MATH 298 Q1-Sem - Winter 2022

Problem solving seminar

Contents

1 Continuous functions

Wednesday, January 5, 2022

Let's recall briefly what it means for a function to be continuous. For most of the functions that you normally encounter in calculus, you can imagine that the function is continuous on the interval $[a, b]$ when you can trace its graph without lifting your pencil. Of course, this is not mathematically rigorous, but there's nothing wrong with keeping this picture in mind. One problem with this is that it may simply not be possible at all to draw the graph of a function, even when it is continuous almost everywhere on its domain.

Continuity of a function at a point can be defined using ϵ and δ. It can also be defined using sequences, which is a point of view that we will use also below.

Definition 1.1. *Suppose that the function f is defined on the open interval (a, b) and $c \in (a, b)$. f is continuous at c if $\forall\, \epsilon > 0\ \exists\, \delta > 0$ such that if $|x - c| < \delta$ and $x \in (a, b)$ then $|f(x) - f(c)| < \epsilon$.*

Definition 1.2. *Suppose that the function f is defined on the open interval (a, b) and $c \in (a, b)$. f is continuous at c if, for every sequence $\{x_n\}_{n=1}^{\infty}$ in (a, b) with $\lim_{n\to\infty} x_n = c$, we have $\lim_{n\to\infty} f(x_n) = f(c)$.*

This definition of continuity may be the most convenient when one wants to show that a function f is not continuous at some point c. Here is a classical example of a function not continuous at 0 and with rather strange behavior near 0: $f(x) = \sin\left(\frac{1}{x}\right)$ for $x \neq 0$ and $f(0) = 0$. As x gets closer to 0, the graph of this function oscillates faster and faster between -1 and 1. To see that $f(x)$ is not continuous at 0, it is enough to find one sequence $\{x_n\}_{n=1}^{\infty}$ that converges to 0 but for which $\{f(x_n)\}_{n=1}^{\infty}$ does not converge to $f(0)$. There are many possibilities. One of them is the sequence such that $\frac{1}{x_n} = 2\pi n + \frac{\pi}{2}$, that is $x_n = \frac{1}{2\pi n + \frac{\pi}{2}}$:

$$\lim_{n\to\infty} \frac{1}{2\pi n + \frac{\pi}{2}} = 0 \text{ but } f(x_n) = \sin\left(2\pi n + \frac{\pi}{2}\right) = 1\ \forall\, n \geqslant 1.$$

Here are two fundamental theorems about continuous functions which you must have seen in your first calculus course.

Theorem 1.3 (Extreme Value Theorem)**.** *If f is a continuous function on the closed interval $[a, b]$, then there exist $c, d \in [a, b]$ such that*

$$f(c) \leqslant f(x) \leqslant f(d)\ \forall\, x \in [a, b].$$

In other words, f reaches its minimum and its maximum values on $[a, b]$.

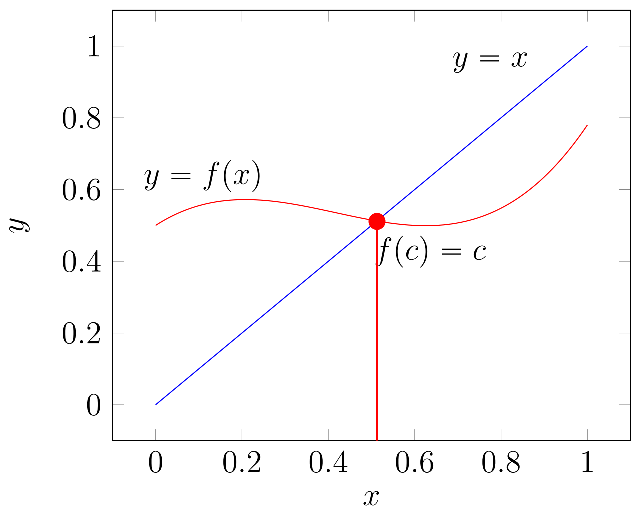
Theorem 1.4 (Intermediate Value Theorem)**.** *If f is a continuous function on $[a, b]$ and $f(a) < y < f(b)$ or $f(a) > y > f(b)$, there exists $c \in (a, b)$ such that $f(c) = y$. It can also be stated with $f(a) \leqslant y \leqslant f(b)$ or $f(a) \geqslant y \geqslant f(b)$, in which case the open interval (a, b) should be replaced by the closed interval $[a, b]$.*

Here is an example of application of the Intermediate Value Theorem which you may have seen before. It has to do with the notion of fixed point which plays an important role in different areas of mathematics: dynamical systems, topology, etc.

Example 1: Let $f : [a, b] \to [a, b]$ be a continuous function. Prove that $\exists\, c \in [a, b]$ such that $f(c) = c$.

A number c such that $f(c) = c$ is called a fixed point of f.

Solution: Before giving the rigorous solution, let's look at a picture in the case $a = 0, b = 1$ which makes it very convincing that such a c exists.



To prove the existence of c, the Intermediate Value Theorem is needed, but it should not be applied to f. Instead, it should be applied to another function, but which one? This requires some thinking and a good idea, which is to let $g(x) = f(x) - x$. This is also a continuous function on $[a, b]$. Then $g(a) = f(a) - a \geqslant 0$ since $f(a) \in [a, b]$ and $g(b) = f(b) - b \leqslant 0$ since $f(b) \in [a, b]$. By the Intermediate Value Theorem applied to g, there exists $c \in [a, b]$ such that $g(c) = 0$. This means that $f(c) - c = 0$, hence $f(c) = c$. $\square$

Example 2 below requires the following observation.

Lemma 1.5. *Let $f : [a, b] \longrightarrow \mathbb{R}$ be a continuous function which is injective (that is, one-to-one). Then f is either strictly increasing or strictly decreasing on $[a, b]$.*

Note that a function is increasing exactly when $x < y \Rightarrow f(x) < f(y)$.

Example 2: Let $f : [0, 1] \to [0, 1]$ be a continuous function such that $f(0) = 0$ and $f(1) = 1$. Fix a positive integer n. Let f^n be the composite of f with itself n times. (For instance, if $n = 3$, then $f^3(x) = f(f(f(x)))$.) Suppose that $f^n(x) = x\ \forall\, x \in [0, 1]$. Prove that $f(x) = x$ $\forall\, x \in [0, 1]$.

Solution: First, let's see that f is injective (one-to-one):

$$\begin{aligned} f(x_1) = f(x_2) &\longrightarrow f^n(x_1) = f^n(x_2) \text{ after applying } f^{n-1} \\ &\longrightarrow x_1 = x_2. \end{aligned}$$

Since f is injective and continuous on $[0,1]$ it follows that it must be either increasing or decreasing according to Lemma 1.5. Since $f(0) = 0$ and $f(1) = 1$, f is increasing.

The rest of the proof is by contradiction. Suppose that $\exists\, x_0 \in [0,1]$ such that $f(x_0) \neq x_0$. This means that either $f(x_0) > x_0$ or $f(x_0) < x_0$.

Case 1: $f(x_0) > x_0$

$$f(x_0) > x_0 \Rightarrow f(f(x_0)) > f(x_0) \Rightarrow f(f(f(x_0))) > f(f(x_0)) \Rightarrow \cdots \Rightarrow f^n(x_0) > f^{n-1}(x_0).$$

Therefore,

$$f^n(x_0) > f^{n-1}(x_0) > \cdots > f^3(x_0) > f^2(x_0) > f(x_0) > x_0,$$

which leads to the contradiction $x_0 > x_0$ because $f^n(x_0) = x_0$.

Case 2: $f(x_0) < x_0$ The same argument works as in the first case, but with all the inequalities reversed. The contradiction comes from:

$$x_0 = f^n(x_0) < f^{n-1}(x_0) < \cdots < f^3(x_0) < f^2(x_0) < f(x_0) < x_0.$$

Since a contradiction has been obtained when $f(x_0) \neq x_0$ for some $x_0 \in [0,1]$, it follows that $f(x) = x \;\forall\, x \in [0,1]$. $\square$

Wednesday, January 12, 2022

Example 3: Langlands runs 10 kilometers on a road in 30 minutes. Prove that, somewhere along, he ran one kilometer in exactly 3 minutes.

First solution: Let x denote the distance in kilometers along the road from where he starts. For $x \in [0,9]$, let $f(x)$ be the time it took Langlands to run from point x to point $x+1$ (so on a distance of one kilometer starting at x). Since it took him 30 minutes to run 10 kilometers,

$$f(0) + f(1) + f(2) + f(3) + f(4) + f(5) + f(6) + f(7) + f(8) + f(9) = 30.$$

It follows that there are integers $a, b \in \{0, 1, 2, \ldots, 9\}$ such that $f(a) \leqslant 3 \leqslant f(b)$. Indeed, if $f(a) > 3$ for $a = 0, 1, 2, \ldots, 9$, then

$$f(0) + f(1) + f(2) + f(3) + f(4) + f(5) + f(6) + f(7) + f(8) + f(9) > 3 \cdot 10 = 30,$$

but this sum is equal to 30, so this is a contradiction unless $f(a) \leqslant 3$ for at least one a in the interval $[0, 9]$. The same argument can be applied to ensure the existence of b such that $3 \leqslant f(b)$.

By the Intermediate Value Theorem, there exists c between a and b such that $f(c) = 3$. This means that Langlands runs from point c along the road to point $c + 1$ in exactly 3 minutes. $\square$

Second solution: Instead of dividing the 10 kilometers into sections of one kilometer each, we can instead partition the 30 minutes into 10 segments of three minutes each. Let d_i be the distance travelled between time $3i$ and time $3i + 3$ for $i = 0, 1, \ldots, 9$. Since

$$d_0 + d_1 + \cdots + d_9 = 10,$$

it follows that, for some i with $0 \leqslant i \leqslant 8$, we have that either $d_i \geqslant 1, d_{i+1} \leqslant 1$ or $d_i \leqslant 1, d_{i+1} \geqslant 1$.

Let's consider only the case $d_i \geqslant 1, d_{i+1} \leqslant 1$, the other one being analogous. For $0 \leqslant t \leqslant 27$, let $f(t)$ be the distance travelled between time t and $t + 3$. By assumption, $f(3i) \geqslant 1$ and $f(3i + 3) \leqslant 1$. By the Intermediate Value Theorem, there exists a time $t_0 \in [3i, 3i + 3]$ such that $f(t_0) = 1$: this means that, from time t_0 to $t_0 + 3$, Langlands ran exactly three kilometers. $\square$

Example 4: Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous and decreasing function. Prove that f has a unique fixed point x_0 (that is, $f(x_0) = x_0$).

Solution: We have to prove first the existence of at least one fixed point and then prove that there is exactly one.

As in Example 1, let $g(x) = f(x) - x$, so we have to see that g vanishes exactly once because $g(x_0) = 0 \iff f(x_0) = x_0$.

Since f is continuous, g is also continuous. Pick any $x_1 \in \mathbb{R}$. If $f(x_1) = x_1$, then we are done, so let's assume that $f(x_1) \neq x_1$. Then either $f(x_1) > x_1$ or $f(x_1) < x_1$.

Case 1: $f(x_1) > x_1$.

Then $g(x_1) = f(x_1) - x_1 > 0$. In order to be able to apply the Intermediate Value Theorem, we need to find x_2 such that $g(x_2) < 0$. Since f and $-\mathrm{id}$ are decreasing functions, g is also a decreasing function, so this suggests that we just need to select a number x_2 which is large enough. Here are three ways to complete the proof.

Subproof 1: Choose x_2 such that $x_2 > x_1$ and $x_2 > f(x_1)$. Since f is decreasing and $x_2 > x_1$, $f(x_2) < f(x_1)$. Then

$$g(x_2) = f(x_2) - x_2 < f(x_1) - x_2 < 0.$$

We have $g(x_1) > 0$ and $g(x_2) < 0$ so, by the Intermediate Value Theorem, the exists x_0 between x_1 and x_2 such that $g(x_0) = 0$, hence $f(x_0) = x_0$. $\square$

Subproof 2: Another choice for x_2 is to set $x_2 = f(x_1)$. Since $f(x_1) > x_1$ and since f is decreasing, applying f to both sides of the previous inequality reverses it: $f(f(x_1)) < f(x_1)$, so $f(x_2) < x_2$. Therefore, $g(x_2) = f(x_2) - x_2 < 0$, and the same argument as in the previous paragraph can be applied. $\square$

Subproof 3: Observe that

$$g(x+1) - g(x) = f(x+1) - (x+1) - (f(x) - x) = f(x+1) - f(x) - 1 < -1$$

since f is decreasing. It follows by induction from $g(x+1) - g(x) < -1$ that $g(x+n) - g(x) < -n$. In particular, $g(x_1 + n) < g(x_1) - n$, so if n is chosen so that $g(x_1) - n < 0$, then $g(x_1 + n) < 0$: x_2 can thus be taken to be $x_1 + n$. The rest of the proof is as in the first subproof. $\square$

Case 2: $f(x_1) < x_1$.

In the second case, $g(x_1) = f(x_1) - x_1 < 0$. That g is a decreasing function means that, as x increases, $g(x)$ decreases: this also means that, as x decreases, $g(x)$ increases. Therefore, to find x_2 such that $g(x_2) > 0$, we need to choose x_2 small enough. Choose x_2 such that $x_2 < x_1$ and $x_2 < f(x_1)$. Since f is decreasing, $f(x_2) > f(x_1)$. Therefore,

$$g(x_2) = f(x_2) - x_2 > f(x_1) - x_2 > 0.$$

We have $g(x_1) < 0$ and $g(x_2) > 0$ so, by the Intermediate Value Theorem, the exists x_0 between x_1 and x_2 such that $g(x_0) = 0$, hence $f(x_0) = x_0$. As in Case 1, we could have used two other approaches to find x_2.

Finally, it remains to prove that the fixed point is unique. Suppose that $f(x_0) = x_0$ and $f(\widetilde{x}_0) = \widetilde{x}_0$. Without loss of generality, suppose that $x_0 \leqslant \widetilde{x}_0$. Since f is decreasing, $f(x_0) \geqslant f(\widetilde{x}_0)$, which implies that $x_0 \geqslant \widetilde{x}_0$. It follows that the only possibility is that $\widetilde{x}_0 = x_0$.

