

# Extra topics

Robin Chapman

26 November 2010

For interest's sake I talk about a few topics that are “beyond the scope of the course” but only just. If I had some more time I might have lectured on these.

## Another definition of continuity

The definition of continuous function I gave in the lectures is not the one found in most texts (although it is in Bryant's *Yet Another Introduction to Analysis*). The definition I gave is that a function  $f : A \rightarrow \mathbf{R}$  is continuous at a point  $a \in A$  if

$$\lim_{n \rightarrow \infty} f(a_n) = f(a) \quad \text{whenever} \quad \lim_{n \rightarrow \infty} a_n = a.$$

This definition is often called *sequential continuity*. The definition in most books is often called  $\varepsilon$ - $\delta$ -*continuity*. Here it is:

the function  $f : A \rightarrow \mathbf{R}$  is continuous at  $a \in A$  if for each  $\varepsilon > 0$  there is  $\delta > 0$  such that if  $x \in A$  and  $|x - a| < \delta$  then  $|f(x) - f(a)| < \varepsilon$ .

A similar definition is often used for limits of functions too.

I used the sequential continuity definition for two reasons:

- (a) the  $\varepsilon$ - $\delta$  definition appears more complicated;
- (b) we have previously done a lot of work on sequences, and it makes sense to exploit that.

If you read other texts or lecture notes, you may worry that the different definition of continuity may matter in some way. The good news is that it doesn't.

**Theorem.** *A function  $f : A \rightarrow \mathbf{R}$  is sequentially continuous at a point  $a \in A$  if and only if it is  $\varepsilon$ - $\delta$ -continuous at  $a$ .*

**Proof** Suppose first that  $f$  is sequentially continuous at  $a$ . In order to derive a contradiction assume that  $f$  is not  $\varepsilon$ - $\delta$ -continuous at  $a$ . Then there is  $\varepsilon > 0$  such that for all  $\delta > 0$  there is some  $x \in A$  such that  $|x - a| < \delta$  but that  $|f(x) - f(a)| \geq \varepsilon$ . For each  $n \in \mathbf{N}$  this is true for  $\delta = 1/n$ . Hence for each  $n \in \mathbf{N}$  there is  $x_n \in A$  such that  $|x_n - a| < 1/n$  and  $|f(x_n) - f(a)| \geq \varepsilon$ . But then the sequence  $(x_n)$  converges to  $a$  and the sequence  $(f(x_n))$  does not converge to  $f(a)$ . So  $f$  is not sequentially continuous at  $a$ , a contradiction. So sequential continuity implies  $\varepsilon$ - $\delta$ -continuity.

Conversely suppose that  $f$  is  $\varepsilon$ - $\delta$ -continuous at  $a$ . Let  $(x_n)$  be a sequence of points in  $A$  converging to  $a$ . To prove that  $f$  is sequentially continuous at  $a$ , it suffices to prove that the sequence  $(f(x_n))$  converges to  $f(a)$ . Let  $\varepsilon > 0$ . Then there is  $\delta > 0$  such that if  $x \in A$  and  $|x - a| < \delta$  then  $|f(x) - f(a)| < \varepsilon$ . Now eventually  $|x_n - a| < \delta$ , so eventually  $|f(x_n) - f(a)| < \varepsilon$ . As this is true for all  $\varepsilon > 0$ , then  $(f(x_n))$  converges to  $f(a)$ . Hence  $f$  is sequentially continuous at  $a$ .  $\square$

## Derivatives of power series

I omitted the proof that a function defined by a power series is differentiable. This was not because the proof was particularly hard, but rather the details are a bit fiddly. I give a proof here.

**Theorem.** *Let the power series*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

*have radius of convergence  $R$ . If  $|x_0| < R$  then  $f$  is differentiable at  $x_0$  with derivative*

$$\sum_{n=1}^{\infty} n a_n x_0^{n-1}.$$

**Proof** We define power series

$$g(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$$

and

$$h(x) = \sum_{n=2}^{\infty} n |a_n| (n-1) x^{n-2}.$$

To prove the theorem, we need to prove that the series for  $g(x)$  converges for  $|x| < R$ , and that

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = g(x_0).$$

The idea is to prove that there is some number  $C$  such that

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - g(x_0) \right| \leq C|x - x_0|$$

at least when  $x$  is close enough to  $x_0$ . The details are, alas, a bit awkward.

