Version of 6.10.98

MA106 Calculus I

Syllabus

1. Real numbers; complex numbers:
   representation, Argand diagram,
   de Moivre's theorem, complex nth roots.
2. Real functions of one variable:
   formulae, graphs, domain and range, inversion;
   differentiation (formal rules, geometric interpretation, tangents).
   Local and global maxima and minima;
   limits (formal rules and pictorial interpretations), l'Hôpital's rule.
   (Examples of functions to include x^α, exp, sin, cos, tan, sinh, cosh, tanh, polynomials, and their inverses.)
3. Sequences and series:
   convergence of sequences, monotonic sequences,
   summability of series, absolutely summable series, comparison and ratio tests.
   (Examples to include geometric and harmonic series, power series for elementary functions.)
   Power series. Radius of convergence
4. Intermediate and mean value theorems,
   Taylor's theorem. Taylor's series;
   Maclaurin's series.

Textbooks There are no textbooks perfectly suited for this course. There is a shelf of approximately relevant books in section QA303 in the library. You should prefer those with lots of pictures and avoid any which use the Greek letters ϵ and δ in incomprehensible ways.

A book which you may find useful not only for this course but also for others is

E.Kreyszig, *Advanced Engineering Mathematics*, Wiley (QA37.K7).

Two others which may be helpful are

E.W.Swokowski, M.Olinick, D.Pence & J.A.Cole, *Calculus*, PWS Publishing (QA303.S7)

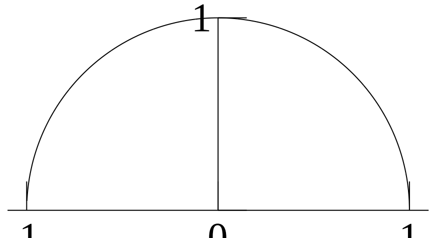
J.Gilbert, *Guide to Mathematical Methods*, McMillan (QA37.G5).

filename ma1063.tex Version of 29.9.97

D. As an example of something a little more complicated, consider the semicircle function, given by saying

$\operatorname{dom} f = \{x : -1 \le x \le 1\} = \{x : x \in \mathbb{R},\ |x| \le 1\}$,
$f(x) = \sqrt{1 - x^2}$,

the graph being



$\sqrt{1-x^2}$

The general rule is this: if you have some notion of what rule-of-calculation you are going to use, pick for your domain the largest set on which this rule works properly. Thus, if the rule-of-calculation is $x \mapsto \sqrt{1-x^2}$, then the set of x for which this works is

$$\{x : 1 - x^2 \geq 0\} = \{x : x^2 \leq 1\} = \{x : |x| \leq 1\},$$

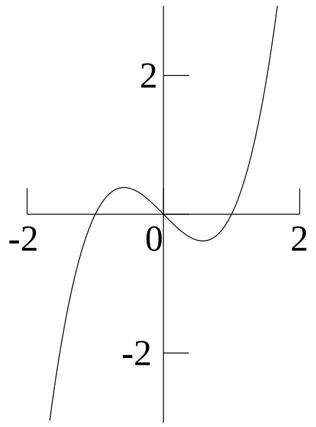
and this is likely to be the most sensible domain to choose. But there might be circumstances in which you wanted (for instance) the domain $\{x : -1 < x < 1\}$ – if, for instance, you were differentiating the function.

I have tried to emphasize (i) that when you come to describe a function precisely, you should normally begin by specifying the domain, because it's not fair to let your reader embark on calculating something that won't come out (ii) that it is for you to decide what domain you want – e.g., whether you're limiting yourself to real numbers in the domain, or allowing complex numbers.

Maxima and minima

Once we have a function properly specified, there's a long list of things to be found out about it, and I'll try to run through the ones you need to be able to do. Let's start with *maxima* and *minima*. These come in two kinds:

local maxima in, for instance, $x \mapsto x^3 - x$.



$x^3 - x$

Here we see a bump to the left of 0. The top point is at $x_0 = -\frac{1}{\sqrt{3}}$, $f(x_0) = \frac{2}{3\sqrt{3}}$. You can see that $f(x) \leq f(x_0)$ if, for instance, $-1 \leq x \leq 0$, even though $f(2) > f(x_0)$.