Another way to prove the uniqueness of the fixed point x_0 is to observe that g is a strictly decreasing function, so g is injective (one-to-one) and thus x_0 is the only number such that $g(x_0) = 0$. $\square$

Wednesday, January 19, 2022

The next example about a continuous function showcases a functional equation, that is, an equation where the unknown is a function which must be identified.

Example 5: Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a continuous function such that

$$f(x_1 + x_2) = f(x_1) + f(x_2). \tag{1}$$

Prove that there exists a real number a such that $f(x) = ax$ for all $x \in \mathbb{R}$.

Solution: Let's play with the equation and see if we can start with a few basic deductions.

$f(0+0) = f(0) + f(0)$, hence $f(0) = 0$.

$f(1+1) = f(1) + f(1)$, so $f(2) = 2f(1)$.

$f(2+1) = f(2) + f(1) = 2f(1) + f(1) = 3f(1)$, so $f(3) = 3f(1)$.

$f(3+1) = f(3) + f(1) = 3f(1) + f(1) = 4f(1)$, hence $f(4) = 4f(1)$.

Looking at these, it is natural to conjecture that $f(m) = mf(1)$ for all integers $m \geqslant 0$. This can be proved by induction. The first step is to check this equation when $m = 0$ and $m = 1$. For the main inductive step, let's assume that $f(m) = mf(1)$ is true and let's prove this for $m+1$:

$$\begin{aligned} f(m+1) &= f(m) + f(1) \text{ from (1);} \\ &= mf(1) + f(1) \text{ (by induction)} \\ &= (m+1)f(1) \end{aligned}$$

Observe that if we set $x_1 = x$ and $x_2 = -x$ in (1), then $f(x-x) = f(x) + f(-x)$, so $0 = f(x) + f(-x)$ and

$$f(-x) = -f(x). \tag{2}$$

Therefore, it follows that, if $-m < 0$, then $f(-m) = -f(m) = -mf(1)$.

The main idea that is needed in this proof is that, since f is continuous, it is enough to show that $f(x) = ax$ for all rational numbers x. Indeed, if this is known and x is now any real number, pick a sequence of rational numbers $\{x_n\}_{n=1}^{\infty}$ that converges to x, that is, $\lim_{n\to\infty} x_n = x$ and each x_n is a fraction. Then

$$\begin{aligned} f(x) &= f\left(\lim_{n\to\infty} x_n\right) \\ &= \lim_{n\to\infty} f(x_n) \text{ (since } f \text{ is continuous)} \\ &= \lim_{n\to\infty} ax_n \text{ (since } f(x_n) = ax_n \text{ for rational numbers)} \\ &= ax \end{aligned}$$

Therefore, we are left to show that $f(x) = ax$ for all rational numbers x. Let's try to see how to prove this for $x = \frac{1}{2}$.

$f\left(\frac{1}{2} + \frac{1}{2}\right) = f\left(\frac{1}{2}\right) + f\left(\frac{1}{2}\right) = 2f\left(\frac{1}{2}\right)$, so $f\left(\frac{1}{2}\right) = \frac{1}{2}f(1)$.

$f\left(\frac{2}{3} + \frac{1}{3}\right) = f\left(\frac{2}{3}\right) + f\left(\frac{1}{3}\right) = f\left(\frac{1}{3} + \frac{1}{3}\right) + f\left(\frac{1}{3}\right) = f\left(\frac{1}{3}\right) + f\left(\frac{1}{3}\right) + f\left(\frac{1}{3}\right) = 3f\left(\frac{1}{3}\right)$, so $f\left(\frac{1}{3}\right) = \frac{1}{3}f(1)$.

This suggests trying to prove by induction the following identity: $f(nx) = nf(x)$ for all $n \in \mathbb{Z}$. This can be done as above, the case $n = 1$ being trivial. For the inductive step, assuming that $f(nx) = nf(x)$ is true, we get:

$$\begin{aligned} f((n+1)x) &= f(nx + x) \\ &= f(nx) + f(x) \\ &= nf(x) + f(x) \text{ (by induction)} \\ &= (n+1)f(x) \end{aligned}$$

The case when $n < 0$ can be dealt with using (2).

The final step is to combine $f(nx) = nf(x)$ with $f(m) = mf(1)$ in the case when x is a rational number. In this case, $x = \frac{m}{n}$, so $f(nx) = nf(x)$ implies that $f\left(n \cdot \frac{m}{n}\right) = nf\left(\frac{m}{n}\right)$, hence $f(m) = nf\left(\frac{m}{n}\right)$ and

$$f\left(\frac{m}{n}\right) = \frac{f(m)}{n} = \frac{m}{n} f(1).$$

This conclude the proof if we set $a = f(1)$. $\square$

Example 6: Prove that there is no continuous function $f : \mathbb{R} \longrightarrow \mathbb{R}$ that takes every value exactly twice. In other words, if a continuous function $f : \mathbb{R} \longrightarrow \mathbb{R}$ takes on no value more than twice, then it must take on some value exactly once.

Solution: Suppose that $f : \mathbb{R} \longrightarrow \mathbb{R}$ takes no value more than twice. Pick a, b such that $a < b$ and $f(a) = f(b)$. Consider the maximum M of f on the interval $[a, b]$. Since f is continuous on the closed interval $[a, b]$, the Extreme Value Theorem can be applied: $M = f(c)$ for some $c \in [a, b]$. If $M = f(a) = f(b)$, then consider the minimum m of f on $[a, b]$. Again, by the Extreme Value Theorem, $m = f(d)$ for some $d \in [a, b]$. If $M = f(a) = f(b)$, then $d \neq a$ and $d \neq b$.

The previous paragraph shows that, without loss of generality, we can assume that $c \in (a, b)$, so $c \neq a$ and $c \neq b$ and $M > f(a)$, $M > f(b)$. (If $M = f(a) = f(b)$, replace the maximum by the minimum m in the following argument.) If f takes the value M only once, then we are done, so let's assume that $M = f(e)$ for some $e \in \mathbb{R}$, $e \neq c$.

If $e \notin [a, b]$, let's assume that $b < e$. (A similar argument would work if $e < a$.) Pick y such that $f(b) < y < M$. By the Intermediate Value Theorem, there must exist x_1, x_2, x_3 such that $a < x_1 < c < x_2 < b < x_3 < e$ and $y = f(x_1) = f(x_2) = f(x_3)$, which contradicts the assumption that f takes no value more than twice.

Otherwise, $e \in (a, b)$. Let's assume, without loss of generality, that $a < c < e < b$. Pick $x_1 \in (c, e)$ in such a way that $f(a) < f(x_1) < M$ and $f(b) < f(x_1) < M$. Set $y = f(x_1)$. By the Intermediate Value Theorem, there exists $x_0 \in (a, c)$ and $x_2 \in (e, b)$ such that $f(x_0) = y$ and $f(x_2) = y$. Again, this contradicts the assumption that f takes no value more than twice.

Having reach a contradiction in all cases, it follows that f takes the value M only once. $\square$

Wednesday, January 26, 2022

The next example features an unusual function on the interval $[0,1]$ which you are not likely to encounter in a regular calculus course. What is unusual is that it is continuous on a dense subset of $[0,1]$ and also discontinuous on a dense subset. Here, dense means that if you consider any small subinterval $(a,b) \subset [0,1]$, then you can find a point in (a,b) where the function is continuous and another point in (a,b) where it is discontinuous. Most of the functions that you have seen in calculus have a finite number of points of discontinuity (or the number of such points is discrete), so perhaps the following example will look interesting.

Example 7: Let $f : [0,1] \to [0,1]$ be the function defined in the following way: $f(1) = 1$ and if $a = 0.a_1a_2a_3\ldots$ is the decimal expansion of a, then $f(a) = 0.0a_10a_20a_3\ldots$

Remember that the decimal expansion is not always unique: for instance, $0.9999\ldots$ (which is denoted $0.\overline{9}$), is equal to 1, since

$$\begin{aligned} 0.9999\ldots &= 9\cdot 0.1111\ldots \\ &= 9\cdot\left(\frac{1}{10}+\frac{1}{10^2}+\frac{1}{10^3}+\cdots\right) \end{aligned}$$

$$\begin{aligned} &= \frac{9}{10}\cdot\left(1+\frac{1}{10}+\frac{1}{10^2}+\frac{1}{10^3}+\cdots\right) \\ &= \frac{9}{10}\cdot\frac{1}{1-\frac{1}{10}} \\ &= \frac{9}{10}\cdot\frac{10}{10-1} = 1. \end{aligned}$$

Here, the following formula was used in the case $t = \frac{1}{10}$:

$$1 + t + t^2 + t^3 + \cdots = \sum_{n=0}^{\infty} t^n = \frac{1}{1-t} \text{ when } |t| < 1.$$

This is an example of a series and you have seen these if you have taken a second course in calculus. (If you haven't yet, don't worry, series will not play much of a role at the beginning of this course.)

More generally, if $a = 0.a_1a_2\cdots a_{k-1}a_k$ with $1 \leqslant a_k \leqslant 9$, then

$$a = 0.a_1a_2\cdots(a_k-1)9999\cdots$$

In this case, set $f(a) = 0.0a_10a_2\cdots 0a_{k-1}0a_k$.

A number a such that $a = \frac{b}{10^k}$ with $b \in \mathbb{Z}$ has a terminating decimal expansion: $a = 0.a_1a_2\cdots a_k$

Let's prove that if a has a terminating decimal expansion, then f is not continuous at a, but it is continuous at all other points on $[0,1]$.

Solution: Suppose that a has a terminating decimal expansion, so $a = 0.a_1a_2\cdots a_k$ with $1 \leqslant a_k \leqslant 9$. To show that f is not continuous at a, it is enough to find a sequence $\{x_n\}_{n=1}^{\infty}$ that converges to a but for which the sequence $\{f(x_n)\}_{n=1}^{\infty}$ does not converge to $f(a)$.

Set

$$\begin{aligned}
x_1 &= 0.a_1a_2\cdots(a_k-1)9\\
x_2 &= 0.a_1a_2\cdots(a_k-1)99\\
x_3 &= 0.a_1a_2\cdots(a_k-1)999\\
&\vdots \qquad\qquad\qquad \vdots\\
x_n &= 0.a_1a_2\cdots(a_k-1)99\cdots9\\
&\vdots \qquad\qquad\qquad \vdots
\end{aligned}$$

Then $\lim_{n\to\infty} x_n = a$ since $|a-x_n| = 0.\underbrace{00000\cdots0}_{k+n-1}1 = 10^{-k-n}$.

However, $f(x_n) = 0.0a_10a_2\cdots0(a_k-1)\underbrace{0909\cdots09}_{n}$, so

$$\begin{aligned}
|f(a)-f(x_n)| &= 0.\underbrace{000\cdots0}_{2k}909090\cdots9091\\
&= \frac{9}{10^{2k+1}} + \frac{9}{10^{2k+3}} + \cdots + \frac{9}{10^{2k+2n-1}} + \frac{1}{10^{2k+2n}}\\
&\geqslant \frac{9}{10^{2k+1}}
\end{aligned}$$

for any $n \geqslant 1$. This shows that $f(x_n)$ cannot converge to $f(a)$ as $n\to\infty$ since its distance from $f(a)$ is always at least $\frac{1}{10^{2k+1}}$. Note that k is fixed: it indicates the position of the last non-zero decimal in the expansion of a: $a = 0.a_1a_2\cdots a_k$.

It remains to prove that if a does not have a terminating decimal expansion, then f is continuous at a. Write a as $a = 0.a_1a_2a_3\cdots$ In this case, it is possible to find $i_1 < i_2 < i_3 < \cdots$ such that $a_{i_n} \neq 9$ for all $n \geqslant 1$. For each $n \geqslant 1$, set

$$b_n = 0.a_1a_2a_3\cdots a_{i_n}00\cdots \quad\text{and}\quad c_n = 0.a_1a_2a_3\cdots a_{i_n-1}(a_{i_n}+1)00\cdots,$$

which is possible since $0 \leqslant a_{i_n} \leqslant 8$.

Then $b_n \leqslant a \leqslant c_n$ and both sequences $\{b_n\}_{n=1}^{\infty}$ and $\{c_n\}_{n=1}^{\infty}$ converge to a. Furthermore, $f([b_n,c_n]) \subset [f(b_n), f(c_n)]$ because of the way the function f is defined (it is increasing) and the sequences $\{f(b_n)\}_{n=1}^{\infty}$ and $\{f(c_n)\}_{n=1}^{\infty}$ converge to $f(a)$.

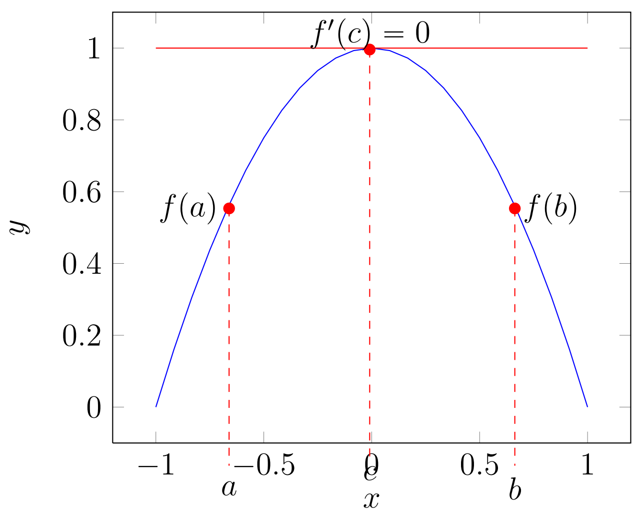
Fix $n \geqslant 1$. If $\{x_m\}_{m=1}^{\infty}$ is any sequence that converges to a, then $x_m \in [b_n, c_n]$ if m is large enough and then $f(x_m) \in [f(b_n), f(c_n)]$ for the same values of m. It follows that $f(x_m)$ also converges to $f(a)$.

This shows that $\lim_{x\to a} f(x) = f(a)$, so f is continuous at a. $\square$

2 Differentiable functions

Let's recall a couple of fundamental theorems about differentiable functions that you have learned in calculus.

