Review Mathematics

Partial Fractions

A way of "breaking apart" fractions with polynomials in them.

What are Partial Fractions?

We can do *this* directly:

$$\frac{2}{x-2} + \frac{3}{x+1} \longrightarrow \frac{5x-4}{x^2-x-2}$$

Like this:

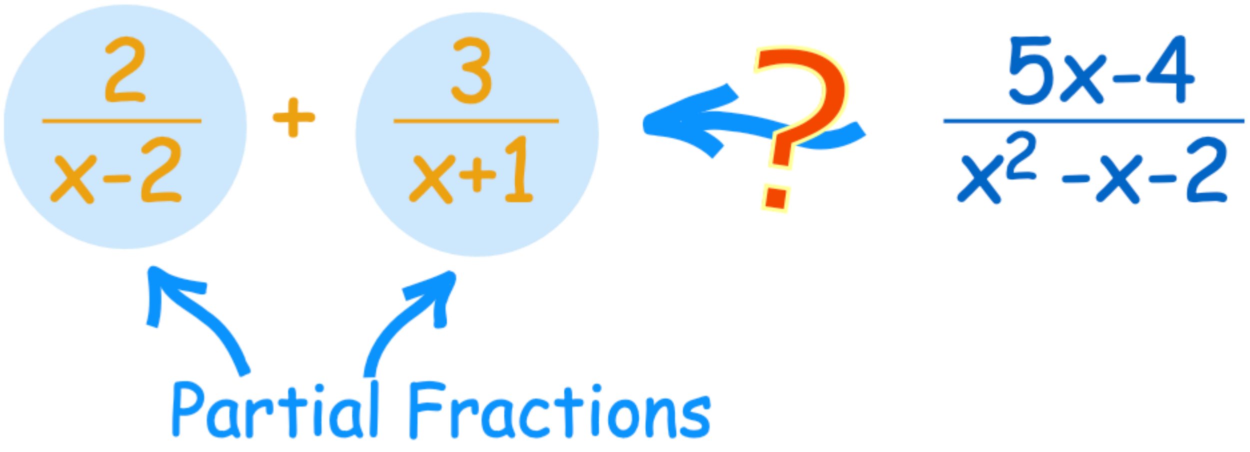
$$\frac{2}{x-2} + \frac{3}{x+1} = \frac{2(x+1) + 3(x-2)}{(x-2)(x+1)}$$

Which can be simplified using Rational Expressions to:

$$= \frac{2x+2 + 3x-6}{x^2+x-2x-2}$$

$$= \frac{5x-4}{x^2-x-2}$$

... but how do we go in the opposite direction?



That is what we are going to discover:

How to find the "parts" that make the single fraction (the "**partial fractions**").

Why Do We Want Them?

First of all ... why do we want them?

Because the partial fractions are each **simpler**.

Partial Fraction Decomposition

So let me show you how to do it.

The method is called *"Partial Fraction Decomposition"*, and goes like this:

Step 1: Factor the bottom

$$\frac{5x-4}{x^2-x-2} = \frac{5x-4}{(x-2)(x+1)}$$

Step 2: Write one partial fraction for each of those factors

$$\frac{5x-4}{(x-2)(x+1)} = \frac{A_1}{x-2} + \frac{A_2}{x+1}$$

Step 3: Multiply through by the bottom so we no longer have fractions

$$5x-4 = A_1(x+1) + A_2(x-2)$$

Step 4: Now find the constants A_1 and A_2

Substituting the roots, or "zeros", of (x−2)(x+1) can help:

Root for (x+1) is x = -1

$$\begin{aligned} 5(-1)-4 &= A_1(-1+1) + A_2(-1-2) \\ -9 &= 0 + A_2(-3) \\ A_2 &= 3 \end{aligned}$$

Root for (x-2) is x = 2

$$\begin{aligned} 5(2)-4 &= A_1(2+1) + A_2(2-2) \\ 6 &= A_1(3) + 0 \\ A_1 &= 2 \end{aligned}$$

And we have our answer:

$$\frac{5x-4}{x^2-x-2} = \frac{2}{x-2} + \frac{3}{x+1}$$

Proper Rational Expressions

Firstly, this only works for **Proper** Rational Expressions, where the degree of the top is **less than** the bottom.

The **degree** is the largest **exponent** the variable has.

- Proper: the degree of the top is less than the degree of the bottom.

Proper: $$\frac{x}{x^3 - 1}$$ degree of top is 1
degree of bottom is 3

- Improper: the degree of the top is greater than, or equal to, the degree of the bottom.

Improper: $$\frac{x^2 - 1}{x + 1}$$ degree of top is 2
degree of bottom is 1

If your expression is Improper, then do polynomial long division first.

Factoring the Bottom

It is up to you to factor the bottom polynomial.

But don't factor them into complex numbers ... you may need to stop some factors at quadratic (called irreducible quadratics because any further factoring leads to complex numbers):

Example: $(x^2-4)(x^2+4)$

- **x^2-4** can be factored into $(x-2)(x+2)$
- But **x^2+4** factors into complex numbers, so don't do it

So the best we can do is:

$$(x-2)(x+2)(x^2+4)$$

So the factors could be a combination of

- linear factors
- irreducible quadratic factors

When you have a quadratic factor you need to include this partial fraction:

$$\frac{B_1x + C_1}{(\text{Your Quadratic})}$$

Factors with Exponents

Sometimes you may get a factor with an exponent, like **$(x-2)^3$** ...

You need a partial fraction for each exponent from 1 up.

Like this:

Example:

$$\frac{1}{(x-2)^3}$$

Has partial fractions

$$\frac{A_1}{x-2} + \frac{A_2}{(x-2)^2} + \frac{A_3}{(x-2)^3}$$

The same thing can also happen to quadratics:

Example:

$$\frac{1}{(x^2+2x+3)^2}$$

Has partial fractions:

$$\frac{B_1x + C_1}{x^2+2x+3} + \frac{B_2x + C_2}{(x^2+2x+3)^2}$$

A Big Example Bringing It All Together

Here is a nice big example for you!

$$\frac{x^2+15}{(x+3)^2\,(x^2+3)}$$

- Because $(x+3)^2$ has an exponent of 2, it needs two terms (A_1 and A_2).
- And (x^2+3) is a quadratic, so it will need $Bx + C$:

$$\frac{x^2+15}{(x+3)^2(x^2+3)} = \frac{A_1}{x+3} + \frac{A_2}{(x+3)^2} + \frac{Bx + C}{x^2+3}$$

Now multiply through by **$(x+3)^2(x^2+3)$**:

$$x^2+15 = (x+3)(x^2+3)A_1 + (x^2+3)A_2 + (x+3)^2(Bx + C)$$

There is a zero at **$x = -3$** (because $x+3=0$), so let us try that:

$$\mathbf{(-3)}^2+15 = 0 + (\mathbf{(-3)}^2+3)A_2 + 0$$

And simplify it to:

$$24 = 12A_2$$

$$\text{so } A_2=2$$

Let us replace A_2 with 2:

$$x^2+15 = (x+3)(x^2+3)A_1 + 2x^2+6 + (x+3)^2(Bx + C)$$

Now expand the whole thing:

$$x^2+15 = (x^3+3x+3x^2+9)A_1 + 2x^2+6 + (x^3+6x^2+9x)B + (x^2+6x+9)C$$

Gather powers of x together:

$$x^2+15 = x^3(A_1+B)+x^2(3A_1+6B+C+2)+x(3A_1+9B+6C)+(9A_1+6+9C)$$

Separate the powers and write as a Systems of Linear Equations:

$$x^3:\quad 0 = A_1+B$$

$$x^2:\quad 1 = 3A_1+6B+C+2$$

$$x:\quad 0 = 3A_1+9B+6C$$

$$\text{Constants:}\quad 15 = 9A_1+6+9C$$

Simplify, and arrange neatly:

$$\begin{aligned}
0 &= A_1 + B \\
-1 &= 3A_1 + 6B + C \\
0 &= 3A_1 + 9B + 6C \\
1 &= A_1 + C
\end{aligned}$$

Now solve.

You can choose your own way to solve this ... I decided to subtract the 4th equation from the 2nd to begin with:

$$\begin{aligned}
0 &= A_1 + B \\
\mathbf{-2} &\mathbf{=} \mathbf{2A_1 + 6B} \\
0 &= 3A_1 + 9B + 6C \\
1 &= A_1 + C
\end{aligned}$$

Then subtract 2 times the 1st equation from the 2nd:

$$\begin{aligned}
0 &= A_1 + B \\
\mathbf{-2} &\mathbf{=} \mathbf{4B} \\
0 &= 3A_1 + 9B + 6C \\
1 &= A_1 + C
\end{aligned}$$

Now I know that **B = −(1/2)**.

We are getting somewhere!

And from the 1st equation I can figure that $\mathbf{A_1 = +(1/2)}$.

And from the 4th equation I can figure that **C = +(1/2)**.

Final Result:

$$A_1=1/2 \qquad A_2=2 \qquad B=-(1/2) \qquad C=1/2$$

And we can now write our partial fractions:

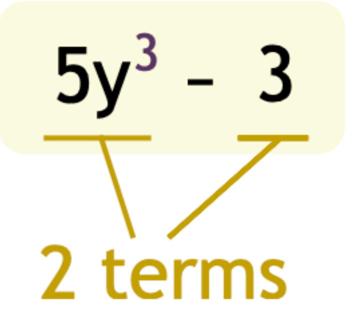
$$\frac{x^2+15}{(x+3)^2(x^2+3)} = \frac{1}{2(x+3)} + \frac{2}{(x+3)^2} + \frac{-x+1}{2(x^2+3)}$$

Summary

- Start with a **Proper** Rational Expressions (if not, do division first)
- Factor the bottom into:
  - linear factors
  - or "irreducible" quadratic factors
- Write out a partial fraction for each factor (and every exponent of each)
- Multiply the whole equation by the bottom
- Solve for the coefficients by
  - substituting zeros of the bottom
  - making a system of linear equations (of each power) and solving
- Write out your answer!

Binomial Theorem

A **binomial** is a polynomial with two terms



example of a binomial

What happens when we multiply a binomial by itself ... many times?

Example: **a+b**

a+b is a binomial (the two terms are **a** and **b**)

Let us multiply **a+b** by itself using Polynomial Multiplication :

$$(a+b)(a+b) = \mathbf{a^2 + 2ab + b^2}$$

Now take that result and multiply by **a+b** again:

$$(a^2 + 2ab + b^2)(a+b) = \mathbf{a^3 + 3a^2b + 3ab^2 + b^3}$$

And again:

$$(a^3 + 3a^2b + 3ab^2 + b^3)(a+b) = \mathbf{a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4}$$

The calculations get longer and longer as we go, but there is some kind of **pattern** developing.

That pattern is summed up by the **Binomial Theorem**:

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

The Binomial Theorem

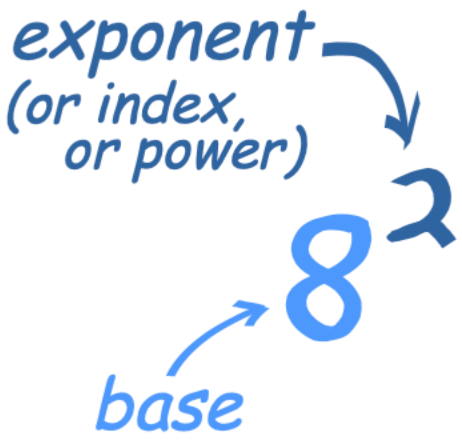
Don't worry ... it will all be explained!

And you will learn lots of cool math symbols along the way.

Exponents

First, a quick summary of Exponents.

An exponent says **how many times** to use something in a multiplication.



Example: $\mathbf{8^2 = 8 \times 8 = 64}$

An exponent of **1** means just to have it appear once, so we get the original value:

Example: $\mathbf{8^1 = 8}$

An exponent of **0** means not to use it at all, and we have only 1:

Example: $\mathbf{8^0 = 1}$

Exponents of (a+b)

Now on to the binomial.

We will use the simple binomial a+b, but it could be any binomial.

Let us start with an exponent of **0** and build upwards.

Exponent of 0

When an exponent is 0, we get **1**:

$$(a+b)^0 = 1$$

Exponent of 1

When the exponent is 1, we get the original value, unchanged:

$$(a+b)^1 = a+b$$

Exponent of 2

An exponent of 2 means to multiply by itself (see

$$(a+b)^2 = (a+b)(a+b) = a^2 + 2ab + b^2$$

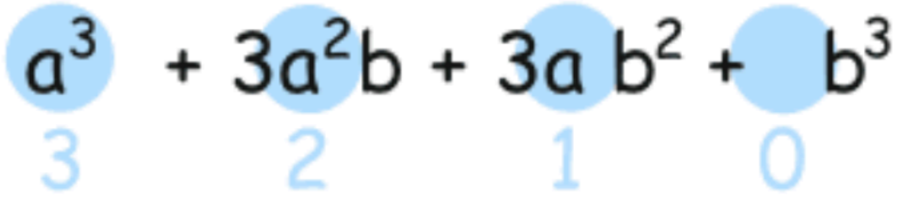
Exponent of 3

For an exponent of 3 just multiply again:

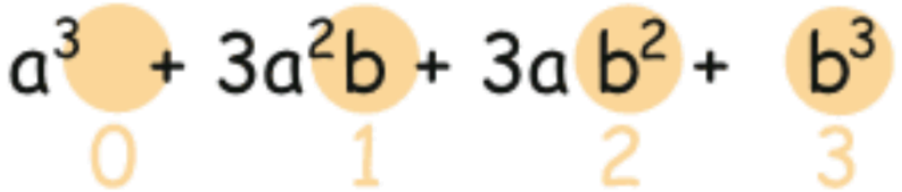
$$(a+b)^3 = (a^2 + 2ab + b^2)(a+b) = a^3 + 3a^2b + 3ab^2 + b^3$$

The Pattern

Now, notice the exponents of a. They start at 3 and go down: 3, 2, 1, 0:



Likewise the exponents of b go upwards: 0, 1, 2, 3:



If we number the terms 0 to *n*, we get this:

k=0	k=1	k=2	k=3
a^3	a^2	a	1
1	b	b^2	b^3

Which can be brought together into this:

$$a^{n-k}b^k$$

Example: When the exponent, n, is 3.

The terms are:

k=0:	**k=1:**	**k=2:**	**k=3:**
$a^{n-k}b^k$ $= a^{3-0}b^0$ $= \mathbf{a^3}$	$a^{n-k}b^k$ $= a^{3-1}b^1$ $= \mathbf{a^2b}$	$a^{n-k}b^k$ $= a^{3-2}b^2$ $= \mathbf{ab^2}$	$a^{n-k}b^k$ $= a^{3-3}b^3$ $= \mathbf{b^3}$

It works like magic!

Coefficients

So far we have: $a^3 + a^2b + ab^2 + b^3$

But we **really** need: $a^3 + 3a^2b + 3ab^2 + b^3$

We are **missing the numbers** (which are called *coefficients*).

Let's look at **all the results** we got before, from $(a+b)^0$ up to $(a+b)^3$:

$$1$$

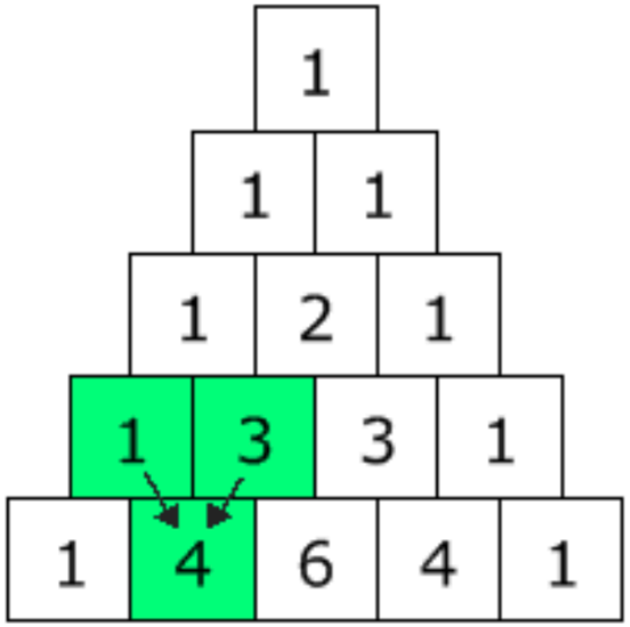
$$a + b$$

$$a^2 + 2ab + b^2$$

$$a^3 + 3a^2b + 3ab^2 + b^3$$

They actually make Pascal's Triangle!

Each number is just the two numbers above it added together (except for the edges, which are all "1")



(Here I have highlighted that **1+3 = 4**)

Armed with this information let us try something new ... an **exponent of 4**:

a exponents go 4,3,2,1,0: $a^4 + a^3 + a^2 + a + 1$

b exponents go 0,1,2,3,4: $a^4 + a^3b + a^2b^2 + ab^3 + b^4$

coefficients go 1,4,6,4,1: $a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$ ✓

And that is the correct answer (compare to the top of the page).

We have success!

We can now use that pattern for exponents of 5, 6, 7, ... 50, ... 112, ... you name it!

That pattern is the essence of the Binomial Theorem.

Now you can take a break.

When you come back see if you can work out **$(a+b)^5$** yourself.

Answer (hover over): $a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5$

As a Formula

Our next task is to write it all as a formula.

We already have the exponents figured out:

$$a^{n-k}b^{k}$$

But how do we write a formula for **"find the coefficient from Pascal's Triangle"** ... ?

Well, there **is** such a formula:

It is commonly called "n choose k" because it is how many ways to choose k elements from a set of n.

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

The "!" means "factorial", for example 4! = 4×3×2×1 = 24

And it matches to Pascal's Triangle like this:

(Note how the top row is row zero
and also the leftmost column is zero!)

$$\begin{array}{ccccccccc}
 & & & & \binom{0}{0} & & & & \\
 & & & \binom{1}{0} & & \binom{1}{1} & & & \\
 & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & \\
 & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & \\
\binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4}
\end{array}$$

Example: Row 4, term 2 in Pascal's Triangle is "6".

Let's see if the formula works:

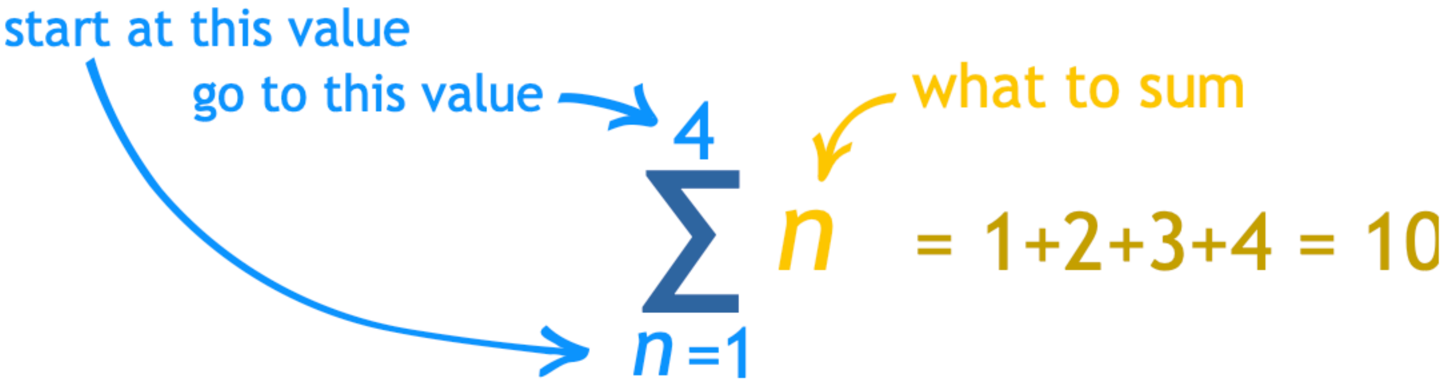
$$\binom{4}{2} = \frac{4!}{2!(4-2)!} = \frac{4!}{2!2!} = \frac{4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1 \cdot 2 \cdot 1} = 6$$

Putting It All Together

The last step is to put all the terms together into **one formula**.

But we are adding lots of terms together ... can that be done using one formula?

Yes! The handy Sigma Notation allows us to sum up as many terms as we want:



Now it can all go into one formula:

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

The Binomial Theorem

So let's try using it for $n = 3$:

$$\begin{aligned}(a+b)^3 &= \sum_{k=0}^{3} \binom{3}{k} a^{3-k} b^k \\ &= \binom{3}{0} a^{3-0} b^0 + \binom{3}{1} a^{3-1} b^1 + \binom{3}{2} a^{3-2} b^2 + \binom{3}{3} a^{3-3} b^3 \\ &= 1 \cdot a^3 b^0 + 3 \cdot a^2 b^1 + 3 \cdot a^1 b^2 + 1 \cdot a^0 b^3 \\ &= a^3 + 3a^2 b + 3ab^2 + b^3\end{aligned}$$

BUT ... it is usually **much easier** just to remember the **patterns**:

- The first term's exponents start at **n and go down**
- The second term's exponents start at **0 and go up**
- Coefficients are from Pascal's Triangle, or by calculation using $\frac{n!}{k!(n-k)!}$

Like this:

Example: What is $(y+5)^4$

Start with exponents:	$y^4 5^0$	$y^3 5^1$	$y^2 5^2$	$y^1 5^3$	$y^0 5^4$
Include Coefficients:	**1**$y^4 5^0$	**4**$y^3 5^1$	**6**$y^2 5^2$	**4**$y^1 5^3$	**1**$y^0 5^4$

Then write down the answer (including all calculations, such as 4×5, 6×5^2, etc):

$$(y+5)^4 = y^4 + 20y^3 + 150y^2 + 500y + 625$$

We may also want to calculate just one term:

Example: What is the coefficient for x^3 in $(2x+4)^8$

The **exponents** for x^3 are **8-5** (=3) for the "2x" and **5** for the "4":

$$(2x)^3 4^5$$

(Why? Because:

2x:	8	7	6	5	4	3	2	1	0
4:	0	1	2	3	4	5	6	7	8
	$(2x)^8 4^0$	$(2x)^7 4^1$	$(2x)^6 4^2$	$(2x)^5 4^3$	$(2x)^4 4^4$	$(2x)^3 4^5$	$(2x)^2 4^6$	$(2x)^1 4^7$	$(2x)^0 4^8$

But we don't need to calculate all the other values if we only want one term.)

And let's not forget "8 choose 5" ... we can use Pascal's Triangle, or calculate directly:

$$\frac{n!}{k!(n-k)!} = \frac{8!}{5!(8-5)!} = \frac{8!}{5!3!} = \frac{8\times7\times6}{3\times2\times1} = 56$$

And we get:

$$56(2x)^3 4^5$$

Which simplifies to:

$$\mathbf{458752}\ x^3$$

And one last, most amazing, example:

Example: A formula for **e** (Euler's Number)

We can use the Binomial Theorem to calculate e (Euler's number).

e = **2.718281828459045...** (the digits go on forever without repeating)

It can be calculated using:

$$(1 + 1/n)^n$$

(It gets more accurate the higher the value of **n**)

That formula is a **binomial**, right? So let's use the Binomial Theorem:

$$(1+\tfrac{1}{n})^n = \sum_{k=0}^{n} \binom{n}{k} 1^{n-k} (\tfrac{1}{n})^k$$

First, we can drop 1^{n-k} as it is always equal to 1:

$$= \sum_{k=0}^{n} \binom{n}{k} (\tfrac{1}{n})^k$$

And, quite magically, most of what is left goes to **1** as n goes to infinity:

$$\sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!} \cdot \frac{1}{n^k}$$

$$\underbrace{n\cdot(n-1)\cdot(n-2)\cdot\ldots\cdot(n-k+1)}_{k\text{ terms}} \cdot \frac{1}{\underbrace{n\cdot n\cdot n\cdot\ldots\cdot n}_{k\text{ terms}}}$$

$$\frac{n}{n}\cdot\frac{n-1}{n}\cdot\frac{n-2}{n}\cdots\frac{n-k+1}{n}$$

as $n \to \infty$ $\quad 1 \cdot 1 \cdot 1 \cdot \ldots \cdot 1$

Which just leaves:

$$\sum_{k=0}^{\infty} \frac{1}{k!} = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \ldots$$

$$= 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \ldots$$

With just those first few terms we get e ≈ 2.7083...

Alternative - much easier

Derivation

$$d(uv) = udv + vdu$$

$$udv = d(uv) - vdu$$

$$\int udv = \int d(uv) - \int vdu$$

$$\int udv = uv - \int vdu$$

Ex 1

$$I = \int x\cos(x)dx$$

$$u = x \qquad dv = \cos(x)dx \rightarrow v = \sin(x)$$

$$I = x\sin(x) - \int \sin(x)dx = x\sin(x) + \cos(x)$$

Ex 2

$$\int u dv = \qquad I = \int \frac{\ln(x)}{x^2} dx$$

$$u = \ln(x) \rightarrow du = \frac{dx}{x^2} \qquad\qquad dv = \frac{1}{x^2} \rightarrow v = -\frac{1}{x}$$

$$\int u dv = uv - \int v du$$

$$\int u dv = \qquad -\frac{1}{x}\ln(x) + \int \frac{dx}{x^2} = -\frac{1}{x}\ln(x) - \frac{1}{x}$$

Derivation of other formula

$$\frac{d}{dx}(fg) = f\frac{dg}{dx} + g\frac{df}{dx}$$

$$\int \frac{d}{dx} fg dx = fg = \int f\frac{dg}{dx}dx + \int g\frac{df}{dx}dx$$

$$\longrightarrow \qquad \int f\frac{dg}{dx}dx = fg - \int g\frac{df}{dx}dx \qquad (1)$$

Let $f(x) = u(x)$ and $\frac{dg}{dx} = v(x)$

$$\longrightarrow \qquad \frac{df}{dx} = u'(x) \quad \text{and} \quad g = \int v dx$$

then (1) becomes

$$\int uvdx = u\int vdx - \int [\int vdx]u'dx$$

or

$$\int uvdx = u\int vdx - \int u'[\int vdx]dx$$

$$\int u\ v\ dx = u \int v\ dx - \int u' \ (\int v\ dx)\ dx$$

—-> other formula

Integration by Parts

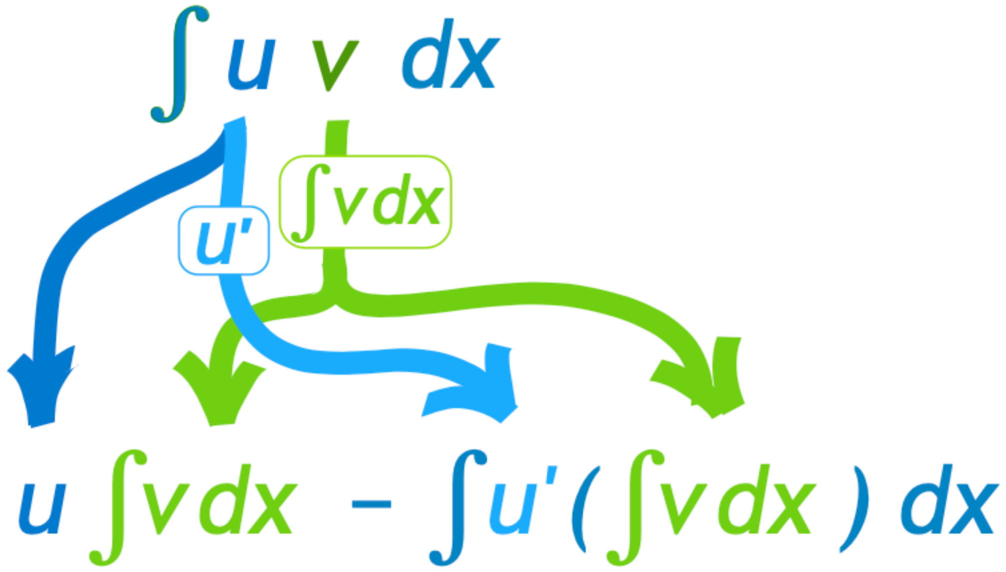
You will see plenty of examples soon, but first let us see the rule:

works but more complicated

$$\int u\, v\, dx = u \int v\, dx - \int u' \left(\int v\, dx \right) dx$$

- **u** is the function u(x)
- **v** is the function v(x)
- **u'** is the derivative of the function u(x)

The rule as a diagram:



Let's get straight into an example:

Example: What is $\int x \cos(x)\, dx$?

OK, we have **x** *multiplied by* **cos(x)**, so integration by parts is a good choice.

First choose which functions for **u** and **v**:

- $u = x$

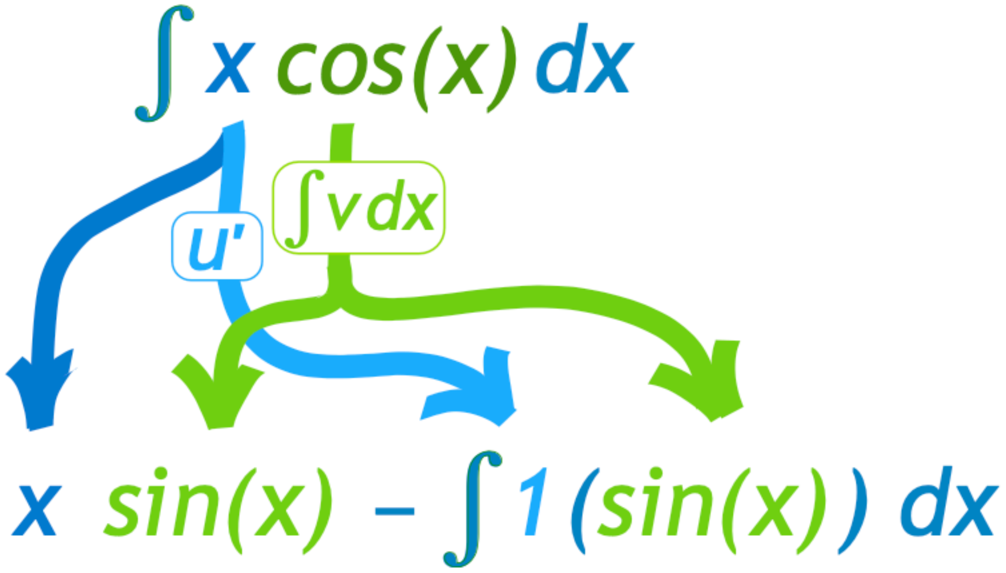
- v = cos(x)

So now it is in the format $\int$**u v dx** we can proceed:

Differentiate **u**: $u' = x' = 1$

Integrate **v**: $\int v\,dx = \int \cos(x)\,dx = \sin(x)$

Now we can put it together:



$$\int x \cos(x)\,dx$$

$$x \sin(x) - \int 1\,(\sin(x))\,dx$$

Simplify and solve:

$$x \sin(x) - \int \sin(x)\,dx$$

$$x \sin(x) + \cos(x) + C$$

Done!

So we followed these steps:

- Choose u and v
- Differentiate u: u'
- Integrate v: $\int v\,dx$
- Put u, u' and $\int v\,dx$ into: $\mathbf{u \int v\,dx - \int u' \left(\int v\,dx \right) dx}$
- Simplify and solve

Let's try some more examples:

Example: What is $\int \ln(x)/x^2\,dx$?

First choose u and v:

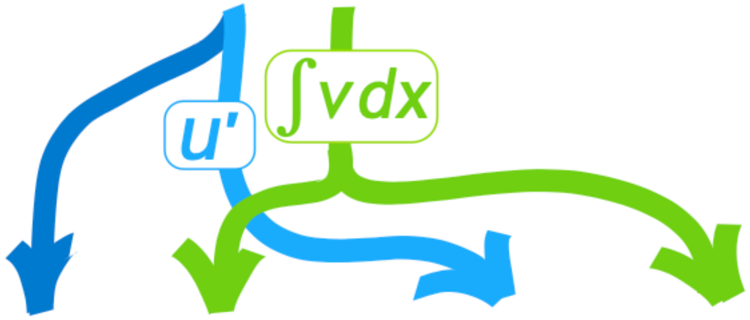
- $u = \ln(x)$
- $v = 1/x^2$

Differentiate u: $\ln(x)' = \frac{1}{x}$

Integrate v: $\int 1/x^2\,dx = \int x^{-2}\,dx = -x^{-1} = \frac{-1}{x}$

Now put it together:

$$\int \ln x \, \frac{1}{x^2} \, dx$$



$$\ln x \, \frac{-1}{x} - \int \frac{1}{x} \left(\frac{-1}{x}\right) dx$$

Simplify:

$$-\ln(x)/x - \int -1/x^2 \, dx$$

$$-\ln(x)/x - 1/x + C$$

$$-\frac{\ln(x) + 1}{x} + C$$

Example: What is $\int \ln(x) \, dx$?

But there is only one function! How do we choose u and v ?