In general, we say that a function f has a **local maximum** at a point $x_0 \in \operatorname{dom} f$ if there is some interval I around x_0 such that $f(x) \leq f(x_0)$ whenever $x \in I \cap \operatorname{dom} f$.

Local minima are the same thing, upside down.

True or **global** or **overall** maxima: a real function f has a **global maximum** at $x_0 \in \operatorname{dom} f$ if $f(x) \le f(x_0)$ for every $x \in \operatorname{dom} f$. Thus $x \mapsto x^3 - x$ has no global maximum (supposing, that is, that we are taking its domain to be $\mathbb{R}$), but cos has global maxima at $0, \pm 2\pi, \pm 4\pi, \ldots$ (with value 1) and global minima at $\pm\pi, \pm 3\pi, \ldots$ (with value -1).

> Remember: every global maximum is a local maximum; not every local maximum is a global maximum.

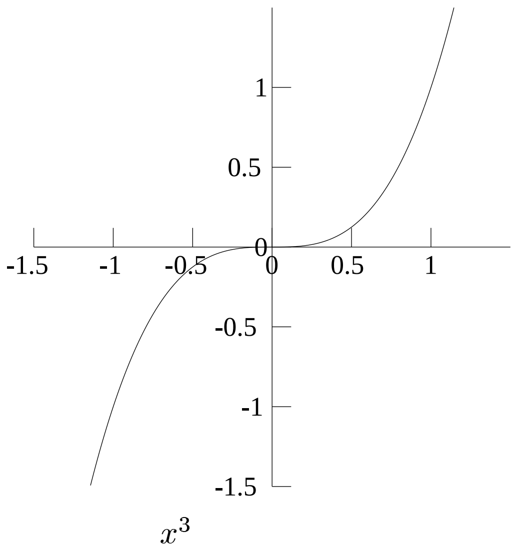
Another special kind of point is a **stationary** point; this is a point at which the graph of the function is running horizontally, that is, the derivative exists and is zero. In fact, when hunting for local and global maxima and minima of a function f the usual method is to start by finding its derivative f' and solving the equation $f'(x) = 0$. This is effective because

> if x is a local maximum or minimum, and $f'(x)$ is defined, then $f'(x) = 0$.

So you can hope to pick up most of your maxima and minima this way, at least when f is differentiable. But

> **Warning!** Not every stationary point is a maximum or minimum.

For instance, consider the function $x \mapsto x^3$, with domain $\mathbb{R}$:

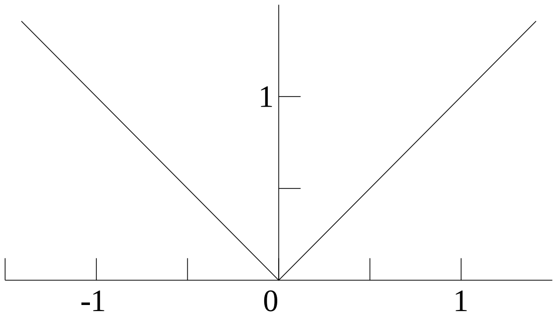


x^3

This has a stationary point at 0 but no local maximum or minimum. What gives most people even more trouble is

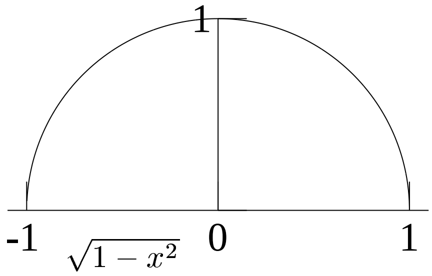
> **Warning!** not every maximum or minimum is a stationary point.

For instance, try $f(x) = |x|$ (with domain $\mathbb{R}$):



$|x|$

This has a global minumum at 0 which is not a stationary point because there's no tangent there. (The function turns a sharp corner.) Finally, you have to check sometimes at the endpoints of domains, when functions stop suddenly:



$\sqrt{1-x^2}$

This has global minima at ± 1, but the only stationary point is at 0, which is a global maximum.

