac{\mathrm{d}y}{\mathrm{d}x} &= f(x)g(y) \\ dy &= [\frac{\mathrm{d}y}{\mathrm{d}x}]\mathrm{d}x = f(x)g(y)\mathrm{d}x \\ \int \frac{\mathrm{d}y}{g(y)} &= \int \frac{\mathrm{d}x}{f(x)} \end{aligned} \tag{59}$$

From here, integrating both sides yields either y as a function of x, or x as a function of y (which one you solve for depends on the context of the question).

For example, take: $\frac{\mathrm{d}y}{\mathrm{d}x} = x^2y - y + x^2 - 1$. This can be rewritten as a separable equation:

$$\begin{aligned} \frac{\mathrm{d}y}{\mathrm{d}x} &= x^2y - y + x^2 - 1 \\ &= y(x^2-1) + 1(x^2-1) \\ &= (y+1)(x^2-1) \\ \int \frac{\mathrm{d}y}{y+1} &= \int (x^2-1)\,\mathrm{d}x \\ \ln(|y+1|) &= \frac{x^3}{3} - x + k \\ y+1 &= \pm e^{(\frac{x^3}{3}-x+k)} = \pm e^{(\frac{x^3}{3}-x)}e^k \\ y &= Ce^{(\frac{x^3}{3}-x)} - 1, C = \pm e^k \end{aligned} \tag{60}$$

From here, if given $y(0)$, you can solve for the value of C and rewrite the equation appropriately. This would then become what is called an **initial value problem** (IVP); you are given the initial value, after all.

In this example, say $y(0) = 3$. We can write:

$$\begin{aligned} 3 &= Ce^{\frac{(0)^3}{3}-0} - 1 \\ 3 &= Ce^0 - 1 \\ C &= 4 \\ y &= 4e^{\frac{x^3}{3}-x} - 1 \end{aligned} \tag{61}$$

8.3 Mathematical Models

Several real-world scenarios can be represented using differential equations, and the following are some of the more common/useful forms, as well as their general solutions and applications.

I: $\alpha y = \frac{\mathrm{d}y}{\mathrm{d}t}, \alpha \in \mathbb{R}$

When solved for a function y of t, this differential equation becomes a (hopefully) familiar form;

$$\begin{aligned} \frac{\mathrm{d}y}{\mathrm{d}t} &= \alpha y \\ \int \frac{\mathrm{d}y}{y} &= \int \alpha \mathrm{d}t \\ \ln|y| &= \alpha t + C \\ y &= \pm e^{(\alpha t + C)} = \pm e^C e^{\alpha t} = ke^{\alpha t}, k = \pm e^C \end{aligned} \tag{62}$$

If we say $y(0) = y_0$, then we can rewrite $y(t) = y_0 e^{\alpha t}$. This is the typical form used to represent simple **exponential growth/decay**; which of these two the function represents depends on α:

- $\alpha > 0$;**growth:** used in scenarios such as simple population models.
- $\alpha < 0$; **decay:** used in scenarios such as radioactive decay.

See figure 14 for a visual representation of these two.

II: $\frac{\mathrm{d}y}{\mathrm{d}t} = k(N - y)$

This general form of the equation is used in a number of scenarios, such as Newton's Law of cooling/heating and modeling deprecation. Solving this differential for y as a function of t goes as follows:

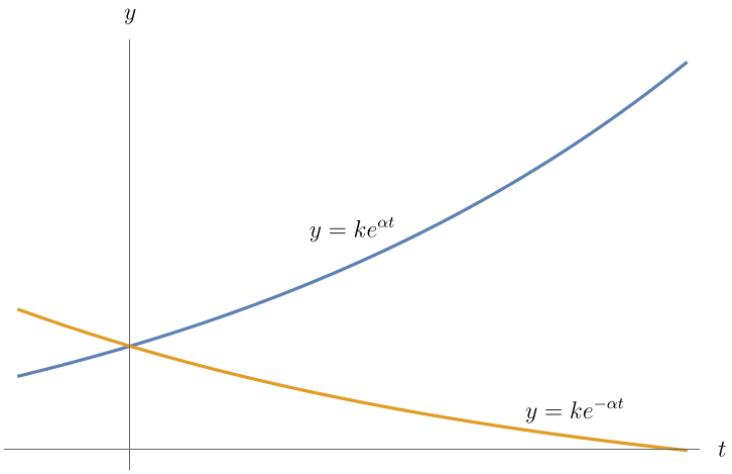


Figure 14: Exponential growth and decay

$$\begin{aligned} \frac{\mathrm{d}y}{\mathrm{d}t} &= k(N-y) \\ \frac{\mathrm{d}y}{N-y} &= k\mathrm{d}t \\ \int \frac{\mathrm{d}y}{N-y} &= \int k\mathrm{d}t \\ -\ln|N-y| &= kt + C \\ N - y &= \pm e^{-C}e^{-kt} = Ke^{-kt} \\ y &= N - Ke^{-kt} \end{aligned} \tag{63}$$

If we state that $y(0) = y_0$, we can rewrite this equation as $y = N + (y_0 - N)e^{-kt}$. This is the same form that Newton's Law of Cooling takes: $T(t) = T_{\text{env}} + (T(0) - T_{\text{env}})e^{-rt}$.

It's important to note that $y(t)$ approaches N as time approaches infinity, ie:

$$\lim_{t\to\infty}[N + (y_0 - N)e^{-kt}] = N$$

If the value of $y_0 - N$ is positive, then this limit is approached from above (ie, the function is *lower bounded* by N), and if $y_0 - N$ is negative, then the limit is approached from below (*upper bounded* by N). See figure 15 for a visual representation of the difference, assuming a constant y_0.

III: $\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{k}{N}y(N-y)$

This differential equation represents a general form of the **logistic equation**, commonly used in population modeling in ecology and other fields. Solving this differential equation for y as a function of x goes

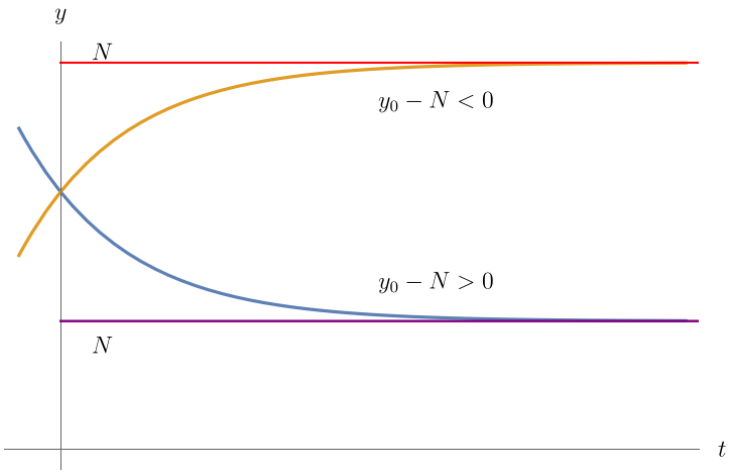


Figure 15: $y = N + (y_0 - N)e^{-kt}, y_0 - N > 0$ vs $y_0 - N < 0$

as follows:

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}t} &= \frac{k}{N}y(N-y) \\
\frac{\mathrm{d}y}{y(N-y)} &= \frac{k\mathrm{d}t}{N} \\
\int \frac{\mathrm{d}y}{y(N-y)} &= \int \frac{k\mathrm{d}t}{N} \\
\int [\tfrac{\frac{1}{N}}{y} + \tfrac{\frac{1}{N}}{N-y}]\,\mathrm{d}y &= \frac{k}{N}t + C \\
\frac{1}{N}(\ln|y| - ln|N-y|) &= \frac{k}{N}t + C \\
\ln|\frac{y}{N-y}| &= kt + D, D = CN \\
\frac{y}{N-y} &= Ke^{kt}, K = \pm e^{D} \\
y = \frac{NKe^{kt}}{1+Ke^{kt}} &= \frac{NKe^{kt}}{1+Ke^{kt}}\frac{\frac{e^{-kt}}{K}}{\frac{e^{-kt}}{K}} \\
&= \frac{N}{\frac{e^{-kt}}{K}+1} = \frac{N}{1+Ae^{-kt}}, A = \frac{1}{K}
\end{aligned} \tag{64}$$

If we say $y(0) = P_0$, or the initial population, then we can simplify further and say:

$$\begin{aligned}
y(0) = P_0 &= \frac{N}{1+b} \\
\implies b &= \frac{N}{P_0} - 1 = \frac{N-P_0}{P_0} \\
\therefore y(t) &= \frac{N}{1+be^{-kt}}, b = \frac{N-P_0}{P_0} \\
y(t) &= \frac{NP_0}{P_0 + (N-P_0)e^{-kt}}
\end{aligned} \tag{65}$$

Just as with Model II, the function approaches N as the time goes to infinity, and as such, we can call N the "carrying capacity" or "max population", depending on the context of the problem.

Another interesting property of the logistic model comes about from exploring its derivatives. Its first derivative is clear, as it is given in the expression of the logistic model as a differential equation. This first derivative ($\frac{dy}{dt}$) is *always positive*, which should be clear from inspection of the original formula for $\frac{dy}{dt}$, as both N and y are always positive, and N must be greater than y.

Recall that the second derivative represents the derivative of the first derivative, and in the context of population growth, for example, represents the change in how fast the population is growing. By extension, the inflection point of the logistic model would represent when the change in the population growth rate is at a maximum (or minimum, but the work below will show that it is indeed a maximum).

We can find this inflection point (IP) by doing the follow:

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}t} &= \frac{k}{N}y(N-y) = ky - \frac{ky^2}{N} \\
\frac{\mathrm{d}^2y}{\mathrm{d}t^2} &= k - \frac{2ky}{N} \\
k - \frac{2ky}{N} &= 0 \\
y &= \frac{N}{2} \\
f(t) &= \frac{N}{2} = \frac{N}{1+be^{-kt}} \\
1 + be^{-kt} &= 2 \\
e^{-kt} &= \frac{1}{b} \\
t &= -\frac{1}{k}\ln\left(\frac{1}{b}\right) = \ln\left((\frac{1}{b})^{-\frac{1}{k}}\right) \\
t &= \ln\left[(\frac{N-P_0}{P_0})^{\frac{1}{k}}\right]
\end{aligned} \tag{66}$$

As such, the point $(\ln[(\frac{N-P_0}{P_0})^{\frac{1}{k}}], \frac{N}{2})$ is an inflection point and represents the maximum change in population growth (proving this is relatively straightforward and is omitted here).

8.4 Linear Equations

A **first-order linear differential equation** is a differential equation of the general form:

$$\frac{\mathrm{d}y}{\mathrm{d}x} + P(x)y = Q(x)$$

,

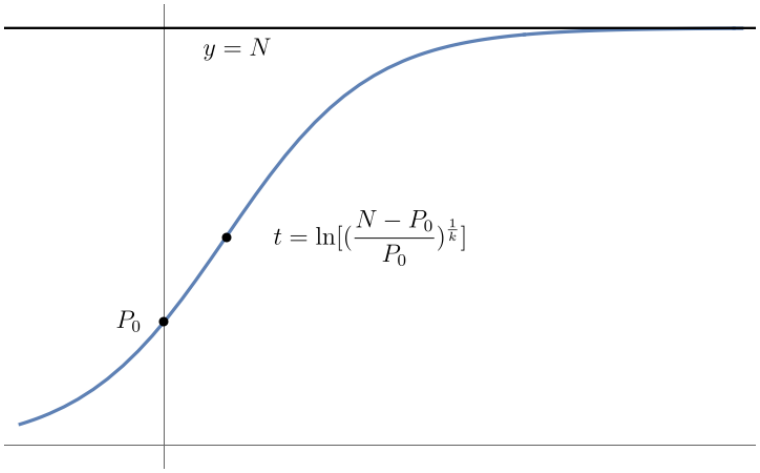


Figure 16: $y(t) = \frac{N}{1+be^{-kt}}, b = \frac{N-P_0}{P_0}$

where P and Q are continuous over a particular interval. Notice that this form *does not* fit the form of a separable equation (see 8.2), and thus is not solvable using the same method.