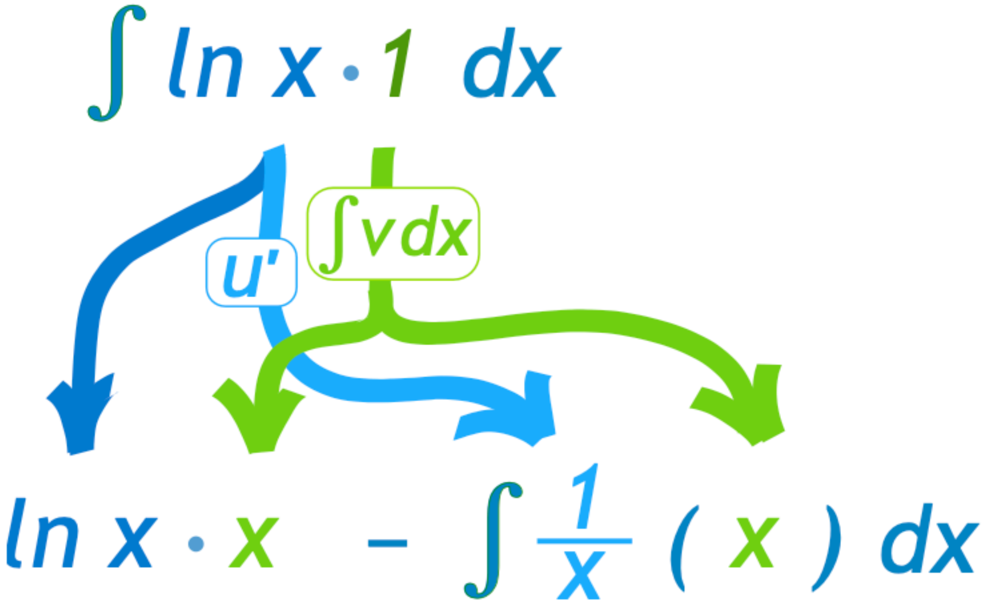
Hey! We can just choose v as being "1":

- $u = \ln(x)$
- $v = 1$

Differentiate u: $\ln(x)' = 1/x$

Integrate v: $\int 1\,dx = x$

Now put it together:



$$\int \ln x \cdot 1\ dx$$

$$\ln x \cdot x \ - \int \frac{1}{x}\,(\,x\,)\,dx$$

Simplify:

$$x\ \ln(x) - \int 1\ dx$$

$$x\ \ln(x) - x + C$$

Example: What is $\int e^x\, x\ dx$?

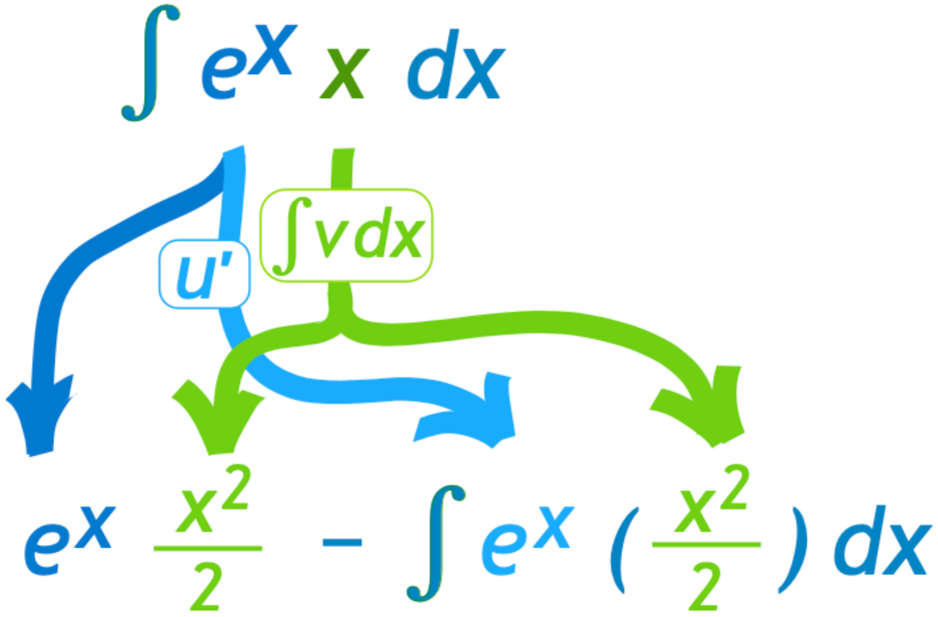
Choose u and v:

- $u = e^x$
- $v = x$

Differentiate u: $(e^x)' = e^x$

Integrate v: $\int x\,dx = x^2/2$

Now put it together:



$$\int e^x\, x\, dx$$

$$e^x \frac{x^2}{2} - \int e^x \left(\frac{x^2}{2}\right) dx$$

It only got worse!

Well, that was a spectacular disaster.

Maybe we could choose a different u and v?

Example: $\int e^x\, x\, dx$ (continued)

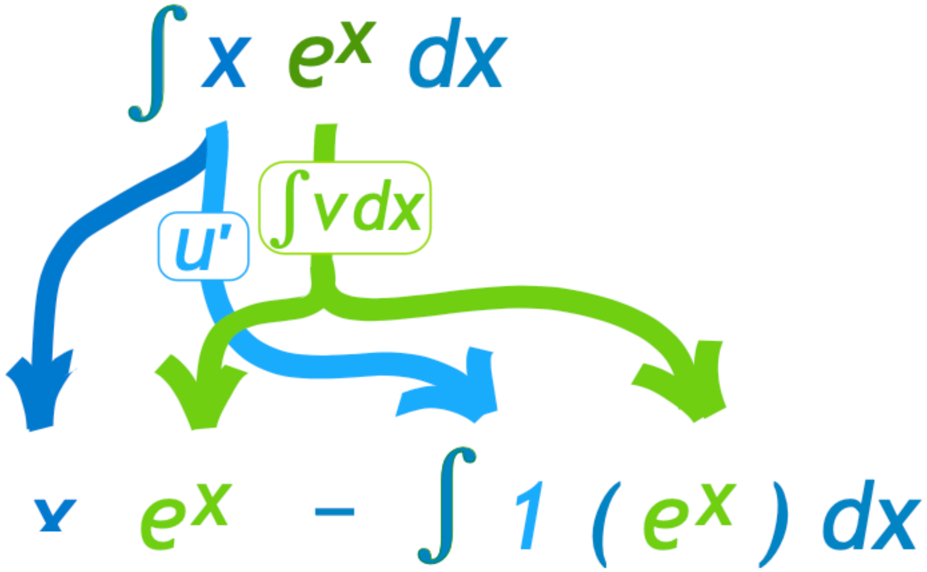
Choose u and v differently:

- $u = x$
- $v = e^x$

Differentiate u: $(x)' = 1$

Integrate v: $\int e^x\,dx = e^x$

Now put it together:



Simplify:

$$x\,e^x - e^x + C$$

$$e^x(x-1) + C$$

The moral of the story: Choose **u** and **v** carefully!

A helpful rule of thumb

Choose **u** based on which of these comes first:

- **I**: Inverse trigonometric functions such as $\sin^{-1}(x)$, $\cos^{-1}(x)$, $\tan^{-1}(x)$
- **L**: Logarithmic functions such as $\ln(x)$, $\log(x)$
- **A**: Algebraic functions such as x^2, x^3
- **T**: Trigonometric functions such as $\sin(x)$, $\cos(x)$, $\tan(x)$
- **E**: Exponential functions such as e^x, 3^x

And here is one last (and tricky) example:

Example: $\int e^x \sin(x)\, dx$

Choose u and v:

- $u = \sin(x)$
- $v = e^x$

Differentiate u: $\sin(x)' = \cos(x)$

Integrate v: $\int e^x\, dx = e^x$

Now put it together:

$$\int e^x \sin(x)\, dx = \sin(x)\, e^x - \int \cos(x)\, e^x\, dx$$

It looks worse, but let us persist! To find $\int \cos(x)\, e^x\, dx$ we can use integration by parts **again**:

Choose u and v:

- $u = \cos(x)$
- $v = e^x$

Differentiate u: $\cos(x)' = -\sin(x)$

Integrate v: $\int e^x\, dx = e^x$

Now put it together:

$$\int e^x \sin(x)\, dx = \sin(x)\, e^x - (\cos(x)\, e^x - \int -\sin(x)\, e^x\, dx)$$

Simplify:

$$\int e^x \sin(x)\, dx = e^x \sin(x) - e^x \cos(x) - \int e^x \sin(x) dx$$

... so we can bring the right hand integral over to the left and we get:

$$2\int e^x \sin(x)\,dx = e^x \sin(x) - e^x \cos(x)$$

Simplify:

$$\int e^x \sin(x)\,dx = \frac{1}{2} e^x (\sin(x) - \cos(x)) + C$$

Alternative - much easier

Derivation

$$d(uv) = udv + vdu$$

$$udv = d(uv) - vdu$$

$$\int udv = \int d(uv) - \int vdu$$

$$\int udv = uv - \int vdu$$

$$I = \int x\cos(x)dx$$

$$u = x \qquad dv = \cos(x)dx \rightarrow v = \sin(x)$$

$$I = x\sin(x) - \int \sin(x)dx = x\sin(x) + \cos(x)$$

Ex 2

$$\int udv = \quad I = \int \frac{\ln(x)}{x^2}dx$$

$$u = \ln(x) \rightarrow du = \frac{dx}{x} \qquad dv = \frac{1}{x^2} \rightarrow v = -\frac{1}{x}$$

$$\int udv = uv - \int vdu$$

$$\int udv = \quad -\frac{1}{x}\ln(x) + \int \frac{dx}{x^2} = -\frac{1}{x}\ln(x) - \frac{1}{x}$$

Derivation of other formula

$$\frac{d}{dx}(fg) = f\frac{dg}{dx} + g\frac{df}{dx}$$

$$\int \frac{d}{dx} fg dx = fg = \int f\frac{dg}{dx}dx + \int g\frac{df}{dx}dx$$

→

$$\int f\frac{dg}{dx}dx = fg - \int g\frac{df}{dx}dx \tag{1}$$

Let $f(x) = u(x)$ and $\dfrac{dg}{dx} = v(x)$

→

$\dfrac{df}{dx} = u'(x)$ and $g = \displaystyle\int v dx$

then (1) becomes

$$\int uvdx = u\int vdx - \int [\int vdx]u'dx$$

or

$$\int uvdx = u\int vdx - \int u'[\int vdx]dx$$

$$\int u\ v\ dx = u \int v\ dx - \int u' \ (\int v\ dx)\ dx$$ —-> **other formula**

Definite Integrals

When the integral has an interval like [a, b] we can use either of these:

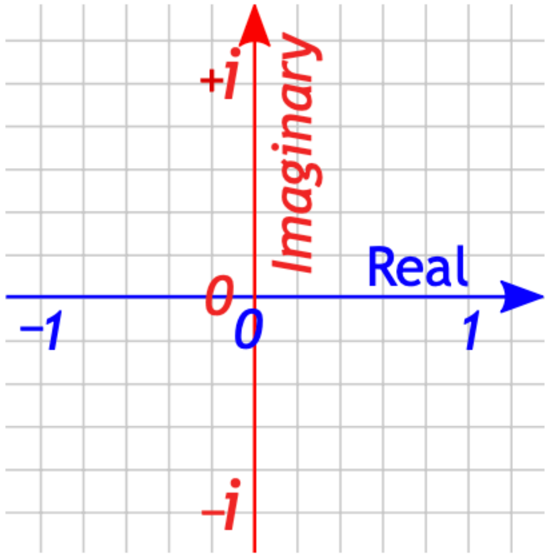
$$\int_a^b u\ v\ dx = \left[u \int v\ dx - \int u'(\int v\ dx)\ dx \right]_a^b$$

$$\int_a^b u\ v\ dx = \left[u \int v\ dx \right]_a^b - \int_a^b u'(\int v\ dx)\ dx$$

Where u and v are functions of x, and a and b are the limits on x.

The second version can help us see the relationship between the left and right integrals.

Complex Plane



(Also called an "Argand Diagram")

A **plane** for **complex** numbers!

Real and Imaginary make Complex

A Complex Number is a combination of a Real Number and an Imaginary Number:

A Real Number is the type of number we use every day.

Examples: 12.38, ½, 0, −2000

When we square a Real Number we get a positive (or zero) result:

$$2^2 = 2 \times 2 = 4$$

$$1^2 = 1 \times 1 = 1$$

$$0^2 = 0 \times 0 = 0$$

What can we square to get −1?

$$?^2 = -1$$

Squaring −1 does not work because multiplying negatives gives a positive: (−1) × (−1) = +1, and no other Real Number works either.

So it seems that mathematics is incomplete ...

... but we can fill the gap by ***imagining*** there is a number that, when multiplied by itself, gives

-1

(call it **i** for imaginary):

$$i^2 = -1$$

An Imaginary Number, when squared gives a negative result

$$\text{imaginary}^2 \rightarrow \text{negative}.$$

Examples: 5**i**, -3.6**i**, **i**/2, 500**i**

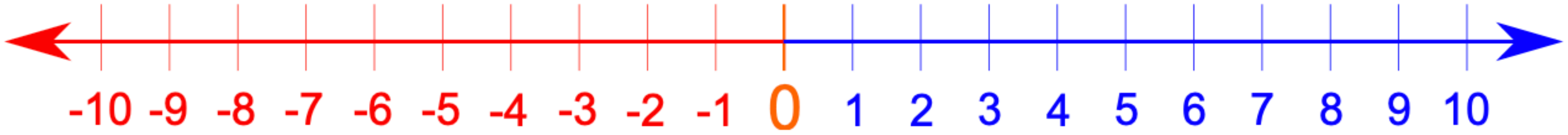
And together:

A Complex Number is a combination of a Real Number and an Imaginary Number

Examples: 3.6 + 4**i**, −0.02 + 1.2**i**, 25 − 0.3**i**, 0 + 2**i**

Putting a Complex Number on a Plane

You may be familiar with the number line:

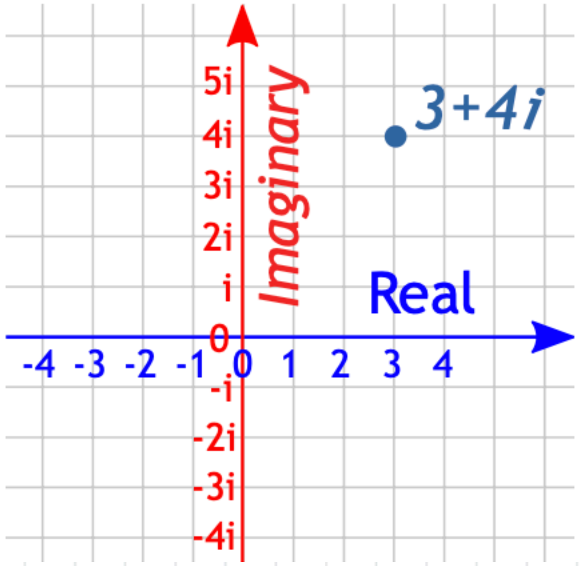


But where do we put a complex number like 3+4**i** ?

Let's have the real number line go left-right as usual, and have the **imaginary number line go up-and-down**:

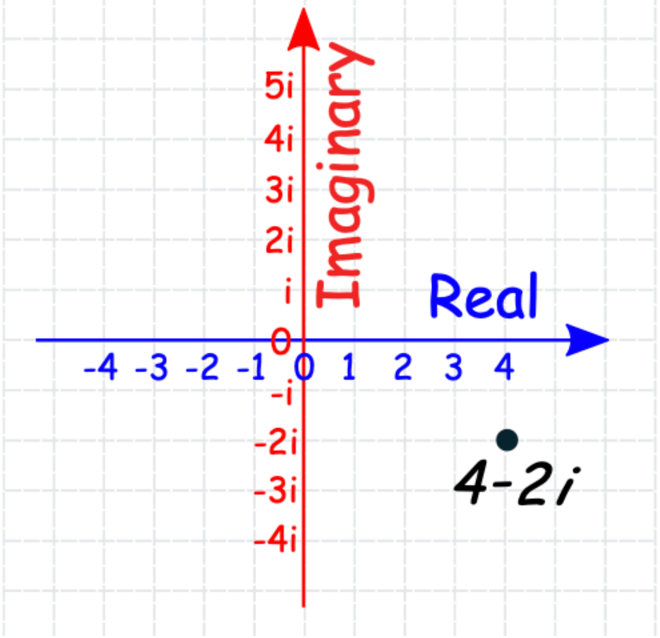
We can then plot a complex number like **3 + 4i** :

- 3 units along (the real axis),
- and 4 units up (the imaginary axis).



And here is **4 - 2i** :

- 4 units along (the real axis),
- and 2 units down (the imaginary axis).



And that is the **complex plane**:

- **complex** because it is a combination of real and imaginary,
- **plane** because it is like a geometric plane (2 dimensional).

Whole New World

Now let's bring the **idea of a plane** (Cartesian coordinates, Polar coordinates, Vectors etc) to complex numbers.

It will open up a whole new world of numbers that are more complete and elegant, as you will see.

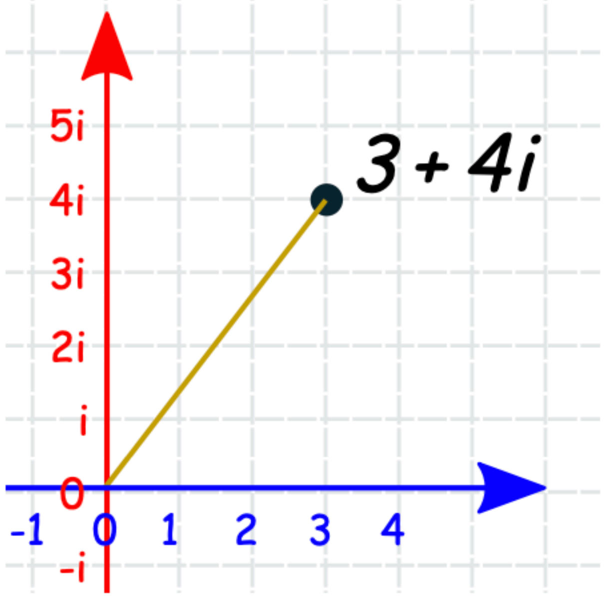
Complex Number as a Vector

We can think of a complex number as a vector.

This is a vector.
It has magnitude (length) and direction.

And here is the complex number **3 + 4i**

as a Vector:



Adding

You can add complex numbers as vectors, too:

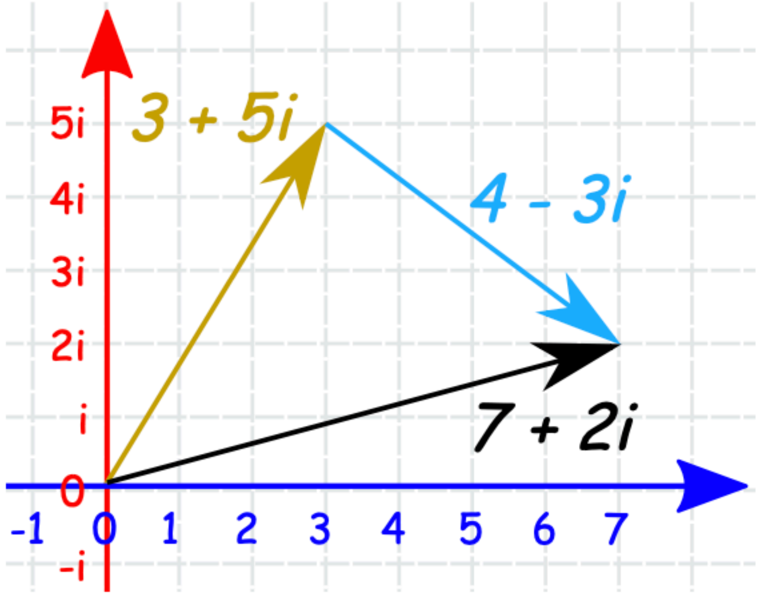
To add the complex numbers **3 + 5i** and **4 − 3i** :

- add the real numbers, and
- add the imaginary numbers

separately, like this:

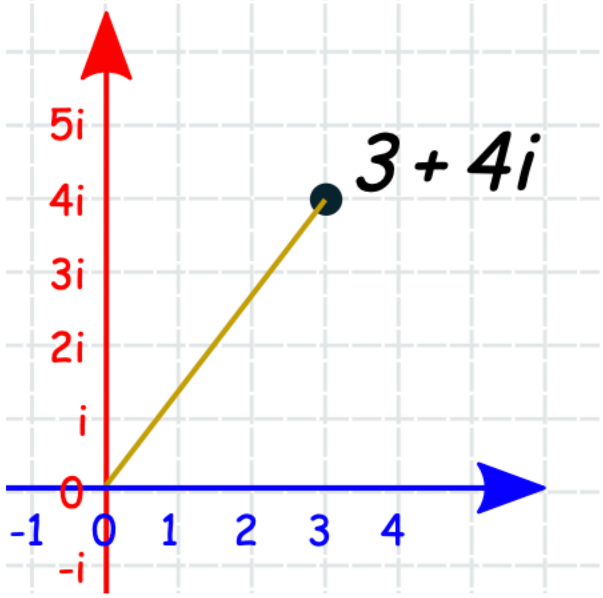
$$(3 + 5\mathbf{i}) + (4 - 3\mathbf{i}) = (3 + 4) + (5 - 3)\mathbf{i}$$

$$= 7 \qquad + 2\mathbf{i}$$

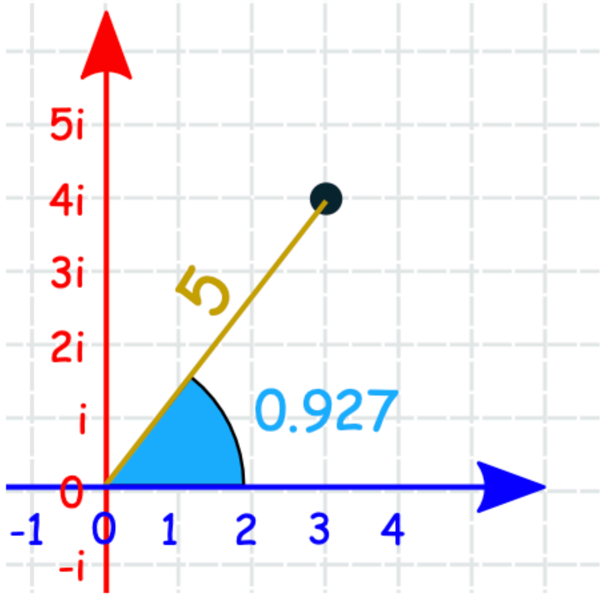


Polar Form

Let's use **3 + 4i** again:



Here it is **in polar form:**



So the complex number **3 + 4i** can also be shown as distance (5) and angle (0.927 radians).

Let's see how to convert from one form to the other using Cartesian to Polar conversion:

Example: the number **3 + 4i**

From **3 + 4i** :

- $\mathbf{r = \sqrt{(x^2 + y^2)}} = \sqrt{(3^2 + 4^2)} = \sqrt{25} = \mathbf{5}$
- $\boldsymbol{\theta = \tan^{-1} (y/x)} = \tan^{-1} (4/3) = \mathbf{0.927}$ (to 3 decimals)

And we get distance (5) and angle (0.927 radians)

Back again:

- $\mathbf{x = r \times \cos(\,\boldsymbol{\theta}\,)} = 5 \times \cos(\,0.927\,) = 5 \times 0.6002... = \mathbf{3}$ (close enough)
- $\mathbf{y = r \times \sin(\,\boldsymbol{\theta}\,)} = 5 \times \sin(\,0.927\,) = 5 \times 0.7998... = \mathbf{4}$ (close enough)

And distance 5 and angle 0.927 becomes 3 and 4 again

In fact a common way to write a complex number in Polar form is

$$\begin{aligned} x + \mathbf{i}y &= r \cos \theta + \mathbf{i}\, r \sin \theta \\ &= r(\cos \theta + \mathbf{i} \sin \theta) \end{aligned}$$

And "$\cos \theta + \mathbf{i} \sin \theta$" is often shortened to "$\text{cis}\ \theta$", so:

$$\mathbf{x + iy = r\ cis\ \boldsymbol{\theta}}$$

cis is just shorthand for **cos *θ* + i sin *θ***

So we can write:

$$3 + 4i = 5 \text{ cis } 0.927$$

In some subjects, like electronics, "cis" is used a lot!

Summary

- The complex plane is a plane with:
  - real numbers running left-right and
  - imaginary numbers running up-down.
- To convert from Cartesian to Polar Form:
  - $r = \sqrt{(x^2 + y^2)}$
  - $\theta = \tan^{-1} (y / x)$
- To convert from Polar to Cartesian Form:
  - $x = r \times \cos(\theta)$
  - $y = r \times \sin(\theta)$
- Polar form $r \cos \theta + i\, r \sin \theta$ is often shortened to $r \text{ cis } \theta$

Hyperbolic Functions

The two basic hyperbolic functions are "sinh" and "cosh":

Hyperbolic Sine:

$$\sinh(x) = \frac{e^x - e^{-x}}{2}$$

(pronounced "shine")

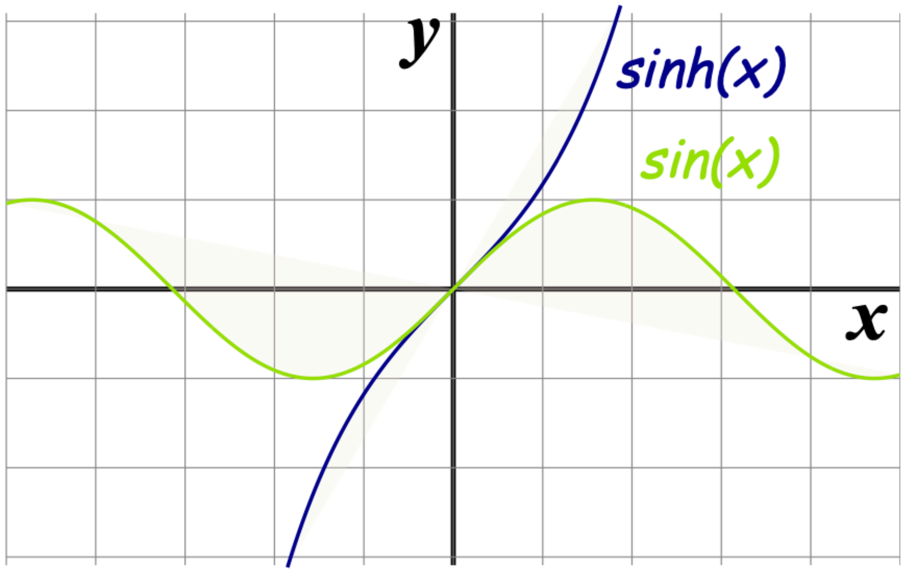
Hyperbolic Cosine:

$$\cosh(x) = \frac{e^x + e^{-x}}{2}$$

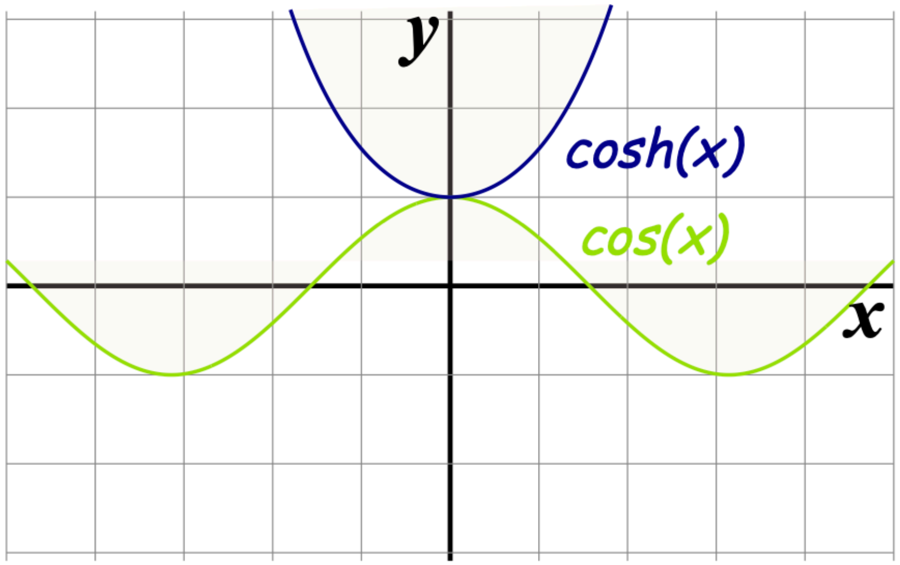
(pronounced "cosh")

They use the natural exponential function **e^x**

And are not the same as sin(x) and cos(x), but a little bit similar:



sinh vs sin



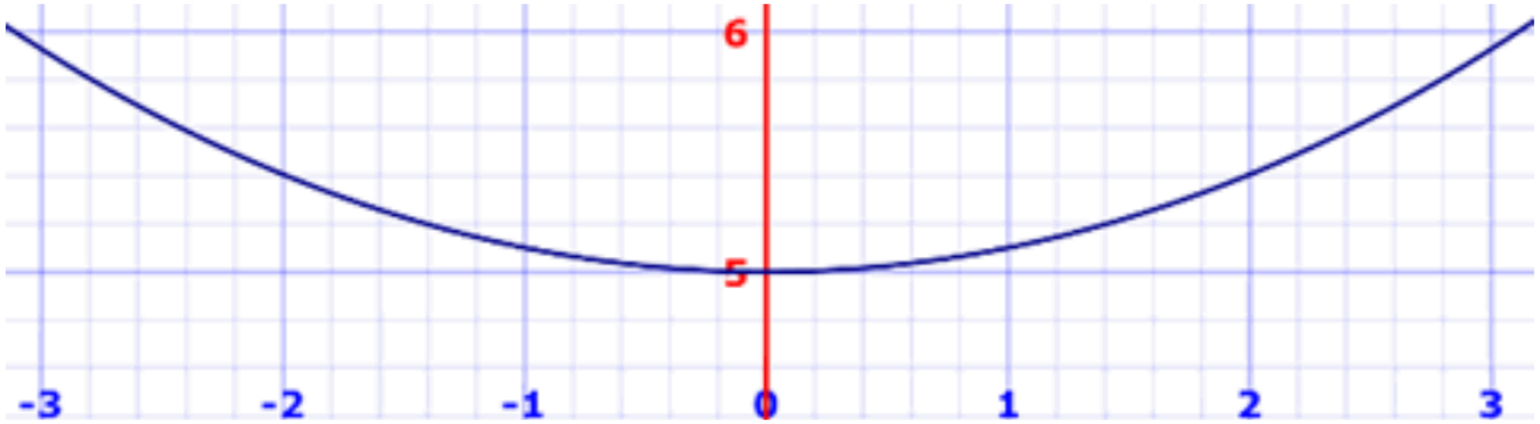
cosh vs cos

Catenary

One of the interesting uses of Hyperbolic Functions is the curve made by suspended cables or chains.