So the best general rule I can give has to be

> Find a formula for f'.
> Look at all solutions of $f'(x) = 0$; see if any of these are maxima or minima.
> Look at any odd points where f' is undefined.
> Look at the ends of the domain of f, if any.
> *Draw the picture!* if you possibly can.

A word on what this course is about. I expect you find the rule above a bit indefinite. In fact it actually asks you to use your brains. I have to say that that is what I think mathematics is about. We do have lots of formulae; places where all you have to do is type the numbers in, press the right buttons and you'll get the right answer. Understanding strict rules which, if exactly followed, will take you safely to the end of a problem, is a very important part of mathematics. But these techniques are important not because they enable us to escape from thinking but because they liberate us to think about something more interesting. I think the biggest difference between 'school' and 'university' mathematics is that from now on we are going to be exploring the edges where things don't quite work as we should like, and finding out what might go wrong and what still goes right. This means that quite a lot of the time – in this course especially – I am going to have to ask you to reconsider something you've been confidently doing, and warn you that in the wrong place it's chancy. But never throw ideas away altogether. The things you have learnt are almost always useful for something, even if they don't reach quite as far as you imagined they did.

Differentiation I've been talking about this for a couple of lectures now. I am sure that you have all done quite a bit of differentiating at one time or another, and I should like to believe that you are all perfectly efficient. But I rather suspect that a bit of revision will be useful, so I will run through the basic rules carefully.

The **derivative** of a real function f at a point x is the slope of the tangent to the graph of f at the point $(x, f(x))$, if there is one (**remember** $|x|$). Call it $f'(x)$.

Basic formulae $\sin' = \cos$, $\cos' = -\sin$, $\exp' = \exp$, $\ln'(x) = \frac{1}{x}$ (only for $x > 0$!); if $f(x) = x^\alpha$ then $f'(x) = \alpha x^{\alpha-1}$.

Oops. Are we quite sure we know what we mean by x^α?

This is of course a matter of convention (just as it's a matter of convention what we mean by the word 'chair'). As with the argument of a complex number, there are some real problems in the definition, which mean that not all authors follow exactly the same rules. However, for $x > 0$ we have a general formula which works very reliably:

$$x^\alpha = \exp(\alpha \ln x) \text{ if } x > 0,\ \alpha \in \mathbb{R}.$$

(This is how most calculators do it. Can you think of a way to find out whether yours does?) It works fine unless we want to look at negative x. For integral n we have a different way to do it:

$$x^0 = 1, \qquad x^{n+1} = x^n \times x \text{ for integers } n \geq 0,$$

$$x^{-n} = \frac{1}{x^n} \text{ for integers } n \geq 1,\ x \neq 0.$$

(Can you find out whether your calculator uses this method?) Of course this means that for things like 2^3 we have two different definitions:

$$2^3 = \exp(3 \ln 2), \quad 2^3 = 2^2 \times 2 = 2^1 \times 2 \times 2 = 2^0 \times 2 \times 2 \times 2 = 1 \times 2 \times 2 \times 2$$

and if I were giving proofs in this course, I should have to take time out to prove that these two rules give the same answer.

Note in particular that on *my* rules, 0^0 comes out as 1. For the things we shall want to do in this course, this rule is what makes the formulae simplest, and I will stick to it.

Rules for differentiation

A. Suppose f and g are two real functions. Then

$$(f + g)' = f' + g';$$

the derivative of the sum is the sum of the derivatives. So, for instance,

$$\tfrac{d}{dx}(x^2 + x^3) = \tfrac{d}{dx}x^2 + \tfrac{d}{dx}x^3 = 2x + 3x^2.$$

But what does this mean if, for instance, we are looking at $\frac{d}{dx}(\ln x + \sqrt{1 - x^2})$? In the next few pages I will give conventions for dealing with combinations of partially-defined functions which, while not universally used, seem to me to be closest to what ordinary mathematicians and scientists actually do when they're thinking about it at all.