We first prove the convergence of  $g$ . Suppose that  $|x| < R$ . Then there is  $x_1$  with  $|x| < x_1 < R$ . The series  $\sum_{n=0}^{\infty} a_n x_1^n$  converges. It follows that the sequence  $(a_n x_1^n)$  is bounded. So there is  $M$  such that  $|a_n x_1^n| \leq M$  for all  $n \in \mathbf{N}$ . Therefore

$$|n a_n x^{n-1}| = n |a_n x_1^{n-1}| |x/x_1|^{n-1} \leq n \frac{M}{x_1} |x/x_1|^{n-1}.$$

As  $|x/x_1| < 1$ , the ratio test shows that the series

$$\sum_{n=1}^{\infty} n |x/x_1|^{n-1}$$

is convergent. By comparison, the series

$$\sum_{n=1}^{\infty} n a_n x^{n-1}$$

is absolutely convergent.

Repeating the above argument with  $g$  playing the rôle of  $f$  proves that the series for  $h(x)$  converges for  $|x| < R$ .

If  $|x_0| < R$ ,  $|x| < R$  and  $x \neq x_0$  then

$$f(x) - f(x_0) = \sum_{n=0}^{\infty} a_n x^n - \sum_{n=0}^{\infty} a_n x_0^n = \sum_{n=0}^{\infty} a_n (x^n - x_0^n) = \sum_{n=1}^{\infty} a_n (x^n - x_0^n).$$

Therefore

$$\frac{f(x) - f(x_0)}{x - x_0} = \sum_{n=1}^{\infty} a_n \frac{x^n - x_0^n}{x - x_0} = \sum_{n=1}^{\infty} a_n \sum_{j=0}^{n-1} x^{n-1-j} x_0^j$$

where I have used the identity

$$x^n - x_0^n = (x - x_0)(x^{n-1} + x^{n-1}x_0 + \cdots + x_0^{n-1}) = (x - x_0) \sum_{j=0}^{n-1} x^{n-1-j} x_0^j.$$

Then

$$\begin{aligned}
\frac{f(x) - f(x_0)}{x - x_0} - g(x_0) &= \sum_{n=1}^{\infty} a_n \sum_{j=0}^{n-1} x^{n-1-j} x_0^j - \sum_{n=1}^{\infty} n a_n x_0^{n-1} \\
&= \sum_{n=1}^{\infty} a_n \left( \sum_{j=0}^{n-1} x^{n-1-j} x_0^j - n x_0^{n-1} \right) \\
&= \sum_{n=2}^{\infty} a_n \sum_{j=0}^{n-1} (x^{n-1-j} x_0^j - x_0^{n-1}) \\
&= \sum_{n=2}^{\infty} a_n \sum_{j=0}^{n-1} x_0^j (x^{n-1-j} - x_0^{n-j-1}) \\
&= (x - x_0) \sum_{n=2}^{\infty} a_n \sum_{j=0}^{n-1} x_0^j \sum_{k=0}^{n-2-j} x^{n-2-j-k} x_0^k.
\end{aligned}$$

(You see what I mean about the details being fiddly). Note that I shifted the start of the summation from  $n = 1$  to  $n = 2$  as the  $n = 1$  term cancels out. There is a number  $y$  such that  $|x_0| < y < R$ . If in addition  $|x| < y$  then

$$\begin{aligned}
\left| \frac{f(x) - f(x_0)}{x - x_0} - g(x_0) \right| &\leq |x - x_0| \sum_{n=2}^{\infty} \left| a_n \sum_{j=0}^{n-1} x_0^j \sum_{k=0}^{n-2-j} x^{n-2-j-k} x_0^k \right| \\
&\leq |x - x_0| \sum_{n=2}^{\infty} |a_n| \sum_{j=0}^{n-1} y^j \sum_{k=0}^{n-2-j} y^{n-2-j-k} y^k \\
&= |x - x_0| \sum_{n=2}^{\infty} |a_n| \sum_{j=0}^{n-1} \sum_{k=0}^{n-2-j} y^{n-2} \\
&= |x - x_0| \sum_{n=2}^{\infty} |a_n| \sum_{j=0}^{n-1} (n - j - 1) y^{n-2} \\
&= |x - x_0| \sum_{n=2}^{\infty} |a_n| \frac{n(n-1)}{2} y^{n-2} \\
&= \frac{|x - x_0|}{2} h(y).
\end{aligned}$$