Theorem 2.1 (Rolle's Theorem)**.** *Suppose that $f : [a, b] \longrightarrow \mathbb{R}$ is a continuous function on the interval $[a, b]$ and is differentiable on (a, b). If $f(a) = f(b)$, then there is a number c in (a, b) such that $f'(c) = 0$.*



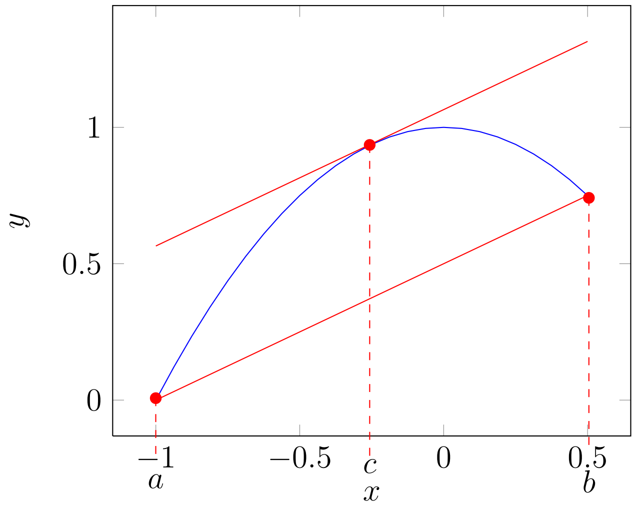
Corollary 2.2. *If f is a differentiable function on $[a, b]$ and if f has n distinct zeros on $[a, b]$, then f' has at least $n-1$ zeros, f'' has at least $n-2$ zeros, etc. In general, $f^{(k)}$ has at least $n-k$ zeros for $0 \leqslant k \leqslant n-1$. (Here, $f^{(k)}$ denotes the k^{th} derivative of f. It is implicit here that this corollary is true for $f^{(k)}$ only if $f^{(k)}$ is known to exist.)*

This corollary can be proved by induction on n using Rolle's Theorem.

Theorem 2.3 (Mean-Value Theorem)**.** *Suppose that $f : [a, b] \longrightarrow \mathbb{R}$ is a continuous function on the interval $[a, b]$ and is differentiable on (a, b). There is a number c in (a, b) such that*

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

Observe that Rolle's Theorem is the special case of the Mean-Value Theorem when $f(a) = f(b)$.



Example 1: Let $a, b, c \in \mathbb{R}$ be some constants. Let $f(x)$ be the polynomial given by

$$f(x) = 4ax^3 + 3bx^2 + 2cx - a - b - c.$$

Prove that $f(x)$ has a root in the interval $(0, 1)$.

Solution: The natural approach here is to hope that $f(0)$ and $f(1)$ have different signs so that the Intermediate Value Theorem can be applied.

$f(0) = -a - b - c$ and $f(1) = 4a + 3b + 2c - a - b - c = 3a + 2b + c$: since we don't have further information about the constants a, b, c, it is not possible to see if $f(0)$ and $f(1)$ have different signs. For instance, a, b and c are not necessarily all positive.

What needs to be observed here is that $f(x)$ contains $4x^3$, $3x^2$ and $2x$: these should make you think about the derivatives of x^4, x^3 and x^2. Since $f(x)$ is a polynomial it is equal to the derivative $g(x)$ of another polynomial, namely $g(x) = ax^4 + bx^3 + cx^2 - ax - bx - cx$, and maybe this plays a role here.

Indeed, $g(0) = 0$ and $g(1) = a + b + c - a - b - c = 0$, so Rolle's Theorem can be applied to $g(x)$: it says that $g'(c) = 0$ for some $c \in (0, 1)$. Since $f(x) = g'(x)$, it follows that $f(c) = 0$. □

Example 2: Check that $x^2 = x\sin(x) + \cos(x)$ for exactly two real values of x.

Solution: Set $f(x) = x^2 - x\sin(x) - \cos(x)$. If $f(x) = 0$ for at least three real values of x, then $f'(x)$ must vanish for at least two real values and $f''(x)$ must equal zero for at least one value according to Corollary 2.2.

$f'(x) = 2x - \sin(x) - x\cos(x) + \sin(x) = 2x - x\cos(x)$, $f''(x) = 2 - \cos(x) + x\sin(x)$. We would like to argue by contradiction and say that $f''(x)$ does not vanish, but this is not true: $f''\left(\frac{\pi}{2}\right) = 2 + \frac{\pi}{2} > 0$ and $f''\left(\frac{3\pi}{2}\right) = 2 - \frac{3\pi}{2} = \frac{4-3\pi}{2} < 0$, so $f''(x) = 0$ for some x between $\frac{\pi}{2}$ and $\frac{3\pi}{2}$.

In this example, it is enough to stop at computing $f'(x)$: observe that

$$f'(x) = x(2 - \cos(x))$$

and thus $f'(x) = 0$ only when $x = 0$ since $2 > \cos(x)$.

Since $f'(x)$ vanishes exactly once, $f(x) = 0$ for at most two values of x. Observe that

$$f(-\pi) = \pi^2 + 1 > 0,\ f(0) = -1,\ f(\pi) = \pi^2 + 1 > 0.$$

By the Intermediate Value Theorem, f has a zero on $(-\pi, 0)$ and another zero on $(0, \pi)$.

In conclusion, $f(x) = 0$ for exactly two real values of x. $\square$

Wednesday, February 2, 2022

Example 3: Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be the function $f(x) = 2^x - 1 - x^2$. How many zeros does f have on $\mathbb{R}$?

Solution: $f(0) = 2^0 - 1 = 0$, $f(1) = 2^1 - 1 - 1 = 0$, so f has at least two zeros. Can f have more than two zeros? If it does, then f' must have at least two and f'' must have at least one.

$f'(x) = \ln(2)2^x - 2x$ and $f''(x) = \ln(2)^2 2^x - 2$, so $f''(x)$ does indeed have a zero: notice that $f''(x) > 0$ if x is large enough (since $\lim_{x\to\infty} 2^x = +\infty$) and $f''(-2) = \ln(2)^2 \cdot \frac{1}{4} - 2 < 0$.

Therefore, perhaps f has three zeros. This is indeed the case, although it is not clear what is the exact value of this third zero: $f(2) = 2^2 - 1 - 2^2 = -1$ and $f(5) = 2^5 - 1 - 5^2 = 6$ so, by the Intermediate Value Theorem, f vanishes at a point between 2 and 5.

Can f have four zeros? If this was the case then, by Corollary 2.2, $f'(x)$ would have at least three zeros and $f''(x)$ would have at least two. However, $f''(x) = \ln(2)^2 2^x - 2$, so $f''(x)$ is a strictly increasing function since 2^x is strictly increasing, hence $f''(x)$ can vanish at most once. Therefore, f cannot have four zeros.

In conclusion, f vanishes at exactly three points on the real line. $\square$

Example 4: Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a function such that

$$|f(x_1) - f(x_2)| \leqslant (x_1 - x_2)^2 \text{ for all } x_1, x_2 \in \mathbb{R}.$$

Prove that f is a constant function.

Solution: If f is differentiable, then it is a constant function if and only if $f'(x) = 0$. In this example, f is not assumed differentiable, but it is possible to show, at the same time, that it is differentiable and $f'(x) = 0$.

$$\begin{aligned}
|f'(x)| &= \left|\lim_{h\to 0} \frac{f(x+h) - f(x)}{h}\right| \\
&= \lim_{h\to 0} \frac{|f(x+h) - f(x)|}{|h|} \\
&\leqslant \lim_{h\to 0} \frac{(x+h-x)^2}{|h|} \\
&= \lim_{h\to 0} \frac{h^2}{|h|} \\
&= \lim_{h\to 0} h \\
&= 0
\end{aligned}$$

□

Example 5: Let $f : [a,b] \longrightarrow \mathbb{R}$ be a real-value function such that its k^{th} derivative $f^{(k)}$ exists for $k = 0, 1, \ldots, n+1$ on (a,b). (Here, $f^{(0)}$ is just the function f itself.) Moreover, suppose that $f^{(k)}(a) = 0 = f^{(k)}(b)$ for $k = 0, 1, 2 \ldots, n$. Prove that there exists $c \in (a,b)$ such that $f^{(n+1)}(c) = f(c)$.

Solution: First, let's consider the case $n = 0$, so we have that $f(a) = 0 = f(b)$ and we have to show that there exists $c \in (a,b)$ such that $f'(c) = f(c)$.

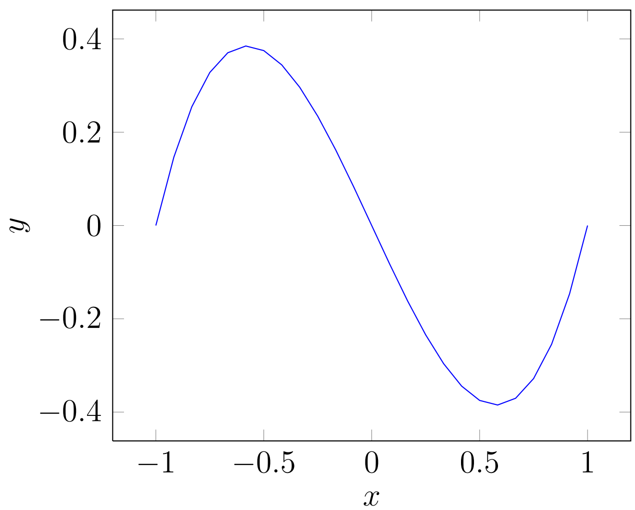
This seems to be related to Rolle's Theorem, but it is not clear how since $f - f'$ is not the derivative of any function. (Actually, this is not quite true since f is the derivative of one of its anti-derivative, but I don't want to talk about integrals and the Fundamental Theorem of Calculus.) If we set $g(x) = f(x) - f'(x)$, then maybe we can try to use the Intermediate Value Theorem to conclude that there exists a $c \in (a,b)$ such that $g(c) = 0$, which is equivalent to $f(c) - f'(c) = 0$.

Let's see if this approach could work.

$$g(a) = f(a) - f'(a) = -f'(a) \text{ and } g(b) = f(b) - f'(b) = -f'(b).$$

Therefore, if $f'(a) > 0$ and $f'(b) < 0$, or if $f'(a) < 0$ and $f'(b) > 0$, then $g(a)$ and $g(b)$ have different signs, so the Intermediate Value Theorem implies that there exists a $c \in (a,b)$ such that $g(c) = 0$, hence $f(c) - f'(c) = 0$.

However, it is possible that $f'(a) > 0$ and $f'(b) > 0$, in which case that argument does not work. What about points where the derivative f' vanishes? Since f is continuous on $[a,b]$, we know (by the Extreme Value Theorem), that is has a global maximum M and a global minimum m on $[a,b]$. Moreover, if $M = f(d)$ and $m = f(e)$ with $a < d, e < b$, then we know also that $f'(d) = 0$ and $f'(e) = 0$. We thus have $g(d) = f(d) - f'(d) = M$ and $g(e) = f(e) - f'(e) = m$, so if $M > 0$ and $m < 0$, we can conclude, by the Intermediate Value Theorem, that there exists a c between d and e such that $g(c) = 0$, hence $f(c) - f'(c) = 0$.



It is possible that both M and m are > 0, in which case this argument does not apply. It could happen, for instance, that $f(x) \geqslant 0$ for all $x \in [a, b]$ and $f(a) = 0 = f'(a)$ and $f(b) = 0 = f'(b)$. What can we do in this case?

What is needed is a trick. It is not at all clear what this trick should be, but it is not difficult to understand why it works and it is good to know it. Consider the function g given instead by $g(x) = e^{-x} f(x)$. Then, by the product rule for the derivative,

$$g'(x) = -e^{-x} f(x) + e^{-x} f'(x).$$

Note that $g(a) = 0 = g(b)$ so, by Rolle's Theorem, there exists $c \in (a, b)$ such that $0 = g'(c) = e^{-c}(f'(c) - f(c))$, so $f'(c) = f(c)$ since $e^{-c} \neq 0$. This completes the solution when $n = 0$.

Wednesday, February 9, 2022

For the general case, one more idea is needed. If $n \geqslant 1$, set

$$h(x) = f(x) + f'(x) + f''(x) + \cdots + f^{(n)}(x)$$

Then $h(a) = 0 = h(b)$ and

$$\begin{aligned} h(x) - h'(x) &= f(x) + f'(x) + f''(x) + \cdots + f^{(n-1)} + f^{(n)}(x) \\ &\quad - (f'(x) + f''(x) + f'''(x) + \cdots + +f^{(n)} + f^{(n+1)}(x)) \\ &= f(x) - f^{(n+1)}(x). \end{aligned}$$

From the case $n = 0$, but with $h(x)$ instead of $f(x)$, we can conclude that $h'(c) = h(c)$ for some $c \in (a, b)$. Since $h(x) - h'(x) = f(x) - f^{(n+1)}(x)$, it follows that $f(c) = f^{(n+1)}(c)$. $\square$

Example 6: [2015 Putnam, B–1] Let f be a three times differentiable function defined on $\mathbb{R}$ and real-valued such that f has at least five distinct real zeros. Prove that $f+6f'+12f''+8f'''$ has at least two distinct real zeros.

Solution: Set $g = f + 6f' + 12f'' + 8f'''$. Observe that $g = \left(1 + 2\frac{d}{dz}\right)^3 f$ because, by the Binomial Theorem,

$$\left(1 + 2\frac{d}{dz}\right)^3 = 1 + 3 \cdot 2\frac{d}{dz} + 3 \cdot 2^2 \left(\frac{d}{dz}\right)^2 + \left(\frac{d}{dz}\right)^3.$$

Here, $\left(\frac{d}{dz}\right)^2$ means: apply twice the derivative, so take the second derivative, $\left(\frac{d}{dz}\right)^2 = \frac{d^2}{dx^2}$. $\left(\frac{d}{dz}\right)^3$ means to take the derivative three times.

Recall that if a function $\widetilde{f}$ has five distinct zeros then, by the more general version of Rolle's Theorem (see Corollary 2.2), $\widetilde{f}'$ must have at least four distinct zeros, $\widetilde{f}''$ must have at least three distinct zeros and $\widetilde{f}'''$ must have at least two distinct zeros. Since $g = \left(1 + 2\frac{d}{dz}\right)^3 f$, we are thus led to try to prove the following variation of Rolle's Theorem.