Given two real functions f and g with domains $\operatorname{dom} f$, $\operatorname{dom} g$ respectively, define the function $f + g$ by saying

$$\begin{aligned}\operatorname{dom}(f+g) &= \{x : f(x) + g(x) \text{ is properly defined}\} \\ &= \{x : x \in \operatorname{dom} f \text{ and } x \in \operatorname{dom} g\} = \operatorname{dom} f \cap \operatorname{dom} g,\end{aligned}$$

and write

$$(f+g)(x) = f(x) + g(x) \text{ for } x \in \operatorname{dom}(f+g) = \operatorname{dom} f \cap \operatorname{dom} g.$$

Now if x is a real number such that $f'(x)$ and $g'(x)$ are both defined, $(f+g)'(x)$ is also defined and is equal to $f'(x) + g'(x)$.

5. Two famous theorems

I want to spend a couple of lectures on two of the fundamental theorems of real analysis. The first is a very straightforward fact:

> **Intermediate Value Theorem: first form** Suppose that f is a continuous real function defined (at least) on the whole of a closed interval $[a, b]$, where $a < b$ in $\mathbb{R}$. If
>
> *either* $f(a) \leq w \leq f(b)$
> *or* $f(b) \leq w \leq f(a)$
>
> then there is some $x \in [a, b]$ such that $f(x) = w$.

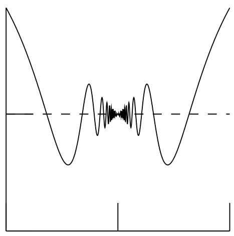
All this is saying is that if, in order to get from one point $(a, f(a))$ on the graph of f to another, $(b, f(b))$, it looks as if you have to cross the line $y = w$, then indeed you do cross that line at some definite point $(x, f(x)) = (x, w)$.

(Diagram not available.)

[Why do we make such a fuss about this? The reason is historical. When, in the nineteenth century, mathematicians were systematically working through the theory of real functions, the very concept of 'function' wasn't properly established, and what 'continuity' meant, or ought to mean, was by no means clear. The 'Intermediate Value Theorem' was one of the important tests by which people were persuaded that they were on the right track as they decided on the modern meanings for these terms.]

What is the theorem good for? Remember that it doesn't *solve* the equation $f(x) = w$. At best it tells you that there is a solution there. (At least one solution, that is – there could be one solution, or two, or three, or even infinitely many.)

(Diagram not available.)



However, this is sometimes very valuable. If you are asked to solve the equation

$$x + \sin x = 2,$$

for instance, you won't be able to find a formula for it. But if you look at the graph of $f(x) = x + \sin x$, it is clear that this crosses the line $y = 2$.

(Diagram not available.)

What is more, you can actually calculate in your head that

$$1 + \sin 1 < 2 < 2 + \sin 2$$

(because $0 < 1 < 2 < \pi$, so $\sin 1$ and $\sin 2$ both lie strictly between 0 and 1). So there has to be at least one solution somewhere in $[1, 2]$. To locate it more precisely, calculate

$$1 \cdot 5 + \sin 1 \cdot 5 \simeq 2 \cdot 5.$$

That tells you that $f(1) < 2 < f(1 \cdot 5)$, so there must be a solution somewhere in $[1, 1 \cdot 5]$. Look at

$$1 \cdot 2 + \sin 1 \cdot 2 \simeq 2 \cdot 13;$$

too big again, try

$$1 \cdot 1 + \sin 1 \cdot 1 \simeq 1 \cdot 99;$$

too low, so now we've got it narrowed down to the interval $[1 \cdot 1, 1 \cdot 2]$. Hunting through that interval, we find that

$$1 \cdot 11 + \sin 1 \cdot 11 \simeq 2 \cdot 01, \qquad 1 \cdot 105 + \sin 1 \cdot 105 \simeq 1 \cdot 998,$$

so we can turn our attention to the interval $[1 \cdot 105, 1 \cdot 11]$. Continuing this way, we can calculate the value of *a* solution to the given equation to as many decimal places as we like. But we already know that there is (at least one) solution which, to two decimal places, is $1 \cdot 11$.