A hanging cable forms a curve called a **catenary** defined using the **cosh** function:

$$f(x) = a \cosh(x/a)$$

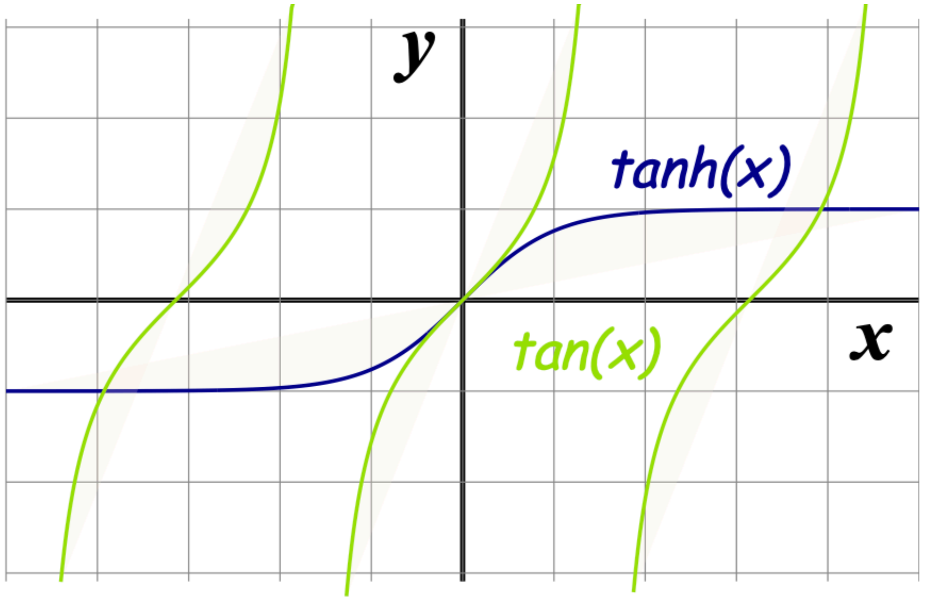


Other Hyperbolic Functions

From **sinh** and **cosh** we can create:

Hyperbolic tangent "tanh" *(pronounced "than")*:

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$



tanh vs tan

Hyperbolic cotangent:

$$\coth(x) = \frac{\cosh(x)}{\sinh(x)} = \frac{e^x + e^{-x}}{e^x - e^{-x}}$$

Hyperbolic secant:

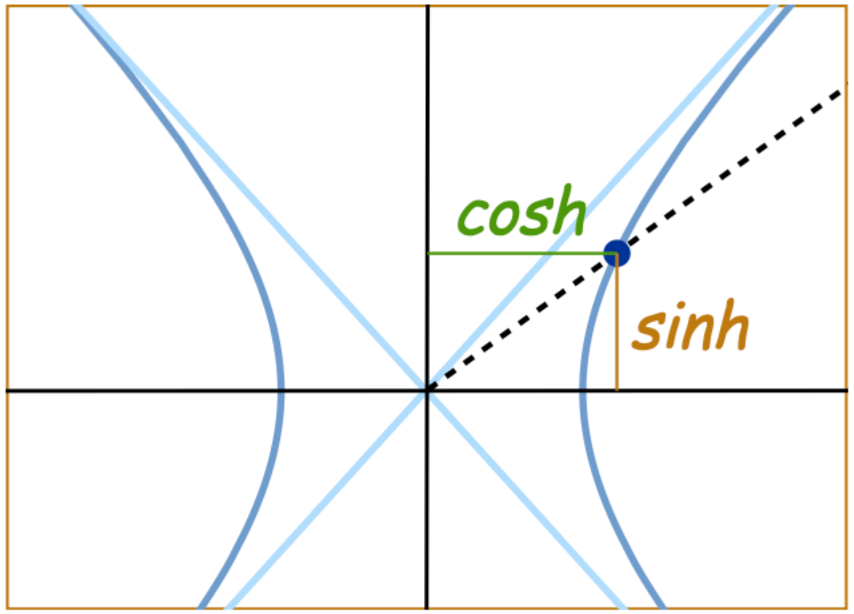
$$\operatorname{sech}(x) = \frac{1}{\cosh(x)} = \frac{2}{e^x + e^{-x}}$$

Hyperbolic cosecant "csch" or "cosech":

$$\operatorname{csch}(x) = \frac{1}{\sinh(x)} = \frac{2}{e^x - e^{-x}}$$

Why the Word "Hyperbolic" ?

Because it comes from measurements made on a Hyperbola:



So, just like the trigonometric functions relate to a circle, the hyperbolic functions relate to a hyperbola.

Identities

- $\sinh(-x) = -\sinh(x)$
- $\cosh(-x) = \cosh(x)$

And

- $\tanh(-x) = -\tanh(x)$
- $\coth(-x) = -\coth(x)$
- $\operatorname{sech}(-x) = \operatorname{sech}(x)$
- $\operatorname{csch}(-x) = -\operatorname{csch}(x)$

Odd and Even

Both **cosh** and **sech** are Even Functions, the rest are Odd Functions.

Derivatives

Derivatives are:

$$\frac{d}{dx}\sinh(x) = \cosh(x)$$

$$\frac{d}{dx}\cosh(x) = \sinh(x)$$

$$\frac{d}{dx}\tanh(x) = 1 - \tanh^2(x)$$

Taylor Series

A Taylor Series is an expansion of some function into an **infinite sum of terms**, where each term has a larger exponent like x, x^2, x^3, etc.

Example: The Taylor Series for e^x

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \ldots$$

says that the function: e^x

is equal to the infinite sum of terms: $1 + x + x^2/2! + x^3/3! + \ldots$ etc

Example: e^x for x=2

Well, we already know the answer is $\mathbf{e^2}$ = 2.71828... × 2.71828... = **7.389056...**

But let's try more and more terms of our infinte series:

Terms	Result
$1+2$	**3**
$1+2+\frac{2^2}{2!}$	**5**
$1+2+\frac{2^2}{2!}+\frac{2^3}{3!}$	**6.3333...**
$1+2+\frac{2^2}{2!}+\frac{2^3}{3!}+\frac{2^4}{4!}$	**7**

$1+2+\frac{2^2}{2!}+\frac{2^3}{3!}+\frac{2^4}{4!}+\frac{2^5}{5!}$	**7.2666...**
$1+2+\frac{2^2}{2!}+\frac{2^3}{3!}+\frac{2^4}{4!}+\frac{2^5}{5!}+\frac{2^6}{6!}$	**7.3555...**
$1+2+\frac{2^2}{2!}+\frac{2^3}{3!}+\frac{2^4}{4!}+\frac{2^5}{5!}+\frac{2^6}{6!}+\frac{2^7}{7!}$	**7.3809...**
$1+2+\frac{2^2}{2!}+\frac{2^3}{3!}+\frac{2^4}{4!}+\frac{2^5}{5!}+\frac{2^6}{6!}+\frac{2^7}{7!}+\frac{2^8}{8!}$	**7.3873...**

It starts out really badly, but it then gets better and better!

Here are some common Taylor Series:

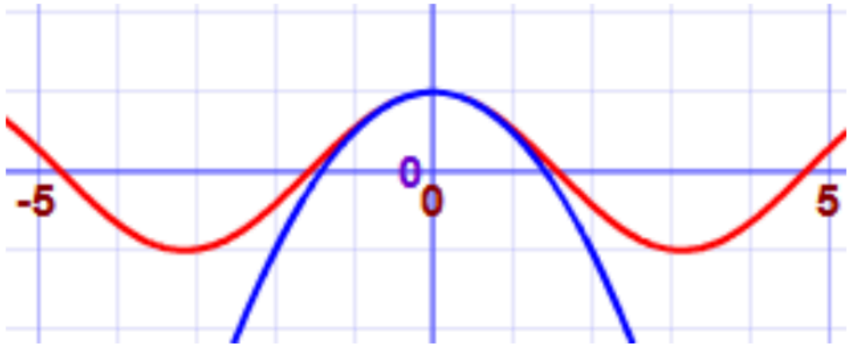
Taylor Series expansion	**As**
$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$	$\sum_{n=0}^{\infty} \frac{x^n}{n!}$
$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots$	$\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$
$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots$	$\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$
$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots \quad \text{for } \lvert x \rvert < 1$	$\sum_{n=0}^{\infty} x^n$

Approximations

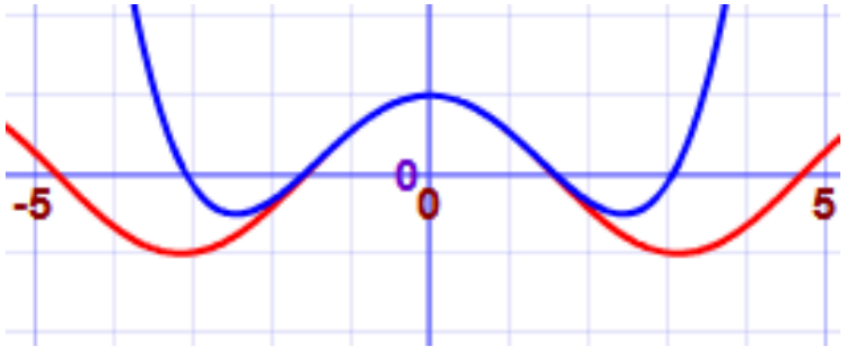
We can use the first few terms of a Taylor Series to get an approximate value for a function.

Here we show better and better approximations for **cos(x)**. The red line is **cos(x)**, the blue is the approximation (try plotting it yourself) :

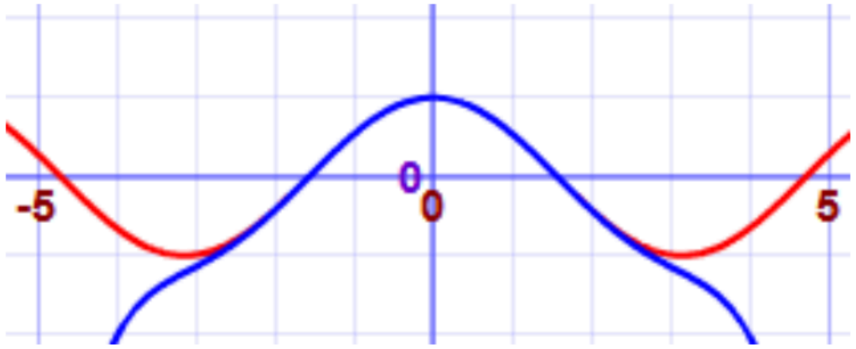
$1 - x^2/2!$



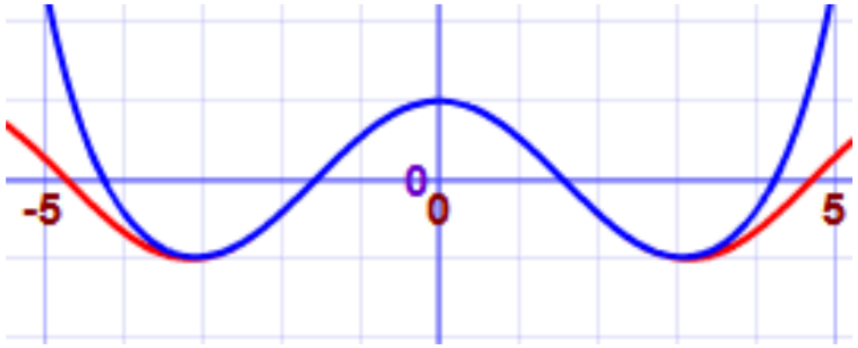
$1 - x^2/2! + x^4/4!$



$1 - x^2/2! + x^4/4! - x^6/6!$



$1 - x^2/2! + x^4/4! - x^6/6! + x^8/8!$



What is this Magic?

How can we turn a function into a series of power terms like this?

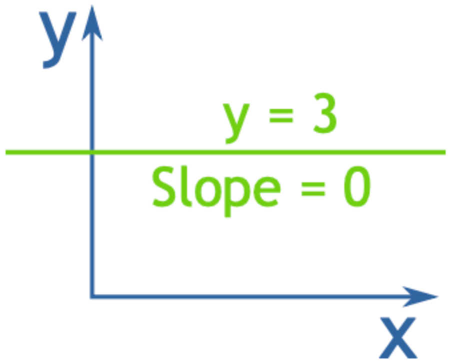
Well, it isn't really magic. First we say we **want** to have this expansion:

$$f(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \dots$$

Then we choose a value "a", and work out the values c_0 , c_1 , c_2 , ... etc

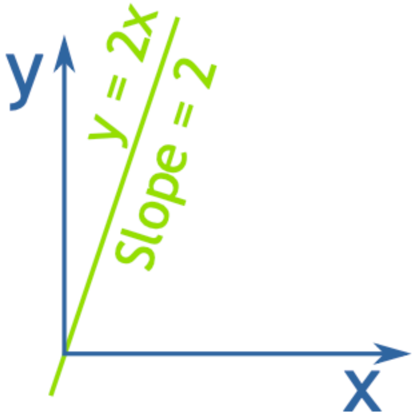
And it is done using **derivatives** (so we must know the derivative of our function)

Quick review: a derivative gives us the slope of a function at any point.



These basic derivative rules can help us:

- The derivative of a constant is **0**
- The derivative of **ax** is **a** (example: the derivative of **2x** is **2**)
- The derivative of **x^n** is **nx^{n-1}** (example: the derivative of **x^3** is **$3x^2$**)



We will use the little mark ′ to mean "derivative of".

To get c_0, choose x=a so all the (x-a) terms become zero, leaving us with:

$$f(a) = c_0$$

$$\text{So } \mathbf{c_0 = f(a)}$$

To get c_1, take the derivative of f(x):

$$f'(x) = c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \ldots$$

With x=a all the (x-a) terms become zero:

$$f'(a) = c_1$$

$$\text{So } \mathbf{c_1 = f'(a)}$$

To get c_2, do the derivative again:

$$f''(x) = 2c_2 + 3 \times 2 \times c_3(x-a) + \ldots$$

With x=a all the (x-a) terms become zero:

$$f''(a) = 2c_2$$

$$\text{So } \mathbf{c_2 = f''(a)/2}$$

In fact, a pattern is emerging. Each term is

- the next higher derivative ...
- ... divided by all the exponents so far multiplied together (for which we can use factorial notation, for example 3! = 3×2×1)

And we get:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x\text{-}a) + \frac{f''(a)}{2!}(x\text{-}a)^2 + \frac{f'''(a)}{3!}(x\text{-}a)^3 + \ldots$$

Now we have a way of finding our own Taylor Series:

For each term: take the next derivative, divide by n!, multiply by $(x\text{-}a)^n$

Example: Taylor Series for cos(x)

Start with:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x\text{-}a) + \frac{f''(a)}{2!}(x\text{-}a)^2 + \frac{f'''(a)}{3!}(x\text{-}a)^3 + \ldots$$

The derivative of **cos** is **−sin**, and the derivative of **sin** is **cos**, so:

- $f(x) = \cos(x)$
- $f'(x) = -\sin(x)$
- $f''(x) = -\cos(x)$
- $f'''(x) = \sin(x)$
- etc...

Note: A **Maclaurin Series** is a Taylor Series where **a=0**, so all the examples we have been using so far can **also** be called Maclaurin Series.

And we get:

$$\cos(x) = \cos(a) - \frac{\sin(a)}{1!}(x\text{-}a) - \frac{\cos(a)}{2!}(x\text{-}a)^2 + \frac{\sin(a)}{3!}(x\text{-}a)^3 + \ldots$$

Now put **a=0**, which is nice because **cos(0)=1** and **sin(0)=0**:

$$\cos(x) = 1 - \frac{0}{1!}(x\text{-}0) - \frac{1}{2!}(x\text{-}0)^2 + \frac{0}{3!}(x\text{-}0)^3 + \frac{1}{4!}(x\text{-}0)^4 + \ldots$$

Simplify:

$$\cos(x) = 1 - x^2/2! + x^4/4! - \ldots$$

Limits *(Evaluating)*

Quick Summary of Limits

Sometimes we can't work something out directly ... but we **can** see what it should be as we get closer and closer!

Example:

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

Let's work it out for x=1:

$$\frac{(1^2 - 1)}{(1 - 1)} = \frac{(1 - 1)}{(1 - 1)} = \frac{0}{0}$$

Now 0/0 is a difficulty! We don't really know the value of 0/0 (it is "indeterminate"), so we need another way of answering this.

So instead of trying to work it out for x=1 let's try **approaching** it closer and closer:

Example Continued:

x	$\frac{(x^2-1)}{(x-1)}$
0.5	1.50000
0.9	1.90000
0.99	1.99000
0.999	1.99900
0.9999	1.99990
0.99999	1.99999
...	...

Now we see that as x gets close to 1, then $\frac{(x^2-1)}{(x-1)}$ gets **close to 2**

We are now faced with an interesting situation:

- When x=1 we don't know the answer (it is **indeterminate**)
- But we can see that it is **going to be 2**

We want to give the answer "2" but can't, so instead mathematicians say exactly what is going on by using the special word "limit"

The **limit** of $\frac{(x^2-1)}{(x-1)}$ as x approaches 1 is **2**

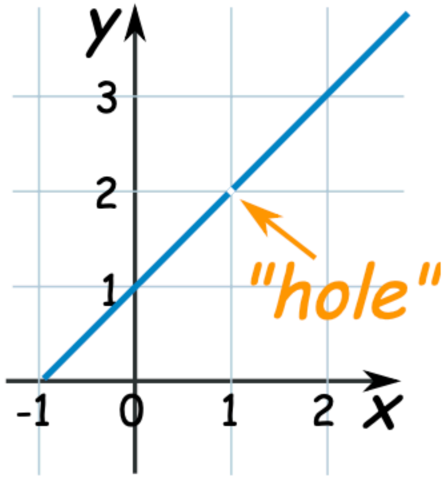
And it is written in symbols as:

$$\lim_{x \to 1} \frac{x^2-1}{x-1} = 2$$

As a graph it looks like this:

So, in truth, we **cannot say what the value at x=1 is.**

But we **can** say that as we approach 1, **the limit is 2.**



Evaluating Limits

"Evaluating" means to find the value of (*think e-"**value**"-ating*)

In the example above we said the limit was 2 because it **looked like it was going to be**. But that is not really good enough!

In fact there are **many ways** to get an accurate answer. Let's look at some:

1. Just Put The Value In

The first thing to try is just putting the value of the limit in, and see if it works (in other words substitution).

Example:

$$\lim_{x\to 10} \frac{x}{2} \Rightarrow \frac{10}{2} = 5$$ ✓

Easy!

Example:

$$\lim_{x\to 1} \frac{x^2-1}{x-1} \Rightarrow \frac{(1-1)}{(1-1)} = \frac{0}{0}$$ ✗

No luck. Need to try something else.

2. Factors

We can try factoring.

Example:

$$\lim_{x \to 1} \frac{x^2 - 1}{x - 1}$$

By factoring (x^2-1) into $(x-1)(x+1)$ we get:

$$\lim_{x \to 1} \frac{x^2 - 1}{x - 1} = \lim_{x \to 1} \frac{(x-1)(x+1)}{(x-1)}$$

$$= \lim_{x \to 1} (x+1)$$

Now we can just substitiute x=1 to get the limit:

$$\lim_{x \to 1} (x+1) = 1+1 = 2$$

3. Conjugate

For some fractions multiplying top and bottom by a conjugate can help.

The conjugate is where we change the sign in the middle of 2 terms like this:

Here is an example where it will help us find a limit:

$\lim\limits_{x\to4} \frac{2-\sqrt{x}}{4-x}$ Evaluating this at x=4 gives 0/0, which is not a good answer!

So, let's try some rearranging:

Multiply top and bottom by the conjugate of the top: $\frac{2-\sqrt{x}}{4-x} \times \frac{2+\sqrt{x}}{2+\sqrt{x}}$

Simplify top using $(a+b)(a-b) = a^2 - b^2$: $\frac{2^2 - (\sqrt{x})^2}{(4-x)(2+\sqrt{x})}$

Simplify top further: $\frac{4-x}{(4-x)(2+\sqrt{x})}$

Cancel (4−x) from top and bottom: $\frac{1}{2+\sqrt{x}}$

So, now we have:

$$\lim_{x\to 4} \frac{2-\sqrt{x}}{4-x} = \lim_{x\to 4} \frac{1}{2+\sqrt{x}} = \frac{1}{2+\sqrt{4}} = \frac{1}{4}$$

L'Hôpital's Rule

L'Hôpital's Rule can help us calculate a limit that may otherwise be hard or impossible.

It says that the **limit** when we divide one function by another is the same after we take the derivative of each function (with some special conditions shown later).

In symbols we can write:

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

The limit as x approaches c of "f-of–x over g-of–x" equals the the limit as x approaches c of "f-dash-of–x over g-dash-of–x"

All we did is add that little dash mark ′ on each function, which means to take the derivative.

Example:

$$\lim_{x \to 2} \frac{x^2+x-6}{x^2-4}$$

At **x=2** we would normally get:

$$\frac{2^2+2-6}{2^2-4} = \frac{0}{0}$$

Which is indeterminate, so we are stuck. Or are we?

Let's try L'Hôpital!

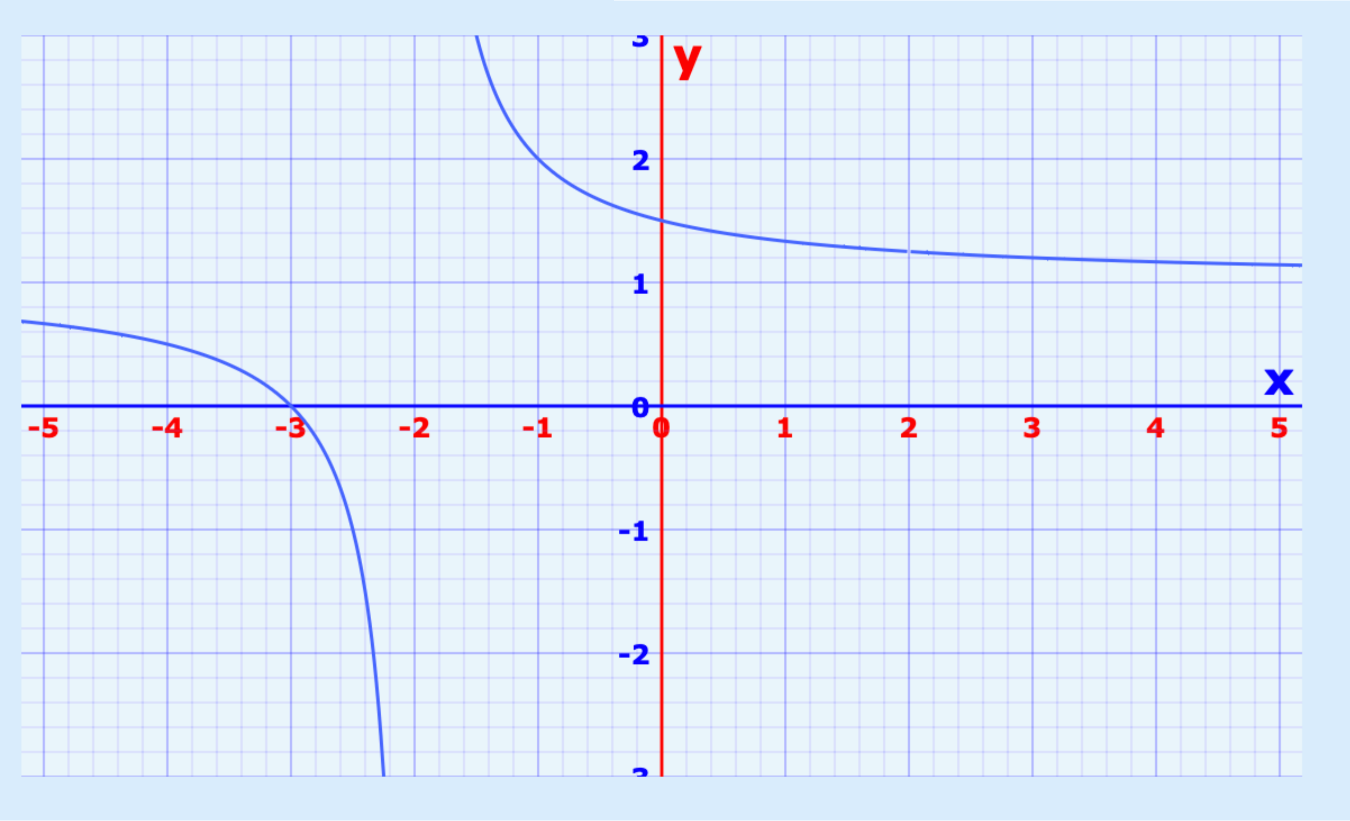
Differentiate both top and bottom (see Derivative Rules):

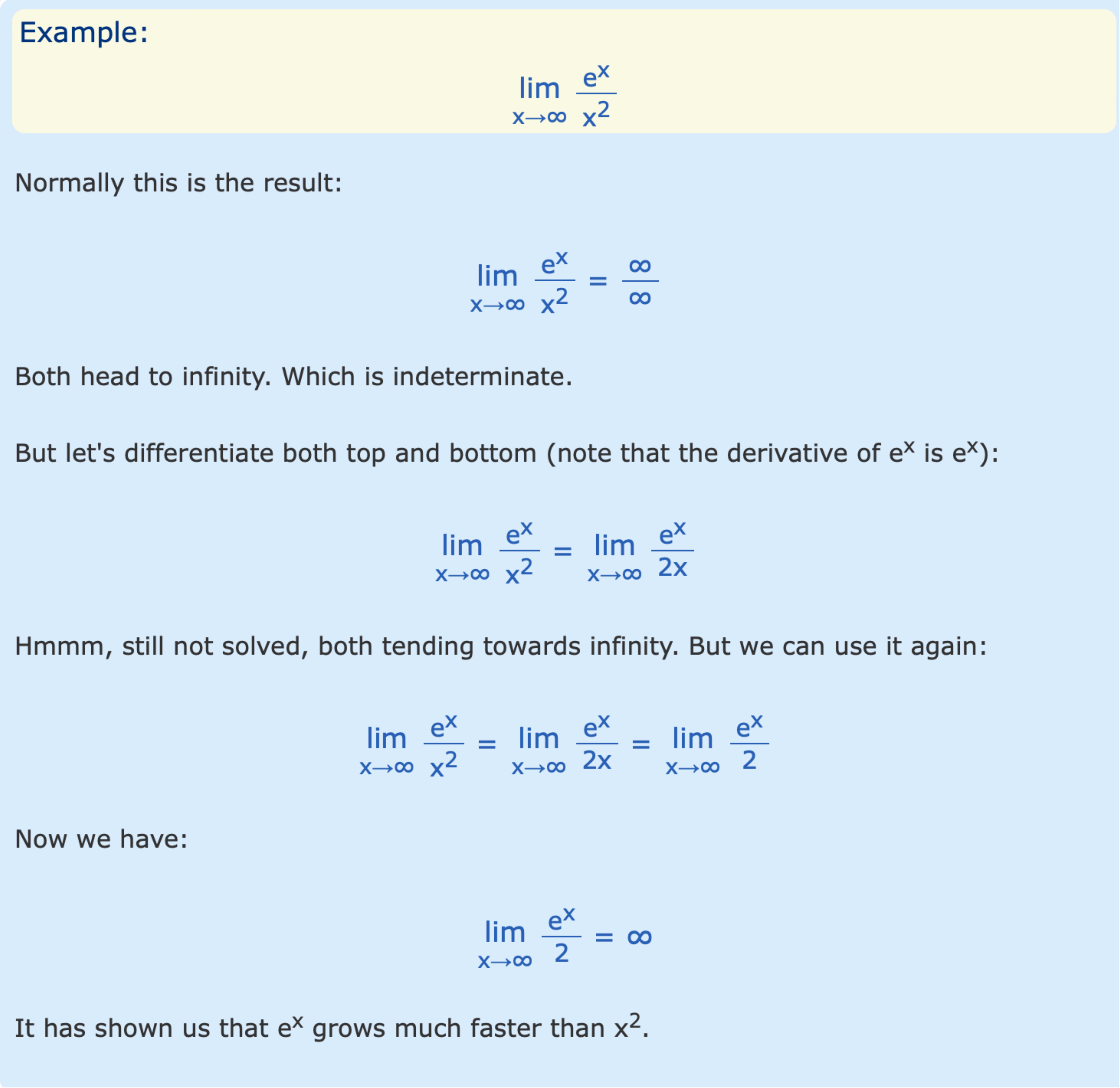
$$\lim_{x \to 2} \frac{x^2+x-6}{x^2-4} = \lim_{x \to 2} \frac{2x+1-0}{2x-0}$$

Now we just substitute **x=2** to get our answer:

$$\lim_{x \to 2} \frac{2x+1-0}{2x-0} = \frac{5}{4}$$

Here is the graph, notice the "hole" at x=2:





Example:

$$\lim_{x\to\infty} \frac{e^x}{x^2}$$

Normally this is the result:

$$\lim_{x\to\infty} \frac{e^x}{x^2} = \frac{\infty}{\infty}$$

Both head to infinity. Which is indeterminate.

But let's differentiate both top and bottom (note that the derivative of e^x is e^x):

$$\lim_{x\to\infty} \frac{e^x}{x^2} = \lim_{x\to\infty} \frac{e^x}{2x}$$

Hmmm, still not solved, both tending towards infinity. But we can use it again:

$$\lim_{x\to\infty} \frac{e^x}{x^2} = \lim_{x\to\infty} \frac{e^x}{2x} = \lim_{x\to\infty} \frac{e^x}{2}$$

Now we have:

$$\lim_{x\to\infty} \frac{e^x}{2} = \infty$$

It has shown us that e^x grows much faster than x^2.

Cases

We have already seen a $\frac{0}{0}$ and $\frac{\infty}{\infty}$ example. Here are all the indeterminate forms that L'Hopital's Rule may be able to help with:

$$\frac{0}{0} \quad \frac{\infty}{\infty} \quad 0\times\infty \quad 1^{\infty} \quad 0^{0} \quad \infty^{0} \quad \infty-\infty$$

Conditions

Differentiable

For a limit approaching c, the original functions must be differentiable either side of c, but not necessarily at c.

Likewise g′(x) is not equal to zero either side of c.

The Limit Must Exist

This limit must exist:

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

Why? Well a good example is functions that never settle to a value.

Example:

$$\lim_{x\to\infty} \frac{x+\cos(x)}{x}$$

Which is a $\frac{\infty}{\infty}$ case. Let's differentiate top and bottom:

$$\lim_{x\to\infty} \frac{1-\sin(x)}{1}$$

And because it just wiggles up and down it never approaches any value.

So that new limit does not exist!

And so L'Hôpital's Rule is not usable in this case.

BUT we can do this:

$$\lim_{x\to\infty} \frac{x+\cos(x)}{x} = \lim_{x\to\infty} \left(1 + \frac{\cos(x)}{x}\right)$$

As x goes to infinity then $\frac{\cos(x)}{x}$ tends to between $\frac{-1}{\infty}$ and $\frac{+1}{\infty}$, and both tend to zero.

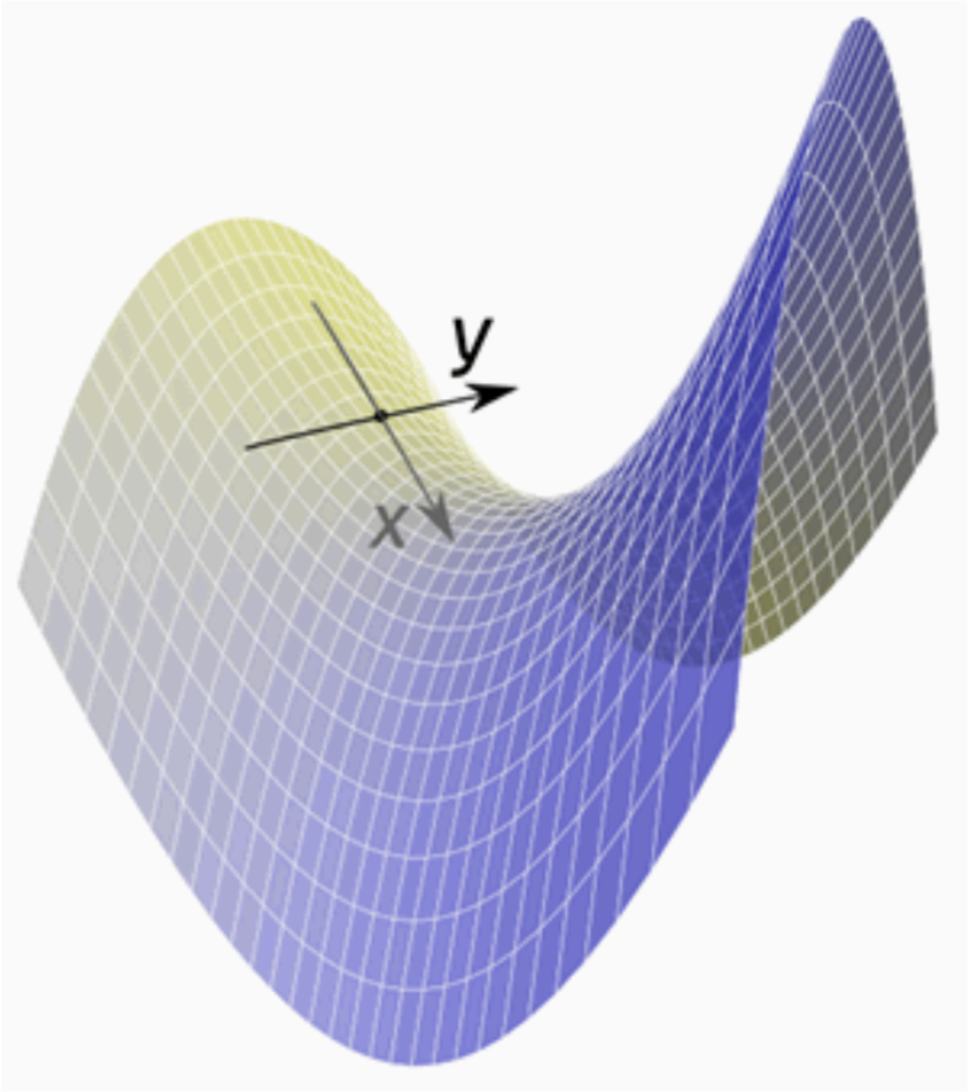
And we are left with just the "1", so:

$$\lim_{x\to\infty} \frac{x+\cos(x)}{x} = \lim_{x\to\infty} \left(1 + \frac{\cos(x)}{x}\right) = 1$$

Partial Derivatives

A Partial Derivative is a derivative where we hold some variables constant. Like in this example:

Example: a function for a surface that depends on two variables **x** and **y**



When we find the slope in the **x** direction (while keeping **y** fixed) we have found a partial derivative.

Or we can find the slope in the **y** direction (while keeping **x** fixed).