Instead, we have to use a new function, called an **integrating factor**, called $I(x)$, which will make the equation solvable for y. To understand the motivation behind $I(x)$, recall the product rule for calculating derivatives:

$$(xy)' = xy' + y$$

We can define $I(x)$ as a factor such that the left-hand side of the general equation becomes the derivative of $I(x)y$, and simplifies as follows:

$$\begin{aligned} I(x)(\frac{\mathrm{d}y}{\mathrm{d}x} + P(x)y) = (I(x)y)' &= I(x)Q(x) \\ \int (I(x)y)' &= \int I(x)Q(x)\,\mathrm{d}x \\ I(x)y &= \int I(x)Q(x)\,\mathrm{d}x + C \\ y(x) &= \frac{1}{I(x)}[\int I(x)Q(x)\,\mathrm{d}x + C] \end{aligned} \tag{67}$$

This provides a general solvable form for the equation, but we still need to find $I(x)$ to make this useful. We can do so by rearranging $I(x)(\frac{\mathrm{d}y}{\mathrm{d}x} + P(x)y) = (I(x)y)'$ as follows:

$$\begin{aligned} I(x)(\frac{\mathrm{d}y}{\mathrm{d}x} + P(x)y) = I(x)(y' + P(x)y) &= (I(x)y)' \\ I(x)y' + I(x)P(x)y &= I'(x)y + I(x)y' \\ I(x)P(x)y &= I'(x)y \\ I(x)P(x) &= I'(x) \end{aligned} \tag{68}$$

This is just a separable equation, which can be solved for I:

$$\begin{aligned} \int \frac{\mathrm{d}I}{I} &= \int P(x)\,\mathrm{d}x \\ \ln|I| &= \int P(x)\,\mathrm{d}x \\ I &= \pm e^{\int P(x)\,\mathrm{d}x + C} = Ae^{\int P(x)\,\mathrm{d}x}, A = \pm e^C \end{aligned} \tag{69}$$

We can typically set $A = 1$ for a general I. Using these general steps, we can thus solve a linear first-order differential equation by multiplying it by the appropriate integrating factor $I(x) = e^{\int P(x)\,\mathrm{d}x}$. For instance, to solve $x^2y' + xy = 1$:

$$\begin{aligned} x^2y' + xy &= 1 \\ y' + \frac{1}{x}y &= \frac{1}{x^2} \\ I(x) &= e^{\int \frac{1}{x}\,\mathrm{d}x} = e^{\ln x} = x \\ I(x)(y' + \frac{1}{x}y) &= I(x)\frac{1}{x^2} \\ xy' + y &= \frac{1}{x} \\ (xy)' &= \frac{1}{x} \\ \int (xy)' &= \int \frac{1}{x}\,\mathrm{d}x \\ xy &= \ln x + C \\ y &= \frac{\ln x + C}{x} \end{aligned} \tag{70}$$

Note that the original equation was not in standard form, but we were able to simply divide each factor by x^2 to get it into standard form.

8.5 Orthogonal Families of Curves

Using differential equations, we can determine the general equation of curves orthogonal (ie "perpendicular") to a given curve (or, typically, a family of curves). Consider the general family of curves described by;

$$ax^2 + by^2 = k,$$

where a, b, k are real constants. The derivative of any perpendicular curve must be the negative reciprocal of the derivative of the original curve (by the very definition of perpendicularity), so we can find the derivative of the original curve as follows, using implicit differentiation:

$$\begin{aligned} ax^2 + by^2 &= k \\ 2ax + 2by\frac{\mathrm{d}y}{\mathrm{d}x} &= 0 \\ \frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{-2ax}{2by} = -\frac{ax}{by} \\ \therefore m_{orth.} &= -(\frac{\mathrm{d}y}{\mathrm{d}x})^{-1} = \frac{by}{ax} \end{aligned} \tag{71}$$

If we define P as the orthogonal curve's y-value, we can then find a general equation for the orthogonal curve as a function of x by rewriting it as a (*separable*) differential equation:

$$\begin{aligned} \frac{\mathrm{d}P}{\mathrm{d}x} &= m_{orth.} = \frac{by}{ax} = \frac{bP}{ax} \\ \frac{\mathrm{d}P}{bP} &= \frac{\mathrm{d}x}{ax} \\ \int \frac{\mathrm{d}P}{bP} &= \int \frac{\mathrm{d}x}{ax} \\ \frac{1}{b}\ln|P| &= \frac{1}{a}\ln|x| + C \\ \ln|P| &= \frac{b}{a}\ln|x| + C \\ P(x) = Ke^{\ln|x^{\frac{b}{a}}|} &= Kx^{\frac{b}{a}}, K = \pm e^C \end{aligned} \tag{72}$$

This general method can be used to find the general equation of any curve orthogonal to a given curve (assuming the given curve is differentiable). See figure 17 for a visualization of this concept.

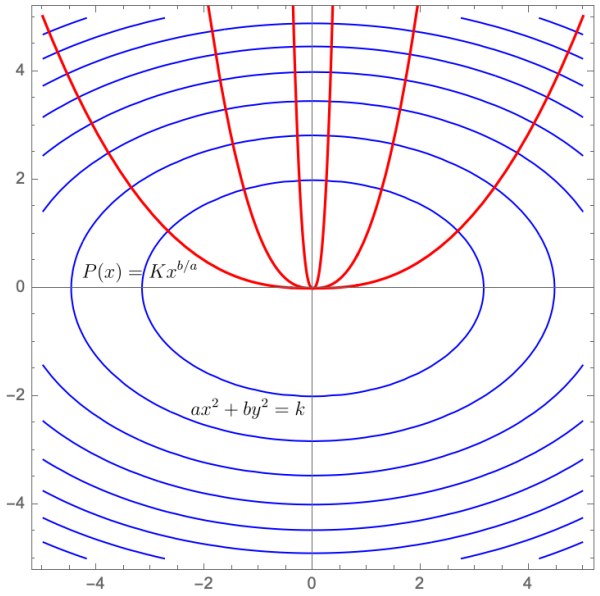


Figure 17: $ax^2 + by^2 = k$ and its orthogonal curves

Part II

Multivariable Calculus

9 Partial Derivatives and Applications

9.1 Introductions/Visualizations

Numerous real-world concepts can be described as function of multiple variables. A function of two variables, for instance, can be defined as a $f(x, y)$.

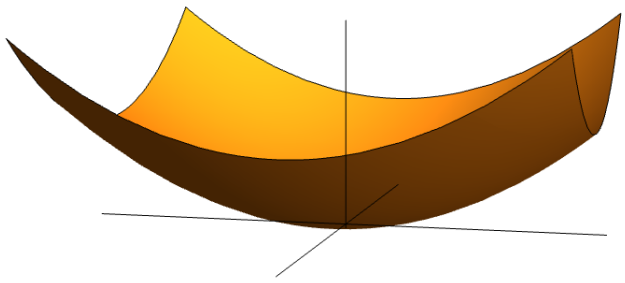


Figure 18: $f(x, y) = x^2 + y^2$

Functions of many variables can often be difficult to visualize (particularly of more than 2 variables), but at least for a function of two variables, we can visualize it as a surface in 3D. For instance, consider the function $f(x, y) = x^2 + y^2$, and figure 18.

While this is helpful, 3D plots aren't always realistic to use to visualize a function. Instead, we can draw **level curves**, by plotting a number of curves of a constant z-value in the standard $x - y$ plane. See figure 19

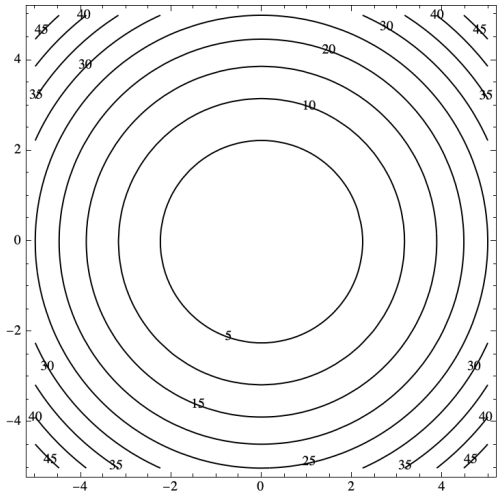


Figure 19: Level curves of $f(x, y) = x^2 + y^2$

for a level curve graph of $f(x, y) = x^2 + y^2$. Note that, visually, this plot is essentially the same as looking at the surface of the function in 3D from above the z-axis.

Note that, while harder to visualize, functions of more than two variables can be described similarly. For instance, a function of three variables can be described as a surface in 4D, and a function of four variables can be described as a surface in 5D, and so on. By extension, we can also describe the level curves of a function of n variables as a family of functions in n (ie a function of 3 variables can be visualized as a level plot in 3D).

9.2 Limits

A limit of a function of multiple variables is described as the value of the said function when all of its variables approach a particular coordinate. For the function $f(x, y)$, for example, we can write its limit as it approaches some point (a, b):

$$\lim_{(x,y)\to(a,b)} f(x, y)$$

Determining limits (and in particular, if a limit exists) is not quite as straightforward as for functions of a single variable. For a function $f(x)$ and a point a, for instance, if the limit approaching from the right (a^+) doesn't equal the limit approaching from the left (a^-), then the limit doesn't exist. For a function of two variables, though, a limit can be approached from an infinite number of directions, assuming the corresponding (x, y) is still in the domain of the function. By extension, we can only say that a limit exists in this case if, no matter which direction we approach the limit from, the limit is the same.

We can use, however, the precise definition of a limit to more accurately define a limit of $f(x, y)$ with domain D as it approaches (a, b):

$$\lim_{(x,y)\to(a,b)} f(x, y) = L$$

if for every $\epsilon > 0$, there exists a $\delta > 0$ such that if $(x, y) \in D$ and $0 < \sqrt{(x-a)^2 + (y-b)^2} < \delta$, then $|f(x, y) - L| < \epsilon$.

By extension of this definition of a limit, we can say a function $f(x, y)$ is continuous at (a, b) if $\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b)$, very similarly to the definition of continuity for a function of a single variable.

9.3 Partial Derivatives

Extending the definition of a derivative of a function of 1 variable, we can find the **partial derivative** of a function of multiple in terms of a particular variable (notated as $\frac{\partial f}{\partial x}$ or $f_x(x, y)$), while treating the other variable(s) as a constant:

$$\begin{aligned} f_x(x, y) = \frac{\partial f}{\partial x} &= \lim_{h\to 0}[\frac{f(x+h, y) - f(x, y)}{h}] \\ f_y(x, y) = \frac{\partial f}{\partial y} &= \lim_{h\to 0}[\frac{f(x, y+h) - f(x, y)}{h}] \end{aligned} \tag{73}$$

This definition can, naturally, be extended to higher derivatives. Just as there were two partial derivatives for $f(x, y)$ depending on which variable f was being differentiated in respect to, there are similarly two second partial derivatives for each first partial derivative (and a similar pattern follows for functions of a higher number of variables). For $f(x, y)$, it can be said:

$$\begin{aligned} (f_x)_x = f_{xx} &= \frac{\partial}{\partial x}(\frac{\partial f}{\partial x}) = \frac{\partial^2 f}{\partial x^2} \\ (f_x)_y = f_{xy} &= \frac{\partial}{\partial y}(\frac{\partial f}{\partial x}) = \frac{\partial^2 f}{\partial y \partial x} \\ (f_y)_x = f_{yx} &= \frac{\partial}{\partial x}(\frac{\partial f}{\partial y}) = \frac{\partial^2 f}{\partial x \partial y} \\ (f_y)_y = f_{yy} &= \frac{\partial}{\partial y}(\frac{\partial f}{\partial y}) = \frac{\partial^2 f}{\partial y^2} \end{aligned} \tag{74}$$

Theorem (Clairaut's Theorem)**.** *If the point (a, b) is in the domain of $f(x, y)$, and f_{xy} and f_{yx} are continuous, then $f_{xy}(a, b) = f_{yx}(a, b)$.*

9.4 Differentials, Errors, Approximations

In functions of a single variable, it can be said:

$$\mathrm{d}y = f'(x)\mathrm{d}x$$

Similarly, in a function of multiple variables (say $z = f(x, y)$), we can say:

$$\begin{aligned} \mathrm{d}z &= f_x(x,y)\mathrm{d}x + f_y(x,y)\mathrm{d}y \\ &= \frac{\partial z}{\partial x}\mathrm{d}x + \frac{\partial z}{\partial y}\mathrm{d}y \end{aligned} \tag{75}$$

Similarly to how the approximate error Δf in a function $f(x)$ can be estimated from the value of Δx by calculating $\Delta f \approx df = f'(x)dx$, we can similarly estimate the error in a function of multiple variables by calculating:

$$\begin{aligned} \mathrm{d}z &= \frac{\partial z}{\partial x}\mathrm{d}x + \frac{\partial z}{\partial y}\mathrm{d}y \\ \Delta z &= \frac{\partial z}{\partial x}\Delta x + \frac{\partial z}{\partial y}\Delta y \end{aligned} \tag{76}$$

9.5 Partial Differential Equations

Just as a *differential equation* relates a function with its derivative(s), a *partial differential equation* relates a multi-variable function with its partial derivatives.

One of the most notable of these is the **Laplace equation**,

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \tag{77}$$

with applications in various fields.