Of course this is pretty tedious if all we have is a simple calculator. But any sort of programmable calculator should be able to do it in a twinkling (and I dare say many of them have the method pre-programmed). Where the mathematics comes in is in assuring us that there's something there for us to find.

I cheated a little when I picked the intervals above; I looked at $1 \cdot 11$ straight away, rather than calculating $f(1 \cdot 15)$, $f(1 \cdot 12)$ first. It would be a rather smart program that knew when to do that. But you don't need a smart program unless you're doing it hundreds of times. The dumbest possible program (the **repeated bisection method**), which just looks at the centre of the interval each time (for the example above, it would calculate $f(1)$ and $f(2)$ to get started, then $f(1 \cdot 5)$, $f(1 \cdot 25)$, $f(1 \cdot 125)$, $f(1 \cdot 0625)$, $f(1 \cdot 09375)$, $f(1 \cdot 109375)$, $\ldots$) would run fast enough to be quite adequate for solving individual problems.

Note that

(a) you do need a look at the picture to get started;

(b) if there's more than one solution the method will find one, but won't tell you about the others;

(c) you really do have to check that f is defined everywhere in $[a, b]$ and is continuous. For instance:

example if $f(x) = \frac{1}{x}$, then $f(-1) < 0 < f(1)$ but there is no x such that $f(x) = 0$

(Diagram not available.)

(what is happening here is that f is undefined at some point of $[-1, 1]$), and

example if $f(x)$ is the integer part of x, then $f(1) = 1 < \frac{3}{2} < 2 = f(2)$ but there is no x such that $f(x) = \frac{3}{2}$

(Diagram not available.)

(what is happening here is that f jumps at integers, and the graph falls apart).

On the other hand, provided the function f really is continuous, and you really can calculate it quickly and as accurately as you need to, the method is guaranteed to work, and you can be sure of getting an extra three decimal places for every ten calculations, because each calculation halves the interval in which you're working, so ten calculations reduce it by a factor of 1024.

There is another version of the theorem which is sometimes what we use:

> **Intermediate Value Theorem: second form** Suppose that f and g are continuous real functions defined (at least) on the whole of a closed interval $[a, b]$, where $a < b$ in $\mathbb{R}$. If
>
> *either* $f(a) \leq g(a)$ and $g(b) \leq f(b)$,
>
> *or* $g(a) \leq f(a)$ and $f(b) \leq g(b)$,
>
> then there is some $x \in [a, b]$ such that $f(x) = g(x)$.

(If it looks as if the graphs of the functions have to cross, they do indeed cross somewhere.) For instance,

$$\ln 1 = 0 < 1 = \frac{1}{1}, \quad \frac{1}{3} < 1 < \ln 3,$$

so there must be some $x \in [1, 3]$ such that $\ln x = \frac{1}{x}$.

(Diagram not available.)

Really this is just the same theorem. If we set $h = f - g$, then all it says is that

> if either $h(a) \leq 0 \leq h(b)$ or $h(a) \geq 0 \geq h(b)$ then there is an $x \in [a, b]$ such that $h(x) = 0$.

Mean Value Theorem

The next big theorem – quite a bit more sophisticated than the last – deals with *differentiable* functions rather than *continuous* ones. It says

> **Mean Value Theorem** If a real function f, defined (at least) on an interval $[a, b]$ where $a < b$ in $\mathbb{R}$, is continuous on the closed interval $[a, b]$ and differentiable on the open interval $]a, b[$, then there is an $x \in]a, b[$ such that
>
> $$f'(x) = \tfrac{f(b)-f(a)}{b-a}.$$

If you look at the right-hand side of the equation, $\frac{f(b)-f(a)}{b-a}$ is just the slope of the chord from $(a, f(a))$ to $(b, f(b))$. So what the theorem is saying is that for any chord joining two points of the graph of f, there is a point in between where the tangent is parallel to that chord.

(Diagram not available.)