Let's first think about a function of **one variable** (x):

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

We can find its derivative using the Power Rule:

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

But what about a function of **two variables** (x and y):

$$f(x, y) = x^2 + y^3$$

We can find its **partial** derivative **with respect to x** when we treat **y as a constant** (imagine y is a number like 7 or something):

$$f'_x = 2x + 0 = 2x$$

Explanation:

- *the derivative of x^2 (with respect to x) is 2x*
- *we **treat y as a constant**, so y^3 is also a constant (imagine y=7, then 7^3=343 is also a constant), and the derivative of a constant is 0*

To find the partial derivative **with respect to y**, we treat **x as a constant**:

$$f'_y = 0 + 3y^2 = 3y^2$$

Explanation:

- *we now **treat x as a constant**, so x^2 is also a constant, and the derivative of a constant is 0*
- *the derivative of y^3 (with respect to y) is $3y^2$*

That is all there is to it. Just remember to treat **all other variables as if they are constants**.

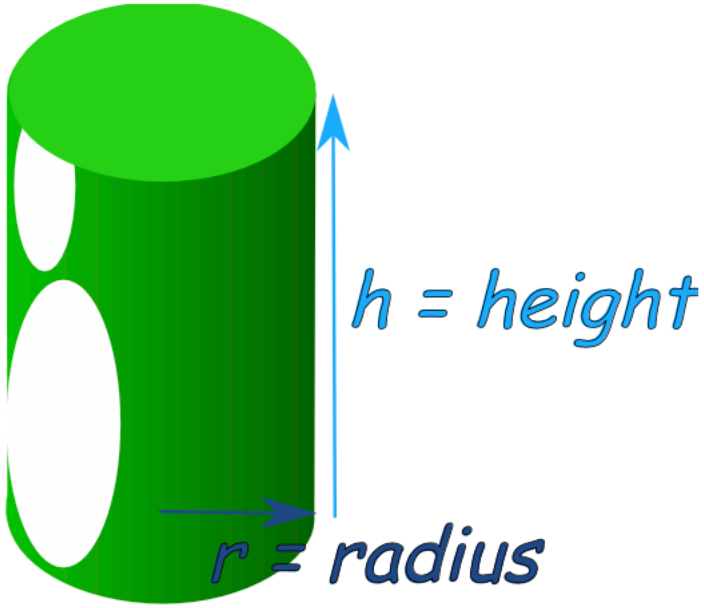
Holding A Variable Constant

So what does "holding a variable constant" look like?

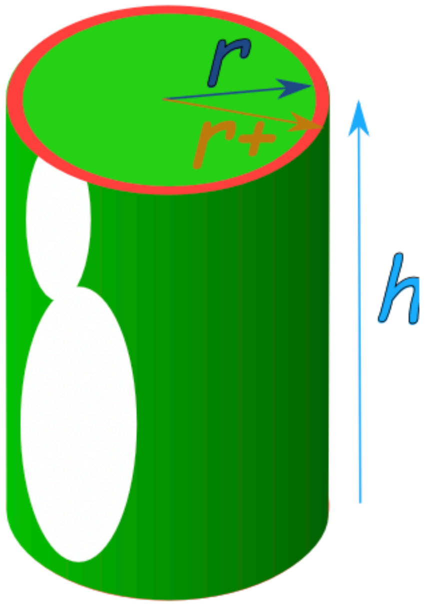
Example: the volume of a cylinder is $V = \pi r^2 h$

We can write that in "multi variable" form as

$$f(r, h) = \pi r^2 h$$



For the partial derivative with respect to r we hold **h constant**, and r changes:



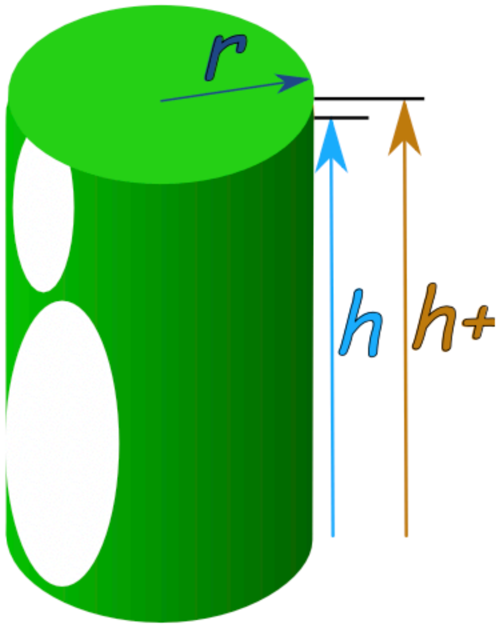
$$f'_r = \pi\ (2r)\ h = 2\pi rh$$

(The derivative of r^2 *with respect to r is 2r, and* **π** *and h are constants)*

It says "as only the radius changes (by the tiniest amount), the volume changes by 2**π**rh"

It is like we add a skin with a circle's circumference (2**π**r) and a height of h.

For the partial derivative with respect to h we hold **r constant**:



$$f'_h = \boldsymbol{\pi}\ r^2\ (1) = \boldsymbol{\pi} r^2$$

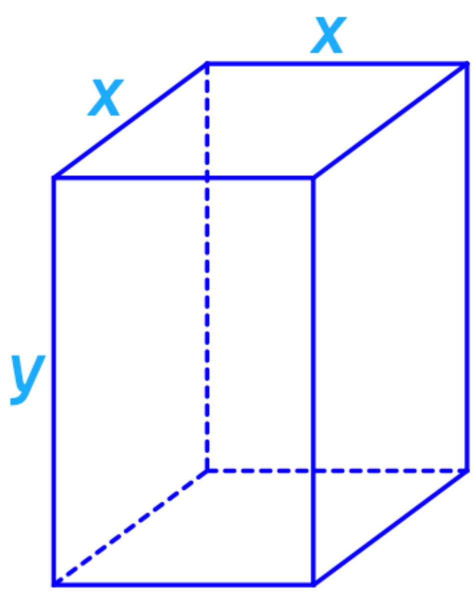
($\boldsymbol{\pi}$ and r^2 are constants, and the derivative of h with respect to h is 1)

It says "as only the height changes (by the tiniest amount), the volume changes by $\boldsymbol{\pi} r^2$"

It is like we add the thinnest disk on top with a circle's area of $\boldsymbol{\pi} r^2$.

Let's see another example.

Example: The surface area of a square prism.



The surface includes the top and bottom with areas of $\mathbf{x^2}$ each, and 4 sides of area **xy** each:

$$f(x, y) = 2x^2 + 4xy$$

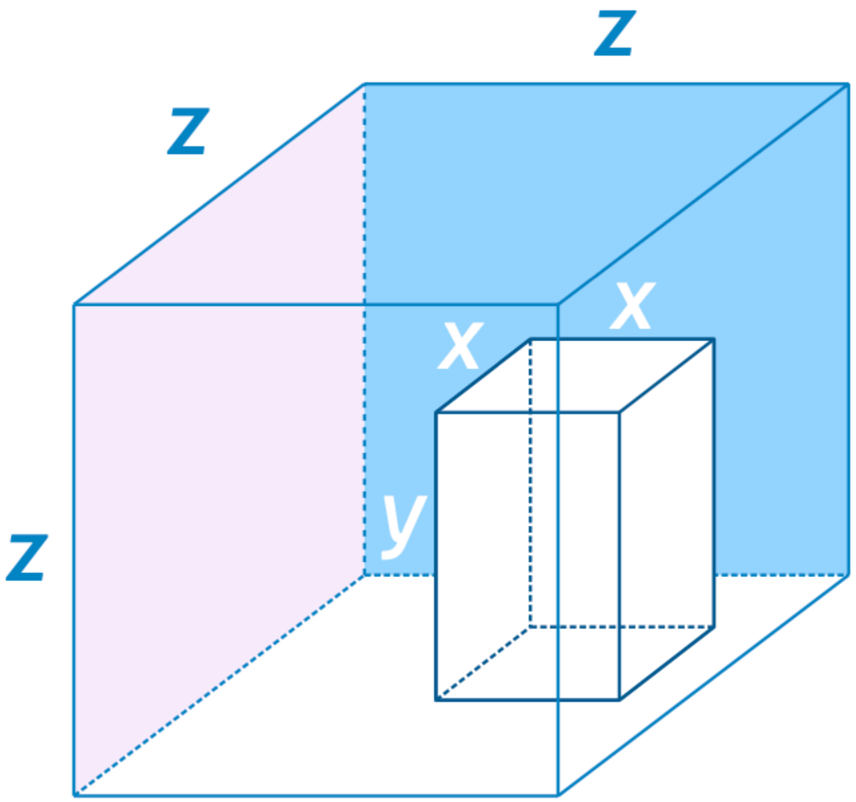
$$f'_x = 4x + 4y$$

$$f'_y = 0 + 4x = 4x$$

Three or More Variables

We can have 3 or more variables. Just find the partial derivative of each variable in turn while treating **all other variables as constants**.

Example: The volume of a cube with a square prism cut out from it.



$$f(x, y, z) = z^3 - x^2y$$

$$f'_x = 0 - 2xy = -2xy$$

$$f'_y = 0 - x^2 = -x^2$$

$$f'_z = 3z^2 - 0 = 3z^2$$

When there are many x's and y's it can get confusing, so a mental trick is to change the "constant" variables into letters like "c" or "k" that *look* like constants.

Example: $f(x, y) = y^3\sin(x) + x^2\tan(y)$

It has x's and y's all over the place! So let us try the letter change trick.

With respect to x we can change "y" to "k":

$f(x, y) = k^3\sin(x) + x^2\tan(k)$

$$f'_x = k^3\cos(x) + 2x\tan(k)$$

But remember to turn it back again!

$$f'_x = y^3\cos(x) + 2x\tan(y)$$

Likewise with respect to y we turn the "x" into a "k":

$f(x, y) = y^3\sin(k) + k^2\tan(y)$

$$f'_y = 3y^2\sin(k) + k^2\sec^2(y)$$

$$f'_y = 3y^2\sin(x) + x^2\sec^2(y)$$

But only do this if you have trouble remembering, as it is a little extra work.

Notation: we have used $\mathbf{f'_x}$ to mean "the partial derivative with respect to x", but another very common notation is to use a funny backwards d (∂) like this:

$$\frac{\partial f}{\partial x} = 2x$$

Which is the same as:

$$f'_x = 2x$$

∂ is called "del" or "dee" or "curly dee"

So $\frac{\partial f}{\partial x}$ can be said "del f del x"

Example: find the partial derivatives of $\mathbf{f(x, y, z) = x^4 - 3xyz}$ using "curly dee" notation

$f(x, y, z) = x^4 - 3xyz$

$$\frac{\partial f}{\partial x} = 4x^3 - 3yz$$

$$\frac{\partial f}{\partial y} = -3xz$$

$$\frac{\partial f}{\partial z} = -3xy$$

Stationary Points

(Definition & How to Find Stationary Points)

A ***stationary point***, or ***critical point***, is a point at which the ***curve's gradient equals to zero***. Consequently if a curve has equation $y = f(x)$ then at a stationary point we'll always have:

$$f'(x) = 0$$

which can also be written:

$$\frac{dy}{dx} = 0$$

In other words the ***derivative function equals to zero at a stationary point***.

Different Types of Stationary Points

There are ***three types of stationary points***:

- ***local (or global) maximum points***
- ***local (or global) minimum points***
- ***horizontal (increasing or decreasing) points of inflexion***.

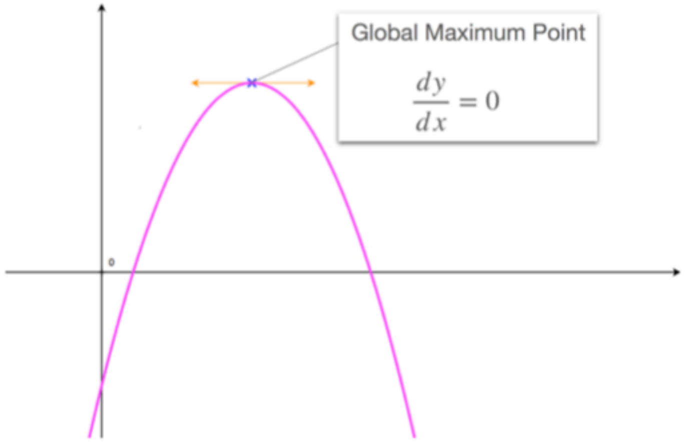
It is worth pointing out that *maximum* and *minimum* points are often called ***turning points***.

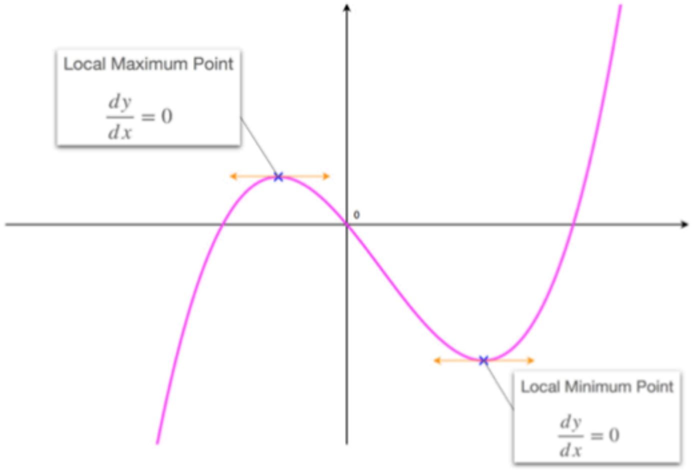
Turning Points

A ***turning point*** is a ***stationary point***, which is either:

- a *local* (or global) *minimum*
- a *local* (or global) *maximum*

each of which are illustared in the graphs shown here, where the *horizontal tangent* is shown in orange:



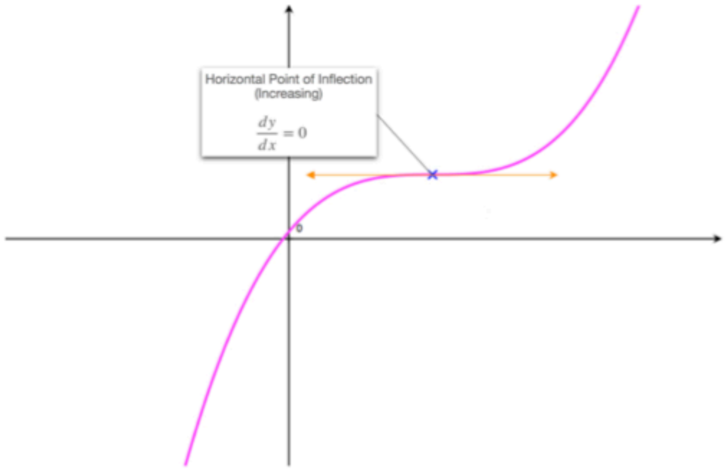


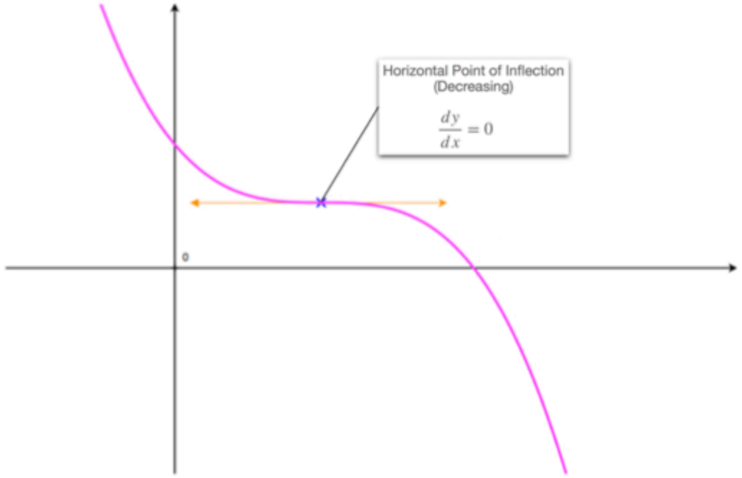
Horizontal Points of Inflection

A ***horizontal point of inflection*** is a ***stationary point***, which is either:

- a *increasing horizontal point of inflection*
- a *decreasing horizontal point of inflection*

each of which are illustared in the graphs shown here, where the *horizontal tangent* is shown in orange:





Method: finding stationary points

Given a function $f(x)$ and its curve $y = f(x)$, to find any *stationary point(s)* we follow ***three steps***:

- **Step 1:** find $f'(x)$
- **Step 2:** solve the equation $f'(x) = 0$, this will give us the x-coordinate(s) of any *stationary point(s)*.
- **Step 3** (if needed/asked)**:** calculate the y-coordinate(s) of the stationary point(s) by plugging the x values found in step 2 into $f(x)$.

Example 1

Given the function defined by the equation:

$$y = x^2 - 4x + 5$$

find the coordinates of any *stationary point(s)*.

Solution

Following our three-step method:

- **Step 1:** find $\frac{dy}{dx}$.
  For $y = x^2 - 4x + 5$, we find:

  $$\frac{dy}{dx} = 2x - 4$$

- **Step 2:** solve the equation $\frac{dy}{dx} = 0$.
  That's:

  $$2x - 4 = 0$$

  Solving this leads to:

  $$2x = 4$$

  Finally:

  $$x = 2$$

At this stage we can state the curve $y = x^2 - 4x + 5$ has one stationary point whose x-coordinate is $x = 2$.

- **Step 3:** calculate the sationary point's y-coordinate.

  To do this we replace x by 2 in the curve's equation $y = x^2 - 4x + 5$.

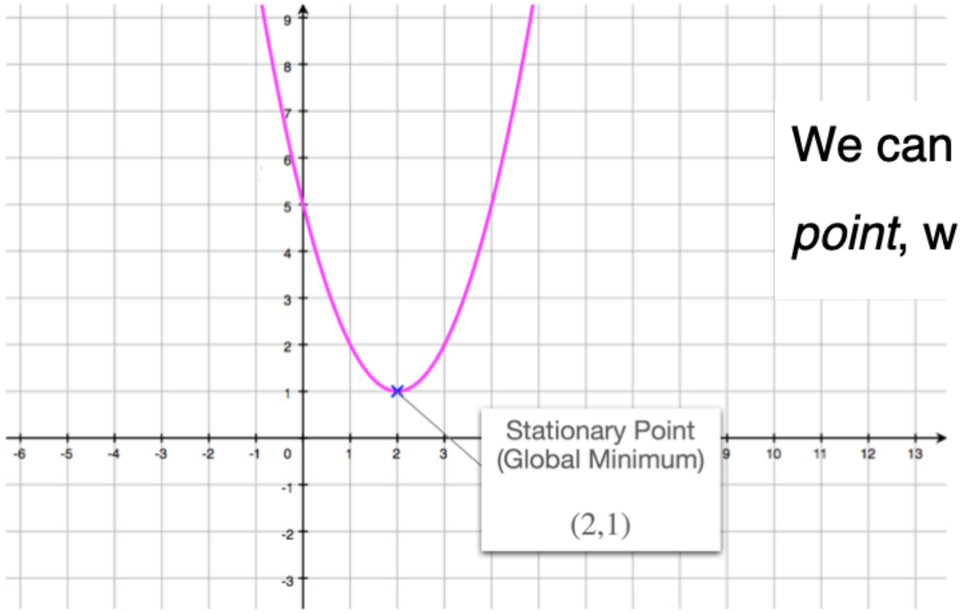
  That's:

$$\begin{aligned} y &= 2^2 - 4 \times 2 + 5 \\ &= 4 - 8 + 5 \\ &= -4 + 5 \\ y &= 1 \end{aligned}$$

  So the *stationary point* has y-coordinate $y = 1$.

We can therefore state that the curve $y = x^2 - 4x + 5$ has one stationary point with coordinates $(2, 1)$.

This result is confirmed, using our graphical calculator and looking at the curve $y = x^2 - 4x + 5$:



We can see quite clearly that the curve has a *global minimum point*, which is a *stationary point*, at $(2, 1)$.

Example 2

Find the coordinates of any stationary point(s) of the function defined by:

$$y = 2x^3 + 3x^2 - 12x + 1$$

Solution

Following our three-step method:

- **Step 1:** find $\frac{dy}{dx}$.

  Using the power rule for differentiation we find:

$$\begin{aligned}\frac{dy}{dx} &= 3 \times 2x^{3-1} + 2 \times 3x^{2-1} - 12 \\ &= 6x^2 + 6x^1 - 12 \\ \frac{dy}{dx} &= 6x^2 + 6x - 12\end{aligned}$$

- **Step 2:** solve the equation $\frac{dy}{dx} = 0.$

  Since $\frac{dy}{dx} = 6x^2 + 6x - 12$ we need to solve the *quadratic equation*:

$$6x^2 + 6x - 12 = 0$$

  we find two solutions:

$$x = -2 \quad \text{and} \quad x = 1$$

So at this stage we can state that the function $y = 2x^3 + 3x^2 - 12x + 1$ has two *stationary points*. One with x-coordinate $x = -2$ and the other with x-coordinate $x = 1$.

- **Step 3:** calculate the sationary point's y-coordinate.

  Since we found two values of x, in step 2, there are two y-coordinates to calculate, one for each value of x.

  - when $x = -2$:

    Replacing x by -2 in $y = 2x^3 + 3x^2 - 12x + 1$, we find:

    $$\begin{aligned} y &= 2 \times (-2)^3 + 3 \times (-2)^2 - 12 \times (-2) + 1 \\ &= 2 \times (-8) + 3 \times 4 - (-24) + 1 \\ &= -16 + 12 + 24 + 1 \\ y &= 21 \end{aligned}$$

    So the function has a *stationary point* at:

    $$\left(-2, 21\right)$$

  - when $x = 1$:

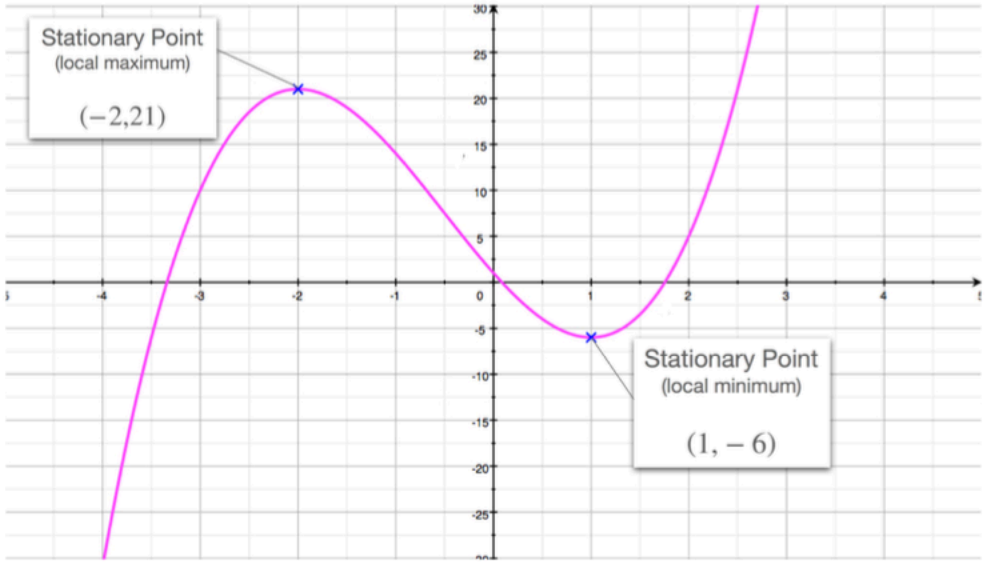
    Replacing x by 1 in $y = 2x^3 + 3x^2 - 12x + 1$, we find:

    $$\begin{aligned} y &= 2 \times 1^3 + 3 \times 1^2 - 12 \times 1 + 1 \\ &= 2 \times 1 + 3 \times 1 - 12 + 1 \\ &= 2 + 3 - 12 + 1 \\ y &= -6 \end{aligned}$$

So the function has its second *stationary point* at:

$$(1, -6)$$

We can see both of these *stationary points* on the graph shown below:



We can see quite clearly that the *stationary point* at $(-2, 21)$ is a *local maximum* and the *stationary point* at $(1, -6)$ is a *local minimum*.

Example 3

Given the function defined by:

$$y = x^3 - 6x^2 + 12x - 12$$

Find the coordinates of any *stationary point(s)* along this function's curve's length.

Solution

Following our three-step method:

- **Step 1:** find $\frac{dy}{dx}$.
  Since $y = x^3 - 6x^2 + 12x - 12$, we use the power rule for differentiation to find this function's *derivative*:

$$\begin{aligned}\frac{dy}{dx} &= 3 \times x^{3-1} - 2 \times 6x^{2-1} + 12x^{1-1} + 0 \\ &= 3x^2 - 12x^1 + 12x^0 \\ \frac{dy}{dx} &= 3x^2 - 12x + 12\end{aligned}$$

- **Step 2:** solve the equation $\frac{dy}{dx} = 0$.
  Since $\frac{dy}{dx} = 3x^2 - 12x + 12$ we have to solve the *quadratic equation*:

$$3x^2 - 12x + 12 = 0$$

We can solve this using the quadratic formula or by factoring

In doing so we find one solution:

$$x = 2$$

So, at this stage, we can state that this function has one *stationary point* whose x-coordinate is $x = 2$.

- **Step 3:** calculate the sationary point's y-coordinate.

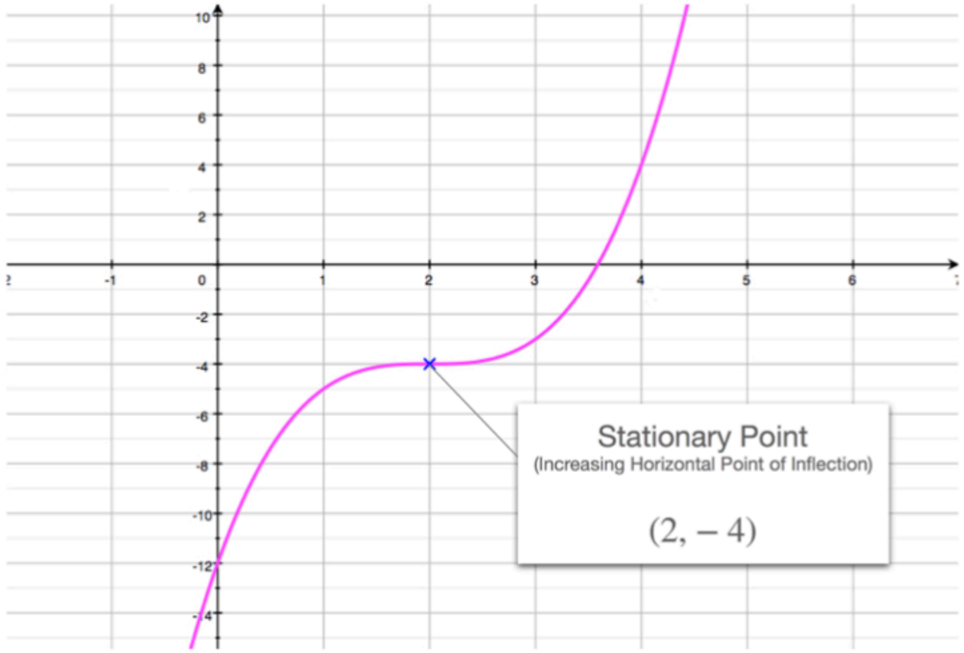
  Since we found that the *stationary point* had x-coordinate $x = 2$, to find its y-coordinate we replace x by 2 in the function's equation $y = x^3 - 6x^2 + 12x - 12$. That's:

$$\begin{aligned} y &= 2^3 - 6 \times 2^2 + 12 \times 2 - 12 \\ &= 8 - 6 \times 4 + 24 - 12 \\ &= 8 - 24 + 24 - 12 \\ y &= -4 \end{aligned}$$

So this function has a *stationary point* with coordinates:

$$\left(2, -4 \right)$$

This result is confirmed when we look at the graph of $y = x^3 - 6x^2 + 12x - 12$:



Looking at this graph, we can see that this curve's *stationary point* at $\left(2, -4 \right)$ is an *increasing horizontal point of inflection.*

Example 4

Given the function defined by:

$$y = x + \frac{4}{x}$$

find the coordinates of any *stationary points* along this curve's length.

Solution

Following our three-step method:

- **Step 1:** find $\frac{dy}{dx}$.
  We can re-write $y = x + \frac{4}{x}$, using *negative exponents*:

  $$y = x + 4.x^{-1}$$

  We can now use the power rule for differentiation to find the *derivative*:

  $$\begin{aligned} \frac{dy}{dx} &= 1 + (-1).4.\,x^{-1-1} \\ &= 1 - 4x^{-2} \\ \frac{dy}{dx} &= 1 - \frac{4}{x^2} \end{aligned}$$

- **Step 2:** solve the equation $\frac{dy}{dx} = 0$.

  Since $\frac{dy}{dx} = 1 - \frac{4}{x^2}$ we need to solve the equation:

  $$1 - \frac{4}{x^2} = 0$$

  To solve this equation by hand we start by writing the entire left hand side over x^2 using the fact that $1 = \frac{x^2}{x^2}$ so that:

  $$1 - \frac{4}{x^2} = 0$$

  can be written:

  $$\frac{x^2}{x^2} - \frac{4}{x^2} = 0$$

  and therefore we have to solve:

  $$\frac{x^2 - 4}{x^2} = 0$$

  This will equal to 0 if and only if the numerator, $x^2 - 4$, equals to 0.

  So all we need to solve is:

  $$x^2 - 4 = 0$$

That's:

$$x^2 = 4$$

which leads to two solutions:

$$x = -2 \quad \text{and} \quad x = 2$$

- **Step 3:** calculate the sationary point's y-coordinate.

  Since we found two values of x, in step 2, there are two y-coordinates to calculate, one for each value of x.
  - when $x = -2$:

    replacing x by -2 in $y = x + \frac{4}{x}$, we find:

    $$\begin{aligned} y &= -2 + \frac{4}{-2} \\ &= -2 + (-2) \\ &= -2 - 2 \\ y &= -4 \end{aligned}$$

    So one of this function's *stationary points* is:

    $$\left(-2, -4\right)$$
  - when $x = 2$:

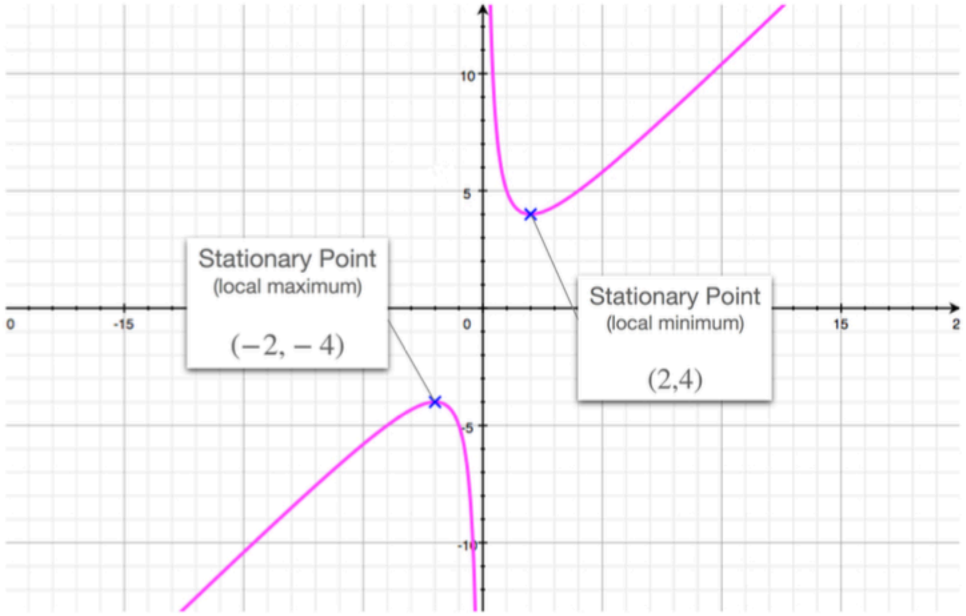
    replacing x by 2 in $y = x + \frac{4}{x}$, we find:

$$
\begin{aligned}
y &= 2 + \frac{4}{2} \\
&= 2 + 2 \\
y &= 4
\end{aligned}
$$

So the function's second *stationary point* has coordinates:

$$(2, 4)$$

We can see both of these *stationary points* on the graph shown below:

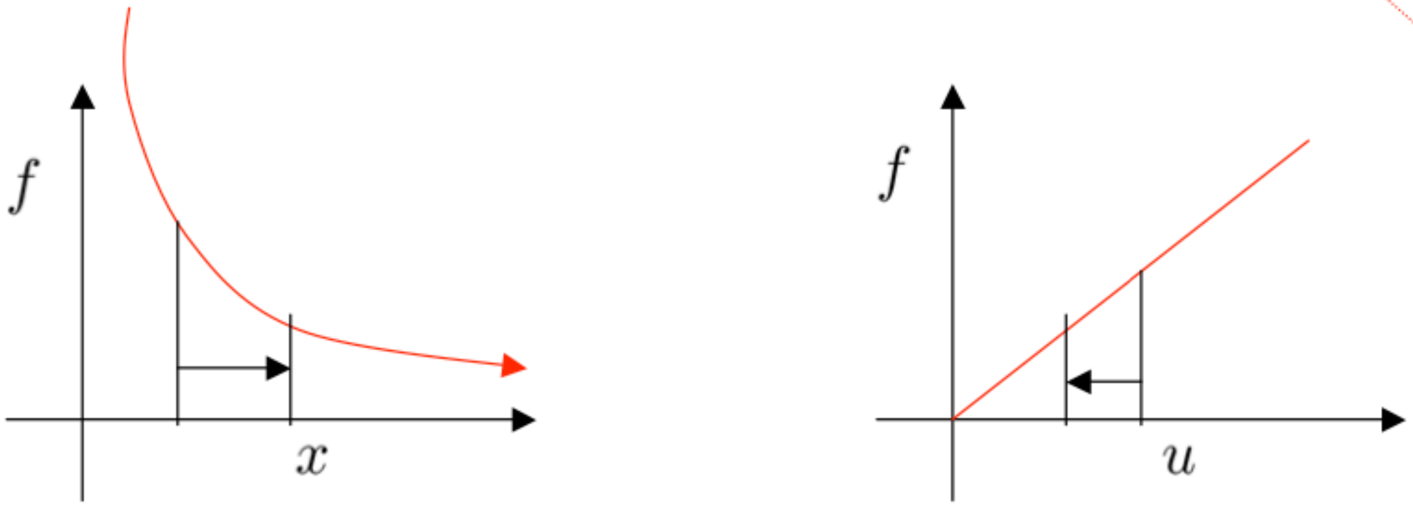


We can see quite clearly that the *stationary point* at $(-2, -4)$ is a *local maximum* and the *stationary point* at $(2, 4)$ is a *local minimum*.