If  $(x_n)$  is a sequence converging to  $x_0$ , with  $x_n \neq x_0$  then eventually  $|x_n| < y$ . Eventually then,

$$\left| \frac{f(x_n) - f(x_0)}{x_n - x_0} - g(x_0) \right| \leq \frac{h(y)}{2} |x_n - x_0|.$$

As  $x_n \rightarrow x_0$  then

$$\frac{f(x_n) - f(x_0)}{x_n - x_0} \rightarrow g(x_0).$$

Hence  $f'(x_0)$  exists and equals  $g(x_0)$ .  $\square$

As an immediate corollary  $f$  is continuous in the interval  $(-R, R)$ . Iterating the argument shows that  $f$  is differentiable any number of times in this interval, and that the  $k$ -th derivative of  $f$  is given by

$$f^{(k)}(x) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} a_n x^{n-k}.$$

## Cauchy's mean value theorem and l'Hôpital's rule

L'Hôpital's rule always seems popular with students. It has what I'll call a weak and a strong version. The weak version is easier to prove.

**Theorem.** *Let  $f$  and  $g$  be continuous and differentiable functions on an interval  $I$ . Let  $a \in I$  and suppose that  $f(a) = g(a) = 0$ . Assume that  $g'(x) \neq 0$  for all  $x \in I$ . Then*

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

**Proof** As  $f(a) = g(a) = 0$  then for  $x \neq a$

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f(x) - f(a)}{x - a} \left( \frac{g(x) - g(a)}{x - a} \right)^{-1}.$$

By the algebra of limits, as

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = f'(a) \quad \text{and} \quad \lim_{x \rightarrow a} \frac{g(x) - g(a)}{x - a} = g'(a)$$

then

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

(To be pedantic, we needn't worry about the possibility  $g(x) - g(a) = 0$ ; this is impossible by Rolle's theorem and the nonvanishing of  $g'$ .)  $\square$

There is a stronger version of l'Hôpital's rule, that can be iterated. The above version can't be applied when  $g'(a) = 0$ .

The proof of the strong version depends on Cauchy's mean value theorem, which is a generalization of the vanilla mean value theorem.

**Theorem.** *Let  $f$  and  $g$  be continuous functions on the interval  $[a, b]$ . Suppose also that  $f$  and  $g$  are differentiable on  $(a, b)$  and that  $g'$  is nonzero on  $(a, b)$ . Then there is  $t \in (a, b)$  with*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(t)}{g'(t)}.$$

The vanilla mean value theorem is the special case with  $g(x) = x$ . As with the mean value theorem, this is proved using a cunning application of Rolle's theorem.

**Proof** We remark that the statement makes sense: we cannot have  $g(b) - g(a) = 0$  for then by Rolle's theorem there would be  $c \in (a, b)$  with  $g'(c) = 0$ , contrary to hypothesis.

Define

$$h(x) = (f(b) - f(a))g(x) - (g(b) - g(a))f(x).$$

Then

$$h(a) = (f(b) - f(a))g(a) - (g(b) - g(a))f(a) = f(b)g(a) - g(b)f(a)$$

and

$$h(b) = (f(b) - f(a))g(b) - (g(b) - g(a))f(b) = -f(a)g(b) + g(a)f(b).$$

Therefore  $h(a) = h(b)$ . As  $h$  is continuous on  $[a, b]$  and differentiable on  $(a, b)$  then by Rolle's theorem there is  $t \in (a, b)$  with  $h'(t) = 0$ . But

$$h'(t) = (f(b) - f(a))g'(t) - (g(b) - g(a))f'(t)$$

and so

$$(f(b) - f(a))g'(t) = (g(b) - g(a))f'(t).$$

Equivalently

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(t)}{g'(t)}.$$

□

One consequence of Cauchy's mean value theorem is L'Hôpital's rule. I give a careful statement and proof.