9.6 Tangents, Linear Approximations

Just as it is possible to find the *tangent line* to a particular function $f(x)$ by finding the slope $m = f'(a)$ at some point a, we can find the *tangent plane* to a function $f(x,y)$ at a particular point (a,b,c). In general, such a plane can be defined as:

$$\begin{aligned} z - c &= m_x(x-a) + m_y(y-b) \\ z - c &= f_x(a,b)(x-a) + f_y(a,b)(y-b) \end{aligned} \tag{78}$$

Note the similarities between using this equation and the linear equation $y = mx + b$ when finding the tangent line to a function of a single variable.

We can also use the tangent plane to a function $f(x,y)$ to define a linear approximation (or, more accurately,

tangent approximation) $L(x, y)$:

$$L(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b) \tag{79}$$

The derivation of this equation is very similar to that of a single-variable linearization. It should be clear that this is simple the equation for the tangent plane at a particular point (a, b), simply rewritten as a function of x and y.

9.6.1 Normal Equation to a Plane

Using the general form of a tangent plane ($z = ax + by + c$), we can find the normal equation to the the plane at which it is tangent to the original function fairly easily. Rewriting this equation, we can say:

$$\begin{aligned} z &= ax + by + c \\ z - ax - by &= c \end{aligned} \tag{80}$$

One property of the function in this form is that a vector that is perpendicular to it can be written using the coefficients of each variable, ie $\vec{n} = \langle -a, -b, 1 \rangle$ (this can fairly easily be proven). From here, we can find a set of functions in terms of (for instance) t that parametrically describe the general normal equation. The slope in each direction x, y, z must be the same as the respective quantity in our normal vector $\vec{n}$, so we can say:

$$\begin{aligned} x &= -at \\ y &= -bt \\ z &= t \end{aligned} \tag{81}$$

This form assumes that the original tangent plane in question is at the origin, but can be easily modified to account for a different point, say at $(x, y, z) = (m, n, p)$:

$$\begin{aligned} x &= m - at \\ y &= n - bt \\ z &= p + t \end{aligned} \tag{82}$$

9.7 The Chain Rule

The chain rule for derivatives of functions of 1 variable describes the derivative of a composition of two functions; say $y = f(x)$ and $x = g(t)$ (assuming f and g are differentiable), then the derivative of y with respect to t can be described as:

$$\begin{aligned} \frac{\mathrm{d}y}{\mathrm{d}t} &= \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t} \\ (f(g(t)))' &= f'(g(t))g'(t) \end{aligned} \tag{83}$$

This is extendable to functions of multiple variables, though with some caveats. In general, though, if u is a function of n variables $x_1, x_2, \ldots, x_n$, and each x is a (differentiable) function of m variables $t_1, t_2, \ldots, t_m$, then u is also a function of $t_1, t_2, \ldots, t_m$, and the derivative of u with respect to t_i can be described as:

$$\begin{aligned} \frac{\partial u}{\partial t_i} &= \frac{\partial u}{\partial x_1}\frac{\partial x_1}{\partial t_i} + \frac{\partial u}{\partial x_2}\frac{\partial x_2}{\partial t_i} + \cdots + \frac{\partial u}{\partial x_n}\frac{\partial x_n}{\partial t_i} \\ &= \sum_{j=1}^{m} \frac{\partial u}{\partial x_j}\frac{\partial x_j}{\partial t_i} \end{aligned} \tag{84}$$

This general form can seem quite daunting, but can be simplified down to a couple of main cases:

Case 1: $u = z = f(g(t), h(t))$

Say $u = z = f(x, y)$, where $x = g(t)$ and $y = h(t)$. We can then write:

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} \tag{85}$$

Case 2: $u = z = f(g(s, t), h(s, t))$

Say $u = z = f(x, y)$, where $x = g(s, t)$ and $y = h(s, t)$. We can then write:

$$\begin{aligned} \frac{\partial z}{\partial s} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} \\ \frac{\partial z}{\partial t} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t} \end{aligned} \tag{86}$$

Notice that Case 2 demonstrates the equation for *two* partial derivatives. By extension, the more variables that u is a function of, the more partial derivatives there will be; in this situation, further "cases" simply follow

the general aforementioned.

For example, if $z = e^x \sin y$, $x = st^2$, and $y = s^2 t$, then:

$$\begin{aligned}
\frac{\partial z}{\partial s} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} \\
&= \frac{\partial}{\partial x}(e^x \sin y)\frac{\partial}{\partial s}(st^2) + \frac{\partial}{\partial y}(e^x \sin y)\frac{\partial}{\partial s}(s^2 t) \\
&= (e^x \sin y)(t^2) + (e^x \cos y)(2st) \\
&= (e^{st^2} \sin\left(s^2 t\right))(t^2) + (e^{st^2} \cos\left(s^2 t\right))(2st) \\
&= t^2 e^{st^2} \sin\left(s^2 t\right) + 2ste^{st^2} \cos\left(s^2 t\right)
\end{aligned} \tag{87}$$

The same method can be used to find $\frac{\partial z}{\partial t}$.

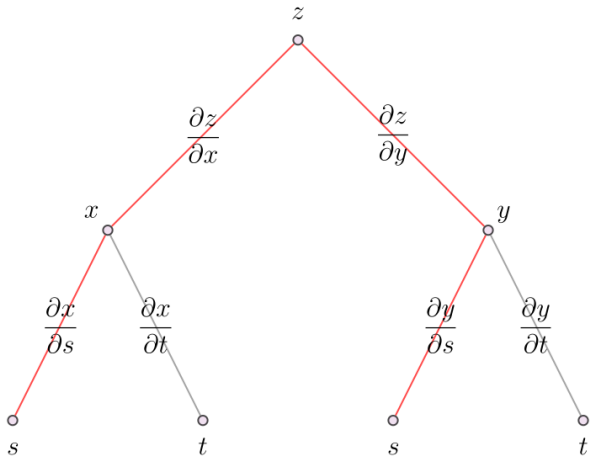


Figure 20: Tree graph for the partial derivative of $z = f(x(s,t), y(s,t))$.

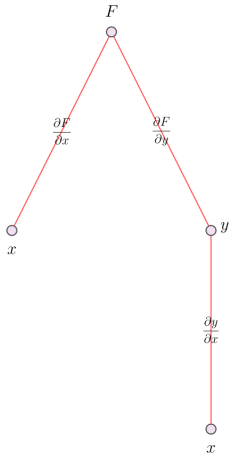


Figure 21: Tree graph for $F(x, f(x))$.

9.7.1 Tree Graphs

Using the chain rule for implicit functions of many variables can easily become very confusing. It can help to create a tree graph to visualize the relationships between the functions at hand. For example, for the example above, where $z = f(x(s,t), y(s,t))$, we create a tree graph as in figure 20.

To find, for instance, $\frac{\partial z}{\partial s}$, you follow the path from z to each s in the tree graph, and multiply the partial derivatives of each function along the way (note the highlighted branches). In this case, $\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s}$.

9.8 Implicit Differentiation

Given a function in the form $F(x,y) = 0$ where y is a function f of x (ie $F = (x, f(x))$), we can differentiate both sides in terms of x, using the chain rule (see figure 21):

$$\begin{aligned} F(x,y) &= 0 \\ \frac{\partial F}{\partial x}\frac{dx}{dx} + \frac{\partial F}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}x} &= 0 \\ \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{dy}{dx} &= 0 \\ \frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} &= -\frac{F_x}{F_y} \end{aligned} \tag{88}$$

Note that $\frac{dx}{dx} = 1$, as we can think of x as a function of x (ie $x = x$). The proof for functions of more variables is fairly similar; for a function of the form $F(x,y,z) = 0$ where $z = f(x,y)$, we can find $\frac{\partial z}{\partial x}$:

$$\begin{aligned} F(x,y,z) &= 0 \\ \frac{\partial F}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} &= 0 \\ \frac{\partial x}{\partial x} = 1, \frac{\partial y}{\partial x} &= 0 \\ \frac{\partial F}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} &= 0 \\ \therefore \frac{\partial z}{\partial x} &= -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} \end{aligned} \tag{89}$$

Note that the third line of this proof is valid as $\frac{\partial x}{\partial x} = 1$ as shown above, and $\frac{\partial y}{\partial x} = 0$ because y is not a function of x. Finding $\frac{\partial z}{\partial y}$ is incredibly similar (and equals $-\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}}$).

Consider a "simpler" function, $x^2 + y^3 = 0$; differentiating in terms of x gives:

$$
\begin{aligned}
F(x,y) = x^2 + y^3 &= 0 \\
\frac{\mathrm{d}y}{\mathrm{d}x} &= -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} \\
&= -\frac{2x}{3y^2}
\end{aligned}
\tag{90}
$$

9.9 Directional Derivatives and the Gradient

The rate of change (derivative) of a function $f(x,y)$ of 2 variables can theoretically be defined in any arbitrary direction, say, in the direction of a unit vector $\vec{u} = \langle a, b \rangle$. By this logic, then $f_x(x,y)$ is the derivative in the direction of $\hat{i} = \langle 1, 0 \rangle$, and $f_y(x,y)$ is the derivative in the direction of $\hat{j} = \langle 0, 1 \rangle$. The derivative in the direction of any arbitrary unit vector $\vec{u}$ is called a *directional derivative*, and is denoted as $D_{\vec{u}} f(x,y)$. Using the limit definition of a derivative, we can write:

$$
D_{\vec{u}} f(x_0, y_0) = \lim_{h \to \infty} [\frac{f(x_0 + ah, y_0 + bh) - f(x_0, y_0)}{h}] \tag{91}
$$

Note that the top line states $f(x_0 + ah, y_0 + bh)$ because the function is increasing by a in the x direction and b in the y direction.

To simplify this, let $g(h) = f(x_0 + ah, y_0 + bh)$, then we can say, using eqn. 91:

$$
g'(0) = \lim_{h \to \infty} [\frac{g(h) - g(0)}{h}] = \lim_{h \to \infty} [\frac{f(x_0 + ah, y_0 + bh) - f(x_0, y_0)}{h}] = D_{\vec{u}} f(x_0, y_0) \tag{92}
$$

Since $g(h) = f(x_0 + ah, y_0 + bh)$, we can define $x = x_0 + ah$ and $y = y_0 + bh$ (ie, x and y are both functions of h), and using the chain rule, write:

$$
\begin{aligned}
g(h) &= f(x_0 + ah, y_0 + bh) = f(x(h), y(h)) \\
g'(h) &= \frac{\partial f}{\partial x}\frac{dx}{dh} + \frac{\partial f}{\partial y}\frac{dy}{dh} \\
\frac{dx}{dh} &= \frac{d}{dh}(x_0 + ah) = a \\
\frac{dy}{dh} &= \frac{d}{dh}(y_0 + bh) = b \\
\therefore g'(h) &= \frac{\partial f}{\partial x}(a) + \frac{\partial f}{\partial y}(b) \\
&= f_x(x, y)a + f_y(x, y)b \\
g'(0) &= f_x(x_0, y_0)a + f_y(x_0, y_0)b \\
&= D_{\vec{u}} f(x_0, y_0)
\end{aligned}
\tag{93}
$$

Note that this last line involves combining the result of calculating $g'(h)$ with the original definition of $g'(0)$ from eqn. 93. In summary, if $\vec{u} = \langle a, b \rangle$, then the derivative of $f(x, y)$ in the direction of $\vec{u}$ is:

$$
D_{\vec{u}} f(x, y) = f_x(x, y)a + f_y(x, y)b \tag{94}
$$

Note that this can be rewritten as a *dot product*:

$$
\begin{aligned}
D_{\vec{u}} f(x, y) &= f_x(x, y)a + f_y(x, y)b \\
&= \langle f_x(x, y), f_y(x, y) \rangle \cdot \langle a, b \rangle \\
&= \langle f_x(x, y), f_y(x, y) \rangle \cdot \vec{u}
\end{aligned}
\tag{95}
$$

The first vector in this equation is the **gradient** (denoted ∇) of $f(x, y)$; specifically:

$$
\begin{aligned}
\nabla f(x, y) &= \langle f_x(x, y), f_y(x, y) \rangle \\
\left(= \frac{\partial f}{\partial x}\hat{i} + \frac{\partial f}{\partial y}\hat{j} \right)
\end{aligned}
\tag{96}
$$

Using this notation, we can simplify the form of the directional derivative to

$$
D_{\vec{u}} f(x, y) = \nabla f(x, y) \cdot \vec{u} \tag{97}
$$

In higher dimensions, the same general concepts apply; for instance, $D_{\vec{u}}f(x,y,z) = \nabla f(x,y,z) \cdot \vec{u}$, where $\vec{u}$ is a vector in $\mathbb{R}^3$.