Jacobians

- In 1D problems we are used to a simple change of variables, e.g. from *x* to *u*

$$\int_a^b f(x)\mathrm{d}x = \int_\alpha^\beta f(x(u))\frac{\mathrm{d}x}{\mathrm{d}u}\mathrm{d}u$$



1D Jacobian

maps strips of width d*x* to strips of width d*u*

- Example: $\int_1^2 \frac{1}{x}\mathrm{d}x = \ln(2)$ Substitute $x = u^{-1} \rightarrow \frac{\mathrm{d}x}{\mathrm{d}u} = -u^{-2}$

$$= -\int_1^{\frac{1}{2}} \frac{u}{u^2}\mathrm{d}u = [\ln u]_{\frac{1}{2}}^{1} = \ln(2)$$

2D Jacobian

- For a continuous 1-to-1 transformation from (*x*,*y*) to (*u*,*v*)
- Then $x = x(u, v)$ and $y = y(u, v)$

$$\int \int_R f(x, y)\mathrm{d}x\mathrm{d}y = \int \int_{R'} f(x(u, v), y(u, v)) \left| \frac{\partial(x,y)}{\partial(u,v)} \right| \mathrm{d}u\mathrm{d}v$$

- Where Region (in the *xy* plane) maps onto region R in the *uv* plane R'

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

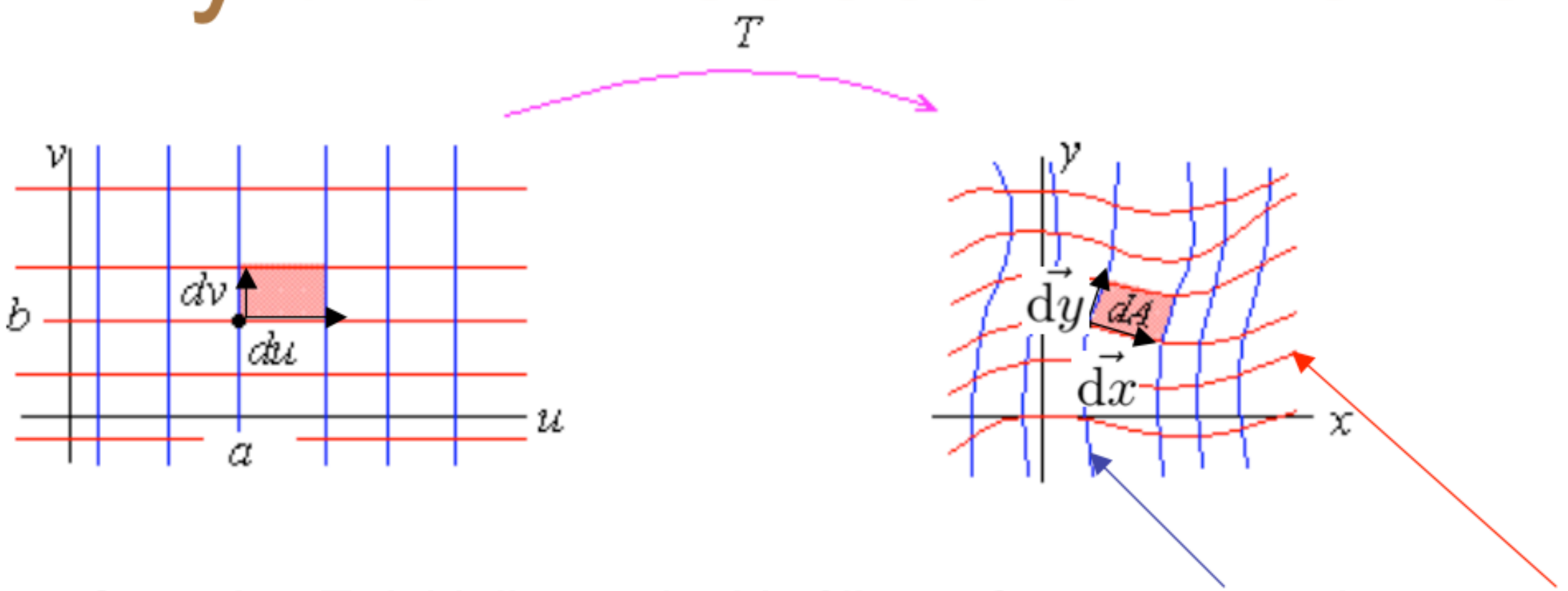
2D Jacobian

maps areas d*x*d*y* to areas d*u*d*v*

$$= \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} = x_u y_v - x_v y_u$$

- Hereafter call such terms x_u

Why the 2D Jacobian works



- Transformation *T* yield distorted grid of lines of constant *u* and constant *v*
- For small d*u* and d*v,* rectangles map onto parallelograms

$$\vec{\mathrm{d}u} = (\mathrm{d}u, 0) \quad \text{and} \quad \vec{\mathrm{d}v} = (0, \mathrm{d}v)$$

$$\vec{\mathrm{d}x} = \begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix} \begin{pmatrix} \mathrm{d}u \\ 0 \end{pmatrix} \qquad \vec{\mathrm{d}y} = \begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix} \begin{pmatrix} 0 \\ \mathrm{d}v \end{pmatrix}$$

$$\mathrm{d}A = |\vec{\mathrm{d}x} \times \vec{\mathrm{d}y}| = \begin{pmatrix} x_u \mathrm{d}u \\ y_u \mathrm{d}u \end{pmatrix} \times \begin{pmatrix} x_v \mathrm{d}v \\ y_v \mathrm{d}v \end{pmatrix} = (x_u y_v - x_v y_u)\mathrm{d}u\mathrm{d}v$$

- This is a ***Jacobian***, i.e. the determinant of the ***Jacobian Matrix***

Relation between Jacobians

- The Jacobian matrix $\frac{\partial(x,y)}{\partial(u,v)}$ is the ***inverse matrix*** of $\frac{\partial(u,v)}{\partial(x,y)}$ i.e.,

$$\begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix} \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

- Because (and similarly for d*y)*

$$\mathrm{d}x = x_u \mathrm{d}u + x_v \mathrm{d}v = x_u \mathrm{d}u + x_v (v_x \mathrm{d}x + v_y \mathrm{d}y)$$

$$\begin{aligned} x \text{ constant } &\Rightarrow \mathrm{d}x = 0 \Rightarrow 0 = x_u u_y + x_v v_y \\ y \text{ constant } &\Rightarrow \mathrm{d}y = 0 \Rightarrow 1 = x_u u_x + x_v v_x \end{aligned}$$

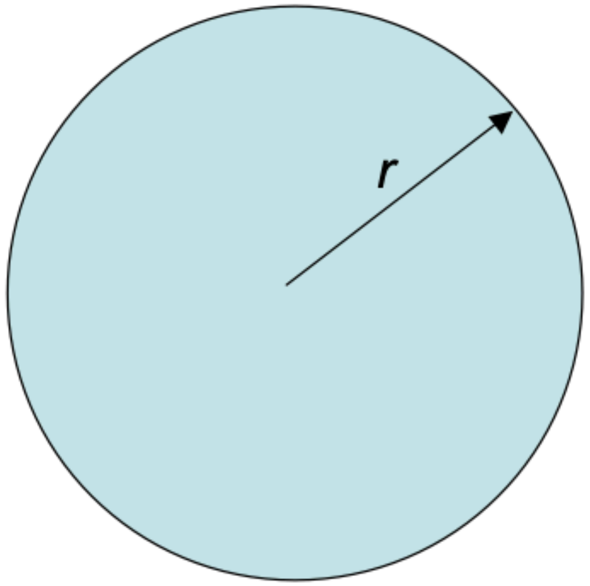
- This makes sense because Jacobians measure the relative areas of d*x*d*y* and d*u*d*v*, i.e

$$\det(AB) = \det(A)\det(B) = 1 \Rightarrow \det(A) = \frac{1}{\det B}$$

- So

$$\left|\frac{\partial(x,y)}{\partial(u,v)}\right| = \left|\frac{\partial(u,v)}{\partial(x,y)}\right|^{-1}$$

Simple 2D Example



Area of circle *A*=

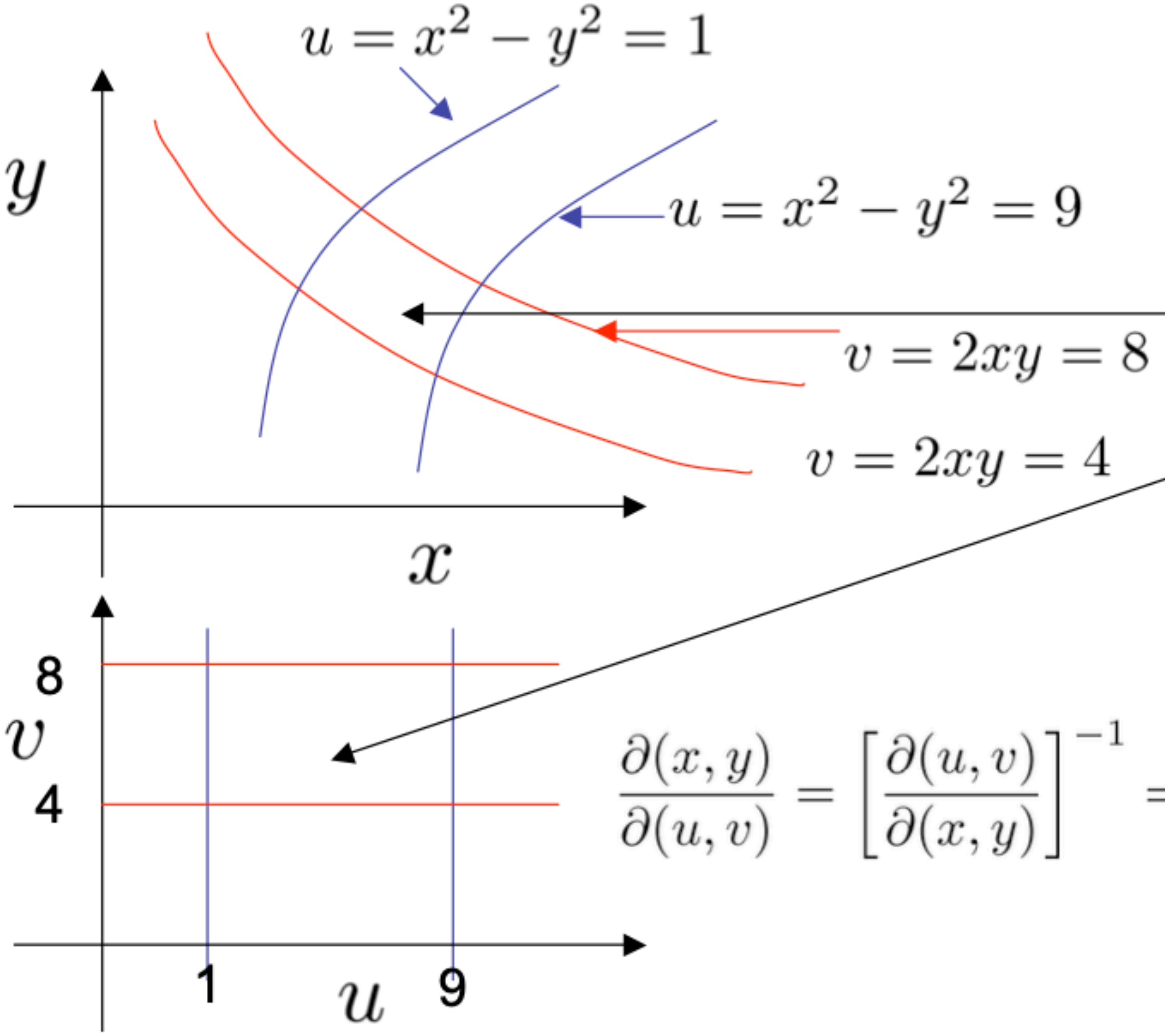
$$\int\int_A \mathrm{d}x\mathrm{d}y$$

$$x = \rho\cos\theta \quad \text{and} \quad y = \rho\sin\theta$$

$$\left|\begin{array}{cc} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \theta} \end{array}\right| = \left|\begin{array}{cc} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{array}\right| = \rho(\cos^2\theta + \sin^2\theta) = \rho$$

$$A = \int_{\rho=0}^{\rho=r}\int_{\theta=0}^{\theta=2\pi} \rho\mathrm{d}\rho\mathrm{d}\theta = \left[\tfrac{1}{2}\rho^2\right]_0^r [\theta]_0^{2\pi} = \pi r^2$$

Harder 2D Example



$$\int \int_R (x^2 + y^2) \mathrm{d}x\mathrm{d}y$$

where *R* is this region of the *xy* plane, which maps to *R'* here

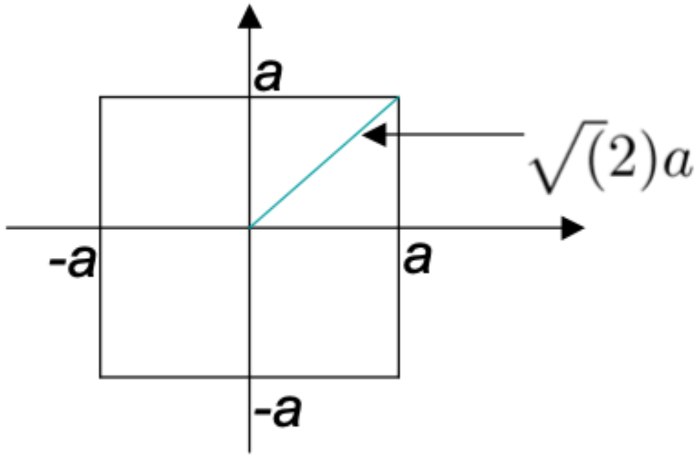
$$\frac{\partial(x,y)}{\partial(u,v)} = \left[\frac{\partial(u,v)}{\partial(x,y)}\right]^{-1} = \begin{vmatrix} 2x & -2y \\ 2y & 2x \end{vmatrix}^{-1} = \frac{1}{4(x^2+y^2)}$$

but $u^2 + v^2 = (x^2 - y^2)^2 + 4x^2y^2 = (x^2 + y^2)^2$ so $\frac{\partial(x,y)}{\partial(u,v)} = \frac{1}{4\sqrt{(u^2+v^2)}}$

$$\int \int_R (x^2 + y^2) \mathrm{d}x\mathrm{d}y = \frac{1}{4} \int_1^9 \int_4^8 \mathrm{d}u\mathrm{d}v = 8$$

An Important 2D Example

- Evaluate $$I = \int_{-\infty}^{\infty} e^{-x^2}\mathrm{d}x$$



- First consider

$$I_a^2 = \int_{-a}^{a} e^{-x^2}\mathrm{d}x \int_{-a}^{a} e^{-y^2}\mathrm{d}y$$

$$I_a^2 = \int_{-a}^{a}\int_{-a}^{a} e^{-(x^2+y^2)}\mathrm{d}y\mathrm{d}x$$

- Put $x = r\cos\phi$ and $y = r\sin\phi$ $\quad \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \cos\phi & -r\cos\phi \\ \sin\phi & r\cos\phi \end{vmatrix} = r$

$$\int_0^a \int_0^{2\pi} re^{-r^2}\mathrm{d}r\mathrm{d}\phi < I_a^2 < \int_0^{\sqrt{2}a}\int_0^{2\pi} re^{-r^2}\mathrm{d}r\mathrm{d}\phi$$

- $\pi(1-e^{-a^2}) < I_a^2 < \pi(1-e^{-2a^2})$ as $a \longrightarrow \infty \Rightarrow I_a = \sqrt{\pi}$

3D Example

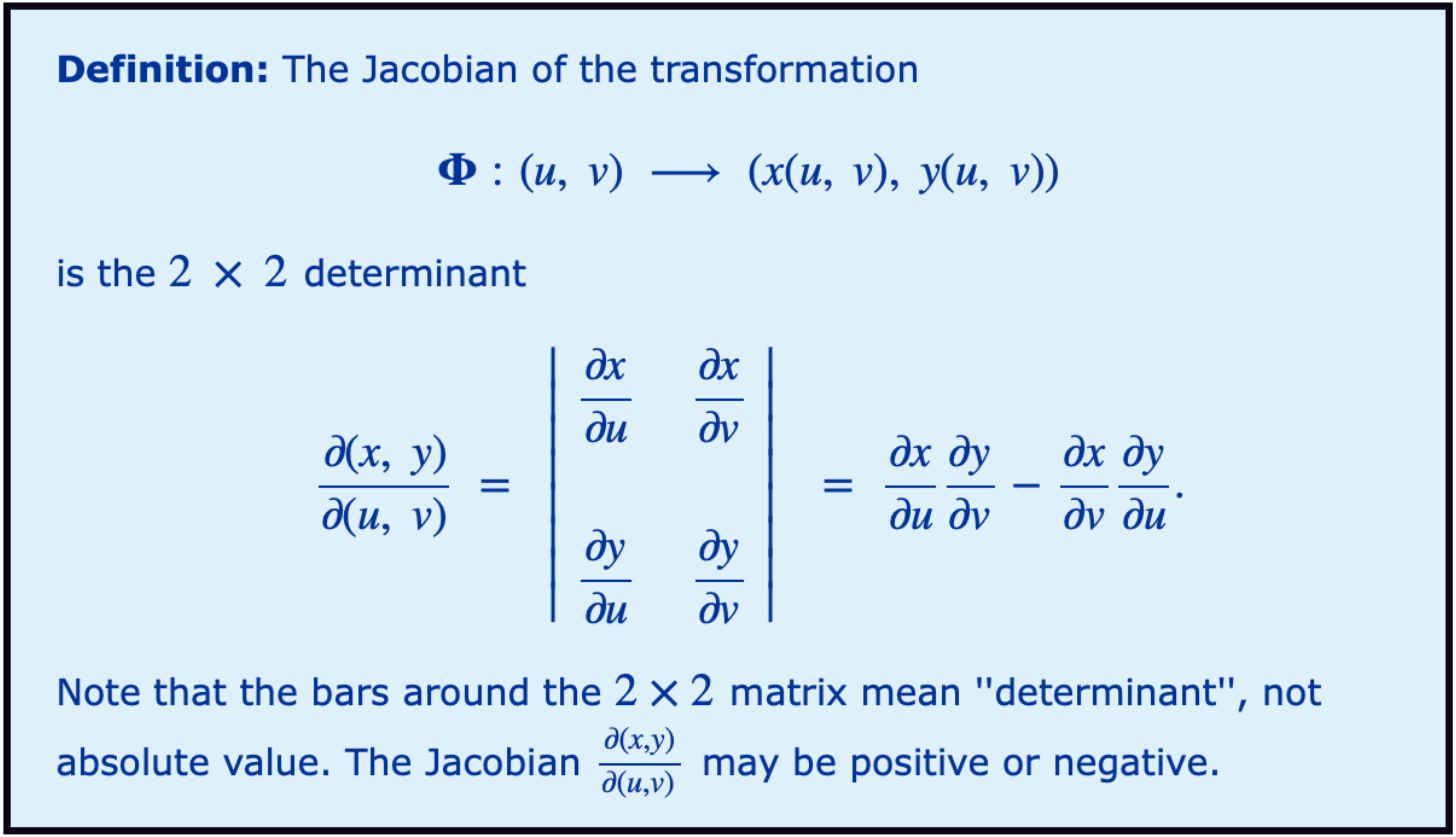
- Transformation of volume elements between Cartesian and spherical polar coordinate systems (see Lecture 4)

$$
\begin{aligned}
x &= r\sin\theta\cos\phi \\
y &= r\sin\theta\sin\phi \\
z &= r\cos\theta
\end{aligned}
$$

$$
\frac{\partial(x,y,z)}{\partial(r,\theta,\phi)} = \begin{vmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{vmatrix}
$$

$$
= r^2\sin\theta
$$

More..........



Definition: The Jacobian of the transformation

$$\mathbf{\Phi} : (u,\ v) \longrightarrow (x(u,\ v),\ y(u,\ v))$$

is the 2×2 determinant

$$\frac{\partial(x,\ y)}{\partial(u,\ v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}.$$

Note that the bars around the 2×2 matrix mean "determinant", not absolute value. The Jacobian $\frac{\partial(x,y)}{\partial(u,v)}$ may be positive or negative.

Change-of-variable formula: If a 1-1 mapping Φ sends a region D^* in uv-space to a region D in xy-space, then

$$\iint_D f(x, y)dx\,dy \;=\; \iint_{D^*} f(\Phi(u, v)) \left|\frac{\partial(x, y)}{\partial(u, v)}\right| du\,dv.$$

Note that this involves the *absolute value* of the Jacobian. Even when the Jacobian is negative, the distortion in volume is positive.

Example 1: *Compute the Jacobian of the polar coordinates transformation*

$$x = r\cos\theta, \qquad y = r\sin\theta\,.$$

Solution: Since

$$\frac{\partial x}{\partial r} = \cos(\theta), \qquad \frac{\partial y}{\partial r} = \sin(\theta),$$
$$\frac{\partial x}{\partial \theta} = -r\sin(\theta), \qquad \frac{\partial y}{\partial \theta} = r\cos(\theta),$$

our Jacobian is

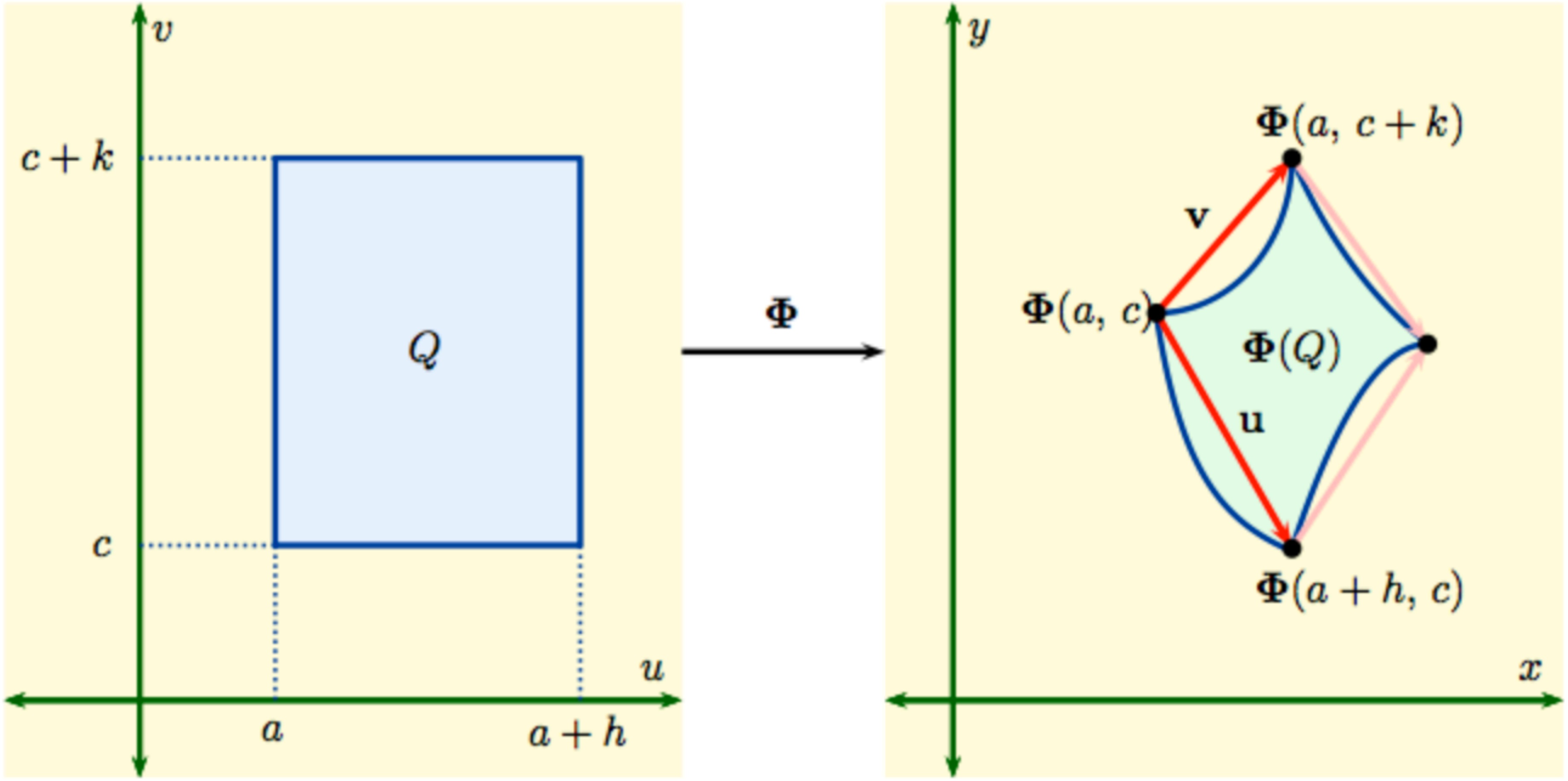
$$\begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r\,.$$

This explains why there's an r factor in polar integrals! The area element $dA = dx\,dy$ is not equal to $dr\,d\theta$. Instead, dA is equal to $r\,dr\,d\theta$.

Let's see why the Jacobian is the distortion factor in general for a mapping

$$\mathbf{\Phi} : (u,\ v) \ \rightarrow\ (x(u,\ v),\ y(u,\ v)) \ =\ x(u,\ v)\,\mathbf{i} + y(u,\ v)\,\mathbf{j}\,,$$

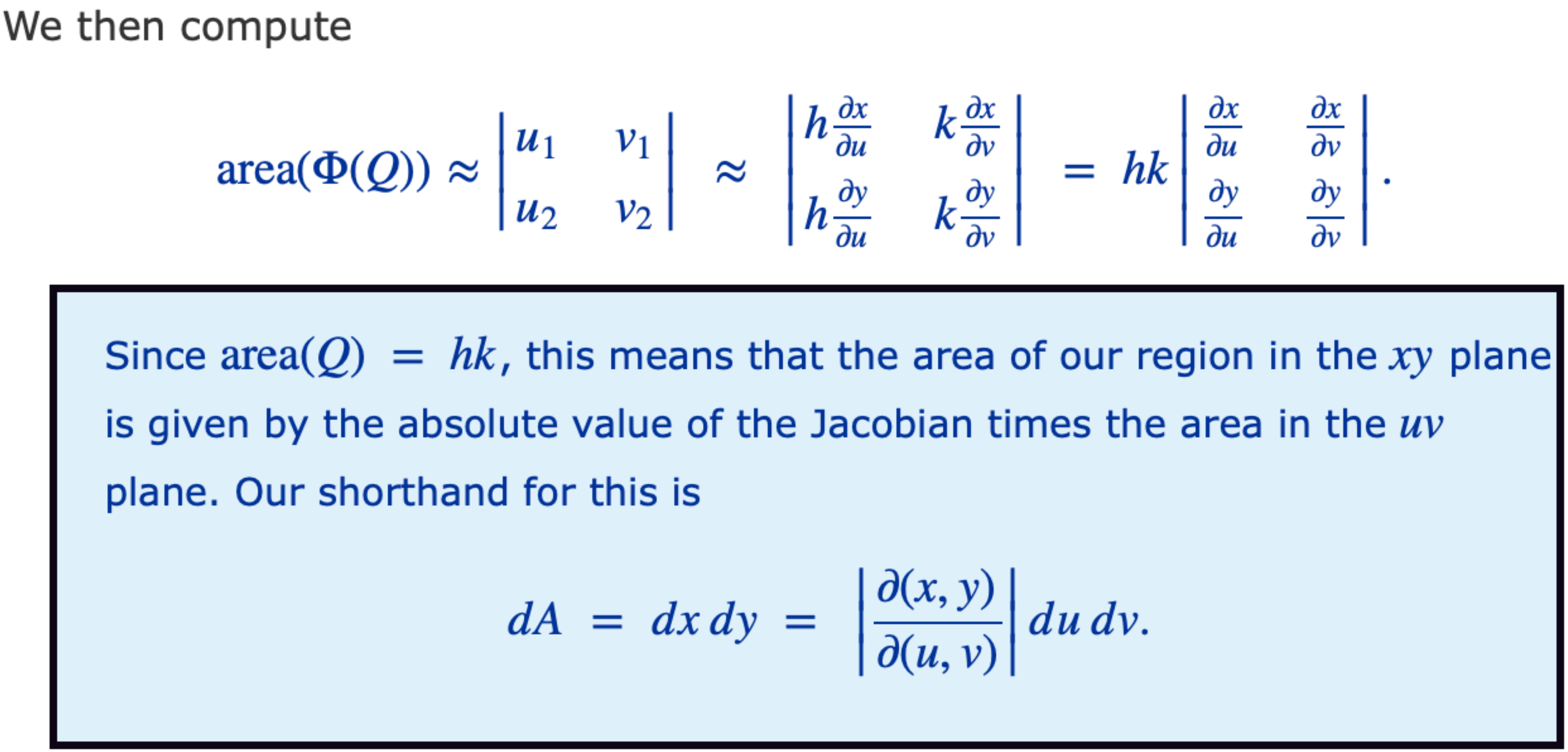
making good use of all the vector calculus we've developed so far. Let $Q = [a,\ a+h] \times [c,\ c+k]$ be a rectangle in the uv-plane and $\mathbf{\Phi}(Q)$ its image in the xy-plane as shown in



Then

$$\mathbf{u} \ =\ \mathbf{\Phi}(a+h,\ c) - \mathbf{\Phi}(a,\ c)\,, \qquad \mathbf{v} \ =\ \mathbf{\Phi}(a,\ c+k) - \mathbf{\Phi}(a,\ c)\,.$$

The area of the parallelogram spanned by $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$ and $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j}$ is the determinant $\begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix}$.



We then compute

$$\operatorname{area}(\Phi(Q)) \approx \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} \approx \begin{vmatrix} h\frac{\partial x}{\partial u} & k\frac{\partial x}{\partial v} \\ h\frac{\partial y}{\partial u} & k\frac{\partial y}{\partial v} \end{vmatrix} = hk \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}.$$

Since $\operatorname{area}(Q) = hk$, this means that the area of our region in the xy plane is given by the absolute value of the Jacobian times the area in the uv plane. Our shorthand for this is

$$dA = dx\,dy = \left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\,dv.$$

Areas are always positive, so the area of a small parallelogram in xy-space is always the **absolute value** of the Jacobian times the area of the corresponding rectangle in uv-space.

Example 2: *Compute* $\int_0^{10} e^{-x/5}\,dx$*.*

Solution: We did this before using $x = g(u) = 5u$. Instead, let's take $x = -5u$, so $g'(u) = -5$ is negative. Now $e^{-x/5} = e^u$ and $dx = -5du$. The map g sends the interval from 0 to -2 in u-space to the interval from 0 to 10 in x-space, and our change-of-variable formula says

$$\int_0^{10} e^{-x/5}\,dx \;=\; \int_0^{-2} -5e^u\,du.$$

Of course, we usually integrate from -2 to 0, not from 0 to -2.