**Theorem.** *Let  $f$  and  $g$  be continuous functions on an interval  $I$ . Let  $a \in I$  and suppose that  $f(a) = g(a) = 0$ . Assume that  $f$  and  $g$  are differentiable at all points of  $I$  save possibly  $a$  and that  $g'(x) \neq 0$  for all  $x \in I$  with  $x \neq a$ . If*

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

*exists, then*

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = L.$$

**Proof** Let  $(x_n)$  be a sequence of numbers in  $I$  converging to  $a$  with  $x_n \neq a$  for all  $n$ . By Cauchy's mean value theorem, for each  $n$  there is some  $t_n$  between  $a$  and  $x_n$  such that

$$\frac{f(x_n) - f(a)}{g(x_n) - g(a)} = \frac{f'(t_n)}{g'(t_n)}.$$

As  $f(a) = g(a) = 0$ , this says that

$$\frac{f(x_n)}{g(x_n)} = \frac{f'(t_n)}{g'(t_n)}.$$

Since  $|t_n - a| < |x_n - a|$  and  $x_n \rightarrow a$  it follows that  $t_n \rightarrow a$ . Hence

$$\frac{f'(t_n)}{g'(t_n)} \rightarrow \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} = L$$

and so

$$\frac{f(x_n)}{g(x_n)} \rightarrow L.$$

We conclude that

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = L.$$

□

The advantage of the strong form of l'Hôpital's rule is that it can be iterated. Suppose  $f$  and  $g$  are nice functions with lots of derivatives. If  $f'(a) = g'(a) = 0$  the weak form is flummoxed, but with the strong form we can turn the handle again:

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

and so on.

However I don't like l'Hôpital's rule. If one needs to iterate the procedure more than once, the higher derivatives of  $f$  and  $g$  can become cumbersome. The l'Hôpital fan is urged to try evaluating the limit

$$\lim_{x \rightarrow 0} \frac{\sin x - \sinh x}{x \log(1+x) \log(1-x)}$$

using l'Hôpital's rule. As I demonstrated in the lectures, this example yields very easily to my favoured method. This is to write

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow 0} \frac{f(x+a)}{g(x+a)}$$

and then expand  $f(x+a)$  and  $g(x+a)$  as power series in  $x$ . In fact, we only need the first nonzero term in the numerator and denominator.

Another annoyance I find is circular reasoning involving l'Hôpital's rule. I spent a while in the lectures using geometric arguments to prove that

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1. \quad (*)$$

Over the years I have been often asked why one can't just use l'Hôpital's rule to prove this. The argument is apparently simple: take  $f(x) = \sin x$ ,  $g(x) = x$  and  $a = 0$  in the weak form of l'Hôpital's rule to get

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = \lim_{x \rightarrow 0} \frac{f(x)}{g(x)} = \frac{f'(0)}{g'(0)} = \frac{\cos 0}{1} = 1.$$

But this uses the fact that  $\sin' 0 = 1$ . That **means** that

$$\lim_{x \rightarrow 0} \frac{\sin x - \sin 0}{x - 0} = 1$$

which is precisely the assertion that

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1.$$

So prove  $(*)$  via l'Hôpital's rule we need to **use** the fact that  $(*)$  is true! This temptation towards implicit circular reasoning is yet another reason I dislike l'Hôpital's rule.

## Decimals

I will only deal with decimal expansions of numbers between 0 and 1. In general you get other real numbers by adding integers to these.