Theorem 2.4 (Variant of Rolle's Theorem)**.** *Suppose that $h : [a,b] \longrightarrow \mathbb{R}$ is a continuous function which is differentiable on (a,b). If $h(a) = 0 = h(b)$, then $h + 2h'$ has a zero on the interval (a,b).*

Proof. The clever trick that is needed is essentially the same as the one used in the previous example, but since we have $h + 2h'$ instead of $f - f'$, we just need to change one constant to make it work. Set $\widetilde{h}(x) = e^{\frac{x}{2}}h(x)$. This functions vanishes also at a and b: $\widetilde{h}(a) = 0$ and $\widetilde{h}(b) = 0$. By Rolle's Theorem, there exists $c \in (a,b)$ such that $\widetilde{h}'(c) = 0$.

By the product rule for the derivative, $\widetilde{h}'(x) = \frac{1}{2}e^{\frac{x}{2}}h(x) + e^{\frac{x}{2}}h'(x)$. Therefore,

$$0 = \widetilde{h}'(c) = \frac{1}{2}e^{\frac{c}{2}}h(c) + e^{\frac{c}{2}}h'(c).$$

Since $e^{\frac{c}{2}} \neq 0$, it follows that $\frac{1}{2}h(c) + h'(c) = 0$, hence $h(c) + 2h'(c) = 0$. $\square$

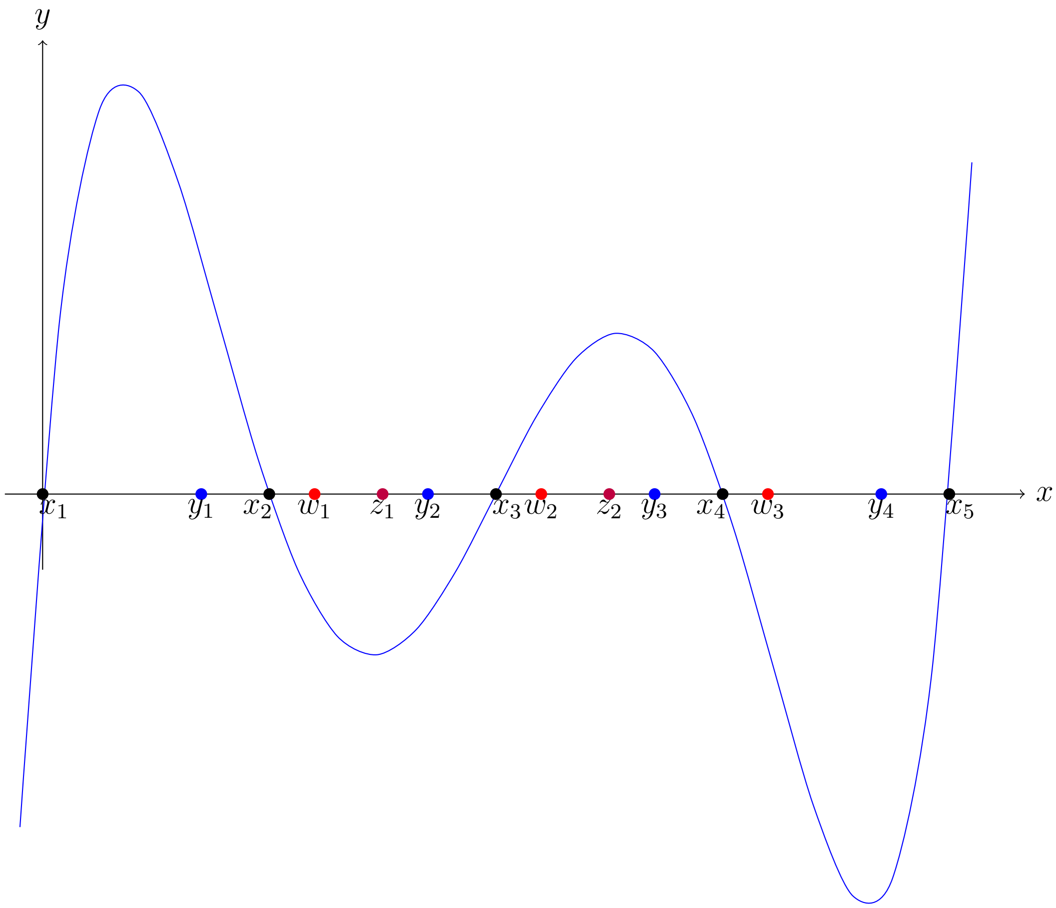
This proof shows that 2 could be replaced by any non-zero real number k and $h + 2h'$ can be replaced by $h + kh'$. In the proof, $e^{\frac{x}{2}}$ should then be replaced by $e^{\frac{x}{k}}$.

Now let's return to the Putnam problem. Suppose that the zeros of f are located at x_1, x_2, x_3, x_4, x_5 and $x_1 < x_2 < x_3 < x_4 < x_5$.

By the previous theorem, $f + 2f'$ has four distinct zeros at y_1, y_2, y_3, y_4 with $x_1 < y_1 < x_2 < y_2 < x_3 < y_3 < x_4 < y_4 < x_5$.

Applying the theorem a second time, we conclude that $\left(1 + 2\frac{d}{dz}\right)^2 f$, which equals $f + 4f' + 4f''$, has at least three zeroes at w_1, w_2, w_3 with $y_1 < w_1 < y_2 < w_2 < y_3 < w_3 < y_4$.

One more application of the theorem says that $\left(1 + 2\frac{d}{dz}\right)^3 f$ has at least two zeros at z_1, z_2 with $w_1 < z_1 < w_2 < z_2 < w_3$.



□

3 Recurrence formulas

Let n be an integer $\geqslant 0$ (or $\geqslant 1$). A recurrence formula is a formula for a number a_n that depends on $a_0, a_1, \ldots, a_{n-1}$ (or on $a_1, a_2, \ldots, a_{n-1}$ if a_0 is not defined).

For instance, $a_n = 2a_{n-1}$ for all $n \geqslant 1$ is a recurrence formula. In this case, we can see that

$$a_n = 2a_{n-1} = 2 \cdot 2 \cdot a_{n-2} = 2^2 a_{n-2} = 2^2 \cdot 2a_{n-3} = 2^3 a_{n-3}.$$

It can be proved by induction that $a_n = 2^n a_0$, so the value of a_n is uniquely determined by the recurrence formula and the value of a_0. For example, if $a_0 = 1$, then $a_n = 2^n$ for all $n \geqslant 0$.

A famous recurrence formula is

$$F_n = F_{n-1} + F_{n-2} \text{ for all } n \geqslant 2.$$

If $F_0 = 0$ and $F_1 = 1$, then we obtain that

$$F_2 = F_1 + F_0 = 1 + 0 = 1,\ F_3 = F_2 + F_1 = 1 + 1 = 2,\ F_4 = F_3 + F_2 = 2 + 1 = 3,\ \text{etc.}$$

The sequence that is generated in this way is

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, \ldots$$

and it is called the Fibonacci sequence (or the sequence of Fibonacci numbers). It is not just a curiosity: this sequence plays a role in certain parts of mathematics and it even appears in nature.

Is it possible to obtain an explicit formula for a_n in terms of only n and the initial value(s) (e.g. a_0, a_1, a_2)? In general, this is not possible, but it is possible when the recurrence formula is of the form

$$a_n = ba_{n-1} + ca_{n-2} \text{ for all } n \geqslant 2.$$

In this case, let r_1 and r_2 be the two distinct non-zero roots of the quadratic equation $x^2 = bx + c$. Then

$$a_n = dr_1^n + er_2^n$$

where the constants d and e are determined by the values of a_0 and a_1 (or a_1 and a_2):

$$a_0 = d + e \ \text{ and } \ a_1 = dr_1 + er_2.$$

(If the equation $x^2 = bx + c$ has only one root of multiplicity two, the formula for a_n has to be modified a little bit: $a_n = dr^n + enr^n$.) The equation $x^2 = bx + c$ is called the characteristic equation of the recurrence relation $a_n = ba_{n-1} + ca_{n-2}$. Even when a_n is an integer, it is possible for the roots r_1 and r_2 to be irrational numbers or even complex numbers.

The characteristic equation of the Fibonacci recurrence relation is $x^2 = x + 1$ (or $x^2 - x - 1 = 0$). Its roots are

$$\frac{1 \pm \sqrt{5}}{2},$$

so

$$F_n = a\left(\frac{1+\sqrt{5}}{2}\right)^n + b\left(\frac{1-\sqrt{5}}{2}\right)^n$$

for some $a, b \in \mathbb{R}$. Since $F_0 = 0$ and $F_1 = 1$, it follows that

$$\begin{aligned} 0 &= a + b \\ 1 &= a\left(\frac{1+\sqrt{5}}{2}\right) + b\left(\frac{1-\sqrt{5}}{2}\right). \end{aligned}$$

The unique solution is $a = \frac{1}{\sqrt{5}}$ and $b = -\frac{1}{\sqrt{5}}$. Therefore,

$$F_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n.$$

Since F_n is an integer, all the square roots on the right-hand sides cancel each other whether n is even or odd.

There are at least a couple of ways to see why the formula for a_n when $a_n = ba_{n-1} + ca_{n-2}$, namely

$$a_n = dr_1^n + er_2^n,$$

involves the roots of the characteristic equation $x^2 = bx + c$. (We will only consider below the case when the two roots are not equal.) Let's assume that $c \neq 0$, for if $c = 0$, then $a_n = ba_{n-1}$ and $a_n = b^n a_0$.

One way involves linear algebra. The recurrence relation $a_n = ba_{n-1} + ca_{n-2}$ is equivalent to the following vector relation:

$$\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} b & c \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ a_{n-2} \end{pmatrix}.$$

Set $A = \begin{pmatrix} b & c \\ 1 & 0 \end{pmatrix}$. By induction, it can be shown that

$$\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} b & c \\ 1 & 0 \end{pmatrix}^{n-1} \begin{pmatrix} a_1 \\ a_0 \end{pmatrix} = A^{n-1} \begin{pmatrix} a_1 \\ a_0 \end{pmatrix}.$$

The matrix A^{n-1} can be computed by diagonalizing A. The characteristic polynomial of A is $x(x-b) - c$, that is, $x^2 - bx - c$. It roots r_1 and r_2 are the eigenvalues of A, so if $r_1 \neq r_2$ then A is diagonalizable, which means that

$$A = PDP^{-1} \text{ and } A^{n-1} = PD^{n-1}P^{-1} \text{ where } D = \begin{pmatrix} r_1 & 0 \\ 0 & r_2 \end{pmatrix} \text{ and } D^{n-1} = \begin{pmatrix} r_1^{n-1} & 0 \\ 0 & r_2^{n-1} \end{pmatrix}.$$

Therefore,

$$\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = PD^{n-1}P \begin{pmatrix} a_1 \\ a_0 \end{pmatrix} = P \begin{pmatrix} r_1^{n-1} & 0 \\ 0 & r_2^{n-1} \end{pmatrix} P \begin{pmatrix} a_1 \\ a_0 \end{pmatrix} = PD^{-1} \begin{pmatrix} r_1^n & 0 \\ 0 & r_2^n \end{pmatrix} P \begin{pmatrix} a_1 \\ a_0 \end{pmatrix}.$$

Note that $r_1 \neq 0$ and $r_2 \neq 0$ since $c \neq 0$, so D is invertible.

If follows from $\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = PD^{-1} \begin{pmatrix} r_1^n & 0 \\ 0 & r_2^n \end{pmatrix} P \begin{pmatrix} a_1 \\ a_0 \end{pmatrix}$ that $a_n = dr_1^n + er_2^n$ for some constants d, e.

Another way to obtain a formula for a_n is to introduce the generating function $f(x) = \sum_{n=0}^{\infty} a_n x^n$. The recurrence relation $a_n = ba_{n-1} + ca_{n-2}$ implies

$$\sum_{n=2}^{\infty} a_n x^n = b \sum_{n=2}^{\infty} a_{n-1} x^n + c \sum_{n=2}^{\infty} a_{n-2} x^n,$$

which can be rewritten as

$$f(x) - a_1 x - a_0 = bx(f(x) - a_0) + cx^2 f(x),$$

so

$$(cx^2 + bx - 1)f(x) = (ba_0 - a_1)x - a_0.$$

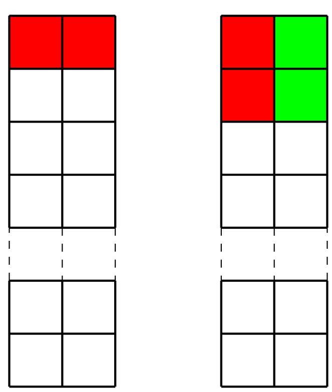
If $\widetilde{r}_1$ and $\widetilde{r}_2$ are the (distinct) roots of $cx^2 + bx - 1$, then $\widetilde{r}_1 = r_1^{-1}$, $\widetilde{r}_2 = r_2^{-1}$ and

$$\begin{aligned}
f(x) &= \frac{(ba_0 - a_1)x - a_0}{cx^2 + bx - 1} = \frac{(ba_0 - a_1)x - a_0}{c(x - \widetilde{r}_1)(x - \widetilde{r}_2)} \\
&= \frac{(ba_0 - a_1)x - a_0}{c(\widetilde{r}_1 - x)(\widetilde{r}_2 - x)} = \frac{r_1 r_2((ba_0 - a_1)x - a_0)}{c(1 - xr_1)(1 - xr_2)} \\
&= \frac{r_1 r_2((ba_0 - a_1)x - a_0)}{cx(r_1 - r_2)} \left(\frac{1}{1 - xr_1} - \frac{1}{1 - xr_2} \right) \\
&= \frac{r_1 r_2((ba_0 - a_1)x - a_0)}{cx(r_1 - r_2)} \left(\sum_{n=0}^{\infty} (r_1^n - r_2^n) x^n \right) \\
&= \frac{r_1 r_2(ba_0 - a_1)}{c(r_1 - r_2)} \left(\sum_{n=0}^{\infty} (r_1^n - r_2^n) x^n \right) - \frac{r_1 r_2 a_0}{cx(r_1 - r_2)} \left(\sum_{n=1}^{\infty} (r_1^n - r_2^n) x^n \right) \\
&= \frac{r_1 r_2(ba_0 - a_1)}{c(r_1 - r_2)} \left(\sum_{n=0}^{\infty} (r_1^n - r_2^n) x^n \right) - \frac{r_1 r_2 a_0}{c(r_1 - r_2)} \left(\sum_{n=1}^{\infty} (r_1^n - r_2^n) x^{n-1} \right) \\
&= \frac{r_1 r_2(ba_0 - a_1)}{c(r_1 - r_2)} \left(\sum_{n=0}^{\infty} (r_1^n - r_2^n) x^n \right) - \frac{r_1^2 r_2 a_0}{c(r_1 - r_2)} \left(\sum_{n=0}^{\infty} r_1^n x^n \right) - \frac{r_1 r_2^2 a_0}{c(r_1 - r_2)} \left(\sum_{n=0}^{\infty} r_2^n x^n \right)
\end{aligned}$$

Comparing the coefficients of x^n on both sides (so in both $f(x)$ and the last expression obtained just above) shows that a_n is a linear combination of r_1^n and r_2^n.