Flipping the limits of integration changes the sign of the answer, so

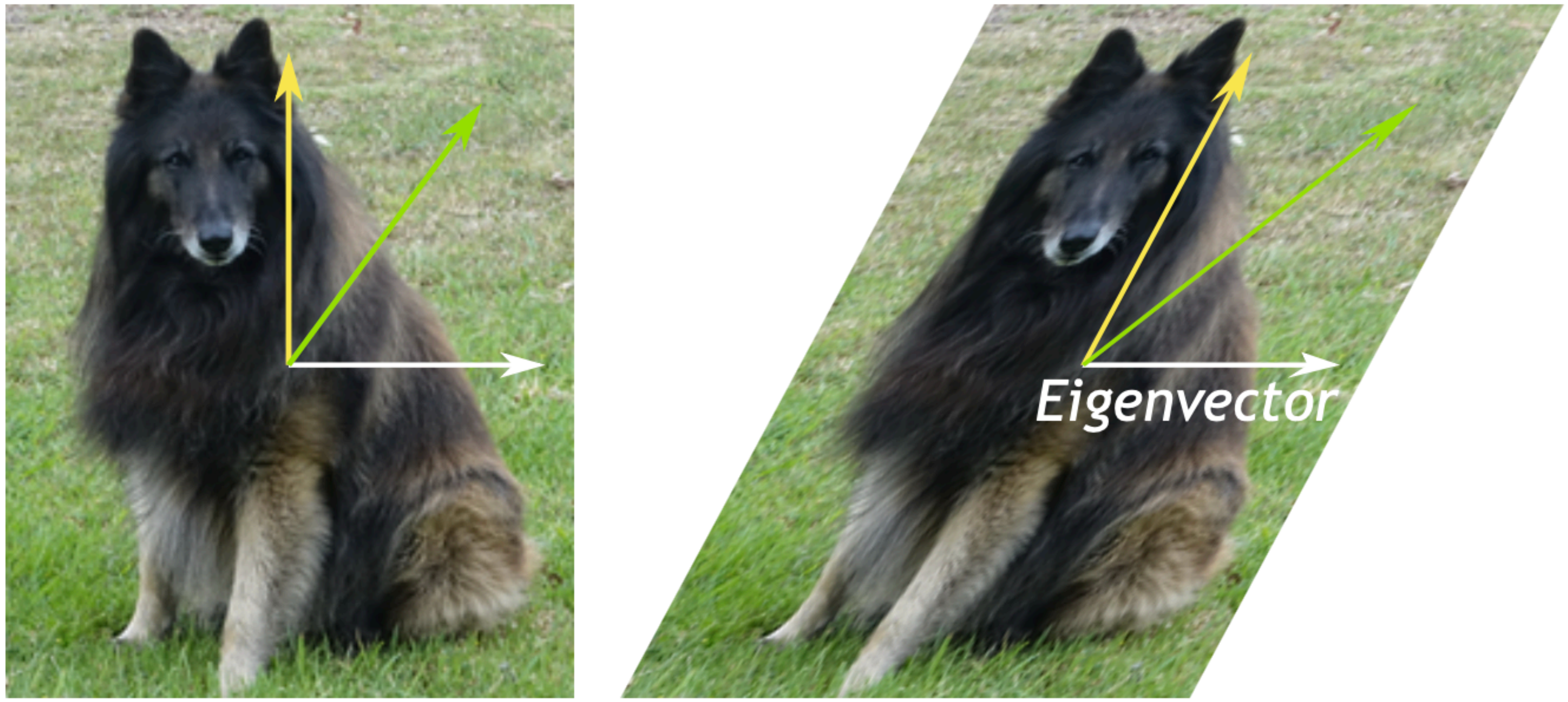
$$\int_0^{10} e^{-x/5}\,dx \;=\; \int_{-2}^{0} +5e^u\,du = 5(1 - e^{-2}).$$

If we had written our 1-dimensional integrals in terms of regions instead in terms of starting points and end points, we would have had a factor of $+5$, rather than -5, all along. The mapping $x = -5u$ sends the region $D^* = [-2, 0]$ to the region $D = [0, 10]$, and

$$\int_D e^{-x/5}\,dx \;=\; \int_{D^*} e^u \left|\frac{dx}{du}\right| du.$$

Eigenvector and Eigenvalue

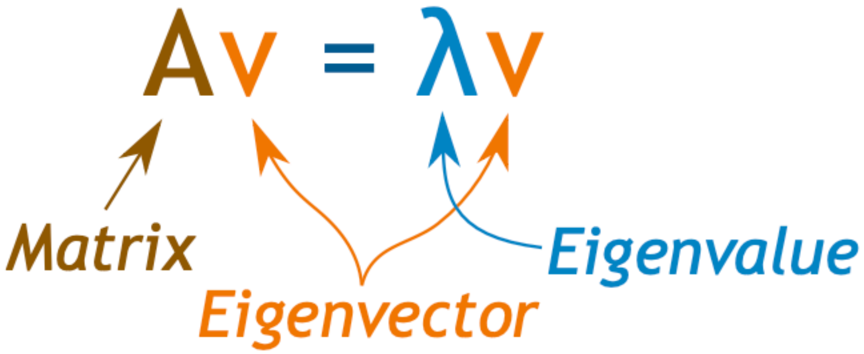
A simple example is that an eigenvector **does not change direction** in a transformation:



How do we find that vector?

The Mathematics Of It

For a square matrix **A**, an Eigenvector and Eigenvalue make this equation true:



$$Av = \lambda v$$

Let us see it in action:

Example: For this matrix

$$\begin{bmatrix} -6 & 3 \\ 4 & 5 \end{bmatrix}$$

an eigenvector is

$$\begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

with a matching eigenvalue of 6

Let's do some matrix multiplies to see if that is true.

Av gives us:

$$\begin{bmatrix} -6 & 3 \\ 4 & 5 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} -6\times1+3\times4 \\ 4\times1+5\times4 \end{bmatrix} = \begin{bmatrix} 6 \\ 24 \end{bmatrix}$$

λv gives us :

$$6\begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ 24 \end{bmatrix}$$

Yes they are equal!

So we get $Av = \lambda v$ as promised.

Notice how we multiply a **matrix** by a **vector** and get the same result as when we multiply a **scalar** (just a number) by that **vector**.

How do we find these eigen things?

We start by finding the **eigenvalue**. We know this equation must be true:

$$Av = \lambda v$$

Next we put in an identity matrix so we are dealing with matrix-vs-matrix:

$$Av = \lambda Iv$$

Bring all to left hand side:

$$Av - \lambda Iv = 0$$

If **v** is non-zero then we can (hopefully) solve for λ using just the determinant:

$$| A - \lambda I | = 0$$

Let's try that equation on our previous example:

Example: Solve for λ

Start with $| A - \lambda I | = 0$

$$\left| \begin{bmatrix} -6 & 3 \\ 4 & 5 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = 0$$

Which is:

$$\begin{vmatrix} -6-\lambda & 3 \\ 4 & 5-\lambda \end{vmatrix} = 0$$

Calculating that determinant gets:

$$(-6-\lambda)(5-\lambda) - 3\times4 = 0$$

Which simplifies to this Quadratic Equation:

$$\lambda^2 + \lambda - 42 = 0$$

And solving it gets:

$$\lambda = -7 \text{ or } 6$$

And yes, there are **two** possible eigenvalues.

Now we know **eigenvalues**, let us find their matching **eigenvectors**.

Example (continued): Find the Eigenvector for the Eigenvalue **λ = 6**:

Start with:

$$Av = \lambda v$$

Put in the values we know:

$$\begin{bmatrix} -6 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 6 \begin{bmatrix} x \\ y \end{bmatrix}$$

After multiplying we get these two equations:

$$\begin{aligned} -6x + 3y &= 6x \\ 4x + 5y &= 6y \end{aligned}$$

Bringing all to left hand side:

$$\begin{aligned} -12x + 3y &= 0 \\ 4x - 1y &= 0 \end{aligned}$$

Either equation reveals that **y = 4x**, so the **eigenvector** is any non-zero multiple of this:

$$\begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

And we get the solution shown at the top of the page:

$$\begin{bmatrix} -6 & 3 \\ 4 & 5 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} -6\times1+3\times4 \\ 4\times1+5\times4 \end{bmatrix} = \begin{bmatrix} 6 \\ 24 \end{bmatrix}$$

... and also ...

$$6\begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ 24 \end{bmatrix}$$

So $Av = \lambda v$, and we have success!

Now it is **your turn** to find the eigenvector for the other eigenvalue of -7

Why?

What is the purpose of these?

One of the cool things is we can use matrices to do transformations in space, which is used a lot in computer graphics.

In that case the eigenvector is "the direction that doesn't change direction" !

And the eigenvalue is the scale of the stretch:

- **1** means no change,
- **2** means doubling in length,
- **−1** means pointing backwards along the eigenvalue's direction
- etc

There are also many applications in physics, etc.

Why "Eigen"

> Eigen is a German word meaning "own" or "typical"
>
> *"das ist ihnen **eigen**"* is German for *"that is **typical** of them"*

Sometimes in English we use the word "characteristic", so an eigenvector can be called a "characteristic vector".

Not Just Two Dimensions

Eigenvectors work perfectly well in 3 and higher dimensions.

Example: find the eigenvalues for this 3x3 matrix:

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 5 \\ 0 & 4 & 3 \end{bmatrix}$$

First calculate $A - \lambda I$:

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 5 \\ 0 & 4 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2-\lambda & 0 & 0 \\ 0 & 4-\lambda & 5 \\ 0 & 4 & 3-\lambda \end{bmatrix}$$

Now the determinant should equal zero:

$$\begin{vmatrix} 2-\lambda & 0 & 0 \\ 0 & 4-\lambda & 5 \\ 0 & 4 & 3-\lambda \end{vmatrix} = 0$$

Which is:

$$(2-\lambda)\ [\ (4-\lambda)(3-\lambda) - 5\times4\ ] = 0$$

This ends up being a cubic equation, but just looking at it here we see one of the roots is **2** (because of $2-\lambda$), and the part inside the square brackets is Quadratic, with roots of **−1** and **8**.

Example (continued): find the Eigenvector that matches the Eigenvalue **−1**

Put in the values we know:

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 5 \\ 0 & 4 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = -1 \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

After multiplying we get these equations:

$$\begin{aligned} 2x &= -x \\ 4y + 5z &= -y \\ 4y + 3z &= -z \end{aligned}$$

Bringing all to left hand side:

$$\begin{aligned} 3x &= 0 \\ 5y + 5z &= 0 \\ 4y + 4z &= 0 \end{aligned}$$

So $\mathbf{x = 0}$, and $\mathbf{y = -z}$ and so the **eigenvector** is any non-zero multiple this:

$$\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

TEST Av:

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 5 \\ 0 & 4 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 4-5 \\ 4-3 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

And λv:

$$-1 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

So Av = λv, yay!

(You can try your hand at the eigenvalues of **2** and **8**)

Rotation

Back in the 2D world again, this matrix will do a rotation by θ:

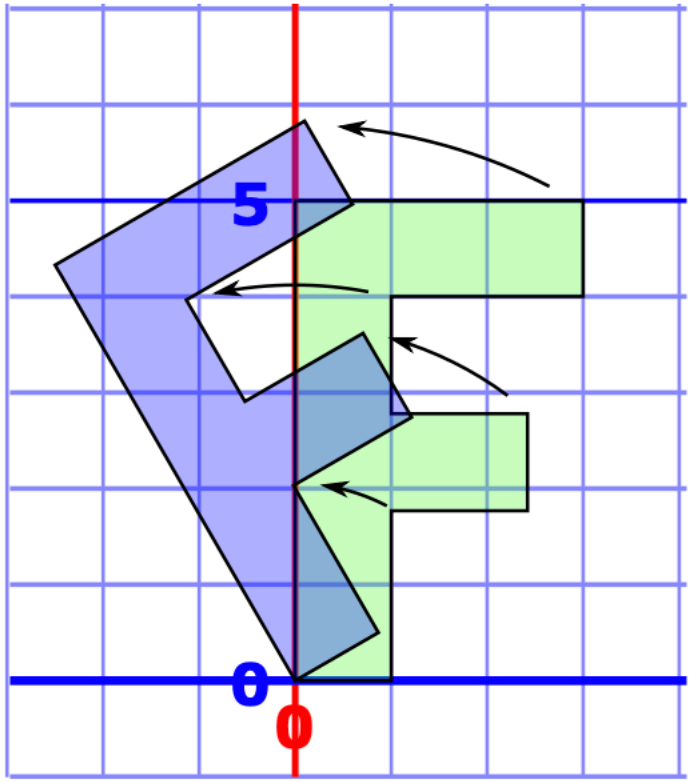
$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

Example: Rotate by 30°

$\cos(30°) = \frac{\sqrt{3}}{2}$ and $\sin(30°) = \frac{1}{2}$, so:

$$\begin{bmatrix} \cos(30°) & -\sin(30°) \\ \sin(30°) & \cos(30°) \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{-1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$$

But if we **rotate all points**, what is the "direction that doesn't change direction"?



Let us work through the mathematics to find out:

First calculate $A - \lambda I$:

$$\begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{-1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{2}-\lambda & \frac{-1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2}-\lambda \end{bmatrix}$$

Now the determinant should equal zero:

$$\begin{vmatrix} \frac{\sqrt{3}}{2}-\lambda & \frac{-1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2}-\lambda \end{vmatrix} = 0$$

Which is:

$$(\frac{\sqrt{3}}{2}-\lambda)(\frac{\sqrt{3}}{2}-\lambda) - (\frac{-1}{2})(\frac{1}{2}) = 0$$

Which becomes this Quadratic Equation:

$$\lambda^2 - (\sqrt{3})\lambda + 1 = 0$$

Whose roots are:

$$\lambda = \frac{\sqrt{3}}{2} \pm \frac{i}{2}$$

The eigenvalues are complex!

I don't know how to show you that on a graph, but we still get a solution.

Eigenvector

So, what is an eigenvector that matches, say, the $\frac{\sqrt{3}}{2} + \frac{i}{2}$ root?

Start with:

$$Av = \lambda v$$

Put in the values we know:

$$\begin{vmatrix} \frac{\sqrt{3}}{2} & \frac{-1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{vmatrix} \begin{vmatrix} x \\ y \end{vmatrix} = \left(\frac{\sqrt{3}}{2} + \frac{i}{2}\right) \begin{vmatrix} x \\ y \end{vmatrix}$$

After multiplying we get these two equations:

$$\frac{\sqrt{3}}{2}x - \frac{1}{2}y = \frac{\sqrt{3}}{2}x + \frac{i}{2}x$$

$$\frac{1}{2}x + \frac{\sqrt{3}}{2}y = \frac{\sqrt{3}}{2}y + \frac{i}{2}y$$

Which simplify to:

$$-y = ix$$

$$x = iy$$

And the solution is any non-zero multiple of:

$$\begin{bmatrix} i \\ 1 \end{bmatrix}$$

or

$$\begin{bmatrix} -i \\ 1 \end{bmatrix}$$

Wow, such a simple answer!

Is this just because we chose 30°? Or does it work for any rotation matrix? I will let you work that out! Try another angle, or better still use "cos(θ)" and "sin(θ)".

Oh, and let us **check** at least one of those solutions:

$$\begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{-1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} i\frac{\sqrt{3}}{2} - \frac{1}{2} \\ \frac{i}{2} + \frac{\sqrt{3}}{2} \end{bmatrix}$$

Does it match this?

$$(\frac{\sqrt{3}}{2} + \frac{i}{2}) \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} i\frac{\sqrt{3}}{2} - \frac{1}{2} \\ \frac{\sqrt{3}}{2} + \frac{i}{2} \end{bmatrix}$$

Oh yes it does!

Solving Systems of Linear Equations Using Matrices

Example: Solve

- $x + y + z = 6$
- $2y + 5z = -4$
- $2x + 5y - z = 27$

But first we need to write the question in Matrix form.

In Matrix Form?

OK. A Matrix is an array of numbers, right?

$$\begin{bmatrix} 6 & 4 & 24 \\ 1 & -9 & 8 \end{bmatrix}$$

A Matrix

Well, think about the equations:

$$\begin{array}{ccccccc} x & + & y & + & z & = & 6 \\ & & 2y & + & 5z & = & -4 \\ 2x & + & 5y & - & z & = & 27 \end{array}$$

They could be turned into a table of numbers like this:

$$\begin{array}{ccccc} 1 & 1 & 1 & = & 6 \\ 0 & 2 & 5 & = & -4 \\ 2 & 5 & -1 & = & 27 \end{array}$$

We could even separate the numbers before and after the "=" into:

$$\begin{array}{ccc} 1 & 1 & 1 \\ 0 & 2 & 5 \\ 2 & 5 & -1 \end{array} \quad \text{and} \quad \begin{array}{c} 6 \\ -4 \\ 27 \end{array}$$

Now it looks like we have 2 Matrices.

In fact we have a third one, which is [x y z]:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 5 \\ 2 & 5 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ -4 \\ 27 \end{bmatrix}$$

Why does [x y z] go there? Because when we Multiply Matrices the left side becomes:

"Dot Product"

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 5 \\ 2 & 5 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + y + z \\ 2y + 5z \\ 2x + 5y - z \end{bmatrix}$$

Which is the original left side of our equations above (you might like to check that).

The Matrix Solution

We can write this:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 5 \\ 2 & 5 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ -4 \\ 27 \end{bmatrix}$$

like this:

$$AX = B$$

where

- A is the 3x3 matrix of x, y and z **coefficients**
- X is **x, y and z**, and
- B is **6, −4 and 27**

Then (as shown on the Inverse of a Matrix page) the solution is this:

$$X = A^{-1}B$$

What does that mean?

It means that we can find the values of x, y and z (the X matrix) by multiplying the **inverse of the A matrix** by the **B matrix**.

So let's go ahead and do that.

First, we need to find the **inverse of the A matrix** (assuming it exists!)

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 5 \\ 2 & 5 & -1 \end{bmatrix}^{-1} = \frac{1}{-21}\begin{bmatrix} -27 & 6 & 3 \\ 10 & -3 & -5 \\ -4 & -3 & 2 \end{bmatrix}$$

Then multiply A^{-1} by B (we can use the Matrix Calculator again):

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{-21}\begin{bmatrix} -27 & 6 & 3 \\ 10 & -3 & -5 \\ -4 & -3 & 2 \end{bmatrix}\begin{bmatrix} 6 \\ -4 \\ 27 \end{bmatrix} = \frac{1}{-21}\begin{bmatrix} -105 \\ -63 \\ 42 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix}$$

And we are done! The solution is:

$x = 5$,
$y = 3$,
$z = -2$

See Linear Algebra notes next for complete discussion.

Basic Linear Algebra

2.1 Matrices and Vectors

Matrices

DEFINITION ■ A **matrix** is any rectangular array of numbers. ■

For example,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} 2 & 1 \end{bmatrix}$$

are all matrices.

If a matrix A has m rows and n columns, we call A an $m \times n$ matrix. We refer to $m \times n$ as the **order** of the matrix. A typical $m \times n$ matrix A may be written as

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

DEFINITION ■ The number in the *i*th row and *j*th column of A is called the ***ij*th element** of A and is written a_{ij}. ■

For example, if

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

then $a_{11} = 1$, $a_{23} = 6$, and $a_{31} = 7$.

Sometimes we will use the notation $A = [a_{ij}]$ to indicate that A is the matrix whose *ij*th element is a_{ij}.

DEFINITION ■ Two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are **equal** if and only if A and B are of the same order and for all i and j, $a_{ij} = b_{ij}$. ■

For example, if

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} x & y \\ w & z \end{bmatrix}$$

then $A = B$ if and only if $x = 1$, $y = 2$, $w = 3$, and $z = 4$.

Vectors

Any matrix with only one column (that is, any $m \times 1$ matrix) may be thought of as a **column vector.** The number of rows in a column vector is the **dimension** of the column vector. Thus,

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

may be thought of as a 2×1 matrix or a two-dimensional column vector. R^m will denote the set of all m-dimensional column vectors.

In analogous fashion, we can think of any vector with only one row (a $1 \times n$ matrix as a **row vector.** The dimension of a row vector is the number of columns in the vector. Thus, $[9 \quad 2 \quad 3]$ may be viewed as a 1×3 matrix or a three-dimensional row vector. In this book, vectors appear in boldface type: for instance, vector **v.** An m-dimensional vector (either row or column) in which all elements equal zero is called a **zero vector** (written **0**). Thus,

$$[0 \quad 0] \qquad \text{and} \qquad \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

are two-dimensional zero vectors.

Any m-dimensional vector corresponds to a directed line segment in the m-dimensional plane. For example, in the two-dimensional plane, the vector

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

corresponds to the line segment joining the point

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

to the point

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

The directed line segments corresponding to

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$$

are drawn in Figure 1.

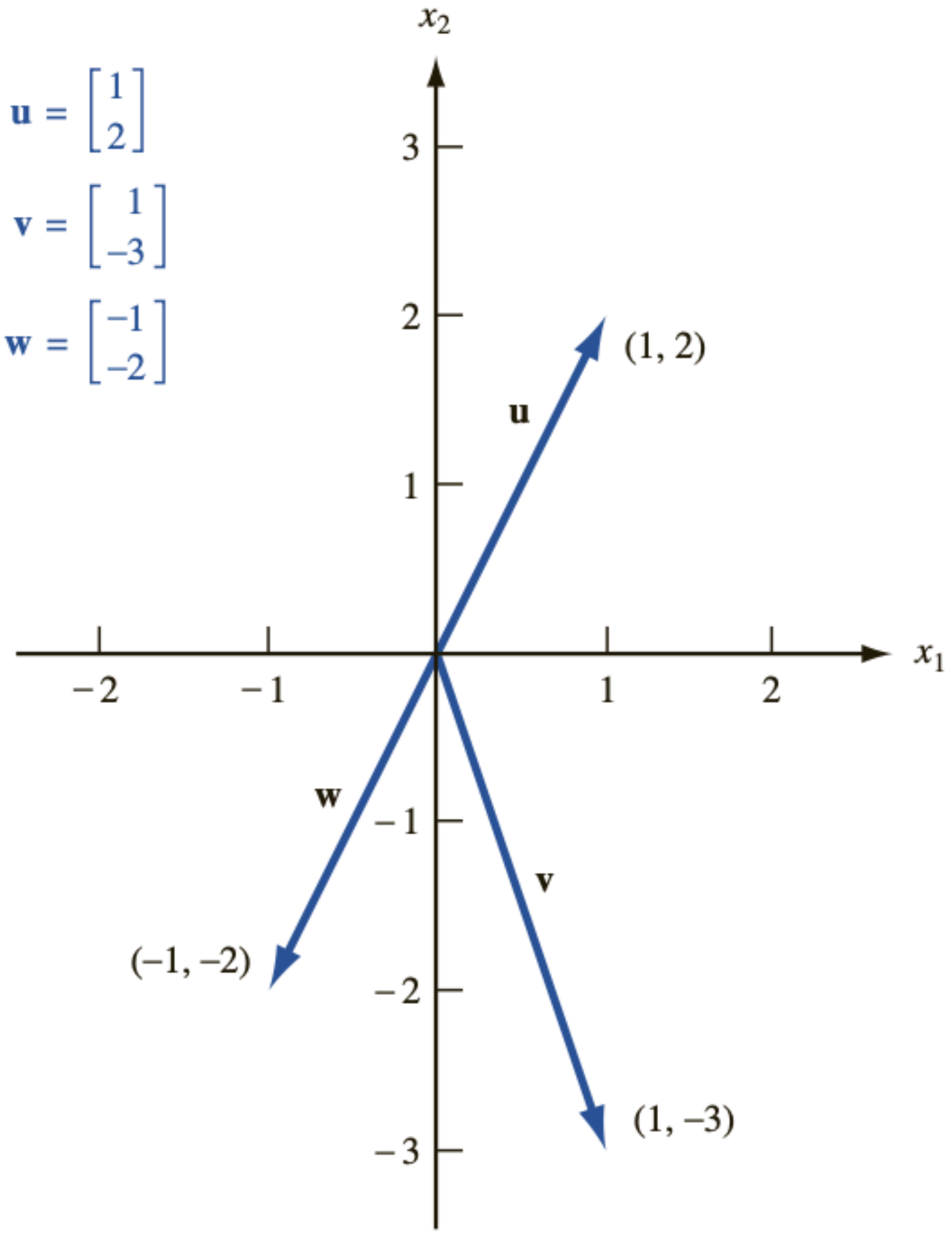


FIGURE 1
Vectors Are Directed Line Segments

The Scalar Product of Two Vectors

An important result of multiplying two vectors is the *scalar product.* To define the scalar product of two vectors, suppose we have a row vector $\mathbf{u} = [u_1 \quad u_2 \quad \cdots \quad u_n]$ and a column vector

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

of the same dimension. The **scalar product** of $\mathbf{u}$ and $\mathbf{v}$ (written $\mathbf{u} \cdot \mathbf{v}$) is the number $u_1v_1 + u_2v_2 + \cdots + u_nv_n$.

For the scalar product of two vectors to be defined, the first vector must be a row vector and the second vector must be a column vector. For example, if

$$\mathbf{u} = [1 \quad 2 \quad 3] \qquad \text{and} \qquad \mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$$

then $\mathbf{u} \cdot \mathbf{v} = 1(2) + 2(1) + 3(2) = 10$. By these rules for computing a scalar product, if

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \text{and} \qquad \mathbf{v} = [2 \quad 3]$$

then $\mathbf{u} \cdot \mathbf{v}$ is not defined. Also, if

$$\mathbf{u} = [1 \quad 2 \quad 3] \qquad \text{and} \qquad \mathbf{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

then $\mathbf{u} \cdot \mathbf{v}$ is not defined because the vectors are of two different dimensions.

Note that two vectors are perpendicular if and only if their scalar product equals 0. Thus, the vectors $[1 \quad -1]$ and $[1 \quad 1]$ are perpendicular.

We note that $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos \theta$, where $\|\mathbf{u}\|$ is the length of the vector $\mathbf{u}$ and θ is the angle between the vectors $\mathbf{u}$ and $\mathbf{v}$.

Matrix Operations

We now describe the arithmetic operations on matrices that are used later in this book.

The Scalar Multiple of a Matrix

Given any matrix A and any number c (a *number* is sometimes referred to as a *scalar*), the matrix cA is obtained from the matrix A by multiplying each element of A by c. For example,

$$\text{if} \qquad A = \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}, \qquad \text{then} \qquad 3A = \begin{bmatrix} 3 & 6 \\ -3 & 0 \end{bmatrix}$$

For $c = -1$, scalar multiplication of the matrix A is sometimes written as $-A$.

Addition of Two Matrices

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be two matrices with the same order (say, $m \times n$). Then the matrix $C = A + B$ is defined to be the $m \times n$ matrix whose ijth element is $a_{ij} + b_{ij}$. Thus, to obtain the sum of two matrices A and B, we add the corresponding elements of A and B. For example, if

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 1 \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} -1 & -2 & -3 \\ 2 & 1 & -1 \end{bmatrix}$$

then

$$A + B = \begin{bmatrix} 1-1 & 2-2 & 3-3 \\ 0+2 & -1+1 & 1-1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}.$$

This rule for matrix addition may be used to add vectors of the same dimension. For example, if $\mathbf{u} = [1 \quad 2]$ and $\mathbf{v} = [2 \quad 1]$, then $\mathbf{u} + \mathbf{v} = [1 + 2 \quad 2 + 1] = [3 \quad 3]$. Vectors may be added geometrically by the parallelogram law (see Figure 2).

We can use scalar multiplication and the addition of matrices to define the concept of a line segment. A glance at Figure 1 should convince you that any point u in the m-dimensional plane corresponds to the m-dimensional vector $\mathbf{u}$ formed by joining the origin to the point u. For any two points u and v in the m-dimensional plane, the **line segment** joining u and v (called the line segment uv) is the set of all points in the m-dimensional plane that correspond to the vectors $c\mathbf{u} + (1 - c)\mathbf{v}$, where $0 \leq c \leq 1$ (Figure 3). For example, if $u = (1, 2)$ and $v = (2, 1)$, then the line segment uv consists of the points corresponding to the vectors $c[1 \quad 2] + (1 - c)[2 \quad 1] = [2 - c \quad 1 + c]$, where $0 \leq c \leq 1$. For $c = 0$ and $c = 1$, we obtain the endpoints of the line segment uv; for $c = \frac{1}{2}$, we obtain the midpoint $(0.5\mathbf{u} + 0.5\mathbf{v})$ of the line segment uv.

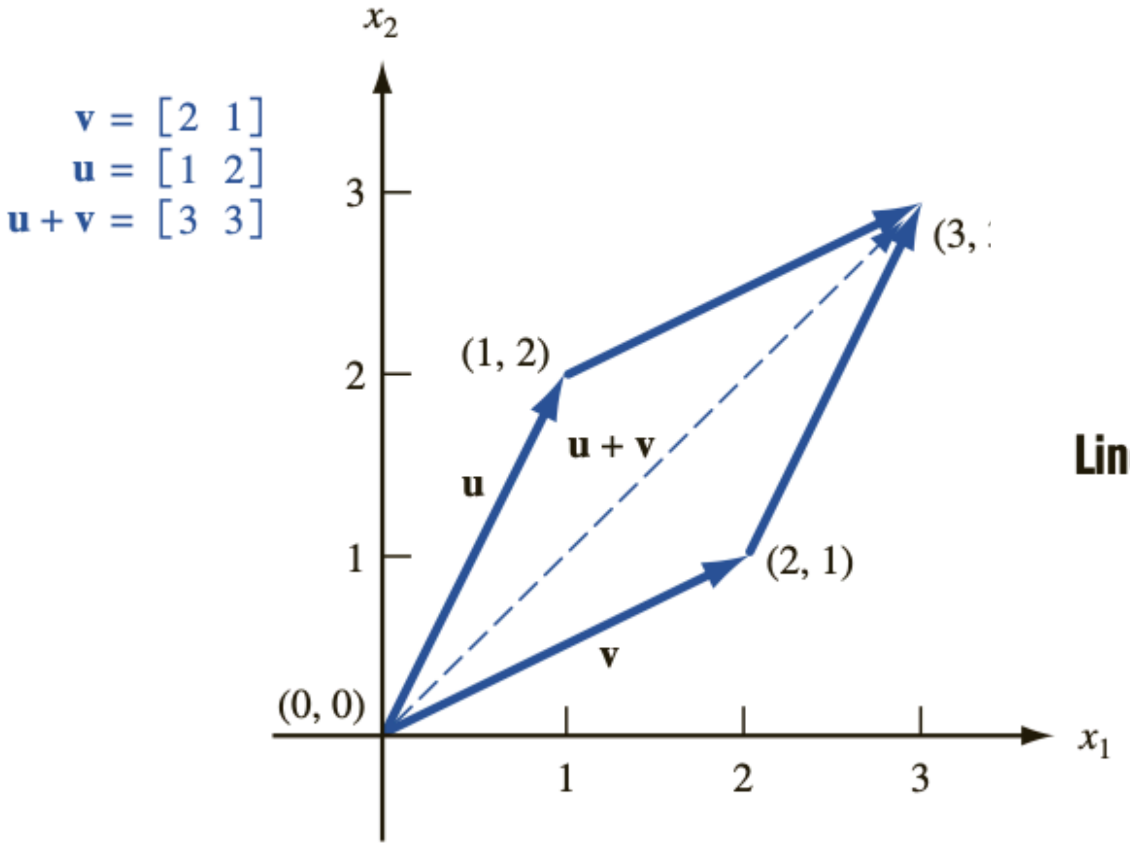


FIGURE 2
Addition of Vectors

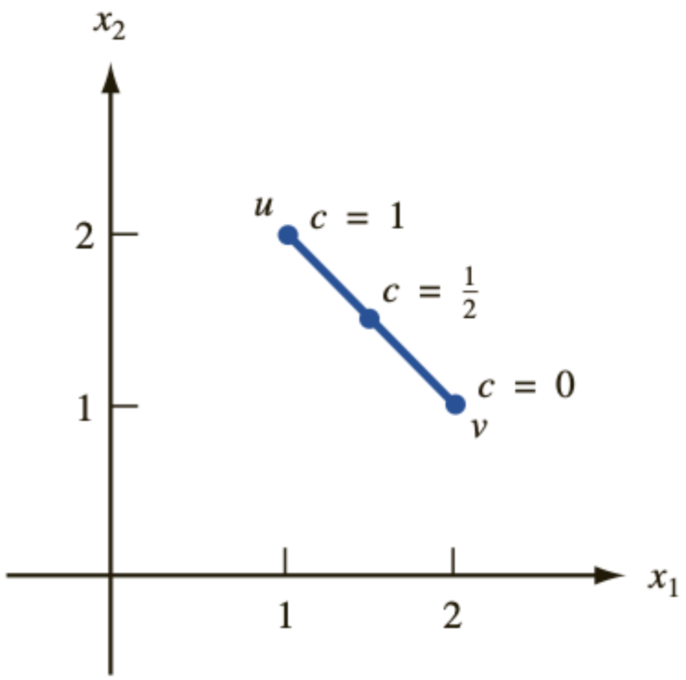


FIGURE 3
Line Segment Joining
$u = (1, 2)$ and
$v = (2, 1)$

Using the parallelogram law, the line segment *uv* may also be viewed as the points corresponding to the vectors $\mathbf{u} + c(\mathbf{v} - \mathbf{u})$, where $0 \leq c \leq 1$ (Figure 4). Observe that for $c = 0$, we obtain the vector $\mathbf{u}$ (corresponding to point u), and for $c = 1$, we obtain the vector $\mathbf{v}$ (corresponding to point v).