I expect you find the phrase 'continuous on the closed interval and differentiable on the open interval' a bit of a mouthful. I put it in because it's the traditional expression, there's a whole family of theorems which use it (don't worry, we're only going to meet three of them altogether) and it means that we can deal with functions like $\sqrt{1 - x^2}$, which is continuous on $[-1, 1]$ and differentiable on $]-1, 1[$ (and has a tangent at $(0, 1)$ parallel to the chord between $(-1, 0)$ and $(1, 0)$).

(Diagram not available.)

Like the Intermediate Value Theorem, this is just an 'existence theorem'; it tells us that the equation $f'(x) = \frac{f(b)-f(a)}{b-a}$ has a solution, but not what that solution is (except, of course, that it's somewhere in $]a, b[$). In fact we rarely use it to solve the equation $f'(x) = w$; there are usually faster methods. (It's rather unusual to have a situation where you know that f is differentiable, you can calculate f efficiently, but you can't calculate f' efficiently; and if you have a direct method of calculating f', then the Intermediate Value Theorem is likely to give you quick solutions to equations of the form $f'(x) = w$, using repeated bisections as above.) In fact we use it oftener for a theoretical analysis. For instance:

Theorem If a, $b \in \mathbb{R}$, then $|\sin b - \sin a| \leq |b - a|$.

proof Suppose first that $a < b$. Because sin is differentiable (and continuous) everywhere, there must be some $x \in]a, b[$ such that $\sin'(x) = \frac{\sin(b)-\sin(a)}{b-a}$, that is, $\sin b - \sin a = (b - a)\cos x$, so that $|\sin b - \sin a| = |b - a||\cos x|$. But we know that $|\cos x| \leq 1$ (even though we have no idea what x is!), so we must have $|\sin b - \sin a| \leq |b - a|$.

So the theorem is true if $a < b$. If $a = b$ then

$$|\sin b - \sin a| = 0 = |b - a|,$$

while if $b < a$ then

$$|\sin b - \sin a| = |\sin a - \sin b| \leq |a - b| = |b - a|,$$

so the theorem is true in these cases also.

The sort of calculation we can make is

Example Suppose we know that f is a differentiable function defined (at least) on $[0, 2]$, that its derivative f' is also continuous, and that $f(0) = 1$, $f(1) = -1$, $f(2) = 0$. What can we tell about the set of values of f'?

answer Because $\frac{f(1)-f(0)}{1-0} = -2$, we know that there is some $x_1 \in]0, 1[$ such that $f'(x_1) = -2$. Because $\frac{f(2)-f(1)}{2-1} = 1$, we know that there is some $x_2 \in]1, 2[$ such that $f'(x_2) = 1$. Now we are supposing that f' is continuous, so the Intermediate Value Theorem tells us that it must take every value between -2 and 1; that is, that the set of values of f' includes $[-2, 1]$.

(Of course it could be much bigger.)

The 'ordinary' Mean Value Theorem above is very useful in this kind of theoretical investigation; but there's another version which goes even farther, the **Cauchy Mean Value Theorem**.

> **Theorem** Suppose that f and g are real functions, both defined (at least) on an interval $[a, b]$ where $a < b$ in $\mathbb{R}$, both continuous on the closed interval $[a, b]$

and differentiable on the open interval $]a, b[$. Suppose moreover that $g'(x) \neq 0$ for every $x \in]a, b[$. Then there is an $x \in]a, b[$ such that

$$\frac{f'(x)}{g'(x)} = \frac{f(b)-f(a)}{g(b)-g(a)}.$$

(This is the formulation you need if you ever want to prove that L'Hôpital's rule works as advertised, for instance.) The ordinary Mean Value Theorem is what you get if you set $g(x) = x$.