Example 1: In how many ways can an $n \times 2$ rectangle be covered by smaller rectangles of size 2×1 (or 1×2)?

Solution: Let a_n be the number of ways to cover the large rectangle by the smaller ones. Imagine that the rectangle of size $n \times 2$ is oriented vertically. There are two cases to consider.



Case 1: A small 1×2 rectangle is placed horizontally at the top. Then we are left to cover a rectangle of size $(n-1) \times 2$, which can be done in a_{n-1} ways.

Case 2: A 2×1 rectangle is placed vertically at the top left. Since the large rectangle has size $n \times 2$, there must be another 2×1 rectangle placed at the top right corner. We are then left to cover a rectangle of size $(n-2) \times 2$, which can be done in a_{n-2} ways.

Therefore, $a_n = a_{n-1} + a_{n-2}$. Moreover, $a_1 = 1$ and $a_2 = 2$, so $a_n = F_{n+1}$ for $n \geqslant 1$, the $(n+1)^{th}$ Fibonacci number. $\square$

Wednesday, February 16, 2022

Example 2: Consider a language whose only letter is the letter A and in which words of any length are allowed. Sentences are built from words separated by a space. How many different sentences in this language can be typed using exactly n keystrokes? (A keystroke is either an A or a space. Let's assume that a sentence starts with an A and ends with an A.)

Solution: Since a sentence begins with an A, there are two completely distinct possibilities:

Case 1: The first A is followed by another A.

In this case, the number of sentences of length n is equal to the number of sentences of length $n-1$.

$$A\,\underbrace{AA\cdots A\cdots A}_{n-1 \text{ keystrokes}}$$

Case 2: The first A is followed by a space.

In this case, the number of sentences of length n is equal to the number of sentences of length $n-2$.

$$A\quad\underbrace{AA\cdots A\cdots A}_{n-2 \text{ keystrokes}}$$

If S_n is the number of sentences that can be typed using exactly n keystrokes, then $S_n = S_{n-1} + S_{n-2}$. Moreover, $S_1 = 1$ and $S_2 = 1$ (the only sentence of length 2 is AA), so $S_n = F_n$, the n^{th} Fibonacci number. $\square$

Example 3: For a set S of integers, let $S+1 = \{x+1 \,|\, x \in S\}$. For how many subsets S of $\{1, 2, \ldots, n\}$ is it the case that $S \cup (S+1)$ contains $\{1, 2, \ldots, n\}$?

Solution: Let a_n be the number of such subsets. Let's consider the following two cases:

Case 1: $n \in S$.

Set $\widetilde{S} = S \backslash \{n\}$, so $\widetilde{S} \subset \{1, 2, \ldots, n-1\}$. Then

$$S \cup (S+1) = \{n, n+1\} \cup \widetilde{S} \cup (\widetilde{S}+1) \supset \{1, 2, \ldots, n\},$$

so $\widetilde{S} \cup (\widetilde{S}+1)$ must contain $\{1, 2, \ldots, n-1\}$ and there are a_{n-1} such subsets.

Case 2: $n \notin S$.

In this case, S must contain $n-1$, otherwise $S \cup (S+1)$ cannot contain n. Set $\widetilde{S} = S \backslash \{n-1\}$, so $\widetilde{S} \subset \{1, 2, \ldots, n-2\}$. Then

$$S \cup (S+1) = \{n-1, n\} \cup \widetilde{S} \cup (\widetilde{S}+1) \supset \{1, 2, \ldots, n\},$$

so $\widetilde{S} \cup (\widetilde{S}+1)$ must contain $\{1, 2, \ldots, n-2\}$ and there are a_{n-2} such subsets.

Therefore, $a_n = a_{n-1} + a_{n-2}$. Since $a_1 = 1$ and $a_2 = 2$ (when $n = 2$, S can be either $\{1\}$ or $\{1, 2\}$), it follows that $a_n = F_{n+1}$, the $(n+1)^{th}$ Fibonacci number. $\square$

Example 4: Let A_n be the matrix with 1's on and above the main diagonal and with -1's below the main diagonal. Let $a_n = \det(A_n)$. Show that $a_n = 2a_{n-1}$ for $n \geqslant 2$.

Solution: The matrix A_n is

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & -1 & 1 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ -1 & -1 & -1 & \cdots & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 & \cdots & -1 & 1 \end{pmatrix}.$$

The row operation $R_1 - R_2 \mapsto R_1$ (replacing the first row by row 1 - row 2) does not change the determinant, so

$$a_n = \begin{vmatrix} 2 & 0 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & -1 & 1 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ -1 & -1 & -1 & \cdots & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 & \cdots & -1 & 1 \end{vmatrix}$$

$$= 2\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & 1 & 1 & \cdots & 1 & 1 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ -1 & -1 & \cdots & -1 & 1 & 1 \\ -1 & -1 & -1 & \cdots & -1 & 1 \end{vmatrix} \text{ by expansion along the first row;}$$

$$= 2a_{n-1}$$

Since $a_1 = 1$, it follows that $a_n = 2^{n-1}$ for all $n \geqslant 1$.

It may help to see how this works when $n = 4$:

$$\det(A_4) = \begin{vmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 0 & 0 \\ -1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{vmatrix}$$

$$= 2\begin{vmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ --1 & -1 & 1 \end{vmatrix} = 2\begin{vmatrix} 2 & 0 & 0 \\ -1 & 1 & 1 \\ -1 & -1 & 1 \end{vmatrix}$$

$$= 2 \cdot 2\begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} = 4 \cdot (1 - (-1)) = 8.$$

□

Wednesday, March 2, 2022

Example 5 (Tower of Hanoi): Consider n rings and three pegs. Initially, all the rings are stacked on the first peg to form a pyramid with the largest ring at the bottom. The rings need to be transferred to the second peg to form an identical pyramid, but we are not permitted to place a larger ring on top of a smaller one. What is the smallest number of moves necessary to complete the transfer?

Solution: Let a_n be this smallest number of moves. The first $n-1$ rings need to be transferred to the third peg, which can be done in a minimum of a_{n-1} moves, after which the largest ring can be moved to the second peg. All the other rings on the third peg can then be moved to the second peg, which takes again a minimum of a_{n-1} moves.

Therefore, $a_n = a_{n-1} + 1 + a_{n-1} = 2a_{n-1} + 1$. Since $a_1 = 1$, it can be shown by induction that $a_n = 2^n - 1$. Another way to see this is to set $\widetilde{a}_n = a_n + 1$ and note that $a_n = a_{n-1} + 1$ is equivalent to $\widetilde{a}_n = 2\widetilde{a}_{n-1}$, hence $\widetilde{a}_n = 2^{n-1}\widetilde{a}_1 = 2^{n-1} \cdot 2 = 2^n$. Therefore, $a_n - 1 = 2^n$ and $a_n = 2^n + 1$. □

Example 6: [1969 Putnam, A–2] Let A_n be the $n \times n$ matrix whose (i, j) entry is $|i - j|$. Show that $\det(A_n) = (-1)^{n-1}(n - 1)2^{n-2}$

First solution:

$$A_n = \begin{pmatrix} 0 & 1 & 2 & 3 & \cdots & n-2 & n-1 \\ 1 & 0 & 1 & 2 & \cdots & n-3 & n-2 \\ 2 & 1 & 0 & 1 & \cdots & n-4 & n-3 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ n-2 & n-3 & n-4 & \cdots & 1 & 0 & 1 \\ n-1 & n-2 & n-3 & n-4 & \cdots & 1 & 0 \end{pmatrix}.$$

Let's replace the first row by $R_1 - 2R_2 + R_3$. The reason for doing this is to put as many zeros as possible in the first row. We obtain the matrix

$$\begin{pmatrix} 0 & 2 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 1 & 2 & \cdots & n-3 & n-2 \\ 2 & 1 & 0 & 1 & \cdots & n-4 & n-3 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ n-2 & n-3 & n-4 & \cdots & 1 & 0 & 1 \\ n-1 & n-2 & n-3 & n-4 & \cdots & 1 & 0 \end{pmatrix}.$$

It is tempting to expand the determinant along the first row, but one problem with doing this is that the $(n-1) \times (n-1)$ matrix obtained by deleting the first row and second column is not A_{n-1}. Instead, since A_n is symmetric, let's apply the same operations to the columns, so let's replace the first column by $C_1 - 2C_2 + C_3$. We thus obtain the matrix

$$\begin{pmatrix} -4 & 2 & 0 & 0 & \cdots & 0 & 0 \\ 2 & 0 & 1 & 2 & \cdots & n-3 & n-2 \\ 0 & 1 & 0 & 1 & \cdots & n-4 & n-3 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & n-3 & n-4 & \cdots & 1 & 0 & 1 \\ 0 & n-2 & n-3 & n-4 & \cdots & 1 & 0 \end{pmatrix}.$$

Let's called this last matrix $\widetilde{A}$. Then $\widetilde{A}$ has $\widetilde{a}_{11} = -4$, $\widetilde{a}_{01} = 2$, $\widetilde{a}_{10} = 2$ and zeros in all the other entries in the first row and first column. Therefore, if $a_n = \det(A_n)$, then expansion of the determinant along the first row shows that

$$
\begin{aligned}
a_n = & -4\begin{vmatrix}
0 & 1 & 2 & \cdots & n-3 & n-2\\
1 & 0 & 1 & \cdots & n-4 & n-3\\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots\\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots\\
n-3 & n-4 & \cdots & 1 & 0 & 1\\
n-2 & n-3 & n-4 & \cdots & 1 & 0
\end{vmatrix}\\
& -2\begin{vmatrix}
2 & 1 & 2 & \cdots & n-3 & n-2\\
0 & 0 & 1 & \cdots & n-4 & n-3\\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots\\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots\\
0 & n-4 & \cdots & 1 & 0 & 1\\
0 & n-3 & n-4 & \cdots & 1 & 0
\end{vmatrix}
\end{aligned}
$$

$$
\begin{aligned}
a_n = & -4a_{n-1} - 2\cdot 2\begin{vmatrix}
0 & 1 & \cdots & n-4 & n-3\\
\ddots & \ddots & \ddots & \vdots & \vdots\\
\ddots & \ddots & \ddots & \vdots & \vdots\\
n-4 & \cdots & 1 & 0 & 1\\
n-3 & n-4 & \cdots & 1 & 0
\end{vmatrix}\\
= & -4a_{n-2} - 4a_{n-1}.
\end{aligned}
$$

The characteristic equation of this recurrence relation is $x^2 = -4x - 4$ or, equivalently, $x^2 + 4x + 4 = 0$. The only root of $x^2 + 4x + 4$ is -2 and has multiplicity 2. It follows that

$$a_n = d(-2)^n + en(-2)^n$$

for some $d, e \in \mathbb{R}$.

$a_1 = 0$ and $a_2 = \begin{vmatrix} 0 & 1\\ 1 & 0\end{vmatrix} = -1$, so

$$
\begin{array}{rcrcr}
0 & = & -2d & - & 2e\\
-1 & = & 4d & + & 8e
\end{array}
$$

The only solution is $d = \frac{1}{4}$ and $e = -\frac{1}{4}$, so

$$a_n = \frac{1}{4}(-2)^n - \frac{n}{4}(-2)^n = (-2)^{n-2} - n(-2)^{n-2} = (-1)^{n-1}(n-1)2^{n-2}.$$

□

Second solution. Let $a_n = \det(A_n)$. Applying the row operations $R_i - R_{i+1}$ with i from 1 to $n-1$ produces the matrix

$$\begin{pmatrix} -1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & -1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & -1 & -1 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ -1 & -1 & -1 & \cdots & -1 & -1 & 1 \\ n-1 & n-2 & n-3 & n-4 & \cdots & 1 & 0 \end{pmatrix}.$$

The row operation $R_1 - R_2$ then yields the matrix

$$\begin{pmatrix} 0 & 2 & 0 & 0 & \cdots & 0 & 0 \\ -1 & -1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & -1 & -1 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ -1 & -1 & -1 & \cdots & -1 & -1 & 1 \\ n-1 & n-2 & n-3 & n-4 & \cdots & 1 & 0 \end{pmatrix}.$$

Therefore,

$$a_n = -2\det \begin{pmatrix} -1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & -1 & 1 & \cdots & 1 & 1 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ -1 & -1 & \cdots & -1 & -1 & 1 \\ n-1 & n-3 & n-4 & \cdots & 1 & 0 \end{pmatrix}.$$

We can use again the row operation $R_1 - R_2$ and obtain that

$$a_n = -2\det \begin{pmatrix} 0 & 2 & 0 & \cdots & 0 & 0 \\ -1 & -1 & 1 & \cdots & 1 & 1 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ -1 & -1 & \cdots & -1 & -1 & 1 \\ n-1 & n-3 & n-4 & \cdots & 1 & 0 \end{pmatrix} = (-2)^2\det \begin{pmatrix} -1 & 1 & \cdots & 1 & 1 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ -1 & \cdots & -1 & -1 & 1 \\ n-1 & n-4 & \cdots & 1 & 0 \end{pmatrix}.$$

Repeating this eventually leads to

$$a_n = (-2)^{n-2}\det \begin{pmatrix} -1 & 1 \\ n-1 & 0 \end{pmatrix} = (-1)^{n-1}2^{n-2}(n-1).$$

$\square$

Third solution. Let $a_n = \det(A_n)$. Applying the row operations $R_i - R_{i+1}$ with i from 1 to $n-1$ produces the matrix

$$\begin{pmatrix} -1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & -1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & -1 & -1 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ -1 & -1 & -1 & \cdots & -1 & -1 & 1 \\ n-1 & n-2 & n-3 & n-4 & \cdots & 1 & 0 \end{pmatrix}$$

and the row operation $R_n + \frac{1}{2}(R_1 + \cdots + R_{n-1})$ then yields the matrix

$$\begin{pmatrix} -1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & -1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & -1 & -1 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ -1 & -1 & -1 & \cdots & -1 & -1 & 1 \\ \frac{n-1}{2} & \frac{n-1}{2} & \frac{n-1}{2} & \frac{n-1}{2} & \cdots & \frac{n-1}{2} & \frac{n-1}{2} \end{pmatrix}.$$