9.9.1 Maximizing the Gradient

Given a function, $f(x,y)$, a common question is to find what direction it changes the fastest, ie, what direction should $\vec{u}$ be pointing to maximize $D_{\vec{u}}f(x,y)$? This is fairly simple to calculate using eqn. 97's definition of the gradient:

$$\begin{aligned} D_{\vec{u}}f(x,y) &= \nabla f(x,y) \cdot \vec{u} \\ &= |\nabla f(x,y)||\vec{u}| \cos(\theta) \\ &= |\nabla f(x,y)| \cos(\theta) \end{aligned} \tag{98}$$

Since $\cos$ is at a max when $\theta = 0$, we can thus maximize $D_{\vec{u}}f(x,y)$ by setting $\vec{u}$ in the same direction as $\nabla f(x,y)$.

9.10 Using a Jacobian Matrix

The Jacobian matrix is a matrix of partial derivatives of a function of multiple variables. It has many more applications beyond the scope of this course, but is still interesting/useful enough to include here. To understand the derivation of the Jacobian matrix, first consider the derivation of Cramer's rule.

9.10.1 Cramer's Rule

Cramer's Rule is a general method for solving systems of linear equations, specifically, when there are as many unknowns as there are equations. Its easiest to understand by deriving it first from a simple system of 2 equations (and thus, 2 unknowns):

$$\begin{aligned} ax + by &= \alpha \\ cx + dy &= \beta \end{aligned} \tag{99}$$

We can cancel out the y terms:

$$
\begin{aligned}
d(ax + by = \alpha) \\
b(cx + dy = \beta) \\
(adx + bdy = d\alpha) \\
-(cbx + bdy = b\beta) \\
\hline
adx - cbx = d\alpha - b\beta \\
x(ad - cb) = d\alpha - b\beta \\
x = \frac{d\alpha - b\beta}{ad - cb}
\end{aligned}
\tag{100}
$$

If we consider A as the matrix of coefficients of the system and B as the matrix with the first column of A replaced with the column vector $\langle \alpha, \beta \rangle$ (ie the solutions), then we can write the above as:

$$
\begin{aligned}
A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad B = \begin{bmatrix} \alpha & b \\ \beta & d \end{bmatrix} \\
x = \frac{d\alpha - b\beta}{ad - cb} = \frac{\det(B)}{\det(A)}
\end{aligned}
\tag{101}
$$

Cramer's Rule in general states that for an equation $Ax = b$, $x_i = \frac{\det(B_i)}{\det(A)}$, where B_i is the matrix with the i^{th} column of A replaced with b and x_i is the i^{th} variable in the equation(s). See here for more Linear Algebra.

9.10.2 Using Cramer's Rule to Derive the Jacobian Matrix

Given two functions $F(x, y, u, v) = 0$ and $G(x, y, u, v) = 0$, where u and v are both functions of x and y, we can find $\frac{\partial u}{\partial x}$ (or $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$, and $\frac{\partial v}{\partial y}$, by the same method) by setting up a system of equations. For both F and G, we can take the derivative of both sides with respect to x:

$$
\begin{aligned}
\frac{\partial F}{\partial x} + \frac{\partial F}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial F}{\partial v}\frac{\partial v}{\partial x} = 0 \\
\frac{\partial F}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial F}{\partial v}\frac{\partial v}{\partial x} = -\frac{\partial F}{\partial x} \\
\frac{\partial G}{\partial x} + \frac{\partial G}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial G}{\partial v}\frac{\partial v}{\partial x} = 0 \\
\frac{\partial G}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial G}{\partial v}\frac{\partial v}{\partial x} = -\frac{\partial G}{\partial x}
\end{aligned}
\tag{102}
$$

In this way, we can consider this as a system of two equations and two variables ($\frac{\partial u}{\partial x}$ and $\frac{\partial v}{\partial x}$), which we can solve using Cramer's Rule, for example, for $\frac{\partial u}{\partial x}$:

$$\frac{\partial u}{\partial x} = \frac{\left|\begin{bmatrix} -\frac{\partial F}{\partial x} & \frac{\partial F}{\partial v} \\ -\frac{\partial G}{\partial x} & \frac{\partial G}{\partial v} \end{bmatrix}\right|}{\left|\begin{bmatrix} \frac{\partial F}{\partial u} & \frac{\partial F}{\partial v} \\ \frac{\partial G}{\partial u} & \frac{\partial G}{\partial v} \end{bmatrix}\right|} = \frac{\frac{-\partial(F,G)}{\partial(x,v)}}{\frac{\partial(F,G)}{\partial(u,v)}} \tag{103}$$

9.11 Optimization

9.11.1 Extrema of Functions of Multiple Variables

For a function $f(x)$ of a single variable, an extrema of the function exists when $f'(x) = 0$. By extension, we say that an extrema exists at some point (call it P_0) on a function $f(x, y)$ of two variables when both $[\frac{\partial f}{\partial x}]_{P_0} = 0$ and $[\frac{\partial f}{\partial y}]_{P_0} = 0$. By extension, P_0 is an extrema if $\vec{\nabla} f(P_0) = \vec{0}$. There exist several proofs for this fact, but consider the following 2:

Proof. Assume that P_0 is a minimum point of $z = f(x, y)$. The tangent plane to this point must thus be horizontal. Now take a new function $F(x, y, z) = z - f(x, y) = 0$. Finding the gradient of F:

$$\vec{\nabla} F = -\frac{\partial F}{\partial x}\hat{i} - \frac{\partial F}{\partial y}\hat{j} + \hat{k}$$

This gradient is normal to P_0. Since P_0 is a minimum, the normal vector to $f(x, y)$ at P_0 is $-\hat{k}$. Thus, $\vec{\nabla} F = \vec{\nabla} f(P_0)$, and we can write:

$$\vec{\nabla} F = -\frac{\partial f}{\partial x}\hat{i} - \frac{\partial f}{\partial y}\hat{j} + \hat{k} = -\hat{k}$$
$$\frac{\partial f}{\partial x}\hat{i} + \frac{\partial f}{\partial y}\hat{j} = \vec{0}$$

Thus, $\frac{\partial f}{\partial x} = 0$ and $\frac{\partial f}{\partial y} = 0$, and so $\nabla \vec{f}(P_0) = \vec{0}$ ■

A similar proof can be made for a maximum point...

Proof. Take a function $f(x, y)$ that has a maximum at P_0. Now, assume that $\vec{\nabla} f(P_0) \neq \vec{0}$. If this were the case, then, by definition, the gradient of the function is pointing to a direction where the function is increasing. This

contradicts the fact that P_0 is a maximum point, and thus, $\vec{\nabla} f(P_0)$ must equal $\vec{0}$. ■

For functions of 3 variables (and higher), similar reasoning to the second proof above can be used to show that $\vec{\nabla} f(P_0) = \vec{0}$ if P_0 is a maximum or minimum point; there is no real geometric proof for this, however.

9.11.2 Second Derivative Test

Just as for functions of a single variable, it is not sufficient to say that a point is an extrema of a function of two variables simply because the gradient is zero. Instead, we must use the **second derivative test**. The motivation/derivation of this test comes from Taylor polynomial expansion of a general function of two variables, but won't be covered just yet.

Consider a function $f(x, y)$, and a point P_0 at which $\nabla f(P_0) = \vec{0}$. We can define a new matrix containing the second derivatives of f at P_0:

$$\begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial xy} \\ \frac{\partial^2 f}{\partial yx} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$$

Note that $\frac{\partial^2 f}{\partial yx} = \frac{\partial^2 f}{\partial xy}$, assuming continuity at P_0, by Clairaut's Theorem. This, somewhat similarly to the Jacobian matrix defined earlier, is called the **Hessian matrix** of f.

The determinant of this matrix, from the definition of the determinant of a 2 by 2 matrix, is $f_{xx}f_{yy} - f_{xy}f_{yx} = f_{xx}f_{yy} - f_{xy}^2$. The value of this determinant can then be used to determine what exactly the point P_0 is:

- $f_{xx}f_{yy} - f_{xy}^2 > 0$
  - $f_{xx} > 0$ and $f_{yy} > 0$: P_0 is a local minimum
  - $f_{xx} < 0$ and $f_{yy} < 0$: P_0 is a local maximum
- $f_{xx}f_{yy} - f_{xy}^2 < 0$: P_0 is a **saddle point**. This is a point where the function decreases in one direction, and increases in the other (think of the central point of a hyperbolic paraboloid).
- $f_{xx}f_{yy} - f_{xy}^2 = 0$: the test is inconclusive. If this situation arises, you'd have to do significantly more work (ie by further expanding the Taylor polynomial of f) to determine what exactly the point is.

Example: consider a rectangular open box with dimensions x, y, and z, with a volume of 32 and a minimal surface area. We can define S as the surface area of the box:

$$S = xy + 2xz + 2yz$$

The volume is given by $xyz = 32 \implies z = \frac{32}{xy}$, so we can rewrite S as a function of x and y:

$$S(x,y) = xy + \frac{64}{y} + \frac{64}{x}$$

Optimizing S:

$$\begin{aligned}
\nabla S &= \langle y - \frac{64}{x^2}, x - \frac{64}{y^2} \rangle \\
\nabla S &= \langle 0, 0 \rangle \\
\implies x &= \frac{64}{y^2} \text{ and } y = \frac{64}{x^2}
\end{aligned}$$

Solving this set of equations, we get $x = y = 4$. To verify that this is indeed a minimum, we can use the second derivative test:

$$\begin{aligned}
H_S &= \begin{bmatrix} \frac{128}{x^3} & 1 \\ 1 & \frac{128}{y^3} \end{bmatrix} \\
&\frac{128}{x^3} \cdot \frac{128}{y^3} - 1^2 \\
&= \frac{16384}{4096} - 1 > 0 \\
S_{xx}(4,4) &= \frac{128}{4^3} > 0 \\
S_{yy}(4,4) &= \frac{128}{4^3} > 0
\end{aligned}$$

Thus, the point $(4,4)$ is a local minimum of S. We can find the length of side $z = \frac{32}{4*4} = 2$, and so the minimum surface area is $S = 4*4 + 2*4*2 + 2*4*2 = 48$.

Testing the types of extrema of a function of more than two variables is slightly trickier, but still uses the aforementioned Hessian matrix (an in fact, the formula above is just a special case of the following "method"). Take a Hessian matrix of a function of n variables:

$$\begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 x_n} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n x_1} & \frac{\partial^2 f}{\partial x_n x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

The ith **principal minor** (denoted Δ_i) of this matrix is the determinant of the matrix formed by the upper left

i by i submatrix of this matrix. For instance, we have:

$$\Delta_1 = \frac{\partial^2 f}{\partial x_1^2}; \quad \Delta_2 = \left|\begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}\right|; \quad \Delta_3 = \left|\begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} & \frac{\partial^2 f}{\partial x_1 x_3} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 x_3} \\ \frac{\partial^2 f}{\partial x_3 x_1} & \frac{\partial^2 f}{\partial x_3 x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{bmatrix}\right|; \quad \dots$$

A particular point is a local *maximum* if $\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \dots$ (alternating positive/negative) *(ie, the matrix is negative definite)*, and a local *minimum* if $\Delta_1 > 0, \Delta_2 > 0, \dots$ (all positive) *(ie, the matrix is positive definite)*.

9.11.3 Least Squares

Consider a set of N points $(x_1, y_1), \dots, (x_N, y_N)$. We can fit a line to these points (of the form $y = mx + b$) by finding the points with "the least error" of the original points, ie $(x_1, mx_1 + b), \dots, (x_N, mx_N + b)$. We define this "error" as the sum of the squares of the differences between the original points and the points on the line, hence the name "least squares". We can thus define a function $S(m, b)$ describing the square of the distance between the given points, and optimize it to find the best fit line:

$$S(m, b) = \sum_{k=1}^{N} (mx_k + b - y_k)^2 \tag{104}$$

We can find the extrema of this function by finding when the partial derivatives of this function are equal to zero, as described in the previous section:

$$\begin{aligned} \frac{\partial S}{\partial m} &= \frac{\partial}{\partial m} \sum_{k=1}^{N} (mx_k + b - y_k)^2 \\ &= \sum_{k=1}^{N} \frac{\partial}{\partial m} (mx_k + b - y_k)^2 \\ &= \sum_{k=1}^{N} 2(mx_k + b - y_k)x_k = 0 \end{aligned}$$

Note that this relies on the fact that the derivative of a sum is equal to the sum of the derivatives. Rearranging terms:

$$m[\sum_{k=1}^{N} x_k^2] + b[\sum_{k=1}^{N} x_k] = \sum_{k=1}^{N} x_k y_k$$

Similarly for $\frac{\partial S}{\partial b}$:

$$\begin{aligned}\frac{\partial S}{\partial b} &= \sum_{k=1}^{N} \frac{\partial}{\partial b}(mx_k + b - y_k)^2 \\ &= \sum_{k=1}^{N} 2(mx_k + b - y_k) \\ &= \sum_{k=1}^{N} (mx_k + b - y_k) = 0 \\ m[\sum_{k=1}^{N} x_k] + Nb &= \sum_{k=1}^{N} y_k\end{aligned}$$

While its not obvious in this form, this is now simply a system of two equations in two unknowns (m and b), which, when given a set of N points, can be solved with any number of methods (ie Cramer's).