Proposition $\cos b \geq 1 - \frac{1}{2}b^2$ for every $b \in \mathbb{R}$.

proof Suppose first that $b > 0$. If we set $f(x) = \cos x$, $g(x) = x^2$, $a = 0$ in Cauchy's Mean Value Theorem, we see that there is some $x \in]0, b[$ such that

$$\frac{\cos b - \cos 0}{b^2 - 0^2} = \frac{\cos' x}{2x} = -\frac{\sin x}{2x},$$

that is,

$$\frac{1 - \cos b}{b^2} = \frac{\sin x}{2x}.$$

Now we know from the result above that

$$|\sin x - \sin 0| \leq |x - 0|, \text{ that is, } |\sin x| \leq |x|,$$

so that

$$\left|\frac{1-\cos b}{b^2}\right| = \frac{1}{2}\left|\frac{\sin x}{x}\right| \leq \frac{1}{2},$$

and

$$1 - \cos b \leq \frac{1}{2}b^2, \text{ that is, } \cos b \geq 1 - \frac{1}{2}b^2,$$

as required.

If $b < 0$, then $-b \geq 0$, so

$$\cos b = \cos(-b) \leq 1 - \frac{1}{2}(-b)^2 = 1 - \frac{1}{2}b^2,$$

while if $b = 0$ then $\cos b = 1 = 1 - \frac{1}{2}b^2$. So the result is true for all $b \in \mathbb{R}$.

filename ma1066.tex Version of 18.12.98

Taylor's Theorem

So far I've been talking of power series as if you just waited for an angel to reveal an interesting sequence $\langle c_r \rangle_{r\in\mathbb{N}}$ in a dream, and then worked out which function it represented. But of course there's a reverse approach which one uses quite as often. Recall the rule

$$\text{if } f(x) = \sum_{r=0}^{\infty} c_r x^r, \text{ then } c_r = \frac{1}{r!}f^{(r)}(0)$$

(provided, at least, that the series $\sum_{r=0}^{\infty} c_r x^r$ has radius of convergence greater than 0.) If we have a function f which we understand well enough to get a formula for its rth derivative, we can turn this backwards:

$f(x) = \sum_{r=0}^{\infty} \frac{1}{r!} f^{(r)}(0) x^r$

for x near 0 (with a little bit of luck). This is **Maclaurin's series**. For instance, starting with $f(x) = \sin x$, we have

$$f'(x) = \cos x, \quad f''(x) = -\sin x, \quad f^{(3)}(x) = -\cos x, \quad f^{(4)}(x) = \sin x = f(x)$$

and so on. This means that we can calculate $f^{(r)}(0)$ easily:

$$f^{(0)}(0) = f(0) = \sin 0 = 0, \quad f^{(1)}(0) = \cos 0 = 1,$$

$$f^{(2)}(0) = -\sin 0 = 0, \quad f^{(3)}(0) = -\cos 0 = -1,$$

$$f^{(4)}(0) = f(0) = 0, \quad f^{(5)}(0) = f'(0) = 1, \quad f^{(6)}(0) = f''(0) = 0$$

and so on. Generally

$$\begin{aligned} f^{(r)}(0) &= 0 \text{ if } r \text{ is even,} \\ &= 1 \text{ if } r = 4k+1, \\ &= -1 \text{ if } r = 4k+3; \end{aligned}$$

or, a little more manageably,

$$\begin{aligned} f^{(r)}(0) &= 0 \text{ if } r \text{ is even,} \\ &= (-1)^k \text{ if } r = 2k+1. \end{aligned}$$

So the corresponding Maclaurin series is

$$\begin{aligned} \sin x &= \sum_{r=0}^{\infty} \frac{1}{r!} f^{(r)}(0) x^r \\ &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^r. \end{aligned}$$

I gave this approach second place because there are theoretical problems with it. Recall once again the function $f(x) = e^{-1/x}$ for $x > 0$, $f(x) = 0$ for $x \le 0$; here all the derivatives $f^{(r)}(0) = 0$ are perfectly well defined but tell us nothing about the function f. In practice, however, it's always worth forming the Maclaurin series first and thinking about what it means afterwards.