The determinant of this last matrix is equal to $-\frac{n-1}{2}$ times the determinant of the matrix

$$\begin{pmatrix} -1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & -1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & -1 & -1 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ -1 & -1 & -1 & \cdots & -1 & -1 & 1 \\ -1 & -1 & -1 & -1 & \cdots & -1 & -1 \end{pmatrix}.$$

This last matrix is equal to $-A_n^T$ where A_n is the matrix in Example 4 whose determinant was 2^{n-1}. Therefore,

$$a_n = -\frac{n-1}{2}(-1)^n 2^{n-1} = (-1)^{n-1}(n-1)2^{n-2}.$$

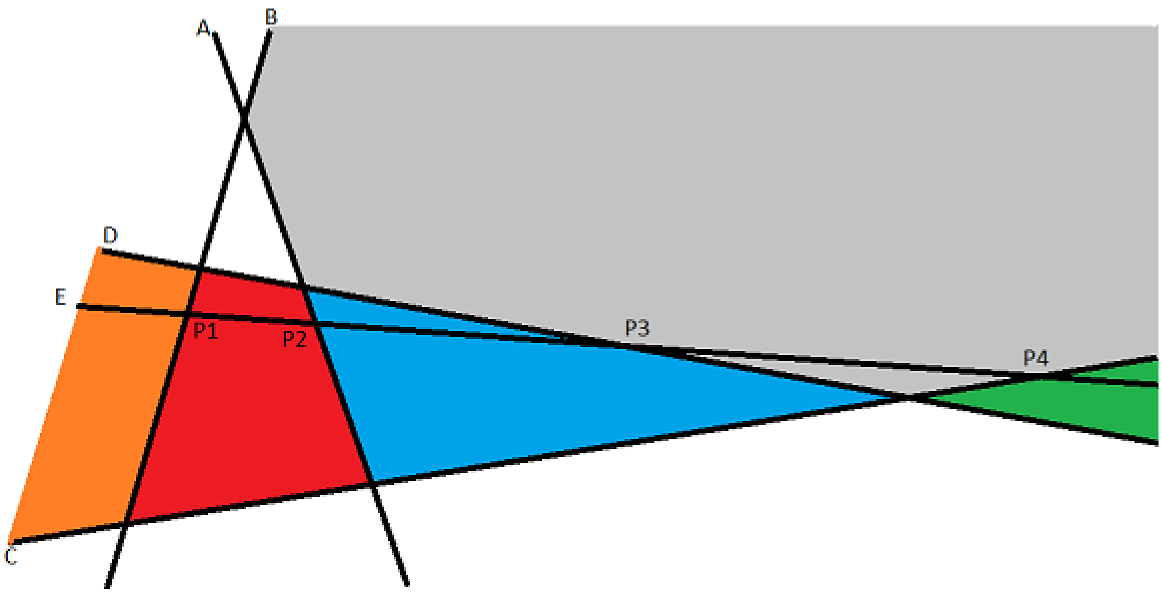
□

Wednesday, March 9, 2022

Example 7: Consider n lines drawn in the plane $\mathbb{R}^2$ such that no two are parallel and no three of them meet at a common point. Let R_n be the number of regions into which the plane is divided by these lines. Explain why $R_{n+1} = R_n + n + 1$.

Solution: Denote the $n+1$ lines by $\ell_1, \ldots, \ell_{n+1}$. Since no two lines are parallel, ℓ_i intersects ℓ_{n+1}. Let $P_i = \ell_i \cap \ell_{n+1}$, that is, P_i is the point of intersection of the lines ℓ_i and ℓ_{n+1}. Let's label the lines in such a way that the points P_i and P_{i+1} are neighbors on the line ℓ_{n+1}, so these points can be read in the order $P_1, P_2, \ldots, P_n$ along the line ℓ_{n+1}, going in one of the two possible directions.

In the following diagram, $n+1 = 5$ and the lines are labelled A, B, C, D, E instead. The coloured regions are the regions created by the lines A, B, C, D that are cut in two by the line E.



Each line segment $[P_i, P_{i+1}]$ separates into two halves a region R that was bounded by some of the lines $\ell_1, \ldots, \ell_n$. This is also true for the infinite half-lines extending past P_1 and P_n. This shows that adding the line ℓ_{n+1} to the previous n lines creates $n+1$ new regions, so $R_{n+1} = R_n + n + 1$.

Since $R_n = R_{n-1} + n$ and $R_1 = 2$, it follows that

$$\begin{aligned}
R_n - R_1 &= (R_n - R_{n-1}) + (R_{n-1} - R_{n-2}) + (R_{n-2} - R_{n-3}) + \cdots + (R_3 - R_2) + (R_2 - R_1) \\
&= \sum_{i=2}^{n} (R_i - R_{i-1}) \\
&= \sum_{i=2}^{n} i = \sum_{i=1}^{n} i - 1 \\
&= \frac{n(n+1)}{2} - 1,
\end{aligned}$$

hence $R_n = \frac{n(n+1)}{2} - 1 + R_1 = \frac{n(n+1)}{2} + 1$. $\square$

Example 8: [1996 Putnam, B–1] Define a selfish set to be a set which has its own cardinality (number of elements) as an element. Find, with proof, the number of subsets of $\{1, 2, \ldots, n\}$ which are *minimal* selfish sets, that is, selfish sets none of whose proper subsets is selfish.

Solution: Let S be a minimal selfish set. If m is the cardinality of S and $k < m$, then $k \notin S$, for otherwise removing $m - k$ integers from S other than k itself would produce a selfish subset, thus contradicting the minimality of S. For instance if $S = \{3, 5, 6, 8, 11\}$, then S is a selfish set, but it is not minimal because the subset $\{3, 5, 6\}$ (or $\{3, 6, 8\}$, etc.) is also selfish.

Let a_n be the number of minimal selfish subsets of $\{1, 2, \ldots, n\}$. Let's consider two cases.

Case 1: $n \notin S$.

Then S is a minimal selfish subset of $\{1, 2, \ldots, n-1\}$ and there are a_{n-1} of these.

Case 2: $n \in S$.

Let m be the cardinality of S. Then $m > 1$ (if $n > 1$) and $m - 1 \notin S$ as explained above with $k = m - 1 < m$. It is also the case that $1 \notin S$ (if $n > 1$). Consider the set $S - 1$ given by

$$S - 1 = \{x - 1 \mid x \in S\}.$$

Then $(S-1)\backslash\{n-1\}$ is a minimal selfish subset of $\{1, 2, \ldots, n-2\}$: it is selfish because $m \in S \Rightarrow m - 1 \in S - 1$, and it is minimal, for if $\widetilde{S}$ is a selfish subset of $(S-1)\backslash\{n-1\}$, then $(\widetilde{S}+1) \cup \{n\}$ is a selfish subset of S, hence $(\widetilde{S}+1) \cup \{n\} = S$ and it follows that $\widetilde{S} = S - 1$.

For instance, if $n = 12$, $S = \{5, 7, 8, 11, 12\}$, then $S - 1\backslash\{11\} = \{4, 6, 7, 10\}$ and this is a minimal selfish subset of $\{1, 2, \ldots, 10\}$.

Conversely, given a minimal selfish subset $\widetilde{S}$ of $\{1, 2, \ldots, n-2\}$, $(\widetilde{S}+1) \cup \{n\}$ is a minimal selfish subset of $\{1, 2, \ldots, n\}$ that contains n. Therefore, we have a bijection

$$\left\{ \begin{array}{c} \text{minimal selfish subsets of} \\ \{1, 2, \ldots, n\} \\ \text{that contain } n \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} \text{minimal selfish subsets of} \\ \{1, 2, \ldots, n-2\} \end{array} \right\}$$

via

$$\begin{aligned} S &\longrightarrow (S-1)\backslash\{n-1\} \\ (\widetilde{S}+1) \cup \{n\} &\longleftarrow \widetilde{S} \end{aligned}$$

Therefore, there are a_{n-2} minimal selfish subsets of $\{1, 2, \ldots, n\}$ with the property that $n \in S$.

Therefore, since Cases 1 and 2 are disjoint,

$$a_n = a_{n-1} + a_{n-2}.$$

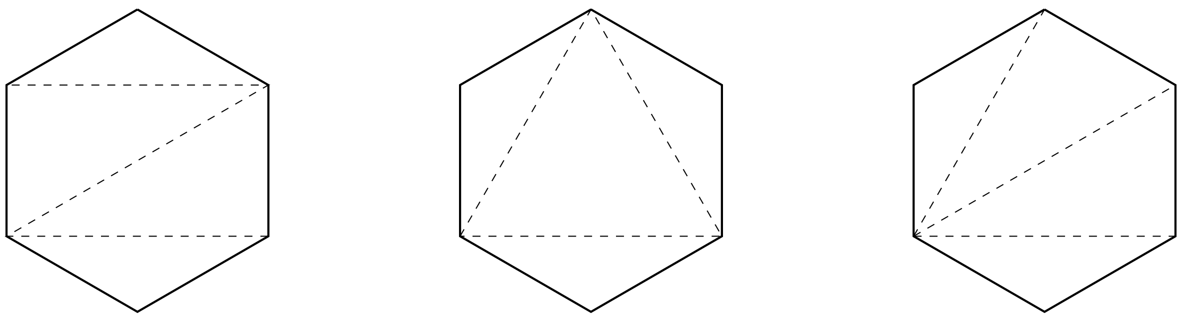
Since $a_1 = 1$ and $a_2 = 1$, it follows that $a_n = F_n$, the n^{th} Fibonacci number. $\square$

Wednesday, March 16, 2022

The next example introduces a new family of integers, the Catalan numbers, which appear in various combinatorial problems.

Definition 3.1. *Consider a convex polygon with n vertices and n sides. A triangulation is a partition of this polygon into triangles, all of whose vertices are vertices of the polygon and which intersect only along a diagonal.*

Here are examples of triangulations of the regular hexagon:



Example 9: What is the number of triangulations of a convex polygon with n vertices and n sides?

Solution: Let T_n be that number of triangulations and let's find a recurrence relation satisfied by T_n. Label the vertices clockwise $P_1, P_2, \ldots, P_n$, starting at some vertex P_1.

Consider the sides $\overline{P_1P_2}$. If we have a triangulation of the regular polygon, then $\overline{P_1P_2}$ is the side of one of the triangles in the triangulation. This triangle is $P_1P_2P_j$ for $3 \leqslant j \leqslant n$.

The triangle $P_1P_2P_j$ divides the interior of the convex polygons into three parts: the triangle $P_1P_2P_j$ itself and the two convex polygons $P_2P_3 \cdots P_j$ and $P_jP_{j+1} \cdots P_nP_1$. The number of triangulations of $P_2P_3 \cdots P_j$ is T_{j-1} since that convex polygon has $j-1$ vertices. The number of triangulations of $P_jP_{j+1} \cdots P_nP_1$ is T_{n-j+2} since that convex polygon has $n-j+2$ vertices.

Therefore,

$$\begin{aligned} T_n &= T_2T_{n-1} + T_3T_{n-2} + T_4T_{n-3} + \cdots + T_{n-1}T_2 \\ &\qquad \text{(with the understanding that } T_2 = 1\text{)} \\ &= \sum_{j=3}^{n} T_{j-1}T_{n-j+2}. \end{aligned}$$

In principle, given enough time, T_n can be determined using this recurrence relation starting with $T_3 = 1$, $T_4 = 2$. An explicit formula for T_n is derived below. $\square$

Let's see how to obtain an explicit and simple formula for T_n. Actually, let's set $C_n = T_{n+2}$ for $n \geqslant 0$ and let's find a formula for C_n. Observe that C_n satisfies a similar recurrence relation, namely

$$C_{n+1} = \sum_{i=0}^{n} C_i C_{n-i},$$

and that $C_0 = 1, C_1 = 1, C_2 = 2$. Let's use the generating function

$$f(x) = C_0 + C_1 x + C_2 x^2 + \cdots = \sum_{n=0}^{\infty} C_n x^n.$$

Then

$$f(x)^2 = C_0^2 + (C_0C_1 + C_1C_0)x + (C_2C_0 + C_1C_1 + C_0C_2)x^2 + \cdots$$

The recurrence relation for the Catalan numbers C_i implies that

$$\begin{aligned} xf(x)^2 &= C_0^2 x + (C_0C_1 + C_1C_0)x^2 + (C_2C_0 + C_1C_1 + C_0C_2)x^3 + \cdots \\ &= C_1 x + C_2 x^2 + C_3 x^3 + \cdots \\ &= f(x) - C_0 \\ &= f(x) - 1. \end{aligned}$$

This means that $f(x)$ is a solution of the equation $xy^2 - y + 1 = 0$ where y should be viewed as the variable that needs to be determined.

It follows from the standard quadratic formula that

$$f(x) = \frac{1 \pm \sqrt{1-4x}}{2x}$$

and the sign is negative since $f(0) = C_0$, so

$$f(x) = \frac{1 - \sqrt{1-4x}}{2x}.$$

In order to obtain an expression for the coefficients of the power series $f(x)$, we need to expand $\sqrt{1-4x}$ as a power series in x. This can be achieved using a generalization of the binomial formula.