9.12 Extrema with Constraints

Say we have a function $f(x, y)$ subject to some constraint $g(x, y) = 0$ that we wish to optimize. Suppose we can represent g parametrically as the curve $C : x = x(t), y = y(t)$. If $P[x(t), y(t)]$ is an extrema of f, we can write $f(x(t), y(t)) = F(t)$ (for some function F). We can then write (via the chain rule):

$$\frac{\mathrm{d}F}{\mathrm{d}t} = 0 = \frac{\mathrm{d}f}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\mathrm{d}f}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}t}$$

This is simply a dot product,

$$\nabla \vec{f} \cdot \langle \frac{\mathrm{d}x}{\mathrm{d}t}, \frac{\mathrm{d}y}{\mathrm{d}t} \rangle,$$

where the second vector is the tangent vector to the curve C at P, which we can call $\vec{T}$. As this dot product is zero, $\nabla \vec{f}$ must be perpendicular to $\vec{T}$. Now, taking the gradient of g (again via the chain rule), we have

$$\frac{\partial g}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial g}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} = 0,$$

and can further say that

$$\nabla \vec{g} \cdot \vec{T} = \nabla \vec{g} \cdot \vec{T} = 0.$$

Here too, (assuming $\nabla \vec{g}(P) \neq 0$), $\nabla \vec{g} \perp \vec{T}$. Since both $\vec{f}$ and $\vec{g}$ are perpendicular to $\vec{T}$, we can thus say

$$\nabla \vec{f}(P) = \pm\lambda \nabla \vec{g}(P),$$

where λ is an arbitrary constant called, in the context of these problems, the **Langrangian multiplier.** In problems involving extrema with constraints, we can also define the **Lagranian**;

$$L(x, y, \lambda) = f + \lambda g,$$

and require that

$$\nabla \vec{L} = \nabla \vec{f} + \lambda \nabla \vec{g} = 0.$$

This gives us a system of three equations for three unknowns; namely

$$\begin{aligned} L_x &= f_x + \lambda g_x = 0 \\ L_y &= f_y + \lambda g_y = 0 \\ L_\lambda &= g(x, y) = 0 \end{aligned}$$

This can be extended to a function, $f(x, y, z)$, subject to two constraint $g(x, y, z) = 0$ and $h(x, y, z) = 0$; we write

$$L(x, y, \lambda, \mu) = f + \lambda g + \mu h,$$

where we require

$$\nabla \vec{L} = \nabla f + \lambda \nabla g + \mu \nabla h = 0.$$

Extending this out yields a system of four equations of four unknowns.

In general, the actual value(s) of the Lagrangian multiplier(s) are irrelevant; simply use the equations that they are a part of to solve for the coordinates of the extrema.

9.12.1 Sufficiency for Extrema

Similar to section 9.11.2; the second derivative test for a Lagrangian of two variables x, y and constraint λ, ie

$$L = f(x, y) \pm \lambda g(x, y)$$

is given

$$H = L_{xx} g_y^2 - 2 L_{xy} g_x g_y + L_{yy} g_x^2;$$

- $H > 0 \implies (x, y)$ a min
- $H < 0 \implies (x, y)$ a max

10 Multiple Integration

10.1 Double Integrals

Consider an arbitrary rectangle D in the x, y-plane, and let $f(x, y)$ be a function in D. Subdividing D into small rectangles, we can say that the total value of f in the smaller rectangle is approximately, taking $P(x^*_{ij}, y^*_{ij})$ to be a point within this rectangle, $f(P)\Delta x_i \Delta y_j$. Summating all of these rectangles yields

$$\begin{aligned}\lim_{m,n\to\infty,\infty} \sum_{i=1}^{n}\sum_{j=1}^{m} f(P)\Delta x_i \Delta y_j \\ = \int_c^d \int_a^b f(x,y)\,\mathrm{d}x\,\mathrm{d}y\end{aligned}$$

Where (a, b) is the interval of x-values and (c, d) is the interval of y-values. This is the simplest case of a double integral, where the region is rectangular and can thus be broken down into a grid of rectangles. A common subcase here is when $f(x, y) = g(x) * h(y)$; in this case, we can write

$$\begin{aligned}\iint_D f(x,y)\,\mathrm{d}A &= \int_c^d \int_a^b f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_c^d h(y) \int_a^b g(x)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_c^d h(y)\,\mathrm{d}y \int_a^b g(x)\,\mathrm{d}x\end{aligned}$$

Another helpful case in evaluating double integrals is the following theorem.

Theorem (Fubini's Theorem)**.** *If an integral has a finite value, then the order of integration may be interchanged. That is,*

$$\int_a^b \int_c^d f(x,y)\,dx\,dy = \int_c^d \int_a^b f(x,y)\,dy\,dx.$$

More complex regions are typically broken down into two types; **Type I**, where the region's y values are bounded by functions of x, and **Type II**, where the region's x values are bounded by functions of y. In short, these can be written as follows:

1. **Type I**: region D bounded by $a \le x \le b$ and $g_1(x) \le y \le g_2(x)$.

$$\iint_D f(x,y)\,\mathrm{d}A = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\,\mathrm{d}y\,\mathrm{d}x$$

2. **Type I**: region D bounded by $h_1(y) \leq x \leq h_2(y)$ and $c \leq y \leq d$.

$$\iint_D f(x,y)\,\mathrm{d}A = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

10.2 Applications of Double Integrals

Other than the obvious (finding the area of an arbitrary region in the x, y-plane), there are a number of other applications;

- **Volume of a Projected Region:** A surface defined by $z = f(x,y)$ can be projected to the region D on the x, y-plane. Considering the surface as the "top" of a pseudo-cylinder, and D as the base, the volume of the resulting solid is
$$V = \iint_D f(x,y)\,\mathrm{d}A$$

- **Volume of a Region Rotated:** Very similarly to finding the volume of a function of a single variable being rotated about an axis, the volume of an arbitrary region can also be found when rotated; about the x-axis:
$$V = \iint_D 2\pi y\,\mathrm{d}A;$$
and the y-axis:
$$V = \iint_D 2\pi x\,\mathrm{d}A.$$

- **Average Value of a Function over a Region:** Given a function $f(x,y)$ over a region D of area A, the average value of the function within this region is, simply,
$$f_{\text{average}} = \frac{\iint_D f(x,y)\,\mathrm{d}A}{A}.$$

- **Center of Mass:** The center of mass of a region R with density function $\rho(x,y) is given$:
$$(x_{\text{cent}}, y_{\text{cent}}) = \left(\frac{\iint_R x\rho(x,y)\mathrm{d}A}{\iint_R \rho(x,y)\mathrm{d}A}, \frac{\iint_R y\rho(x,y)\mathrm{d}A}{\iint_R \rho(x,y)\mathrm{d}A}\right)$$
Notice that when ρ is constant, it simply cancels out, and the center of mass is simply the average of the x and y values of the region.

10.3 Coordinate Transformations

Consider a vector function $\vec{R} = \vec{R}(\vec{u}, \vec{v})$, as well as figure 22. The gradient of this function, $\vec{\nabla R}$, is given by

$$\vec{\nabla R} = \frac{\partial R}{\partial u}\hat{u} + \frac{\partial R}{\partial v}\hat{v},$$

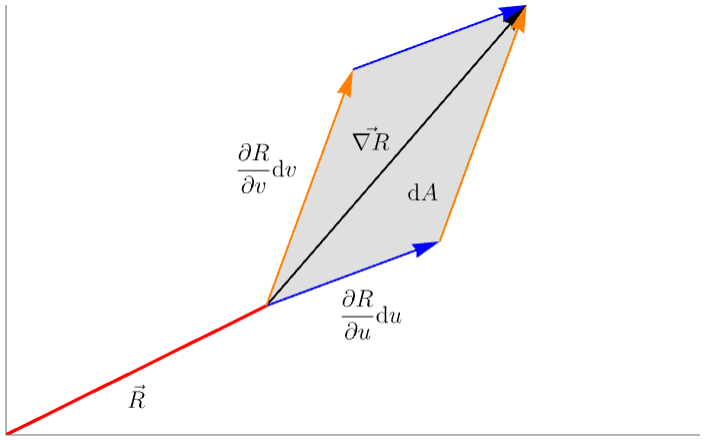


Figure 22: Coordinate transformation

in the base $\{\hat{u}, \hat{v}\}$. The region formed by the $\hat{u}$ and $\hat{v}$ vectors to "add up" to $\vec{\nabla R}$ can be extended (as shown). Consider that the vectors $\hat{u}$ and $\hat{v}$ are changing at rates $\mathrm{d}u$ and $\mathrm{d}v$, respectively; thus, the area (A) of this parallelogram would also be changing accordingly (using the general formula for the area of a parallelogram defined via vectors);

$$\begin{aligned}
\mathrm{A} &= ||\frac{\partial R}{\partial u}\mathrm{d}u \times \frac{\partial R}{\partial v}\mathrm{d}v|| \\
&= ||\frac{\partial R}{\partial u} \times \frac{\partial R}{\partial v}||\mathrm{d}u\mathrm{d}v \\
&= \det \begin{bmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & 0 \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & 0 \end{bmatrix} = \hat{k}(\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u})\,\mathrm{d}u\mathrm{d}v \\
&= |\frac{\partial(x,y)}{\partial(u,v)}|\,\mathrm{d}u\mathrm{d}v
\end{aligned}$$

Thus, $\mathrm{d}A$ is simply (or not so simply?) the determinant of the Jacobian matrix of the relevant variables of transformation. In practice, this means

$$\iint_D f(x,y)\mathrm{d}A = \iint_{D'} f(x(u,v), y(u,v))\,|\frac{\partial(x,y)}{\partial(u,v)}|\,\mathrm{d}u\mathrm{d}v,$$

where D' is a region in the u, v-plane.

One of the most common coordinate transformations is the polar coordinate system, where $x(r,\theta) = r\cos(\theta)$ and $y(r,\theta) = r\sin(\theta)$. We can also more "formally" say that $\vec{R}(r,\theta) = r\cos\theta\hat{i} + r\sin\theta\hat{j}$. Working out the

Jacobian, we get

$$\begin{aligned}\mathrm{d}A = \left|\frac{\partial(x,y)}{\partial(r,\theta)}\right| &= \det\begin{bmatrix}\cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta\end{bmatrix}\mathrm{d}r\mathrm{d}\theta \\ &= (r\cos^2\theta + r\sin^2\theta)\mathrm{d}r\mathrm{d}\theta \\ &= r\,\mathrm{d}r\mathrm{d}\theta\end{aligned}$$

An often-helpful theorem when working with coordinate transformations is as follows;

Theorem. *Two functions $u(x,y)$ and $v(x,y)$ are functionally related as*

$$\Phi(u,v) = 0 \iff \frac{\partial(u,v)}{\partial(x,y)} = 0.$$

Note too that $\frac{\partial(x,y)}{\partial(u,v)} = 1/\frac{\partial(u,v)}{\partial(x,y)}$; practically, this means that (depending on what is easiest to do), we can either rewrite u and v in terms of x and y then find the Jacobian, or we can find the Jacobian and then rewrite u and v in terms of x and y.

10.3.1 Gaussian Integral

Consider the integral $I = \int_0^\infty e^{-x^2}\,\mathrm{d}x$. We can also state that $I = \int_0^\infty e^{-y^2}\,\mathrm{d}y$, and write:

$$\begin{aligned}I^2 &= \int_0^\infty e^{-x^2}\,\mathrm{d}x \int_0^\infty e^{-y^2}\,\mathrm{d}y \\ &= \int_0^\infty \int_0^\infty e^{-x^2}e^{-y^2}\,\mathrm{d}y\mathrm{d}x \\ &= \int_0^\infty \int_0^\infty e^{-(x^2+y^2)}\,\mathrm{d}y\mathrm{d}x\end{aligned}$$

We can then rewrite this in polar coordinates ($x = r\cos\theta$, $y = r\sin\theta$);

$$\begin{aligned}&= \int_0^\infty \int_0^{\frac{\pi}{2}} e^{-(r^2\cos^2\theta + r^2\sin^2\theta)} r\,\mathrm{d}\theta\mathrm{d}r \\ &= \frac{\pi}{2}\int_0^\infty re^{-r^2}\,\mathrm{d}r \\ &= -\frac{\pi}{4}\left[e^{-r^2}\right]_0^\infty = \frac{\pi}{4}\end{aligned}$$

Thus, $I^2 = \frac{\pi}{4} \implies I = \frac{\sqrt{\pi}}{2}$. You can alternatively solve the integral $\int_{-\infty}^\infty e^{-x^2}\,\mathrm{d}x$ by the same method, which would result in more straightforward change-of-bounds when converting to polar coordinates.