The Transpose of a Matrix

Given any $m \times n$ matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

the **transpose** of A (written A^T) is the $n \times m$ matrix

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

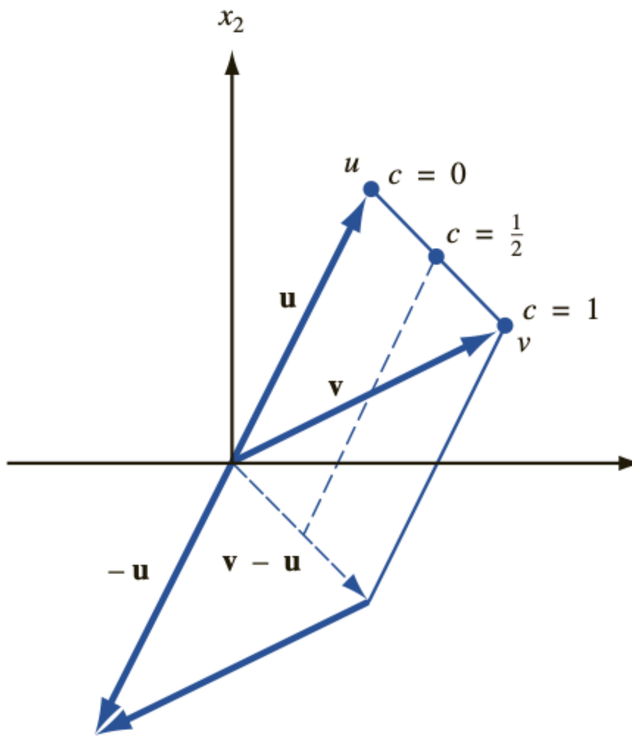


FIGURE 4
Representation of Line Segment *uv*

Thus, A^T is obtained from A by letting row 1 of A be column 1 of A^T, letting row 2 of A be column 2 of A^T, and so on. For example,

$$\text{if} \qquad A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \qquad \text{then} \qquad A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Observe that $(A^T)^T = A$. Let $B = [1 \quad 2]$; then

$$B^T = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \text{and} \qquad (B^T)^T = [1 \quad 2] = B$$

As indicated by these two examples, for any matrix A, $(A^T)^T = A$.

Matrix Multiplication

Given two matrices A and B, the matrix product of A and B (written AB) is defined if and only if

$$\text{Number of columns in } A = \text{number of rows in } B \tag{1}$$

For the moment, assume that for some positive integer r, A has r columns and B has r rows. Then for some m and n, A is an $m \times r$ matrix and B is an $r \times n$ matrix.

DEFINITION ■ The **matrix product** $C = AB$ of A and B is the $m \times n$ matrix C whose ijth element is determined as follows:

$$ij\text{th element of } C = \text{scalar product of row } i \text{ of } A \times \text{column } j \text{ of } B \quad ■ \qquad (2)$$

If Equation (1) is satisfied, then each row of A and each column of B will have the same number of elements. Also, if (1) is satisfied, then the scalar product in Equation (2) will be defined. The product matrix $C = AB$ will have the same number of rows as A and the same number of columns as B.

EXAMPLE 1 Matrix Multiplication

Compute $C = AB$ for

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 1 & 2 \end{bmatrix}$$

Solution Because A is a 2×3 matrix and B is a 3×2 matrix, AB is defined, and C will be a 2×2 matrix. From Equation (2),

$$c_{11} = [1 \quad 1 \quad 2] \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = 1(1) + 1(2) + 2(1) = 5$$

$$c_{12} = [1 \quad 1 \quad 2] \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = 1(1) + 1(3) + 2(2) = 8$$

$$c_{21} = [2 \quad 1 \quad 3] \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = 2(1) + 1(2) + 3(1) = 7$$

$$c_{22} = \begin{bmatrix} 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = 2(1) + 1(3) + 3(2) = 11$$

$$C = AB = \begin{bmatrix} 5 & 8 \\ 7 & 11 \end{bmatrix}$$

EXAMPLE 2 Column Vector Times Row Vector

Find AB for

$$A = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 \end{bmatrix}$$

Solution Because A has one column and B has one row, $C = AB$ will exist. From Equation (2), we know that C is a 2×2 matrix with

$$c_{11} = 3(1) = 3 \qquad c_{21} = 4(1) = 4$$
$$c_{12} = 3(2) = 6 \qquad c_{22} = 4(2) = 8$$

Thus,

$$C = \begin{bmatrix} 3 & 6 \\ 4 & 8 \end{bmatrix}$$

EXAMPLE 3 Row Vector Times Column Vector

Compute $D = BA$ for the A and B of Example 2.

Solution In this case, D will be a 1×1 matrix (or a scalar). From Equation (2),

$$d_{11} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = 1(3) + 2(4) = 11$$

Thus, $D = [11]$. In this example, matrix multiplication is equivalent to scalar multiplication of a row and column vector.

Recall that if you multiply two real numbers a and b, then $ab = ba$. This is called the *commutative property of multiplication.* Examples 2 and 3 show that for matrix multiplication, it may be that $AB \neq BA$. Matrix multiplication is not necessarily commutative. (In some cases, however, $AB = BA$ will hold.)

EXAMPLE 4 Undefined Matrix Product

Show that AB is undefined if

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix}$$

Solution This follows because A has two columns and B has three rows. Thus, Equation (1) is not satisfied.

TABLE 1

Gallons of Crude Oil Required to Produce 1 Gallon of Gasoline

Crude Oil	Premium Unleaded	Regular Unleaded	Regular Leaded
1	$\frac{3}{4}$	$\frac{2}{3}$	$\frac{1}{4}$
2	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{3}{4}$

Many computations that commonly occur in operations research (and other branches of mathematics) can be concisely expressed by using matrix multiplication. To illustrate this, suppose an oil company manufactures three types of gasoline: premium unleaded, regular unleaded, and regular leaded. These gasolines are produced by mixing two types of crude oil: crude oil 1 and crude oil 2. The number of gallons of crude oil required to manufacture 1 gallon of gasoline is given in Table 1.

From this information, we can find the amount of each type of crude oil needed to manufacture a given amount of gasoline. For example, if the company wants to produce 10 gallons of premium unleaded, 6 gallons of regular unleaded, and 5 gallons of regular leaded, then the company's crude oil requirements would be

$$\text{Crude 1 required} = (\tfrac{3}{4})\,(10) + (\tfrac{2}{3})\,(6) + (\tfrac{1}{4})\,5 = 12.75 \text{ gallons}$$

$$\text{Crude 2 required} = (\tfrac{1}{4})\,(10) + (\tfrac{1}{3})\,(6) + (\tfrac{3}{4})\,5 = 8.25 \text{ gallons}$$

More generally, we define

p_U = gallons of premium unleaded produced

r_U = gallons of regular unleaded produced

r_L = gallons of regular leaded produced

c_1 = gallons of crude 1 required

c_2 = gallons of crude 2 required

Then the relationship between these variables may be expressed by

$$\begin{aligned} c_1 &= \left(\tfrac{3}{4}\right) p_U + \left(\tfrac{2}{3}\right) r_U + \left(\tfrac{1}{4}\right) r_L \\ c_2 &= \left(\tfrac{1}{4}\right) p_U + \left(\tfrac{1}{3}\right) r_U + \left(\tfrac{3}{4}\right) r_L \end{aligned}$$

Using matrix multiplication, these relationships may be expressed by

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} & \frac{2}{3} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{3} & \frac{3}{4} \end{bmatrix} \begin{bmatrix} p_U \\ r_U \\ r_L \end{bmatrix}$$

Properties of Matrix Multiplication

To close this section, we discuss some important properties of matrix multiplication. In what follows, we assume that all matrix products are defined.

1 Row i of AB = (row i of A)B. To illustrate this property, let

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 1 & 2 \end{bmatrix}$$

Then row 2 of the 2 × 2 matrix AB is equal to

$$\begin{bmatrix} 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 7 & 11 \end{bmatrix}$$

This answer agrees with Example 1.

2 Column j of $AB = A$(column j of B). Thus, for A and B as given, the first column of AB is

$$\begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

Properties 1 and 2 are helpful when you need to compute only *part* of the matrix AB.

3 Matrix multiplication is associative. That is, $A(BC) = (AB)C$. To illustrate, let

$$A = [1 \quad 2], \qquad B = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}, \qquad C = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Then $AB = [10 \quad 13]$ and $(AB)C = 10(2) + 13(1) = [33]$.

On the other hand,

$$BC = \begin{bmatrix} 7 \\ 13 \end{bmatrix}$$

so $A(BC) = 1(7) + 2(13) = [33]$. In this case, $A(BC) = (AB)C$ does hold.

4 Matrix multiplication is distributive. That is, $A(B + C) = AB + AC$ and $(B + C)D = BD + CD$.

2.2 Matrices and Systems of Linear Equations

Consider a system of linear equations given by

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array} \tag{3}$$

In Equation (3), $x_1, x_2, \ldots, x_n$ are referred to as **variables,** or unknowns, and the a_{ij}'s and b_i's are **constants.** A set of equations such as (3) is called a linear system of m equations in n variables.

DEFINITION ■ A **solution** to a linear system of m equations in n unknowns is a set of values for the unknowns that satisfies each of the system's m equations. ■

To understand linear programming, we need to know a great deal about the properties of solutions to linear equation systems. With this in mind, we will devote much effort to studying such systems.

We denote a possible solution to Equation (3) by an n-dimensional column vector **x,** in which the ith element of **x** is the value of x_i. The following example illustrates the concept of a solution to a linear system.

EXAMPLE 5 Solution to Linear System

Show that

$$\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

is a solution to the linear system

$$\begin{aligned} x_1 + 2x_2 &= 5 \\ 2x_1 - x_2 &= 0 \end{aligned} \tag{4}$$

and that

$$\mathbf{x} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

is not a solution to linear system (4).

Solution To show that

$$\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

is a solution to Equation (4), we substitute $x_1 = 1$ and $x_2 = 2$ in both equations and check that they are satisfied: $1 + 2(2) = 5$ and $2(1) - 2 = 0$.

The vector

$$\mathbf{x} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

is not a solution to (4), because $x_1 = 3$ and $x_2 = 1$ fail to satisfy $2x_1 - x_2 = 0$.

Using matrices can greatly simplify the statement and solution of a system of linear equations. To show how matrices can be used to compactly represent Equation (3), let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Then (3) may be written as

$$A\mathbf{x} = \mathbf{b} \tag{5}$$

Observe that both sides of Equation (5) will be $m \times 1$ matrices (or $m \times 1$ column vectors). For the matrix $A\mathbf{x}$ to equal the matrix $\mathbf{b}$ (or for the vector $A\mathbf{x}$ to equal the vector $\mathbf{b}$), their corresponding elements must be equal. The first element of $A\mathbf{x}$ is the scalar product of row 1 of A with $\mathbf{x}$. This may be written as

$$[a_{11} \quad a_{12} \quad \cdots \quad a_{1n}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n$$

This must equal the first element of $\mathbf{b}$ (which is b_1). Thus, (5) implies that $a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$. This is the first equation of (3). Similarly, (5) implies that the scalar product of row i of A with $\mathbf{x}$ must equal b_i, and this is just the ith equation of (3). Our discussion shows that (3) and (5) are two different ways of writing the same linear system. We call (5) the **matrix representation** of (3). For example, the matrix representation of (4) is

$$\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$$

Sometimes we abbreviate (5) by writing

$$A|\mathbf{b} \tag{6}$$

If A is an $m \times n$ matrix, it is assumed that the variables in (6) are $x_1, x_2, \ldots, x_n$. Then (6) is still another representation of (3). For instance, the matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 2 \\ 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \end{array}\right]$$

represents the system of equations

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 2 \\ x_2 + 2x_3 &= 3 \\ x_1 + x_2 + x_3 &= 1 \end{aligned}$$

2.3 The Gauss–Jordan Method for Solving Systems of Linear Equations

We develop in this section an efficient method (the Gauss–Jordan method) for solving a system of linear equations. Using the Gauss–Jordan method, we show that any system of linear equations must satisfy one of the following three cases:

Case 1 The system has no solution.

Case 2 The system has a unique solution.

Case 3 The system has an infinite number of solutions.

The Gauss–Jordan method is also important because many of the manipulations used in this method are used when solving linear programming problems by the simplex algorithm (see Chapter 4).

Elementary Row Operations

Before studying the Gauss–Jordan method, we need to define the concept of an **elementary row operation** (ERO). An ERO transforms a given matrix A into a new matrix A' via one of the following operations.

Type 1 ERO

A' is obtained by multiplying any row of A by a nonzero scalar. For example, if

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 5 & 6 \\ 0 & 1 & 2 & 3 \end{bmatrix}$$

then a Type 1 ERO that multiplies row 2 of A by 3 would yield

$$A' = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 9 & 15 & 18 \\ 0 & 1 & 2 & 3 \end{bmatrix}$$

Type 2 ERO

Begin by multiplying any row of A (say, row i) by a nonzero scalar c. For some $j \neq i$, let row j of $A' = c$(row i of A) + row j of A, and let the other rows of A' be the same as the rows of A.

For example, we might multiply row 2 of A by 4 and replace row 3 of A by 4(row 2 of A) + row 3 of A. Then row 3 of A' becomes

$$4\,[1 \quad 3 \quad 5 \quad 6] + [0 \quad 1 \quad 2 \quad 3] = [4 \quad 13 \quad 22 \quad 27]$$

and

$$A' = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 5 & 6 \\ 4 & 13 & 22 & 27 \end{bmatrix}$$

Type 3 ERO

Interchange any two rows of A. For instance, if we interchange rows 1 and 3 of A, we obtain

$$A' = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 3 & 5 & 6 \\ 1 & 2 & 3 & 4 \end{bmatrix}$$

Type 1 and Type 2 EROs formalize the operations used to solve a linear equation system. To solve the system of equations

$$\begin{aligned} x_1 + x_2 &= 2 \\ 2x_1 + 4x_2 &= 7 \end{aligned} \tag{7}$$

we might proceed as follows. First replace the second equation in (7) by -2(first equation in (7)) + second equation in (7). This yields the following linear system:

$$\begin{aligned} x_1 + x_2 &= 2 \\ 2x_2 &= 3 \end{aligned} \tag{7.1}$$

Then multiply the second equation in (7.1) by $\frac{1}{2}$, yielding the system

$$\begin{aligned} x_1 + x_2 &= 2 \\ x_2 &= \tfrac{3}{2} \end{aligned} \tag{7.2}$$

Finally, replace the first equation in (7.2) by -1[second equation in (7.2)] + first equation in (7.2). This yields the system

$$\begin{aligned} x_1 \quad &= \tfrac{1}{2} \\ x_2 &= \tfrac{3}{2} \end{aligned} \tag{7.3}$$

System (7.3) has the unique solution $x_1 = \frac{1}{2}$ and $x_2 = \frac{3}{2}$. The systems (7), (7.1), (7.2), and (7.3) are *equivalent* in that they have the same set of solutions. This means that $x_1 = \frac{1}{2}$ and $x_2 = \frac{3}{2}$ is also the unique solution to the original system, (7).

If we view (7) in the augmented matrix form $(A|\mathbf{b})$, we see that the steps used to solve (7) may be seen as Type 1 and Type 2 EROs applied to $A|\mathbf{b}$. Begin with the augmented matrix version of (7):

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 2 & 4 & 7 \end{array}\right] \tag{7$'$}$$

Now perform a Type 2 ERO by replacing row 2 of (7′) by −2(row 1 of (7′)) + row 2 of (7′). The result is

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 2 & 3 \end{array}\right] \tag{7.1$'$}$$

which corresponds to (7.1). Next, we multiply row 2 of (7.1′) by $\frac{1}{2}$ (a Type 1 ERO), resulting in

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & \frac{3}{2} \end{array}\right] \tag{7.2$'$}$$

which corresponds to (7.2). Finally, perform a Type 2 ERO by replacing row 1 of (7.2′) by −1(row 2 of (7.2′)) + row 1 of (7.2′). The result is

$$\left[\begin{array}{cc|c} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{3}{2} \end{array}\right] \tag{7.3$'$}$$

which corresponds to (7.3). Translating (7.3′) back into a linear system, we obtain the system $x_1 = \frac{1}{2}$ and $x_2 = \frac{3}{2}$, which is identical to (7.3).

Finding a Solution by the Gauss–Jordan Method

The discussion in the previous section indicates that if the matrix $A'|\mathbf{b}'$ is obtained from $A|\mathbf{b}$ via an ERO, the systems $A\mathbf{x} = \mathbf{b}$ and $A'\mathbf{x} = \mathbf{b}'$ are equivalent. Thus, any sequence of EROs performed on the augmented matrix $A|\mathbf{b}$ corresponding to the system $A\mathbf{x} = \mathbf{b}$ will yield an equivalent linear system.

The Gauss–Jordan method solves a linear equation system by utilizing EROs in a systematic fashion. We illustrate the method by finding the solution to the following linear system:

$$\begin{aligned} 2x_1 + 2x_2 + x_3 &= 9 \\ 2x_1 - x_2 + 2x_3 &= 6 \\ x_1 - x_2 + 2x_3 &= 5 \end{aligned} \tag{8}$$

The augmented matrix representation is

$$A|\mathbf{b} = \left[\begin{array}{ccc|c} 2 & 2 & 1 & 9 \\ 2 & -1 & 2 & 6 \\ 1 & -1 & 2 & 5 \end{array}\right] \tag{8'}$$

Suppose that by performing a sequence of EROs on (8′) we could transform (8′) into

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right] \tag{9'}$$

We note that the result obtained by performing an ERO on a system of equations can also be obtained by multiplying both sides of the matrix representation of the system of equations by a particular matrix. This explains why EROs do not change the set of solutions to a system of equations.

Matrix (9′) corresponds to the following linear system:

$$\begin{aligned} x_1 \qquad &= 1 \\ x_2 \quad &= 2 \\ x_3 &= 3 \end{aligned} \tag{9}$$

System (9) has the unique solution $x_1 = 1$, $x_2 = 2$, $x_3 = 3$. Because (9′) was obtained from (8′) by a sequence of EROs, we know that (8) and (9) are equivalent linear systems. Thus, $x_1 = 1$, $x_2 = 2$, $x_3 = 3$ must also be the unique solution to (8). We now show how we can use EROs to transform a relatively complicated system such as (8) into a relatively simple system like (9). This is the essence of the Gauss–Jordan method.

We begin by using EROs to transform the first column of (8′) into

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Then we use EROs to transform the second column of the resulting matrix into

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Finally, we use EROs to transform the third column of the resulting matrix into

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

As a final result, we will have obtained (9′). We now use the Gauss–Jordan method to solve (8). We begin by using a Type 1 ERO to change the element of (8′) in the first row and first column into a 1. Then we add multiples of row 1 to row 2 and then to row 3 (these are Type 2 EROs). The purpose of these Type 2 EROs is to put zeros in the rest of the first column. The following sequence of EROs will accomplish these goals.

Step 1 Multiply row 1 of (8′) by $\frac{1}{2}$. This Type 1 ERO yields

$$A_1|\mathbf{b}_1 = \left[\begin{array}{rrr|r} 1 & 1 & \frac{1}{2} & \frac{9}{2} \\ 2 & -1 & 2 & 6 \\ 1 & -1 & 2 & 5 \end{array}\right]$$

Step 2 Replace row 2 of $A_1|\mathbf{b}_1$ by -2(row 1 of $A_1|\mathbf{b}_1$) + row 2 of $A_1|\mathbf{b}_1$. The result of this Type 2 ERO is

$$A_2|\mathbf{b}_2 = \left[\begin{array}{rrr|r} 1 & 1 & \frac{1}{2} & \frac{9}{2} \\ 0 & -3 & 1 & -3 \\ 1 & -1 & 2 & 5 \end{array}\right]$$

Step 3 Replace row 3 of $A_2|\mathbf{b}_2$ by -1(row 1 of $A_2|\mathbf{b}_2$ + row 3 of $A_2|\mathbf{b}_2$. The result of this Type 2 ERO is

$$A_3|\mathbf{b}_3 = \left[\begin{array}{rrr|r} 1 & 1 & \frac{1}{2} & \frac{9}{2} \\ 0 & -3 & 1 & -3 \\ 0 & -2 & \frac{3}{2} & \frac{1}{2} \end{array}\right]$$

The first column of (8′) has now been transformed into

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

By our procedure, we have made sure that the variable x_1 occurs in only a single equation and in that equation has a coefficient of 1. We now transform the second column of $A_3|\mathbf{b}_3$ into

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

We begin by using a Type 1 ERO to create a 1 in row 2 and column 2 of $A_3|\mathbf{b}_3$. Then we use the resulting row 2 to perform the Type 2 EROs that are needed to put zeros in the rest of column 2. Steps 4–6 accomplish these goals.

Step 4 Multiply row 2 of $A_3|\mathbf{b}_3$ by $-\frac{1}{3}$.The result of this Type 1 ERO is

$$A_4|\mathbf{b}_4 = \left[\begin{array}{ccc|c} 1 & 1 & \frac{1}{2} & \frac{9}{2} \\ 0 & 1 & -\frac{1}{3} & 1 \\ 0 & -2 & \frac{3}{2} & \frac{1}{2} \end{array}\right]$$

Step 5 Replace row 1 of $A_4|\mathbf{b}_4$ by -1(row 2 of $A_4|\mathbf{b}_4$) + row 1 of $A_4|\mathbf{b}_4$. The result of this Type 2 ERO is

$$A_5|\mathbf{b}_5 = \left[\begin{array}{ccc|c} 1 & 0 & \frac{5}{6} & \frac{7}{2} \\ 0 & 1 & -\frac{1}{3} & 1 \\ 0 & -2 & \frac{3}{2} & \frac{1}{2} \end{array}\right]$$

Step 6 Replace row 3 of $A_5|\mathbf{b}_5$ by 2(row 2 of $A_5|\mathbf{b}_5$) + row 3 of $A_5|\mathbf{b}_5$. The result of this Type 2 ERO is

$$A_6|\mathbf{b}_6 = \left[\begin{array}{ccc|c} 1 & 0 & \frac{5}{6} & \frac{7}{2} \\ 0 & 1 & -\frac{1}{3} & 1 \\ 0 & 0 & \frac{5}{6} & \frac{5}{2} \end{array}\right]$$

Column 2 has now been transformed into

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Observe that our transformation of column 2 did not change column 1.

To complete the Gauss–Jordan procedure, we must transform the third column of $A_6|\mathbf{b}_6$ into

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

We first use a Type 1 ERO to create a 1 in the third row and third column of $A_6|\mathbf{b}_6$. Then we use Type 2 EROs to put zeros in the rest of column 3. Steps 7–9 accomplish these goals.

Step 7 Multiply row 3 of $A_6|\mathbf{b}_6$ by $\frac{6}{5}$. The result of this Type 1 ERO is

$$A_7|\mathbf{b}_7 = \left[\begin{array}{ccc|c} 1 & 0 & \frac{5}{6} & \frac{7}{2} \\ 0 & 1 & -\frac{1}{3} & 1 \\ 0 & 0 & 3 & 3 \end{array}\right]$$

Step 8 Replace row 1 of $A_7|\mathbf{b}_7$ by $-\frac{5}{6}$(row 3 of $A_7|\mathbf{b}_7$) + row 1 of $A_7|\mathbf{b}_7$. The result of this Type 2 ERO is

$$A_8|\mathbf{b}_8 = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & -\frac{1}{3} & 1 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Step 9 Replace row 2 of $A_8|\mathbf{b}_8$ by $\frac{1}{3}$(row 3 of $A_8|\mathbf{b}_8$) + row 2 of $A_8|\mathbf{b}_8$. The result of this Type 2 ERO is

$$A_9|\mathbf{b}_9 = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

$A_9|\mathbf{b}_9$ represents the system of equations

$$\begin{aligned} x_1 \phantom{{}+x_2+x_3} &= 1 \\ x_2 \phantom{{}+x_3} &= 2 \\ x_3 &= 3 \end{aligned} \tag{9}$$

Thus, (9) has the unique solution $x_1 = 1$, $x_2 = 2$, $x_3 = 3$. Because (9) was obtained from (8) via EROs, the unique solution to (8) must also be $x_1 = 1$, $x_2 = 2$, $x_3 = 3$.

The reader might be wondering why we defined Type 3 EROs (interchanging of rows). To see why a Type 3 ERO might be useful, suppose you want to solve

$$\begin{aligned} 2x_2 + x_3 &= 6 \\ x_1 + x_2 - x_3 &= 2 \\ 2x_1 + x_2 + x_3 &= 4 \end{aligned} \tag{10}$$

To solve (10) by the Gauss–Jordan method, first form the augmented matrix

$$A|\mathbf{b} = \left[\begin{array}{ccr|c} 0 & 2 & 1 & 6 \\ 1 & 1 & -1 & 2 \\ 2 & 1 & 1 & 4 \end{array}\right]$$

The 0 in row 1 and column 1 means that a Type 1 ERO cannot be used to create a 1 in row 1 and column 1. If, however, we interchange rows 1 and 2 (a Type 3 ERO), we obtain

$$\left[\begin{array}{ccr|c} 1 & 1 & -1 & 2 \\ 0 & 2 & 1 & 6 \\ 2 & 1 & 1 & 4 \end{array}\right] \tag{10$'$}$$

Now we may proceed as usual with the Gauss–Jordan method.

Special Cases: No Solution or an Infinite Number of Solutions

Some linear systems have no solution, and some have an infinite number of solutions. The following two examples illustrate how the Gauss–Jordan method can be used to recognize these cases.

EXAMPLE 6 Linear System with No Solution

Find all solutions to the following linear system:

$$\begin{aligned} x_1 + 2x_2 &= 3 \\ 2x_1 + 4x_2 &= 4 \end{aligned} \tag{11}$$

Solution We apply the Gauss–Jordan method to the matrix

$$A|\mathbf{b} = \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 2 & 4 & 4 \end{array}\right]$$

We begin by replacing row 2 of $A|\mathbf{b}$ by -2(row 1 of $A|\mathbf{b}$) + row 2 of $A|\mathbf{b}$. The result of this Type 2 ERO is

$$\left[\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & 0 & -2 \end{array}\right] \tag{12}$$

We would now like to transform the second column of (12) into

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

but this is not possible. System (12) is equivalent to the following system of equations:

$$\begin{aligned} x_1 + 2x_2 &= 3 \\ 0x_1 + 0x_2 &= -2 \end{aligned} \tag{12'}$$

Whatever values we give to x_1 and x_2, the second equation in (12′) can never be satisfied. Thus, (12′) has no solution. Because (12′) was obtained from (11) by use of EROs, (11) also has no solution.

Example 6 illustrates the following idea: *If you apply the Gauss–Jordan method to a linear system and obtain a row of the form* $[0 \quad 0 \quad \cdots \quad 0|c]$ $(c \neq 0)$, *then the original linear system has no solution.*

EXAMPLE 7 Linear System with Infinite Number of Solutions

Apply the Gauss–Jordan method to the following linear system:

$$\begin{aligned} x_1 + x_2 \qquad &= 1 \\ x_2 + x_3 &= 3 \\ x_1 + 2x_2 + x_3 &= 4 \end{aligned} \tag{13}$$

Solution The augmented matrix form of (13) is

$$A|\mathbf{b} = \left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 3 \\ 1 & 2 & 1 & 4 \end{array}\right]$$

We begin by replacing row 3 (because the row 2, column 1 value is already 0) of $A|\mathbf{b}$ by -1(row 1 of $A|\mathbf{b}$) + row 3 of $A|\mathbf{b}$. The result of this Type 2 ERO is

$$A_1|\mathbf{b}_1 = \left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 3 \\ 0 & 1 & 1 & 3 \end{array}\right] \tag{14}$$

Next we replace row 1 of $A_1|\mathbf{b}_1$ by -1(row 2 of $A_1|\mathbf{b}_1$) + row 1 of $A_1|\mathbf{b}_1$. The result of this Type 2 ERO is

$$A_2|\mathbf{b}_2 = \left[\begin{array}{ccc|c} 1 & 0 & -1 & -2 \\ 0 & 1 & 1 & 3 \\ 0 & 1 & 1 & 3 \end{array}\right]$$

Now we replace row 3 of $A_2|\mathbf{b}_2$ by -1(row 2 of $A_2|\mathbf{b}_2$) + row 3 of $A_2|\mathbf{b}_2$. The result of this Type 2 ERO is

$$A_3|\mathbf{b}_3 = \left[\begin{array}{rrr|r} 1 & 0 & -1 & -2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

We would now like to transform the third column of $A_3|\mathbf{b}_3$ into

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

but this is not possible. The linear system corresponding to $A_3|\mathbf{b}_3$ is

$$x_1 - x_3 = -2 \tag{14.1}$$

$$x_2 + x_3 = 3 \tag{14.2}$$

$$0x_1 + 0x_2 + 0x_3 = 0 \tag{14.3}$$

Suppose we assign an arbitrary value k to x_3. Then (14.1) will be satisfied if $x_1 - k = -2$, or $x_1 = k - 2$. Similarly, (14.2) will be satisfied if $x_2 + k = 3$, or $x_2 = 3 - k$. Of course, (14.3) will be satisfied for any values of x_1, x_2, and x_3. Thus, for any number k, $x_1 = k - 2$, $x_2 = 3 - k$, $x_3 = k$ is a solution to (14). Thus, (14) has an infinite number of solutions (one for each number k). Because (14) was obtained from (13) via EROs, (13) also has an infinite number of solutions. A more formal characterization of linear systems that have an infinite number of solutions will be given after the following summary of the Gauss–Jordan method.

Summary of the Gauss–Jordan Method

Step 1 To solve $A\mathbf{x} = \mathbf{b}$, write down the augmented matrix $A|\mathbf{b}$.

Step 2 At any stage, define a current row, current column, and current entry (the entry in the current row and column). Begin with row 1 as the current row, column 1 as the current column, and a_{11} as the current entry. (**a**) If a_{11} (the current entry) is nonzero, then use EROs to transform column 1 (the current column) to

$$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Then obtain the new current row, column, and entry by moving down one row and one column to the right, and go to step 3. (**b**) If a_{11} (the current entry) equals 0, then do a Type 3 ERO involving the current row and any row that contains a nonzero number in the current column. Use EROs to transform column 1 to

$$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Then obtain the new current row, column, and entry by moving down one row and one column to the right. Go to step 3. (**c**) If there are no nonzero numbers in the first column, then obtain a new current column and entry by moving one column to the right. Then go to step 3.

Step 3 (**a**) If the new current entry is nonzero, then use EROs to transform it to 1 and the rest of the current column's entries to 0. When finished, obtain the new current row, column, and entry. If this is impossible, then stop. Otherwise, repeat step 3. (**b**) If the current entry is 0, then do a Type 3 ERO with the current row and any row that contains a nonzero number in the current column. Then use EROs to transform that current entry to 1 and the rest of the current column's entries to 0. When finished, obtain the new current row, column, and entry. If this is impossible, then stop. Otherwise, repeat step 3. (**c**) If the current column has no nonzero numbers below the current row, then obtain the new current column and entry, and repeat step 3. If it is impossible, then stop.

This procedure may require "passing over" one or more columns without transforming them (see Problem 8).

Step 4 Write down the system of equations $A'\mathbf{x} = \mathbf{b}'$ that corresponds to the matrix $A'|\mathbf{b}'$ obtained when step 3 is completed. Then $A'\mathbf{x} = \mathbf{b}'$ will have the same set of solutions as $A\mathbf{x} = \mathbf{b}$.

Basic Variables and Solutions to Linear Equation Systems

To describe the set of solutions to $A'\mathbf{x} = \mathbf{b}'$ (and $A\mathbf{x} = \mathbf{b}$), we need to define the concepts of basic and nonbasic variables.

DEFINITION ■ After the Gauss–Jordan method has been applied to any linear system, a variable that appears with a coefficient of 1 in a single equation and a coefficient of 0 in all other equations is called a **basic variable** (BV). ■

Any variable that is not a basic variable is called a **nonbasic variable** (NBV). ■

Let *BV* be the set of basic variables for $A'\mathbf{x} = \mathbf{b}'$ and NBV be the set of nonbasic variables for $A'\mathbf{x} = \mathbf{b}'$. The character of the solutions to $A'\mathbf{x} = \mathbf{b}'$ depends on which of the following cases occurs.