Lemma 3.2 (General binomial formula)**.** *If $r \in \mathbb{R}$ and $|x| < 1$, then*

$$(1+x)^r = 1 + \binom{r}{1}x + \binom{r}{2}x^2 + \binom{r}{3}x^3 + \cdots$$

where

$$\binom{r}{k} = \frac{r(r-1)(r-2)\cdots(r-k+1)}{k!} \text{ for } k = 1, 2, 3, \ldots$$

This can be applied to $f(x)$ (with x replaced by $-4x$ and $r = \frac{1}{2}$) to obtain

$$f(x) = \frac{1}{2x} - \frac{1}{2x}\left(1 + \sum_{k=1}^{\infty} \frac{\frac{1}{2} \cdot \left(\frac{1}{2} - 1\right) \cdot \left(\frac{1}{2} - 2\right) \cdots \frac{\frac{1}{2} - k + 1}{2}}{k!}\right)(-4x)^k$$

$$\begin{aligned}
&= \frac{1}{2x} - \frac{1}{2x}\left(1 + \sum_{k=1}^{\infty} (-1)^{k-1} \frac{1 \cdot 1 \cdot 3 \cdot 5 \cdots (2k-3)}{2^k \cdot k!}\right)(-4x)^k \\
&= \frac{1}{2x} - \frac{1}{2x}\left(1 - 2\sum_{k=1}^{\infty} \frac{1 \cdot 3 \cdot 5 \cdots (2k-3) \cdot 2^{k-1}(k-1)!}{k!(k-1)!}\right) x^k \\
&= \frac{1}{2x} - \frac{1}{2x}\left(1 - 2\sum_{k=1}^{\infty} \frac{1 \cdot 3 \cdot 5 \cdots (2k-3) \cdot 2 \cdot 4 \cdot 6 \cdots (2k-2)}{k!(k-1)!}\right) x^k \\
&= \frac{1}{2x} - \frac{1}{2x}\left(1 - 2\sum_{k=1}^{\infty} \frac{(2k-2)!}{k!(k-1)!}\right) x^k \\
&= \sum_{k=1}^{\infty} \frac{1}{k}\binom{2k-2}{k-1} x^{k-1}.
\end{aligned}$$

The coefficient of x^k is

$$\frac{1}{k+1}\binom{2(k+1)-2}{(k+1)-1},$$

so

$$C_k = \frac{1}{k+1}\binom{2k}{k}.$$

Catalan numbers have also the following combinatorial interpretations:

- C_n is the number of words in the alphabet that consists of only the two letters A and B and for which the number of B in an initial segment of a word never exceeds the number of A. For instance, ABAAABBAAB is such an example, but not ABBABAA.
- C_n is the number of non-decreasing lattice paths from $(0,0)$ to (n,n) which stay on or below the diagonal $y = x$. A lattice path consists of horizontal and vertical line segments connecting points (x,y) and $(x+1,y)$ or $(x,y+1)$ where x and y are integers.
- C_n is the number of path starting at $(0,0)$ and ending at $(n,0)$ that do not go below the x-axis and which consists of line segments connecting (x,y) to $(x+1,y+1)$ or $(x+1,y-1)$ where x and y are integers.

Wednesday, March 23, 2022

4 Series

Recall the geometric series:

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \cdots = \frac{1}{1-x} \text{ for } |x| < 1 \tag{3}$$

and its finite version:

$$\sum_{n=0}^{m} x^n = 1 + x + x^2 + \cdots + x^m = \frac{1-x^{m+1}}{1-x} \text{ for } x \neq 1 \tag{4}$$

Note that (4) follows from

$$(1+x+x^2+\cdots+x^m)(1-x) = (1+x+x^2+\cdots+x^m)-(x+x^2+x^3+\cdots+x^{m+1}) = 1-x^{m+1}.$$

A telescoping series is one in which most of the neighboring terms cancel each other, thus leaving only a few terms that can be easily evaluated. For instance,

$$\begin{aligned}\sum_{n=1}^{m}\left(\frac{1}{n}-\frac{1}{n+1}\right) &= \left(\frac{1}{1}-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)+\left(\frac{1}{5}-\frac{1}{5}\right)\\ &\quad+\cdots+\left(\frac{1}{m-1}-\frac{1}{m}\right)+\left(\frac{1}{m}-\frac{1}{m+1}\right)\\ &= 1-\frac{1}{m+1}\end{aligned}$$

Observe that it follows that

$$\sum_{n=1}^{\infty}\frac{1}{n(n+1)} = \sum_{n=1}^{\infty}\left(\frac{1}{n}-\frac{1}{n+1}\right) = 1.$$

Example 1: Determine the sum of the series

$$\sum_{n=1}^{\infty}\frac{1}{(2n+1)(2n-1)}$$

Solution: The fraction $\frac{1}{(2n-1)(2n+1)}$ can be written in the form

$$\frac{A}{2n-1}+\frac{B}{2n+1}.$$

To find the values of A and B, multiply both sides by $(2n-1)(2n+1)$ to obtain

$$A(2n+1)+B(2n-1) = 1.$$

$$\text{If } n = -\frac{1}{2},\ \text{then } 2n+1 = 0 \text{ and } B(-2\cdot\frac{1}{2}-1) = 1,\ \text{so } B = -\frac{1}{2};$$
$$\text{If } n = \frac{1}{2},\ \text{then } 2n-1 = 0 \text{ and } A(2\cdot\frac{1}{2}+1) = 1,\ \text{so } A = \frac{1}{2}.$$

Therefore,

$$\frac{1}{(2n+1)(2n-1)} = \frac{1}{2(2n-1)} - \frac{1}{2(2n+1)} = \frac{1}{2}\left(\frac{1}{2n-1} - \frac{1}{2n+1}\right)$$

and

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{1}{(2n+1)(2n-1)} &= \sum_{n=1}^{\infty}\frac{1}{2}\left(\frac{1}{2n-1} - \frac{1}{2n+1}\right)\\ &= \frac{1}{2}\left(\left(\frac{1}{1}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{5}\right)+\left(\frac{1}{5}-\frac{1}{7}\right)+\left(\frac{1}{7}-\frac{1}{9}\right)+\cdots\right)\\ &= \frac{1}{2}\end{aligned}$$

$\square$

Example 2: Determine the sum

$$\frac{1}{2!}+\frac{2}{3!}+\frac{3}{4!}+\cdots+\frac{n-1}{n!}$$

Solution: We have to determine the sum

$$\sum_{k=2}^{n}\frac{k-1}{k!}.$$

Observe that

$$\frac{k-1}{k!} = \frac{k}{k!} - \frac{1}{k!} = \frac{1}{(k-1)!} - \frac{1}{k!}.$$

Therefore,

$$\begin{aligned}\sum_{k=2}^{n}\frac{k-1}{k!} &= \sum_{k=2}^{n}\left(\frac{1}{(k-1)!} - \frac{1}{k!}\right)\\ &= \left(\frac{1}{1!}-\frac{1}{2!}\right)+\left(\frac{1}{2!}-\frac{1}{3!}\right)+\left(\frac{1}{3!}-\frac{1}{4!}\right)\\ &\quad+\cdots+\left(\frac{1}{(n-2)!}-\frac{1}{(n-1)!}\right)+\left(\frac{1}{(n-1)!}-\frac{1}{n!}\right)\\ &= 1-\frac{1}{n!}\end{aligned}$$

$\square$

It follows that

$$\sum_{k=2}^{\infty} \frac{k-1}{k!} = \lim_{n\to\infty} \sum_{k=2}^{n} \frac{k-1}{k!} = 1.$$

Example 3: Prove the following identity:

$$F_1 + F_3 + F_5 + \cdots + F_{2n-3} + F_{2n-1} = F_{2n}$$

where F_m is the m^{th} Fibonacci number.

Solution: The sum to be determined is $\sum_{k=0}^{n-1} F_{2k+1}$. The Fibonacci recurrence relation says that $F_{2k+1} = F_{2k} + F_{2k-1}$, but if we use it, we don't obtain a telescoping series and it is not clear how this would help to simplify the series. The idea is to use instead the Fibonacci recurrence relation that starts with F_{2k+2}, namely $F_{2k+2} = F_{2k+1} + F_{2k}$, which implies that $F_{2k+1} = F_{2k+2} - F_{2k}$.

$$\begin{aligned}
\sum_{k=0}^{n-1} F_{2k+1} &= \sum_{k=0}^{n-1} (F_{2k+2} - F_{2k}) \\
&= (F_2 - F_0) + (F_4 - F_2) + (F_6 - F_4) + (F_8 - F_6) + \ldots + (F_{2n} - F_{2n-2}) \\
&= -F_0 + F_{2n} = F_{2n}
\end{aligned}$$

$\square$

Series are infinite sums. A natural question to ask is whether it is possible to talk about infinite products. The answer is yes: if $a_n \in \mathbb{R}$, then $\prod_{n=1}^{\infty} a_n$ exists if the sequences of partial products $\prod_{n=1}^{m} a_n$ converges as $m \mapsto \infty$ to a non-zero value. When $a_n > 0$ for all $n \geqslant 1$,

$$\log\left(\prod_{n=1}^{m} a_n\right) = \sum_{n=1}^{m} \log(a_n),$$

so convergence of the infinite product is equivalent to convergence of the series $\sum_{n=1}^{\infty} \log(a_n)$.

Example 4: Evaluate

$$\prod_{n=2}^{\infty} \left(1 - \frac{1}{n^2}\right).$$

Solution: Note that $1 - \frac{1}{n^2} = \frac{n^2-1}{n^2}$. This means that, in the infinite product, we keep multiplying numbers that are in the interval $(0,1)$ and $\frac{n^2-1}{n^2}$ becomes closer and closer to 1 as n increases. Therefore, the product keeps getting smaller and smaller, although it always remain positive, so it will converge, but it's not clear for the moment if it will converge to 0 or to a positive value in $(0,1)$. (Actually, when the limit of the product is 0, we say that the product diverges, note that it converges. This may seem strange, but there is a reason for this.)

Let's transform this product into a series (although this is not necessary, but you are more familiar with series):

$$\begin{aligned}
\log\left(\prod_{n=2}^{\infty}\frac{n^2-1}{n^2}\right) &= \sum_{n=2}^{\infty}\log\left(\frac{n^2-1}{n^2}\right) \\
&= \sum_{n=2}^{\infty}(\log(n^2-1)-\log(n^2)) \\
&= \sum_{n=2}^{\infty}(\log(n-1)+\log(n+1)-2\log(n)) \text{ since } n^2-1=(n-1)(n+1) \\
&= (\log(1)+\log(3)-2\log(2))+(\log(2)+\log(4)-2\log(3)) \\
&\quad +(\log(3)+\log(5)-2\log(4))+(\log(4)+\log(6)-2\log(5)) \\
&\quad +(\log(5)+\log(7)-2\log(6))+(\log(6)+\log(8)-2\log(7))+\cdots
\end{aligned}$$

From the pattern above, one can see that all the terms will cancel except $\log(1)-2\log(2)+\log(2)$, which simplify to $-\log(2)$. Therefore, since $x=e^{\log(x)}$ when $x>0$, it follows, after setting $x=\prod_{n=2}^{\infty}\left(1-\frac{1}{n^2}\right)$, that

$$\prod_{n=2}^{\infty}\left(1-\frac{1}{n^2}\right)=e^{-\log(2)}=e^{\log\left(\frac{1}{2}\right)}=\frac{1}{2}.$$

□

You may have seen that $\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi}{6}$. It is also possible to show that $\prod_{n=1}^{\infty}\frac{4n^2}{4n^2-1}=\frac{\pi}{2}$.

Power series are infinite sums involving powers of a variable x. Many important functions can be represented using power series around 0. These are called Taylor (or Maclaurin) series and you should have seen the following:

$$\sum_{n=0}^{\infty}x^n=1+x+x^2+x^3+\cdots=\frac{1}{1-x} \text{ for } |x|<1 \tag{5}$$

$$e^x=\sum_{n=0}^{\infty}\frac{x^n}{n!}=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\cdots \tag{6}$$

$$\log(1+x)=\sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n}=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\cdots \text{ for } |x|<1 \tag{7}$$

$$\sin(x)=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!}=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\cdots \tag{8}$$

$$\cos(x)=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n)!}=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\cdots \tag{9}$$

Note that taking the derivative term-by-term of the Taylor series of $\sin(x)$ gives the Taylor series for $\cos(x)$ (which is compatible with the fact that $\frac{d}{dx}\sin(x) = \cos(x)$), and taking the derivative term-by-term of the Taylor series of $\log(1+x)$ gives the Taylor series for $\frac{1}{1+x}$, which is compatible with the fact that $\frac{d}{dx}\log(1+x) = \frac{1}{1+x}$. Similar remarks can be made about anti-derivatives.

I hope that you have all seen a proof that $\sqrt{2}$ is an irrational number (that is, it is not a fraction). This proofs relies on the unique factorization of integers as products of prime numbers. It is also true that the number π is irrational: there are a few proofs of this result and some of them do not require very advanced mathematics, but they are all more difficult than for $\sqrt{2}$. Surprisingly, the proof that e is irrational is not very difficult.

Example 5: Prove that e is an irrational number.

Solution: The proof proceeds by contradiction, so suppose that e is rational, that is, $e = \frac{a}{b}$ with $a, b \in \mathbb{Z}$ and $b \neq 0$. In the power expansion of e^x, set $x = 1$. We obtain

$$e - \left(1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{b!}\right) = \frac{1}{(b+1)!} + \frac{1}{(b+2)!} + \cdots$$

Multiplying by $b!$ and using the assumption that $e = \frac{a}{b}$ yields

$$b!\left(\frac{a}{b} - \left(1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{b!}\right)\right) = b!\left(\frac{1}{(b+1)!} + \frac{1}{(b+2)!} + \cdots\right).$$

Therefore,

$$\begin{aligned} a(b-1)! - \left(2b! + \frac{b!}{2!} + \cdots + \frac{b!}{(b-1)!} + 1\right) &= \frac{1}{b+1} + \frac{1}{(b+1)(b+2)} \qquad (10)\\ &\quad + \frac{1}{(b+1)(b+2)(b+3)} + \cdots \\ &= \sum_{n=0}^{\infty} \frac{1}{(b+1)\cdots(b+1+n)} \qquad (11) \end{aligned}$$

The left-hand side is an integer since, if $1 \leqslant k \leqslant b$, then $\frac{b!}{k!} = b(b-1)(b-2)\cdots(k+1)$. The right-hand side of (11) is positive, hence it is a positive integer and is $\geqslant 1$.