10.3.2 Generalizing Triginometric Integrals

Take the Gamma function, $\Gamma(n) = \int_0^\infty e^{-x}x^{n-1}\,\mathrm{d}x$. Substituting $x = w^2, \mathrm{d}x = 2w\,\mathrm{d}w$ gives us

$$\Gamma(n) = 2\int_0^\infty e^{-w^2}w^{2n-1}\,\mathrm{d}w = 2\int_0^\infty e^{-x^2}x^{2n-1}\,\mathrm{d}x$$

Multiplying this result by $\Gamma(m)$ (with the same substitution, but in terms of y rather than x) gives:

$$\begin{aligned}\Gamma(n)\Gamma(m) &= 4\int_0^\infty e^{-x^2}x^{2n-1}\,\mathrm{d}x\int_0^\infty e^{-y^2}y^{2m-1}\,\mathrm{d}y\\ &= 4\int_0^\infty\int_0^\infty e^{-(x^2+y^2)}x^{2n-1}y^{2m-1}\,\mathrm{d}y\,\mathrm{d}x\end{aligned}$$

Rewriting this in terms of polar coordinates:

$$\begin{aligned}\Gamma(n)\Gamma(m) &= 4\int_0^\infty\int_0^{\frac{\pi}{2}} e^{-r^2}r^{2n-1}\cos^{2n-1}r^{2m-1}\sin^{2m-1}\,\mathrm{d}\theta\mathrm{d}r\\ &= \mathbf{2}\int_{\mathbf{0}}^\infty \mathbf{e^{-r^2}r^{2(m+n)-1}}\,\mathbf{dr}\,2\int_0^{\frac{\pi}{2}}\cos^{2n-1}\theta\sin^{2m-1}\theta\,\mathrm{d}\theta\\ &= \mathbf{\Gamma(n+m)}\int_0^{\frac{\pi}{2}}\cos^{2n-1}\theta\sin^{2m-1}\theta\,\mathrm{d}\theta\\ \implies \int_0^{\frac{\pi}{2}}\cos^{2n-1}\theta\sin^{2m-1}\theta\,\mathrm{d}\theta &= \frac{\Gamma(n)\Gamma(m)}{2\Gamma(n+m)}\end{aligned}$$

10.3.3 Laplace Transform

The Laplace Transform, denoted $\mathcal{L}$, is the integral $\mathcal{L}(f(s)) = \int_0^\infty f(t)e^{-st}\,\mathrm{d}t$. A particularly useful transform is

$$\mathcal{L}(t^n) = \int_0^\infty e^{-st}t^n\,\mathrm{d}t,$$

where $s > 0$. Taking $st = x \implies \mathrm{d}t = \frac{\mathrm{d}x}{s}$, we have

$$\begin{aligned}\mathcal{L}(t^n) &= \int_0^\infty e^{-x}\frac{x^n}{s^n}\frac{\mathrm{d}x}{s}\\ &= \frac{1}{s^{n+1}}\int_0^\infty e^{-x}x^n\,\mathrm{d}x\\ &= \frac{\Gamma(n+1)}{s^{n+1}}\end{aligned}$$

Thus, $\mathcal{L}(t^n) = \frac{n!}{s^{n+1}}$ where n is a positive integer.

10.4 Triple Integrals

Consider a function $f(x, y, z)$, continuous in a domain D. The sum of the function over the domain can be written as

$$\lim_{p,n,m\to\infty} \sum_{k=1}^{p} \sum_{j=1}^{n} \sum_{i=1}^{m} f(x^*_{ijk}, y^*_{ijk}, z^*_{ijk}) \Delta V_{ijk} = \iiint_D f(x, y, z)\, \mathrm{d}V$$

10.5 Applications of Triple Integrals

- **Function Evaluated over Region Between Two Surfaces**: consider the region D, enclosed by the two surfaces described by $z = \psi_B(x, y)$ and $z = \psi_T(x, y)$, where $\psi_T > \psi_B$. The value of $f(x, y, z)$ over this region can thus be written as

  $$\iiint_D f(x, y, z)\, \mathrm{d}V = \iint_{R_{xy}} \left[\int_{\psi_B}^{\psi_T} f(x, y, z)\, \mathrm{d}z \right] \mathrm{d}A,$$

  where R_{xy} is the projection of D on the xy-plane. When $f(x, y, z) = 1$, this is equivalent to the volume of the region D:

  $$V = \iiint_D \mathrm{d}V = \iint_{R_{xy}} \left[\int_{\psi_B}^{\psi_T} \mathrm{d}z \right] \mathrm{d}A = \iint_{R_{xy}} [\psi_T(x, y) - \psi_B(x, y)]\, \mathrm{d}A$$

- **Same Applications as in Double Integrals**: just this time, with 3D surfaces instead of 2D curves/areas.

10.5.1 Coordinate Transformations

Similarly to in double integrals, we have

$$\mathrm{d}V = \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| \mathrm{d}u\, \mathrm{d}v\, \mathrm{d}w.$$

Some notable coordinate systems include:

- **Spherical**:

  $$\begin{cases} x = \rho \sin\phi \cos\theta \\ y = \rho \sin\phi \sin\theta \\ z = \rho \cos\phi \end{cases},$$

  where the natural bounds are $0 \le r \le \infty$, $0 \le \phi \le \pi$, and $0 \le \theta \le 2\pi$. The Jacobian is $\frac{\partial(x,y,z)}{\partial(\rho,\phi,\theta)} = \rho^2 \sin\phi$.

  Also note that $x^2 + y^2 + z^2 = \rho^2$, a helpful fact for some integrals.

- **Cylindrical**:

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \\ z = z \end{cases},$$

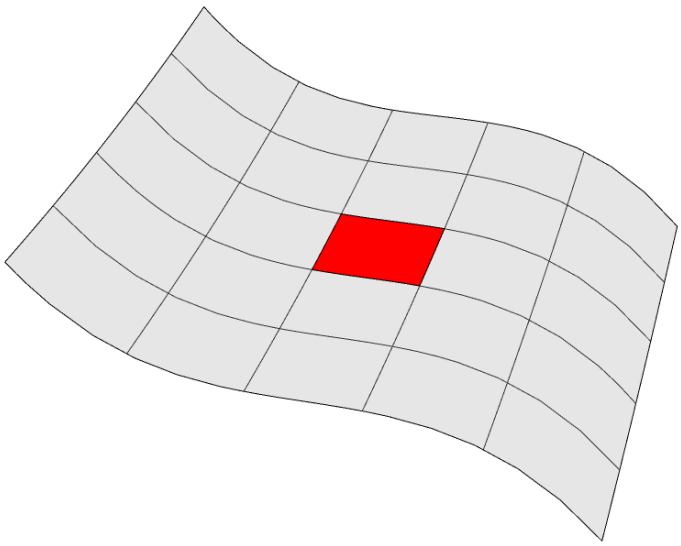
where the natural bounds are $0 \leq r \leq \infty$, $0 \leq \theta \leq 2\pi$, and $-\infty \leq z \leq \infty$. The Jacobian is $\frac{\partial(x,y,z)}{\partial(r,\theta,z)} = r$.

Part III

Vector Calculus (Brief)

11 Vector Calculus

11.1 Surface Integration



11.1.1 Motivations

Consider the following descriptions of surfaces as vector functions in 3D:

- **Cartesians**, $z = f(x, y)$ We have:

$$\vec{R}(x, y) = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + f(x, y)\hat{\mathbf{k}}$$

- **Cylindricals**, height H radius a
  - Bottom:

$$\vec{R}_B(\theta, r) = r\cos\theta\hat{\mathbf{i}} + r\sin\theta\hat{\mathbf{j}} + 0\hat{\mathbf{k}}$$

- Top:
$$\vec{R}_T(\theta, r) = r\cos\theta\hat{\mathbf{i}} + r\sin\theta\hat{\mathbf{j}} + H\hat{\mathbf{k}}$$
i.e. a circle
- Lateral:
$$\vec{R}_L(\theta, z) = a\cos\theta\hat{\mathbf{i}} + a\sin\theta\hat{\mathbf{j}} + z\hat{\mathbf{k}}$$

- **Flat Cone**, radius a angle α with $x-y$ plane
$$\vec{R}_c(\rho, \theta) = \rho\sin\alpha\cos\theta\hat{\mathbf{i}} + \rho\sin\alpha\sin\theta\hat{\mathbf{j}} + \rho\cos\alpha\hat{\mathbf{k}}$$

- **Spherical**, radius a
$$\vec{R}(\phi, \theta) = a\sin\phi\cos\theta\hat{\mathbf{i}} + a\sin\phi\sin\theta\hat{\mathbf{j}} + a\cos\phi\hat{\mathbf{k}}$$

Notice, all of surfaces can be described as a vectorial function of two variables, ie,

$$\vec{R} = \vec{R}(u, v).$$

As was shown in the derivation of Jacobians, it follows that the cross product of the change in surface $\mathrm{d}\vec{R}$ represents the change in area along the surface (viz the area of a parallelogram) Specifically, we have

$$\mathrm{d}S = \left|\left|\frac{\partial\vec{R}}{\partial u}\mathrm{d}u \times \frac{\partial\vec{R}}{\partial v}\mathrm{d}v\right|\right|,$$

the infinitesimal surface area change. It follows that the total surface area of a surface described by $\vec{R}(u, v)$ over a domain D_{uv} would be

$$S = \iint_{D_{uv}} \left|\left|\frac{\partial\vec{R}}{\partial u} \times \frac{\partial\vec{R}}{\partial v}\right|\right| \mathrm{d}u\mathrm{d}v.$$

11.1.2 "Special" Cases

We can "specify" the surface integral formula

$$S = \iint_{D_{uv}} \left|\left|\frac{\partial\vec{R}}{\partial u} \times \frac{\partial\vec{R}}{\partial v}\right|\right| \mathrm{d}u\mathrm{d}v$$

for a number of "special" cases.

- **Cartesians**, $z = f(x, y) \implies \vec{R}(x, y) = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + f(x, y)\hat{\mathbf{k}}$. We thus have $u = x$, $v = y$, and we can

compute:

$$\frac{\partial \vec{R}}{\partial u} = \frac{\partial \vec{R}}{\partial x} = \hat{\mathbf{i}} + 0\hat{\mathbf{j}} + f_x\hat{\mathbf{k}}$$
$$\frac{\partial \vec{R}}{\partial v} = \frac{\partial \vec{R}}{\partial y} = 0\hat{\mathbf{i}} + \hat{\mathbf{j}} + f_y\hat{\mathbf{k}}$$

Taking the cross product of these two:

$$\begin{aligned}\frac{\partial \vec{R}}{\partial x} \times \frac{\partial \vec{R}}{\partial y} &= \det \begin{bmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ 1 & 0 & f_x \\ 0 & 1 & f_y \end{bmatrix} \\ &= -f_x\hat{\mathbf{u}} - f_y\hat{\mathbf{j}} + \hat{\mathbf{k}} \\ ||...|| &= \sqrt{(f_x)^2 + (f_y)^2 + 1}\end{aligned}$$

Thus, the surface area for a surface described with Cartesians is

$$S = \iint_{D_{xy}} \sqrt{(f_x)^2 + (f_y)^2 + 1}\, \mathrm{d}x\mathrm{d}y.$$

- **Sphericals**, $\vec{R}(\theta, \phi) = a\sin\phi\cos\theta\hat{\mathbf{i}} + a\sin\phi\sin\theta\hat{\mathbf{j}} + a\cos\phi\hat{\mathbf{k}}$. We thus have the derivatives

$$\frac{\partial \vec{R}}{\partial \theta} = -a\sin\phi\sin\theta\hat{\mathbf{i}} + a\sin\phi\cos\theta\hat{\mathbf{j}} + 0\hat{\mathbf{k}}$$
$$\frac{\partial \vec{R}}{\partial \phi} = a\cos\phi\cos\theta\hat{\mathbf{i}} + a\cos\phi\sin\theta\hat{\mathbf{j}} - a\cos\phi\hat{\mathbf{k}}.$$

And their cross products:

$$\begin{aligned}\frac{\partial \vec{R}}{\partial \theta} \times \frac{\partial \vec{R}}{\partial \phi} &= \det \begin{bmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ -a\sin\phi\sin\theta & a\sin\phi\cos\theta & 0 \\ a\cos\phi\cos\theta & a\cos\phi\sin\theta & -a\cos\phi \end{bmatrix} \\ &= \cdots \\ ||...|| &= a^2\sin\phi\end{aligned}$$

This gives us a surface area of

$$S = \iint_{D_{\theta\phi}} a^2\sin\phi\, \mathrm{d}\theta\mathrm{d}\phi.$$

NB: the similarity with the Jacobian.