Taylor series All the power series I've discussed so far have been centered on 0. This is much the easiest place to start and gives the prettiest formulae. But of course you can start anywhere else you like. If you want to understand the behaviour of a function f near a point x_0, the natural approach, in the present context, is to look at the **Taylor series**

$$f(x_0) + (x - x_0) f'(x_0) + \frac{(x-x_0)^2}{2!} f''(x_0) + \frac{(x-x_0)^3}{3!} f^{(3)}(x_0) + \dots + \frac{(x-x_0)^r}{r!} f^{(r)}(x_0) + \dots,$$

hoping that this will be summable and equal to $f(x)$ at least for x close to x_0. (The Maclaurin series is just the special case $x_0 = 0$.) For instance, instead of writing

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots \text{ for } -1 < x < 1$$

we could write

$$\ln x = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \dots \text{ when } 0 < x < 2.$$

We still have a notion of 'radius of convergence'. For a Taylor series $\sum_{r=0}^{\infty} c_r(x-x_0)^r$ there is an $R \in [0,\infty]$ such that the series is (absolutely) summable whenever $|x-x_0| < R$ and non-summable whenever $|x-x_0| > R$; we still call R the radius of convergence; we still have

$$f'(x) = \sum_{r=1}^{\infty} r c_r (x-x_0)^{r-1} = \sum_{r=0}^{\infty} (r+1)c_{r+1}(x-x_0)^r$$

when $|x-x_0| < R$.

Of course the fundamental question remains: when is this going to work? I have already said that you ought always to look at the radius of convergence of any power series; this will at least tell you where your series provides an honest function, and you can then try to determine its properties. But if we are building our series from a given function, we have another approach, called **Taylor's theorem with remainder**:

> **Theorem** Suppose that f is a function such that f and all its derivatives up to the nth are defined and continuous on (at least) a closed interval $[x_0, x]$ or $[x, x_0]$, and that the $n+1$st derivative is defined on the open interval $]x_0, x[$ or $]x, x_0[$. Then there is some ξ in the open interval such that
>
> $$f(x) - f(x_0) - (x-x_0)f'(x_0) - \frac{(x-x_0)^2}{2!}f''(x_0) - \dots - \frac{(x-x_0)^n}{n!}f^{(n)}(x_0) = \frac{(x-x_0)^{n+1}}{(n+1)!}f^{(n+1)}(\xi).$$

The sort of function for which this kind of estimation is useful is $f(x) = \cos x$. In this case, we know that (for any n) $f^{(n+1)}(\xi)$ is one of $\cos\xi$, $-\sin\xi$, $-\cos\xi$ or $\sin\xi$, so that $|f^{(n+1)}(\xi)| \le 1$ (regardless of what ξ is). So we get

$$\begin{aligned}
|\cos x - \cos x_0 + (x-x_0)\sin x_0 + \frac{(x-x_0)^2}{2!}\cos x_0 - \dots - \frac{(x-x_0)^n}{n!}\cos^{(n)} x_0| \\
= |\frac{(x-x_0)^{n+1}}{(n+1)}\cos^{(n+1)}\xi| \\
\le \frac{|x-x_0|^{n+1}}{(n+1)!}.
\end{aligned}$$

Since we can be sure that $\lim_{n\to\infty} \frac{(x-x_0)^n}{n!} = 0$ for every x, this tells us that the series always sums to the right answer. But for practical analysis it's just as important to be able to estimate the error, and that is what Taylor's theorem with remainder enables us to do. If $|x-x_0| \le 0\cdot 1$, for instance, the error in calculating $\cos x$ from its Taylor series, centered on x_0, up to the term in $(x-x_0)^n$, is at most $10^{-n-1}/(n+1)!$. So one way of programming a calculator to calculate $\cos x$, for $0 \le x \le \frac{\pi}{2}$, up to ten decimal places, would be to store values of $\cos x_0$ and $\sin x_0$ for $x_0 = 0,\ 0\cdot 2,\ 0\cdot 4, \ldots, 1\cdot 6$ and then program your calculator to

pick the nearest x_0 in the table,

calculate $\cos x_0 - (x - x_0)\sin x_0 - \frac{(x-x_0)^2}{2}\cos x_0 + \ldots - \frac{(x-x_0)^6}{6!}\cos x_0$,

since the error will be at most $10^{-7}/7! \simeq 2 \cdot 10^{-11}$ (provided, of course, that you do all your additions and subtractions and multiplications accurately enough).