Case 1 $A'\mathbf{x} = \mathbf{b}'$ has at least one row of form $[0 \quad 0 \quad \cdots \quad 0|c]$ $(c \neq 0)$. Then $A\mathbf{x} = \mathbf{b}$ has no solution (recall Example 6). As an example of Case 1, suppose that when the Gauss–Jordan method is applied to the system $A\mathbf{x} = \mathbf{b}$, the following matrix is obtained:

$$A'|\mathbf{b}' = \left[\begin{array}{cccc|r} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 1 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right]$$

In this case, $A'\mathbf{x} = \mathbf{b}'$ (and $A\mathbf{x} = \mathbf{b}$) has no solution.

Case 2 Suppose that Case 1 does not apply and NBV, the set of nonbasic variables, is empty. Then $A'\mathbf{x} = \mathbf{b}'$ (and $A\mathbf{x} = \mathbf{b}$) will have a unique solution. To illustrate this, we recall that in solving

$$\begin{aligned} 2x_1 + 2x_2 + x_3 &= 9 \\ 2x_1 - x_2 + 2x_3 &= 6 \\ x_1 - x_2 + 2x_3 &= 5 \end{aligned}$$

the Gauss–Jordan method yielded

$$A'|\mathbf{b}' = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

In this case, BV $= \{x_1, x_2, x_3\}$ and NBV is empty. Then the unique solution to $A'\mathbf{x} = \mathbf{b}'$ (and $A\mathbf{x} = \mathbf{b}$) is $x_1 = 1$, $x_2 = 2$, $x_3 = 3$.

Case 3 Suppose that Case 1 does not apply and NBV is nonempty. Then $A'\mathbf{x} = \mathbf{b}'$ (and $A\mathbf{x} = \mathbf{b}$) will have an infinite number of solutions. To obtain these, first assign each nonbasic variable an arbitrary value. Then solve for the value of each basic variable in terms of the nonbasic variables. For example, suppose

$$A'|\mathbf{b}' = \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & 1 & 3 \\ 0 & 1 & 0 & 2 & 0 & 2 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \tag{15}$$

Because Case 1 does not apply, and BV = $\{x_1, x_2, x_3\}$ and NBV = $\{x_4, x_5\}$, we have an example of Case 3: $A'\mathbf{x} = \mathbf{b}'$ (and $A\mathbf{x} = \mathbf{b}$) will have an infinite number of solutions. To see what these solutions look like, write down $A'\mathbf{x} = \mathbf{b}'$:

$$x_1 + x_4 + x_5 = 3 \tag{15.1}$$

$$x_2 + 2x_4 = 2 \tag{15.2}$$

$$x_3 + x_5 = 1 \tag{15.3}$$

$$0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 = 0 \tag{15.4}$$

Now assign the nonbasic variables (x_4 and x_5) arbitrary values c and k, with $x_4 = c$ and $x_5 = k$. From (15.1), we find that $x_1 = 3 - c - k$. From (15.2), we find that $x_2 = 2 - 2c$. From (15.3), we find that $x_3 = 1 - k$. Because (15.4) holds for all values of the variables, $x_1 = 3 - c - k$, $x_2 = 2 - 2c$, $x_3 = 1 - k$, $x_4 = c$, and $x_5 = k$ will, for any values of c and k, be a solution to $A'\mathbf{x} = \mathbf{b}'$ (and $A\mathbf{x} = \mathbf{b}$).

Our discussion of the Gauss–Jordan method is summarized in Figure 6. We have devoted so much time to the Gauss–Jordan method because, in our study of linear programming, examples of Case 3 (linear systems with an infinite number of solutions) will occur repeatedly. Because the end result of the Gauss–Jordan method must always be one of Cases 1–3, we have shown that any linear system will have no solution, a unique solution, or an infinite number of solutions.

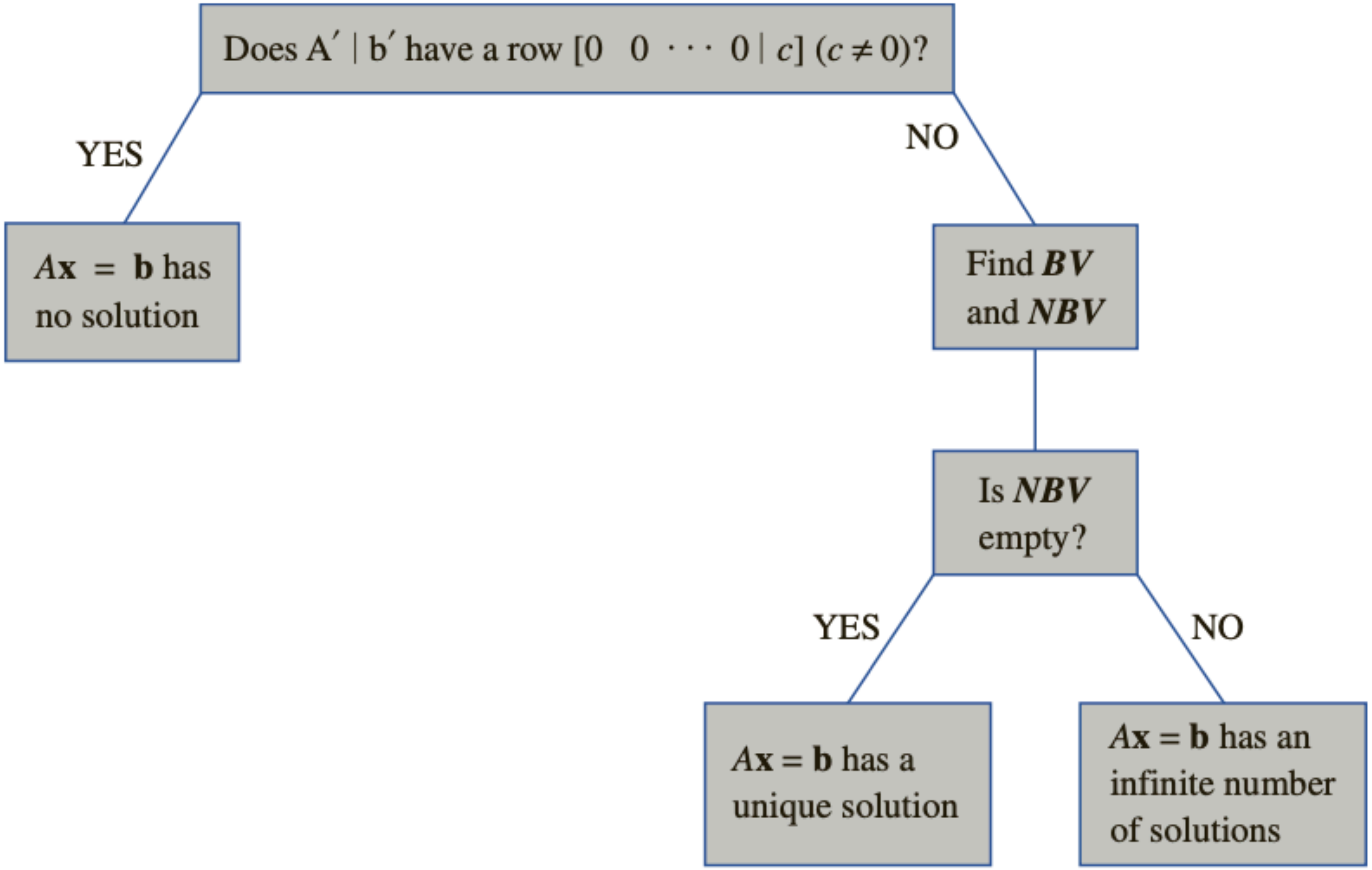


FIGURE 6
Description of Gauss–Jordan Method for Solving Linear Equations

2.5 The Inverse of a Matrix

To solve a single linear equation such as $4x = 3$, we simply multiply both sides of the equation by the multiplicative inverse of 4, which is 4^{-1}, or $\frac{1}{4}$. This yields $4^{-1}(4x) = (4^{-1})3$, or $x = \frac{3}{4}$. (Of course, this method fails to work for the equation $0x = 3$, because zero has no multiplicative inverse.) In this section, we develop a generalization of this technique that can be used to solve "square" (number of equations = number of unknowns) linear systems. We begin with some preliminary definitions.

DEFINITION ■ A **square matrix** is any matrix that has an equal number of rows and columns. ■

The **diagonal elements** of a square matrix are those elements a_{ij} such that $i = j$. ■

A square matrix for which all diagonal elements are equal to 1 and all nondiagonal elements are equal to 0 is called an **identity matrix.** ■

The $m \times m$ identity matrix will be written as I_m. Thus,

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \cdots$$

If the multiplications $I_m A$ and AI_m are defined, it is easy to show that $I_m A = AI_m = A$. Thus, just as the number 1 serves as the unit element for multiplication of real numbers, I_m serves as the unit element for multiplication of matrices.

Recall that $\frac{1}{4}$ is the multiplicative inverse of 4. This is because $4(\frac{1}{4}) = (\frac{1}{4})4 = 1$. This motivates the following definition of the inverse of a matrix.

DEFINITION ■ For a given $m \times m$ matrix A, the $m \times m$ matrix B is the **inverse** of A if

$$BA = AB = I_m \tag{16}$$

(It can be shown that if $BA = I_m$ or $AB = I_m$, then the other quantity will also equal I_m.) ■

Some square matrices do not have inverses. If there does exist an $m \times m$ matrix B that satisfies Equation (16), then we write $B = A^{-1}$. For example, if

$$A = \begin{bmatrix} 2 & 0 & -1 \\ 3 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

the reader can verify that

$$\begin{bmatrix} 2 & 0 & -1 \\ 3 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ -5 & 1 & -7 \\ 1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & 0 & 1 \\ -5 & 1 & -7 \\ 1 & 0 & 2 \end{bmatrix}\begin{bmatrix} 2 & 0 & -1 \\ 3 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Thus,

$$A^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ -5 & 1 & -7 \\ 1 & 0 & 2 \end{bmatrix}$$

To see why we are interested in the concept of a matrix inverse, suppose we want to solve a linear system $A\mathbf{x} = \mathbf{b}$ that has m equations and m unknowns. Suppose that A^{-1} exists. Multiplying both sides of $A\mathbf{x} = \mathbf{b}$ by A^{-1}, we see that any solution of $A\mathbf{x} = \mathbf{b}$ must also satisfy $A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{b}$. Using the associative law and the definition of a matrix inverse, we obtain

$$
\begin{aligned}
& & (A^{-1}A)\mathbf{x} &= A^{-1}\mathbf{b} \\
&\text{or} & I_m\mathbf{x} &= A^{-1}\mathbf{b} \\
&\text{or} & \mathbf{x} &= A^{-1}\mathbf{b}
\end{aligned}
$$

This shows that knowing A^{-1} enables us to find the unique solution to a square linear system. This is the analog of solving $4x = 3$ by multiplying both sides of the equation by 4^{-1}.

The Gauss–Jordan method may be used to find A^{-1} (or to show that A^{-1} does not exist). To illustrate how we can use the Gauss–Jordan method to invert a matrix, suppose we want to find A^{-1} for

$$A = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$$

This requires that we find a matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = A^{-1}$$

that satisfies

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{17}$$

From Equation (17), we obtain the following pair of simultaneous equations that must be satisfied by a, b, c, and d:

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} a \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \qquad \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Thus, to find

$$\begin{bmatrix} a \\ c \end{bmatrix}$$

(the first column of A^{-1}), we can apply the Gauss–Jordan method to the augmented matrix

$$\left[\begin{array}{cc|c} 2 & 5 & 1 \\ 1 & 3 & 0 \end{array}\right]$$

Once EROs have transformed

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$$

to I_2,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

will have been transformed into the first column of A^{-1}. To determine

$$\begin{bmatrix} b \\ d \end{bmatrix}$$

(the second column of A^{-1}), we apply EROs to the augmented matrix

$$\left[\begin{array}{cc|c} 2 & 5 & 0 \\ 1 & 3 & 1 \end{array}\right]$$

When

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$$

has been transformed into I_2,

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

will have been transformed into the second column of A^{-1}. Thus, to find each column of A^{-1}, we must perform a sequence of EROs that transform

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$$

into I_2. This suggests that we can find A^{-1} by applying EROs to the 2×4 matrix

$$A|I_2 = \left[\begin{array}{cc|cc} 2 & 5 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{array}\right]$$

When

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$$

has been transformed to I_2,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

will have been transformed into the first column of A^{-1}, and

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

will have been transformed into the second column of A^{-1}. Thus, *as* A *is transformed into* I_2, I_2 *is transformed into* A^{-1}. The computations to determine A^{-1} follow.

Step 1 Multiply row 1 of $A|I_2$ by $\frac{1}{2}$. This yields

$$A'|I_2' = \left[\begin{array}{cc|cc} 1 & \frac{5}{2} & \frac{1}{2} & 0 \\ 1 & 3 & 0 & 1 \end{array}\right]$$

Step 2 Replace row 2 of $A'|I_2'$ by -1(row 1 of $A'|I_2'$) + row 2 of $A'|I_2'$. This yields

$$A''|I_2'' = \left[\begin{array}{cc|cc} 1 & \frac{5}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 1 \end{array}\right]$$

Step 3 Multiply row 2 of $A''|I_2''$ by 2. This yields

$$A'''|I_2''' = \left[\begin{array}{cc|cc} 1 & \frac{5}{2} & \frac{1}{2} & 0 \\ 0 & 1 & -1 & 2 \end{array}\right]$$

Step 4 Replace row 1 of $A'''|I_2'''$ by $-\frac{5}{2}$(row 2 of $A'''|I_2'''$) + row 1 of $A'''|I_2'''$. This yields

$$\left[\begin{array}{cc|cc} 1 & 0 & 3 & -5 \\ 0 & 1 & -1 & 2 \end{array}\right]$$

Because A has been transformed into I_2, I_2 will have been transformed into A^{-1}. Hence,

$$A^{-1} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}$$

The reader should verify that $AA^{-1} = A^{-1}A = I_2$.

A Matrix May Not Have an Inverse

Some matrices do not have inverses. To illustrate, let

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad \text{and} \quad A^{-1} = \begin{bmatrix} e & f \\ g & h \end{bmatrix} \tag{18}$$

To find A^{-1} we must solve the following pair of simultaneous equations:

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} e \\ g \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{18.1}$$

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} f \\ h \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{18.2}$$

When we try to solve (18.1) by the Gauss–Jordan method, we find that

$$\left[\begin{array}{cc|c} 1 & 2 & 1 \\ 2 & 4 & 0 \end{array}\right]$$

is transformed into

$$\left[\begin{array}{cc|r} 1 & 2 & 1 \\ 0 & 0 & -2 \end{array}\right]$$

This indicates that (18.1) has no solution, and A^{-1} cannot exist.

Observe that (18.1) fails to have a solution, because the Gauss–Jordan method transforms A into a matrix with a row of zeros on the bottom. This can only happen if rank $A < 2$. If $m \times m$ matrix A has rank $A < m$, then A^{-1} will not exist.

The Gauss–Jordan Method for Inverting an $m \times m$ Matrix A

Step 1 Write down the $m \times 2m$ matrix $A|I_m$.

Step 1 Use EROs to transform $A|I_m$ into $I_m|B$. This will be possible only if rank $A = m$. In this case, $B = A^{-1}$. If rank $A < m$, then A has no inverse.

Using Matrix Inverses to Solve Linear Systems

As previously stated, matrix inverses can be used to solve a linear system $A\mathbf{x} = \mathbf{b}$ in which the number of variables and equations are equal. Simply multiply both sides of $A\mathbf{x} = \mathbf{b}$ by A^{-1} to obtain the solution $\mathbf{x} = A^{-1}\mathbf{b}$. For example, to solve

$$\begin{aligned} 2x_1 + 5x_2 &= 7 \\ x_1 + 3x_2 &= 4 \end{aligned} \tag{19}$$

write the matrix representation of (19):

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \end{bmatrix} \tag{20}$$

Let

$$A = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$$

We found in the previous illustration that

$$A^{-1} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}$$

Multiplying both sides of (20) by A^{-1}, we obtain

$$\begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 7 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Thus, $x_1 = 1$, $x_2 = 1$ is the unique solution to system (19).

2.6 Determinants

Associated with any square matrix A is a number called the *determinant* of A (often abbreviated as det A or $|A|$). Knowing how to compute the determinant of a square matrix will be useful in our study of nonlinear programming.

For a 1×1 matrix $A = [a_{11}]$,

$$\det A = a_{11} \tag{21}$$

For a 2×2 matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{22}$$

$$\det A = a_{11}a_{22} - a_{21}a_{12}$$

For example,

$$\det \begin{bmatrix} 2 & 4 \\ 3 & 5 \end{bmatrix} = 2(5) - 3(4) = -2$$

Before we learn how to compute det A for larger square matrices, we need to define the concept of the *minor* of a matrix.

DEFINITION ■ If A is an $m \times m$ matrix, then for any values of i and j, the ijth **minor** of A (written A_{ij}) is the $(m - 1) \times (m - 1)$ submatrix of A obtained by deleting row i and column j of A. ■

For example,

$$\text{if} \qquad A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \qquad \text{then} \qquad A_{12} = \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix} \qquad \text{and} \qquad A_{32} = \begin{bmatrix} 1 & 3 \\ 4 & 6 \end{bmatrix}$$

Let A be any $m \times m$ matrix. We may write A as

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

To compute det A, pick any value of i $(i = 1, 2, \ldots, m)$ and compute det A:

$$\det A = (-1)^{i+1}a_{i1}(\det A_{i1}) + (-1)^{i+2}a_{i2}(\det A_{i2}) + \cdots + (-1)^{i+m}a_{im}(\det A_{im}) \qquad (23)$$

Formula (23) is called the expansion of det A by the cofactors of row i. The virtue of (23) is that it reduces the computation of det A for an $m \times m$ matrix to computations involving only $(m - 1) \times (m - 1)$ matrices. Apply (23) until det A can be expressed in terms of 2×2 matrices. Then use Equation (22) to find the determinants of the relevant 2×2 matrices.

To illustrate the use of (23), we find det A for

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

We expand det A by using row 1 cofactors. Notice that $a_{11} = 1$, $a_{12} = 2$, and $a_{13} = 3$. Also

$$A_{11} = \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix}$$

so by (22), det $A_{11} = 5(9) - 8(6) = -3$;

$$A_{12} = \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix}$$

so by (22), det $A_{12} = 4(9) - 7(6) = -6$; and

$$A_{13} = \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix}$$

so by (22), det $A_{13} = 4(8) - 7(5) = -3$. Then by (23),

$$\begin{aligned} \det A &= (-1)^{1+1}a_{11}(\det A_{11}) + (-1)^{1+2}a_{12}(\det A_{12}) + (-1)^{1+3}a_{13}(\det A_{13}) \\ &= (1)(1)(-3) + (-1)(2)(-6) + (1)(3)(-3) = -3 + 12 - 9 = 0 \end{aligned}$$

The interested reader may verify that expansion of det A by either row 2 or row 3 cofactors also yields det $A = 0$.

We close our discussion of determinants by noting that they can be used to invert square matrices and to solve linear equation systems. Because we already have learned to use the Gauss–Jordan method to invert matrices and to solve linear equation systems, we will not discuss these uses of determinants.

Gram-Schmidt Orthogonalization

We will do this stuff in class and homeworks.

We have seen that it can be very convenient to have an orthonormal basis for a given vector space, in order to compute expansions of arbitrary vectors within that space. Therefore, given a *non*-orthonormal basis, it is desirable to have a process for obtaining an orthonormal basis from it.

Fortunately, we have such a process, known as *Gram-Schmidt orthogonalization.* Suppose that we have a linearly independent, but not orthonormal, set of functions $\{\chi_1, \chi_2, \ldots\}$ that span a given vector space V. To construct an orthonormal set $\{\varphi_1, \varphi_2, \ldots\}$ from this set, we proceed as follows. First, to obtain φ_1, we simply normalize χ_1:

$$\varphi_1 = \frac{\chi_1}{\|\chi_1\|}.$$

Next, to obtain φ_2, we need to ensure that it is orthogonal to φ_1, and then normalize it.

As an intermediate step, we seek a function ψ_2 of the form

$$\psi_2 = \chi_2 + c_{12}\varphi_1$$

such that $\langle\varphi_1|\psi_2\rangle = 0$. Then, we can set $\varphi_2 = \psi_2/\|\psi_2\|$. Taking the scalar product of both sides of the above equation with φ_1, we obtain

$$0 = \langle\varphi_1|\psi_2\rangle = \langle\varphi_1|\chi_2\rangle + c_{12}\langle\varphi_1|\varphi_1\rangle.$$

Because the φ_j are orthonormal, it follows that

$$c_{12} = -\langle\varphi_1|\chi_2\rangle.$$

We conclude that φ_2 can be obtained as follows:

$$\begin{aligned} \psi_2 &= \chi_2 - \langle\varphi_1|\chi_2\rangle\varphi_1 \\ \varphi_2 &= \frac{\psi_2}{\|\psi_2\|}. \end{aligned}$$

We now have a set of two functions that is orthonormal.

Now, to obtain φ_3, we must ensure that it is orthogonal to φ_1 *and* φ_2, and then normalized. To that end, we seek a function ψ_3 of the form

$$\psi_3 = \chi_3 + c_{13}\varphi_1 + c_{23}\varphi_2$$

such that $\langle\varphi_1|\psi_3\rangle = \langle\varphi_2|\psi_3\rangle = 0$. Then, we can set $\varphi_3 = \psi_3/\|\psi_3\|$. Taking the scalar product of both sides of the above equation with φ_1, and then separately, φ_2, we obtain

$$\begin{aligned} 0 &= \langle\varphi_1|\psi_3\rangle = \langle\varphi_1|\chi_3\rangle + c_{13}\langle\varphi_1|\varphi_1\rangle + c_{23}\langle\varphi_1|\varphi_2\rangle \\ 0 &= \langle\varphi_2|\psi_3\rangle = \langle\varphi_2|\chi_3\rangle + c_{13}\langle\varphi_2|\varphi_1\rangle + c_{23}\langle\varphi_2|\varphi_2\rangle. \end{aligned}$$

Because the φ_j are orthonormal, it follows that

$$c_{13} = -\langle \varphi_1 | \chi_3 \rangle, \quad c_{23} = -\langle \varphi_2 | \chi_3 \rangle.$$

We conclude that φ_3 can be obtained as follows:

$$\begin{aligned} \psi_3 &= \chi_3 - \langle \varphi_1 | \chi_3 \rangle \varphi_1 - \langle \varphi_2 | \chi_3 \rangle \varphi_2 \\ \varphi_3 &= \frac{\psi_3}{\|\psi_3\|}. \end{aligned}$$

We now have a set of three functions that are orthonormal.

Continuing this process, we see that we can obtain each function φ_j as follows:

$$\begin{aligned} \psi_j &= \chi_j - \sum_{k=0}^{j-1} \langle \varphi_k | \chi_j \rangle \varphi_k \\ \varphi_j &= \frac{\psi_j}{\|\psi_j\|}. \end{aligned}$$

This yields a set of functions $\{\varphi_1, \varphi_2, \ldots\}$ that is an orthonormal basis of the space spanned by $\{\chi_1, \chi_2, \ldots\}$, with respect to the scalar product that is used.

Example We wish to obtain a set of orthonormal polynomials with respect to the scalar product

$$\langle f | g \rangle = \int_{-1}^{1} f^*(s) g(s)\, ds.$$

This will be accomplished by applying Gram-Schmidt orthogonalization to the set $\{1, x, x^2, x^3, \ldots\}$. Setting $\chi_j(x) = x^j$ for $j = 0, 1, 2, \ldots$, our orthogonal set $\{\varphi_j\}$, $j = 0, 1, 2, \ldots,$, is obtained as follows:

$$\begin{aligned} \psi_0(x) &= \chi_0(x) \\ &= 1, \\ \varphi_0(x) &= \frac{\psi_0(x)}{\|\psi_0\|} \\ &= \frac{1}{\langle 1 | 1 \rangle^{1/2}} \\ &= \frac{1}{\left[\int_{-1}^{1} 1\, ds \right]^{1/2}} \\ &= \frac{1}{\sqrt{2}}, \\ \psi_1(x) &= \chi_1(x) - \langle \varphi_0 | \chi_1 \rangle \varphi_0(x) \\ &= x - \left\langle \frac{1}{\sqrt{2}} \middle| x \right\rangle \frac{1}{\sqrt{2}} \\ &= x - \frac{1}{2} \int_{-1}^{1} s\, ds \\ &= x - \frac{1}{2} 0 \\ &= x, \end{aligned}$$

$$
\begin{aligned}
\varphi_1(x) &= \frac{\psi_1(x)}{\|\psi_1\|} \\
&= \frac{x}{\langle x|x\rangle^{1/2}} \\
&= \frac{x}{\left[\int_{-1}^{1} s^2\,ds\right]^{1/2}} \\
&= \sqrt{\frac{3}{2}}x, \\
\psi_2(x) &= \chi_2(x) - \langle\varphi_0|\chi_2\rangle\varphi_0(x) - \langle\varphi_1|\chi_2\rangle\varphi_1(x) \\
&= x^2 - \left\langle \frac{1}{\sqrt{2}} \middle| x^2 \right\rangle \frac{1}{\sqrt{2}} - \left\langle \sqrt{\frac{3}{2}}x \middle| x^2 \right\rangle \sqrt{\frac{3}{2}}x \\
&= x^2 - \frac{1}{2}\int_{-1}^{1} s^2\,ds - \frac{3}{2}x\int_{-1}^{1} s^3\,ds \\
&= x^2 - \frac{1}{2}\frac{2}{3} - \frac{3}{2}0 \\
&= x^2 - \frac{1}{3}, \\
\varphi_2(x) &= \frac{\psi_2(x)}{\|\psi_2\|} \\
&= \frac{x^2 - \frac{1}{3}}{\left\langle x^2 - \frac{1}{3} \middle| x^2 - \frac{1}{3} \right\rangle^{1/2}} \\
&= \frac{x^2 - \frac{1}{3}}{\left[\int_{-1}^{1} \left(x - \frac{1}{3}\right)^2\,ds\right]^{1/2}} \\
&= \sqrt{\frac{5}{2}}\left(\frac{3}{2}x^2 - \frac{1}{2}\right).
\end{aligned}
$$

Continuing this process, we obtain

$$\varphi_3(x) = \sqrt{\frac{7}{2}}\left(\frac{5}{2}x^3 - \frac{3}{2}x\right).$$

In general,

$$\varphi_n(x) = \sqrt{\frac{2n+1}{2}}P_n(x),$$

where $P_n(x)$ is the *Legendre polynomial* of nth degree. We will learn more about these orthogonal (but not orthonormal) polynomials later in this course. □

While Gram-Schmidt orthogonalization can be applied to the *monomial bais* $\{1, x, x^2, x^3, \ldots\}$ to obtain an orthonormal sequence of polynomials, it can be quite cumbersome, as can be seen from the preceding example. However, a modification of this procedure can yield a much more efficient approach.

Suppose that we have already generated a sequence of n orthonormal polynomials $\varphi_0, \varphi_1, \varphi_2, \ldots, \varphi_{n-1}$ with respect to some scalar product

$$\langle f|g\rangle = \int_a^b f^*(x)g(x)w(x)\,dx,$$

where φ_j is of degree j for $j = 0, 1, 2, \ldots, n-1$. Then, to obtain φ_n, which of degree n, we orthogonalize $x\varphi_{n-1}(x)$, which is of degree n, against $\varphi_0, \varphi_1, \ldots, \varphi_{n-1}$ using the same approach as in Gram-Schmidt orthogonalization. That is, we compute

$$\begin{aligned}\psi_n(x) &= x\varphi_{n-1}(x) - \sum_{j=0}^{n-1} \langle \varphi_j | x\varphi_{n-1} \rangle \varphi_j(x), \\ \varphi_n(x) &= \frac{\psi_n(x)}{\|\psi_n\|}.\end{aligned}$$

Now, consider the scalar product $\langle \varphi_j | x\varphi_{n-1} \rangle$. Using the properties of the scalar product, we have

$$\langle \varphi_j | x\varphi_{n-1} \rangle = \langle x\varphi_j | \varphi_{n-1} \rangle.$$

However, because $\varphi_0, \varphi_1, \ldots, \varphi_{n-1}$ are orthonormal, $\langle p | \varphi_{n-1} \rangle = 0$ if $p(x)$ is *any* polynomial of degree less than $n-1$. Because $x\varphi_j(x)$ is of degree $j+1$, it follows that $\langle \varphi_j | x\varphi_{n-1} \rangle = 0$ whenever $j+1 < n-1$, or $j < n-2$. Therefore, our orthogonalization procedure simplifies to

$$\begin{aligned}\psi_n(x) &= x\varphi_{n-1}(x) - \langle \varphi_{n-2} | x\varphi_{n-1} \rangle \varphi_{n-2}(x) - \langle \varphi_{n-1} | x\varphi_{n-1} \rangle \varphi_{n-1}(x), \\ \varphi_n(x) &= \frac{\psi_n(x)}{\|\psi_n\|}.\end{aligned}$$

That is, any family of orthogonal polynomials satisfies a *three-term recurrence relation*, in which each polynomial depends on the previous two. Table lists several families of orthogonal polynomials that can be generated from such a recurrence relation; we will see some of these families later in the course.

Polynomials	Scalar Product
Legendre	$\int_{-1}^{1} P_n(x)P_m(x)\,dx = 2\delta_{mn}/(2n+1)$
Shifted Legendre	$\int_{0}^{1} P_n^*(x)P_m^*(x)\,dx = \delta_{mn}/(2n+1)$
Chebyshev, first kind	$\int_{-1}^{1} T_n(x)T_m(x)(1-x^2)^{-1/2}\,dx = \delta_{mn}\pi/(2-\delta_{n0})$
Shifted Chebyshev, first kind	$\int_{0}^{1} T_n^*(x)T_m^*(x)[x(1-x)]^{-1/2}\,dx = \delta_{mn}\pi/(2-\delta_{n0})$
Chebyshev, second kind	$\int_{-1}^{1} U_n(x)U_m(x)(1-x^2)^{1/2}\,dx = \delta_{mn}\pi/2$
Leguerre	$\int_{0}^{\infty} L_n(x)L_m(x)e^{-x}\,dx = \delta_{mn}$
Associated Laguerre	$\int_{0}^{\infty} L_n^k(x)L_m^k(x)e^{-x}\,dx = \delta_{mn}(n+k)!/n!$
Hermite	$\int_{-\infty}^{\infty} H_n(x)H_m(x)e^{-x^2}\,dx = 2^n\delta_{mn}\sqrt{\pi}n!$

As can be seen in the following example, Gram-Schmidt orthogonalization can be applied to vectors in *any* inner product space, such as vectors in $\mathbb{R}^n$.

Example Given the vectors in $\mathbb{R}^3$,

$$|\mathbf{a}_1\rangle = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \quad |\mathbf{a}_2\rangle = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}, \quad |\mathbf{a}_3\rangle = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

we will use Gram-Schmidt orthogonalization to obtain an orthonormal set of vectors, $\{|\mathbf{b}_1\rangle, |\mathbf{b}_2\rangle, |\mathbf{b}_3\rangle\}$. We have

$$|\mathbf{b}_1\rangle = \frac{|\mathbf{a}_1\rangle}{\langle \mathbf{a}_1 | \mathbf{a}_1 \rangle^{1/2}}$$

$$
\begin{aligned}
&= \frac{1}{6^{1/2}}|\mathbf{a}_1\rangle \\
&= \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \\
|\mathbf{b}_2'\rangle &= |\mathbf{a}_2\rangle - \langle\mathbf{b}_1|\mathbf{a}_2\rangle|\mathbf{b}_1\rangle \\
&= \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} - \frac{9}{\sqrt{6}}\frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \\
&= \begin{pmatrix} -1/2 \\ 1/2 \\ 0 \end{pmatrix}, \\
|\mathbf{b}_2\rangle &= \frac{|\mathbf{b}_2'\rangle}{\langle\mathbf{b}_2'|\mathbf{b}_2'\rangle^{1/2}} \\
&= \frac{1}{\left(\frac{1}{2}\right)^{1/2}}|\mathbf{b}_2'\rangle \\
&= \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \\
|\mathbf{b}_3'\rangle &= |\mathbf{a}_3\rangle - \langle\mathbf{b}_1|\mathbf{a}_3\rangle|\mathbf{b}_1\rangle - \langle\mathbf{b}_2|\mathbf{a}_3\rangle|\mathbf{b}_2\rangle \\
&= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \left(-\frac{1}{\sqrt{6}}\right)\frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} - \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \\
&= \frac{2}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \\
|\mathbf{b}_3\rangle &= \frac{|\mathbf{b}_3'\rangle}{\langle\mathbf{b}_3'|\mathbf{b}_3'\rangle^{1/2}} \\
&= \frac{1}{\left(\frac{4}{3}\right)^{1/2}}|\mathbf{b}_3'\rangle \\
&= \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.
\end{aligned}
$$

We conclude that our orthonormal set of vectors is

$$
|\mathbf{b}_1\rangle = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \quad |\mathbf{b}_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad |\mathbf{b}_3\rangle = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.
$$