11.1.3 Applications

Many general applications from multiple integration can be applied to surface integrals; instead of evaluating over a volume, we evaluate over a surface now. For instance, the mass of a "sheet" of variable density $\rho(x, y, z)$ over a domain S can be evaluated as

$$\iint_S \rho(x, y, z)\, \mathrm{d}S.$$

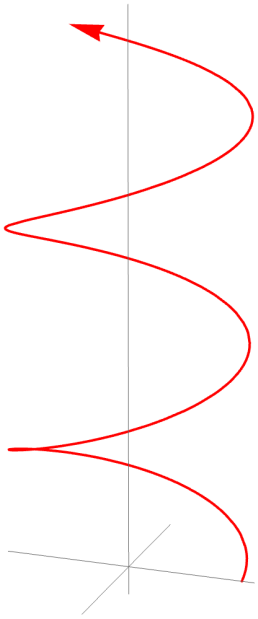
It follows that the center of mass of a particular surface would be

$$\bar{x} = \frac{1}{m} \iint_S x\rho(x, y, z)\, \mathrm{d}S, \quad \text{etc} \cdots .$$

More broadly, the sum of a function over a particular surface is rather straightforward:

$$\iint_S f(x, y)\mathrm{d}S$$

11.2 Differential Geometry



11.2.1 Motivations: Motion on Curves

We can describe the motion of an object in space by parameterizing it with respect to some variable t (often representing time):

$$\vec{R} = x(t)\hat{\mathbf{i}} + y(t)\hat{\mathbf{j}} + z(t)\hat{\mathbf{k}}.$$

It follows naturally that

$$\frac{\mathrm{d}\vec{R}}{\mathrm{d}t} = \dot{x}(t)\hat{\mathbf{i}} + \dot{y}(t)\hat{\mathbf{j}} + \dot{z}(t)\hat{\mathbf{k}} = \vec{v}(t)$$
$$\frac{\mathrm{d}^2\vec{R}}{\mathrm{d}t^2} = \ddot{x}(t)\hat{\mathbf{i}} + \ddot{y}(t)\hat{\mathbf{j}} + \ddot{z}(t)\hat{\mathbf{k}} = \vec{a}(t).$$

Consider some small change in $\vec{R}$ of $\Delta\vec{R}$; the norm of this vector is approximately equal to the change in arc length, Δs. It follows:

$$\begin{aligned} ||\Delta\vec{R}|| \approx \Delta s \implies ||\frac{\Delta\vec{R}}{\Delta s}|| &\approx 1 \\ \implies \frac{\mathrm{d}\vec{R}}{\mathrm{d}s} &= \lim_{\Delta s \to 0} \frac{\Delta\vec{R}}{\Delta s} = \hat{T} \end{aligned}$$

where $\hat{T}$ is the **unit tangent vector**, ie the "vector of change" of the curve with respect to the curve length. We can thus write:

$$\begin{aligned} \vec{v} = \frac{\mathrm{d}\vec{R}}{\mathrm{d}t} &= \frac{\mathrm{d}\vec{R}}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}t} \\ &= \frac{\mathrm{d}s}{\mathrm{d}t}\hat{T} \\ \implies \vec{a} = \frac{\mathrm{d}\vec{v}}{\mathrm{d}t} &= \frac{\mathrm{d}^2 s}{\mathrm{d}t^2}\hat{T} + \frac{\mathrm{d}s}{\mathrm{d}t}\frac{\mathbf{d}\hat{\mathbf{T}}}{\mathbf{dt}} \end{aligned}$$

NB: $\frac{\mathrm{d}s}{\mathrm{d}t}$ is simply the norm of $\frac{\mathrm{d}\vec{R}}{\mathrm{d}t}$, ie the speed of the object; it points, naturally, in the direction of $\hat{T}$.

Unfortunately, we don't know how $\hat{T}$ changes wrt time, t, directly; we can still compute, however, by the following:

$$\frac{\mathrm{d}\hat{T}}{\mathrm{d}t} = \frac{\mathrm{d}\hat{T}}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}t}$$

$$\begin{aligned} \textit{Solving for } \frac{\mathrm{d}\hat{T}}{\mathrm{d}s}: \quad \hat{T}\cdot\hat{T} &= 1 \\ \implies \frac{\mathrm{d}}{\mathrm{d}s}[\hat{T}\cdot\hat{T}] = 0 &= \frac{\mathrm{d}\hat{T}}{\mathrm{d}s}\cdot\hat{T} + \frac{\mathrm{d}\hat{T}}{\mathrm{d}s}\hat{T} \\ \implies 2[\frac{\mathrm{d}\hat{T}}{\mathrm{d}s}\cdot\hat{T}] &= 0 \end{aligned}$$

The dot product of these two vectors can be zero under the following cases:

1. $\hat{T} = \vec{0}$: under the assumption that we are working with smooth ($\iff$ differentiable) curves, this is not possible

2. $\frac{\mathrm{d}\hat{T}}{\mathrm{d}s} = \vec{0}$: this is only possible if the curve is a straight line, which we will assume for now is not the case.

3. $\hat{T} \perp \frac{\mathrm{d}\hat{T}}{\mathrm{d}s}$: this is the "standard case" we will consider; it should also make intuitive sense; the motion that a tangent vector to a curve is changing is perpendicular to the curve itself (Consider the most "extreme" case, of a motion about a circle).

We introduce a new notation here:

$$\frac{\mathrm{d}\hat{T}}{\mathrm{d}s} = \kappa\hat{N},$$

where $\kappa = ||\frac{\mathrm{d}\hat{T}}{\mathrm{d}s}||$, the **curvature** of R, and $\hat{N}$ is the **unit principal normal**. This "derives" rather clearly: if $\frac{\mathrm{d}\hat{T}}{\mathrm{d}s}$ is normal to $\hat{T}$, then it is a scalar multiple of a unit vector perpendicular to $\hat{T}$ (ie $\hat{N}$), where the scalar is the magnitude of $\hat{T}$.

Returning to the earlier definition of $\vec{a}$, we can write;

$$\begin{aligned}
\frac{\mathrm{d}\hat{T}}{\mathrm{d}t} = \frac{\mathrm{d}s}{\mathrm{d}t}\frac{\mathrm{d}\hat{T}}{\mathrm{d}t} &= \frac{\mathrm{d}s}{\mathrm{d}t}\frac{\mathrm{d}s}{\mathrm{d}t}\frac{\mathrm{d}\hat{T}}{\mathrm{d}s} \\
&= \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 \frac{\mathrm{d}\hat{T}}{\mathrm{d}s} \\
&= \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 \kappa\hat{N} \\
\implies \vec{a} &= \frac{\mathrm{d}^2 s}{\mathrm{d}t^2}\hat{T} + \kappa\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 \hat{N} \\
&\equiv a_T\hat{T} + a_N\hat{N}
\end{aligned}$$

where, in the final line, a_T and a_N represent the tangential and normal components of acceleration respectively.

NB: the norm of $\vec{a}$ can be written as $||\vec{a}||^2 = a_T^2 + a_N^2$, as per the definition of the norm of a vector. This is often a helpful result to find either of the components of $\vec{a}$ when given the other, as often a_T is fairly easy to compute as it only relies on $\frac{\mathrm{d}s}{\mathrm{d}t}$ while a_N relies on knowing κ.

11.2.2 The $\hat{T}, \hat{N}, \hat{B}$ frame

Previously, we defined the unit tangent and normal vectors. In order to "complete" our span of $\mathbb{R}^3$, we must define a third vector perpendicular to both of these vectors, viz.,

$$\hat{B} = \hat{T} \times \hat{N}.$$

$\hat{B}$ is called the **binormal unit vector**. Together, these three vectors form a frame that can be used to describe the motion of along a curve (specifically, a smooth curve...) in space. It differs from the "standard" $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ frame as it is not fixed in space; the vectors "follow" the motion of the curve. In addition to this curve, we also briefly define the *radius of curvature*,

$$\rho = \frac{1}{\kappa}.$$

How these vectors further interact (or, more accurately, their derivatives interact) is described in the next section.

11.2.3 The Frenet-Serret Equations

The Frenet-Serret Equations describe the relationship between the derivatives of the $\hat{T}$, $\hat{N}$, and $\hat{B}$ vectors. Previously, we defined

$$\frac{\mathrm{d}\hat{T}}{\mathrm{d}s} = \kappa\hat{N}. \qquad (1)$$

Our goal is to similarly define each of the derivatives of the vectors in the frame $\{\hat{T}, \hat{N}, \hat{B}\}$.

Consider $\hat{B} = \hat{T} \times \hat{N}$. Taking the derivative wrt arc length:

$$\begin{aligned}
\frac{\mathrm{d}\hat{B}}{\mathrm{d}s} &= \frac{\mathrm{d}\hat{T}}{\mathrm{d}s} \times \hat{N} + \hat{T} \times \frac{\mathrm{d}\hat{N}}{\mathrm{d}s} \\
&= \kappa\hat{N} \times \hat{N} + \hat{T} \times \frac{\mathrm{d}\hat{N}}{\mathrm{d}s} \\
&= 0 + \hat{T} \times \frac{\mathrm{d}\hat{N}}{\mathrm{d}s} \\
\implies \frac{\mathrm{d}\hat{B}}{\mathrm{d}s} &\perp \hat{T}
\end{aligned}$$

We now know that $\frac{\mathrm{d}\hat{B}}{\mathrm{d}s}$ is perpendicular to $\hat{T}$; however, this is rather ambiguous, as this could really be in any of four directions (albeit some we can intuitively rule out). To clear this up, we can use a similar "trick" as before, and consider the derivative of the dot product of $\hat{B}$:

$$\begin{aligned}
\hat{B} \cdot \hat{B} = 1 \implies \frac{\mathrm{d}}{\mathrm{d}s}[\hat{B} \cdot \hat{B}] &= 0 \\
2\frac{\mathrm{d}\hat{B}}{\mathrm{d}s} \cdot \hat{B} &= 0
\end{aligned}$$

Under the same assumptions of smoothness, etc, as discussed previously, we can thus conclude that $\frac{\mathrm{d}\hat{B}}{\mathrm{d}s} \perp \hat{B}$. Thus, as $\frac{\mathrm{d}\hat{B}}{\mathrm{d}s} \perp \hat{B}$ *and* $\perp \hat{T}$, the only direction left for $\frac{\mathrm{d}\hat{B}}{\mathrm{d}s}$ to point is along $\hat{N}$. We thus write:

$$\frac{\mathrm{d}\hat{B}}{\mathrm{d}s} = -\tau\hat{N}, \qquad (2)$$

where τ is the **torsion** of the curve, and is negative by convention.

Finally, we seek to find $\frac{\mathrm{d}\hat{N}}{\mathrm{d}s}$. We know $\hat{N} = \hat{B} \times \hat{T}$, thus

$$\begin{aligned}\frac{\mathrm{d}\hat{N}}{\mathrm{d}s} &= \frac{\mathrm{d}\hat{B}}{\mathrm{d}s} \times \hat{T} + \hat{B} \times \frac{\mathrm{d}\hat{T}}{\mathrm{d}s} \\ &= (2) \times \hat{T} + \hat{B} \times (1) \\ &= -\tau\hat{N} \times \hat{T} + \hat{B} \times \kappa\hat{N} \\ &= \tau\hat{B} - \kappa\hat{T}, \qquad (3)\end{aligned}$$

where each step is justified by the earlier definitions as well as the fundamental workings of the cross product of two vectors.

From (1), (2), and (3), we typically write the following (for ease of computation/memory):

$$\begin{bmatrix} \frac{\mathrm{d}\hat{T}}{\mathrm{d}s} \\ \frac{\mathrm{d}\hat{N}}{\mathrm{d}s} \\ \frac{\mathrm{d}\hat{B}}{\mathrm{d}s} \end{bmatrix} = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \begin{bmatrix} \hat{T} \\ \hat{N} \\ \hat{B} \end{bmatrix}.$$

Using these definitions, any curve R described in terms of $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ can be rewritten in terms of the $\hat{T}, \hat{N}, \hat{B}$ frame.

11.2.4 Notes: Helpful Results

- **Arc Length:**

  Recall that $\frac{\mathrm{d}s}{\mathrm{d}t}$, the derivative of arc length with respect to time, can be fairly easily found by finding the norm of the velocity vector. Thus, to find the arc length of a curve;

  $$s(t) = \int_0^t \frac{\mathrm{d}s}{\mathrm{d}\tau}\,\mathrm{d}\tau,$$

  assuming the curve "starts" at $t = 0$. By extension, the upper/lower limits can be changed to find the arc lengths between any particular points in time.