To me, a decimal expansion  $0 \cdot a_1 a_2 a_3 \dots$  is just a sequence  $(a_n)$  of *digits*, that is, each  $a_n \in \mathcal{D} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ . Thus  $0 \cdot a_1 a_2 a_3 \dots$  amounts to a mapping  $n \mapsto a_n$  from  $\mathbf{N}$  to  $\mathcal{D}$ . The *value* of  $0 \cdot a_1 a_2 a_3 \dots$  is defined to be the sum of the series

$$\sum_{n=1}^{\infty} \frac{a_n}{10^n}.$$

This series of nonnegative terms is convergent by comparison with the geometric series

$$\sum_{n=1}^{\infty} \frac{9}{10^n}.$$

We use  $0 \cdot a_1 a_2 a_3 \dots$  as notation for its value, that is

$$0 \cdot a_1 a_2 a_3 \dots = \sum_{n=1}^{\infty} \frac{a_n}{10^n}.$$

In particular when all  $a_n = 9$  we have

$$0 \cdot 999 \dots = \sum_{n=1}^{\infty} \frac{9}{10^n} = \frac{9}{10} \frac{1}{1 - 1/10} = 1.$$

In general

$$0 = 0 \cdot 000 \dots \leq 0 \cdot a_1 a_2 a_3 \dots \leq 0 \cdot 999 \dots = 1.$$

Any number  $x$  with  $0 < x < 1$  and having a finite decimal expansion, has two decimal expansions, one ending in a sequence of zeroes and the other in a sequence of nines. For example

$$0 \cdot 436 = 0 \cdot 436000 \dots = 0 \cdot 435999 \dots.$$

We claim that every number in  $[0, 1]$  has a decimal expansion. We have seen this is the case for 0 and 1 so let  $0 < x < 1$ . We define a sequence  $(a_n)$  recursively such that  $a_n \in \mathcal{D}$  and

$$0 \leq x - \sum_{k=1}^n \frac{a_k}{10^k} < 1. \quad (*)$$

To define  $a_1$  consider  $10x$ . Then  $0 < 10x < 10$  so let  $a_1$  be the integer part of  $10x$ , namely the largest integer with  $a_1 \leq 10x$ . Then  $a_1 \in \mathcal{D}$  and  $a_1 \leq 10x < a_1 + 1$ , that is  $0 \leq 10x - a_1 < 1$  or

$$0 \leq x - \frac{a_1}{10} < \frac{1}{10}.$$

Suppose that  $a_1, \dots, a_n$  have been defined. Let  $b_n = \sum_{k=1}^n a_k / 10^k$ . Then

$$0 \leq 10^n x - 10^n b_n < 1$$

so that

$$0 \leq 10^{n+1} x - 10^{n+1} b_n < 10.$$

Let  $a_{n+1}$  be the integer part of  $10^{n+1} x - 10^{n+1} b_n$ . Then  $a_{n+1} \in \mathcal{D}$  and

$$0 \leq 10^{n+1} x - 10^{n+1} b_n - a_{n+1} < 1.$$

Hence

$$0 \leq x - \left( b_n + \frac{a_{n+1}}{10^{n+1}} \right) < \frac{1}{10^{n+1}}.$$

But

$$b_n + \frac{a_{n+1}}{10^{n+1}} = \sum_{k=1}^n \frac{a_k}{10^k} + \frac{a_{n+1}}{10^{n+1}} = \sum_{k=1}^{n+1} \frac{a_k}{10^k}$$

so that

$$0 \leq x - \sum_{k=1}^{n+1} \frac{a_k}{10^k} < \frac{1}{10^{n+1}}.$$

This gives the induction step, so there is a decimal expansion  $0 \cdot a_1 a_2 a_3 \dots$  such that  $(*)$  holds for all  $n$ .

From  $(*)$  it is apparent that

$$\left| x - \sum_{k=1}^n \frac{a_k}{10^k} \right| < \frac{1}{10^n}.$$

Hence the sequence  $(x - \sum_{k=1}^n a_k/10^k)$  is a null sequence and so

$$x = \sum_{k=1}^{\infty} \frac{a_k}{10^k} = 0 \cdot a_1 a_2 a_3 \dots$$

has a decimal expansion.

Of course there is nothing special about decimals. The obsession with the base ten stems from an accidental fact of primate physiology. One could repeat the above with ten replaced by any integer  $b \geq 2$ . The set of digits would be  $\mathcal{D}_b = \{0, 1, \dots, b-1\}$  and the result is that every number  $x \in [0, 1]$  has a representation

$$x = \sum_{n=1}^{\infty} \frac{a_n}{b^n}$$

with each  $n \in \mathcal{D}_b$ . Again for largely accidental reasons, there is a minor cultural obsession with the case  $b = 2$ ,