- **Obtaining Curvature:**

  We often wish to describe the curvature, κ, of a curve without necessarily finding all of the derivatives of

the vectors in the $\hat{T}, \hat{N}, \hat{B}$ frame. Consider the following:

$$\begin{aligned}
\frac{\mathrm{d}\vec{R}}{\mathrm{d}t} \times \frac{\mathrm{d}^2\vec{R}}{\mathrm{d}t^2} &= \vec{v} \times \vec{a} \\
&= \left[\frac{\mathrm{d}s}{\mathrm{d}t}\hat{T}\right] \times \left[\frac{\mathrm{d}^2 s}{\mathrm{d}t^2}\hat{T} + \kappa\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 \hat{N}\right] \\
&= \frac{\mathrm{d}s}{\mathrm{d}t}\hat{T} \times \frac{\mathrm{d}^2 s}{\mathrm{d}t^2}\hat{T} + \kappa\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^3 \hat{N} \times \hat{T} \\
&= 0 + \kappa\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^3 \hat{B} \\
&= \kappa\left|\left|\frac{\mathrm{d}\vec{R}}{\mathrm{d}s}\right|\right|^3 \hat{B} \\
&\overset{\text{norm}}{\Longrightarrow} \kappa = \frac{||\vec{v} \times \vec{a}||}{||\vec{v}||^3}
\end{aligned}$$

Rewriting this more generally, we have

$$\kappa = \frac{||\frac{\mathrm{d}\vec{R}}{\mathrm{d}t} \times \frac{\mathrm{d}^2\vec{R}}{\mathrm{d}t^2}||}{||\frac{\mathrm{d}R}{\mathrm{d}t}||^3}$$

From here, we can consider a number of special cases:

- **Planar Curves**; ie, no movement in $\hat{\mathbf{k}}$. We thus have $\vec{R}(t) = x(t)\hat{\mathbf{i}} + y(t)\hat{\mathbf{j}}$ (and natural higher-order derivatives). It follows naturally that:

  $$\kappa = \frac{|\dot{x}\ddot{y} - \dot{y}\ddot{x}|}{(\dot{x}^2 + \dot{y}^2)^{3/2}}$$

  Consider a circle, for instance, of $x = h + a\cos t$, $y = k + a\sin t$. The curvature would be give

  $$\begin{aligned}
  \kappa &= \frac{|(-a\sin t)(-a\sin t) - (a\cos t)(-a\cos t)|}{(a^2\sin^2 t + a^2\cos^2 t)^{3/2}} \\
  &= \cdots = \frac{1}{a}
  \end{aligned}$$

  It follows that $\rho = a$, ie the radius of curvature of a circle is its radius.

- **Function of** x; say we have $y = f(x)$. We can parametrize this curve as $x = t$, $y = f(t)$, and

compute the curvature:

$$\begin{aligned}\kappa &= \frac{|f''(t) - 0|}{(1 + f'(t)^2)^{3/2}} \\ &= \frac{\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}}{(1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2)^{3/2}}\end{aligned}$$

i.e., the curvature in "Cartesian form".

- **Direction of $\hat{N}$:**

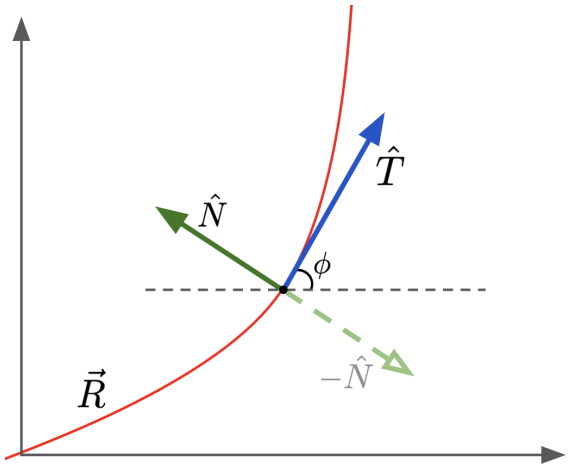


Figure 23: The direction of $\hat{N}$ is always along the "inside" of a curve.

To determine the direction of $\hat{N}$ along a planar curve, one must consider the curvature; $\hat{N}$ always points towards the "inside" of curve, ie, along increasing angle of $\hat{T}$ with the horizontal. See Figure 23.

- **"Planes":**

Given the definitions of the $\hat{T}$, $\hat{N}$, and $\hat{B}$ vectors, we can take the span of the combination of any two of these vectors to form a particular plane to a curve. Namely:

  - $\text{span}(\hat{B}, \hat{N}), \perp \hat{T}$: **Normal Plane**
  - $\text{span}(\hat{T}, \hat{B}), \perp \hat{N}$: **Rectifying Plane**
  - $\text{span}(\hat{T}, \hat{N}), \perp \hat{B}$: **Osculating Plane**

These can be most easily obtained by finding the appropriate orthogonal vector, then finding the plane in question.

- **Area of a Parametrically Defined Curve:**

Say we have a function defined parametrically as $x = x(t)$, $y = y(t)$. The area under a function $F(x)$ is

given:

$$A = \int_a^b F(x)\,\mathrm{d}x.$$

We can substitute $x(t)$ for x, meaning $\mathrm{d}x = \frac{\mathrm{d}x}{\mathrm{d}t}\mathrm{d}t$. We thus have $F(x(t))$, meaning F can simply be written as a function of t, which we will take as our $y(t)$. All together, this gives us:

$$A = \int_a^b y(t)x'(t)\,\mathrm{d}t$$

A similar result can be used to obtain a formula for the volume of a revolved parametrically defined surface.

11.2.5 Osculating Circle

An **osculating circle** is one that touches a curve at a point, such that it has the same curvature and tangent vector at that point.

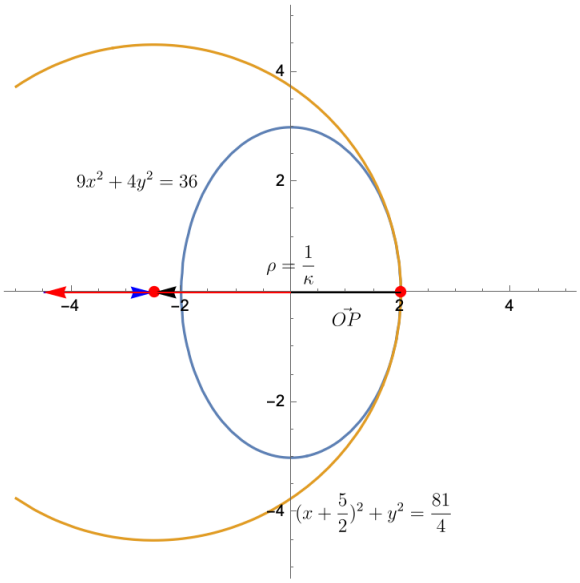


Figure 24: An osculating circle

Generally, to find the osculating circle of a curve:

1. Parametrize the curve as $x = x(t)$, $y = y(t)$.
2. Find the curvature at the point of interest, κ; from here, the radius of the circle is simply $\rho = \frac{1}{\kappa}$.
3. Find the normal vector ($\hat{N}$) at the point of interest.
4. The vector function of the osculating circle is then given by $\vec{R} = \vec{OP} + \rho\hat{N}$, where $\vec{OP}$ is the vector from the origin to the point of interest.

11.2.6 Polar Coordinates

We can describe the unit vectors of a polar coordinate system as follows:

$$\begin{aligned}\hat{\mu}_r &= \cos\theta\hat{\mathbf{i}} + \sin\theta\hat{\mathbf{j}} \\ \hat{\mu}_\theta &= \cos\Big(\theta + \frac{\pi}{2}\Big)\hat{\mathbf{i}} + \sin\Big(\theta + \frac{\pi}{2}\Big)\hat{\mathbf{j}} = -\sin\theta\hat{\mathbf{i}} + \cos\theta\hat{\mathbf{j}}\end{aligned}$$

We thus take

$$\begin{aligned}\vec{R} &= r\hat{\mu}_r \\ \frac{\mathrm{d}\vec{R}}{\mathrm{d}t} &= \frac{\mathrm{d}r}{\mathrm{d}t}\hat{\mu}_r + r\frac{\mathrm{d}\hat{\mu}_r}{\mathrm{d}t} \\ \frac{\mathrm{d}\hat{\mu}_r}{\mathrm{d}t} &= \frac{\mathrm{d}\hat{\mu}_r}{\mathrm{d}\theta}\frac{\mathrm{d}\theta}{\mathrm{d}t} \\ &= (-\sin\theta\hat{\mathbf{i}} + \cos\theta\hat{\mathbf{j}})\frac{\mathrm{d}\theta}{\mathrm{d}t} = \hat{\mu}_\theta\frac{\mathrm{d}\theta}{\mathrm{d}t} \\ \implies \vec{v} &= \frac{\mathrm{d}\vec{R}}{\mathrm{d}t} = \frac{\mathrm{d}r}{\mathrm{d}t}\hat{\mu}_r + r\hat{\mu}_\theta\frac{\mathrm{d}\theta}{\mathrm{d}t} \\ &\equiv v_r\hat{\mu}_r + r\dot{\theta}\hat{\mu}_\theta = v_r\hat{\mu}_r + v_\theta\hat{\mu}_\theta, \text{ where } v_r = \frac{\mathrm{d}r}{\mathrm{d}t}, v_\theta = r\frac{\mathrm{d}\theta}{\mathrm{d}t}\end{aligned}$$

For acceleration, we have:

$$\begin{aligned}\vec{a} &= \frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = \ddot{r}\hat{\mu}_r + \frac{\mathrm{d}\hat{\mu}_r}{\mathrm{d}t}\dot{r} + \dot{r}\dot{\theta}\hat{\mu}_\theta + r\ddot{\theta}\hat{\mu}_\theta + r\dot{\theta}\frac{\mathrm{d}\hat{\mu}_\theta}{\mathrm{d}t} \\ \frac{\mathrm{d}\hat{\mu}_r}{\mathrm{d}t} &= \dot{\theta}\hat{\mu}_\theta, \frac{\mathrm{d}\hat{\mu}_\theta}{\mathrm{d}t} = -\dot{\theta}\hat{\mu}_r \\ \implies \vec{a} &= \left[\ddot{r} - r\dot{\theta}^2\right]\hat{\mu}_r + \left[2\dot{r}\dot{\theta} + r\ddot{\theta}\right]\hat{\mu}_\theta \\ &\equiv a_r\hat{\mu}_r + a_\theta\hat{\mu}_\theta, \text{ where } a_r = \ddot{r} - r\dot{\theta}^2, a_\theta = 2\dot{r}\dot{\theta} + r\ddot{\theta}\end{aligned}$$

NB: to find the radial/circumferential components of velocity/acceleration, we can simply take the dot product of the vector with the appropriate unit vector. For instance:

$$\begin{aligned}v_r &= \vec{v}\cdot\hat{\mu}_r; \quad v_\theta = \vec{v}\cdot\hat{\mu}_\theta \\ a_r &= \vec{a}\cdot\hat{\mu}_r; \quad a_\theta = \vec{a}\cdot\hat{\mu}_\theta\end{aligned}$$

11.2.7 Arc Length and Area of Curves with Polars

From the previous section, we had $\vec{v} = \frac{\mathrm{d}r}{\mathrm{d}t}\hat{\mu}_r + r\frac{\mathrm{d}\theta}{\mathrm{d}t}\hat{\mu}_\theta$. We can thus write:

$$\begin{aligned} ||\vec{v}|| &= \sqrt{\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 + r^2\left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^2} = \frac{\mathrm{d}s}{\mathrm{d}t} \\ \mathrm{d}s &= \mathrm{d}t\sqrt{\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 + r^2\left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^2} \\ &= \sqrt{\left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2 + r^2}\,\mathrm{d}\theta \end{aligned}$$

The length of the curve between two points (well, angles, specifically) α and β would thus be the sum of the changes of arc length between the points, namely

$$L = s = \int ds = \int_\alpha^\beta \sqrt{r^2 + \left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)}\,\mathrm{d}\theta \equiv \text{ length of curve } r = f(\theta)$$

As shown in previous sections, the area between two curves can be described as a double integral over the region described by the relevant area. In polars, we can write

$$A = \int_\alpha^\beta \int_{g(\theta)}^{f(\theta)} r\,\mathrm{d}r\mathrm{d}\theta = \int_\alpha^\beta \frac{[f(\theta)]^2 - [g(\theta)]^2}{2}\,\mathrm{d}\theta$$

End