However,

$$
\begin{aligned}
\sum_{n=0}^{\infty} \frac{1}{(b+1)\cdots(b+1+n)} &= \frac{1}{b+1} + \frac{1}{(b+1)(b+2)} + \frac{1}{(b+1)(b+2)(b+3)} + \cdots \\
&= \frac{1}{b+1}\left(1 + \frac{1}{b+2} + \frac{1}{(b+2)(b+3)} + \cdots\right) \\
&< \frac{1}{b+1}\left(1 + \frac{1}{b+1} + \frac{1}{(b+1)^2} + \cdots\right) \\
&= \frac{1}{b+1} \cdot \left(\frac{1}{1 - \frac{1}{b+1}}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{b+1} \cdot \frac{b+1}{b} \\
&= \frac{1}{b} < 1.
\end{aligned}
$$

This contradicts with what was established above, namely that the right-hand side of (11) is $\geqslant 1$. □

Example 6: Evaluate the series

$$
\frac{1}{0!} + \frac{2}{1!} + \frac{3}{2!} + \frac{4}{3!} + \cdots
$$

Solution: That series is equal to $\sum_{n=0}^{\infty} \frac{n+1}{n!}$ and

$$
\begin{aligned}
\sum_{n=0}^{\infty} \frac{n+1}{n!} &= \sum_{n=0}^{\infty} \frac{n}{n!} + \sum_{n=0}^{\infty} \frac{1}{n!} \\
&= \sum_{n=1}^{\infty} \frac{n}{n!} + \sum_{n=0}^{\infty} \frac{1}{n!} \\
&= \sum_{n=1}^{\infty} \frac{1}{(n-1)!} + \sum_{n=0}^{\infty} \frac{1}{n!} \\
&= \sum_{m=0}^{\infty} \frac{1}{m!} + \sum_{n=0}^{\infty} \frac{1}{n!} \\
&= 2e.
\end{aligned}
$$

The penultimate equality follows by setting $m = n - 1$.

Wednesday, March 30, 2022

There is another way to compute the value of that series. It is more complicated, but it illustrates an idea which will be needed in the next example.

The series looks a bit like the power series of e^x except that the numerators are $1, 2, 3, \ldots$ instead of $1, x, x^2, x^3, \ldots$. Notice that $1, 2, 3, \ldots$ are the exponents of $x, x^2, x^3, \ldots$ and the derivative of x^n is nx^{n-1}, so exponents become coefficients when taking the derivative. This suggest that, to evaluate the series above, we should consider the derivative of e^x. However, $\frac{d}{dx}e^x = e^x$, so this will not bring us any new information.

One new idea is needed here. The general term of the series above is $\frac{n+1}{n!}$, which is what is obtained after evaluating $\frac{d}{dx}\left(\frac{x^{n+1}}{n!}\right)$ at $x = 1$, that is:

$$\frac{d}{dx}\left(\frac{x^{n+1}}{n!}\right) = \frac{(n+1)x^n}{n!} \text{ and } \left.\frac{(n+1)x^n}{n!}\right|_{x=1} = \frac{n+1}{n!}.$$

How do we obtain a power series with general term $\frac{x^{n+1}}{n!}$? We multiply the one for e^x by x:

$$xe^x = x + x^2 + \frac{x^3}{2!} + \frac{x^4}{3!} + \frac{x^5}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{x^{n+1}}{n!}.$$

Now let's take the derivative on both sides:

$$e^x + xe^x = 1 + 2x + \frac{3x^2}{2!} + \frac{4x^3}{3!} + \frac{5x^4}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{(n+1)x^n}{n!}.$$

Setting $x = 1$ on both sides of the previous equality shows that

$$2e = 1 + 2 + \frac{3}{2!} + \frac{4}{3!} + \frac{5}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{n+1}{n!}.$$

$\square$

Example 7: Evaluate the series

$$\frac{1^2}{0!} + \frac{2^2}{1!} + \frac{3^2}{2!} + \frac{4^2}{3!} + \cdots$$

Solution: It was established in the previous example that

$$(1+x)e^x = \sum_{n=0}^{\infty} \frac{(n+1)x^n}{n!}.$$

In the numerator, we need $(n+1)^2$, not just $n+1$: this suggests that we should repeat what was done in the second solution to Example 6, namely multiply by x both sides of the previous equality:

$$x(1+x)e^x = \sum_{n=0}^{\infty} \frac{(n+1)x^{n+1}}{n!}.$$

We can compute the derivative on both sides to obtain

$$(1+x)e^x + xe^x + x(1+x)e^x = (x^2+3x+1)e^x = \sum_{n=0}^{\infty} \frac{(n+1)^2 x^n}{n!}.$$

Setting $x = 1$ gives

$$5e = \sum_{n=0}^{\infty} \frac{(n+1)^2}{n!} = \frac{1^2}{0!} + \frac{2^2}{1!} + \frac{3^2}{2!} + \frac{4^2}{3!} + \cdots$$

□

Example 8: Evaluate the series

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots$$

Solution: The numerical series to evaluate is convergent by the Alternating Series test. Moreover, it is equal to the value at $x = 1$ of the power series $f(x)$ given by

$$f(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - \cdots$$

Observe that

$$f'(x) = 1 - x^2 + x^4 - x^6 + x^8 - \cdots = \frac{1}{1+x^2} \text{ when } |x| < 1.$$

This means that, for $0 \leqslant x < 1$,

$$\int_0^x \frac{1}{1+t^2} dt = f(x),$$

so

$$\arctan(x) = f(x) \text{ for } |x| < 1.$$

Therefore,

$$f(1) = \lim_{x \mapsto 1^-} f(x) = \lim_{x \mapsto 1^-} \arctan(x) = \arctan(1) = \frac{\pi}{4}.$$

□

Note that, as a consequence of the result of the last example,

$$\pi = 4\left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots\right) = \sum_{n=0}^{\infty} (-1)^n \frac{4}{n},$$

so partial sums of this series can be used to approximate π up to any degree of accuracy.

Example 9: Let $f_0(x) = e^x$ and $f_{n+1}(x) = x f_n'(x)$ for $n = 0, 1, 2, \ldots$ Show that

$$\sum_{n=0}^{\infty} \frac{f_n(1)}{n!} = e^e. \tag{12}$$

Solution: Since the constants $f_n(1)$ are defined inductively in terms of derivatives, the left-hand side of (12) is perhaps equal to the value of a Taylor series at $x = 1$. Considering that the right-hand side of (12) is equal to e^e, we are led to consider the function $g(x)$ given by $g(x) = e^{e^x}$.

If the left-hand side of (12) is the Taylor series at 0 of $g(x)$ evaluated at $x = 1$, then we must have that $g^{(n)}(0) = f_n(1)$ for all $n \geqslant 0$, where $g^{(n)}(x)$ is the n^{th}-derivative of $g(x)$. This will follow from the following claim after setting $x = 0$.

Claim: $f_n(e^x) = g^{(n)}(x)$ for all $n \geqslant 0$.

Let's prove the claim by induction. When $n = 0$, $f_0(e^x) = e^{e^x} = g(x) = g^{(0)}(x)$, so the claim is true.

Let's assume that $f_n(e^x) = g^{(n)}(x)$ and let's prove this for $n+1$. Taking the derivative with respect to x on both sides and using the chain rule for the left-hand side, we obtain $e^x f_n'(e^x) = g^{(n+1)}(x)$, so $f_{n+1}(e^x) = g^{(n+1)}(x)$ since $f_{n+1}(x) = x f_n'(x)$. This proves that the claim is also true for $n+1$ when it is true for n. By the Principle of Mathematical Induction, it must be true for all $n \geqslant 0$.

Since $e^0 = 1$, the claim implies that $f_n(1) = g^{(n)}(0)$. We can use this to evaluate (12):

$$\sum_{n=0}^{\infty} \frac{f_n(1)}{n!} = \sum_{n=0}^{\infty} \frac{g^{(n)}(0)}{n!} = \sum_{n=0}^{\infty} \frac{g^{(n)}(0)x^n}{n!}\Big|_{x=1} = g(1) = e^e.$$

$\square$

Wednesday, April 6, 2022

Let's conclude this course by studying examples of problems about series which have appeared before on the Putnam Mathematical Competition.

Example 10 [2014 Putnam, A–1]: Prove that every nonzero coefficient of the Taylor series of

$$(1 - x + x^2)e^x$$

about $x = 0$ is a rational number whose numerator (in lowest terms) is either 1 or a prime number.

Solution: Using the Taylor series at $x = 0$ of e^x, we get the Taylor series of $(1 - x + x^2)e^x$:

$$\begin{aligned}
\sum_{n=0}^{\infty}(1 - x + x^2)\frac{x^n}{n!} &= \sum_{n=0}^{\infty}\frac{x^n}{n!} - \sum_{n=0}^{\infty}\frac{x^{n+1}}{n!} + \sum_{n=0}^{\infty}\frac{x^{n+2}}{n!} \\
&= \sum_{n=0}^{\infty}\frac{x^n}{n!} - \sum_{n=1}^{\infty}\frac{x^n}{(n-1)!} + \sum_{n=2}^{\infty}\frac{x^n}{(n-2)!} \\
&= \sum_{n=0}^{\infty}\left(\frac{1}{n!} - \frac{1}{(n-1)!} + \frac{1}{(n-2)!}\right)x^n
\end{aligned}$$

The coefficient of x^n in the Taylor series of $(1 - x + x^2)e^x$ about $x = 0$ is $\frac{1}{n!} - \frac{1}{(n-1)!} + \frac{1}{(n-2)!}$ for $n \geqslant 2$, which equals $\frac{1-n+n(n-1)}{n!}$. (When $n = 0$, it is 1, and when $n = 1$, it is 0, so we don't have to consider these two cases from now on.) Observe that

$$\frac{1 - n + n(n-1)}{n!} = \frac{n^2 - 2n + 1}{n!} = \frac{(n-1)^2}{n \cdot (n-1) \cdot (n-2)!} = \frac{n-1}{n \cdot (n-2)!}.$$

If $n - 1$ is prime, there is nothing else to say. If not, then $n - 1 = ab$ with $a, b \in \mathbb{Z}$ and $2 \leqslant a, b \leqslant n - 2$. Moreover, a and b can be chosen so that $a \neq b$ unless $n - 1 = p^2$ for some prime number p, in which case the only possibility is $a = b = p$.

If $a \neq b$, then both appear in $(n-2)!$ and can be simplified so that

$$\frac{n-1}{n \cdot (n-2)!} = \frac{ab}{n \cdot (n-2)!} = \frac{1}{c} \quad \text{for some } c \in \mathbb{Z}.$$

If $a = b = p$, a prime number, then let's cancel p in $(n-2)!$, so we can write

$$\frac{n-1}{n \cdot (n-2)!} = \frac{ab}{n \cdot (n-2)!} = \frac{p}{n \cdot c \cdot (a-1)!}$$

for some $c \in \mathbb{Z}$. (More precisely, $c = (n-2)(n-3)\cdots(p+2)(p+1)$ so that $pc\cdot(p-1)! = (n-2)!$.) When written in lowest term, the numerator of $\frac{p}{n\cdot c\cdot(a-1)!}$ is either the prime integer p or 1. $\square$

Example 11 [2016 Putnam, B–1]: Let $x_0, x_1, x_2, \ldots$ be the sequence such that $x_0 = 1$ and for $n \geqslant 0$,

$$x_{n+1} = \ln(e^{x_n} - x_n)$$

(as usual, the function ln is the natural logarithm). Show that the infinite series

$$x_0 + x_1 + x_2 + \cdots$$

converges and find its sum.

Solution: Let's start by making some observation about the elements of the sequence $\{x_n\}_{n=0}^{\infty}$. Observe that $x_n > 0 \; \forall \, n \geqslant 0$ since $e^x - x > 1 \; \forall \, x > 0$ and hence $\ln(e^x - x) > \ln(1) = 0$ when $x > 0$. (This last inequality is equivalent to $e^x > 1 + x \; \forall \, x > 0$, which follows either from the Taylor expansion of e^x or from the fact that if f and g are two functions on $\mathbb{R}$ such that $f(0) = g(0)$ and $f'(x) > g'(x)$ for all $x > 0$, then $f(x) > g(x) \; \forall \, x > 0$.) Moreover, since $x_n > 0$, $x_{n+1} < \ln(e^{x_n}) = x_n$: this means that $\{x_n\}_{n=0}^{\infty}$ is a decreasing sequence of positive integers (hence bounded below by 0). It follows that it must converge to a limit L.

It is even possible to say what L is using the recurrence relation that defines this sequence:

$$x_{n+1} = \ln(e^{x_n} - x_n) \Rightarrow \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \ln(e^{x_n} - x_n) \Rightarrow L = \ln(e^L - L) \Rightarrow e^L = L - L \Rightarrow L = 0.$$

We still have to check that the series $\sum_{n=0}^{\infty} x_n$ converges and evaluate it. Observe that $e^{x_{n+1}} - x_{n+1} = e^{x_n} - x_n - x_{n+1}$. This suggests setting $y_n = e^{x_n} - x_n$, so that $y_n - y_{n+1} = x_{n+1}$. The series $x_0 + x_1 + x_2 + \cdots$ is thus equal to the telescoping series

$$1 + (y_0 - y_1) + (y_1 - y_2) + (y_2 - y_3) + \cdots$$

which equals $1 + y_0$ and $1 + y_0 = e$.

Wait! Since $\lim_{n\to\infty} x_n = 0$, it follows that

$$\lim_{n\to\infty} y_n = \lim_{n\to\infty} (e^{x_n} - x_n) = e^0 - 0 = 1.$$

This means that we cannot simply say that the series $1 + (y_0 - y_1) + (y_1 - y_2) + (y_2 - y_3) + \cdots$ is a telescoping series that converges to $1 + y_0$. Actually, the partial sums S_m of the series $\sum_{n=0}^{\infty} x_n$ is

$$S_m = \sum_{n=0}^{m} x_n = 1 + (y_0 - y_1) + (y_1 - y_2) + (y_2 - y_3) + \cdots + (y_{m-1} - y_m) + (y_m - y_{m+1}),$$

so $S_m = 1 + y_0 - y_{m+1}$. Therefore,

$$\lim_{m\to\infty} S_m = \lim_{m\to\infty} 1 + y_0 - y_{m+1} = y_0 = e - 1.$$

The series $\sum_{n=0}^{\infty} x_n$ thus converges to $e - 1$, not to e as was first stated above.

It is also possible to reach the same conclusion by setting $z_n = e^{x_n}$ so that $x_n = z_n - z_{n+1}$. The same argument as above applies. □

Example 12 [1961 Putnam, A–3]: Find

$$\lim_{n\to\infty} \sum_{i=1}^{N} \frac{n}{N + i^2} \text{ where } N = n^2.$$

Solution: Observe that $\frac{n}{N+i^2} = \frac{1}{n} \cdot \frac{1}{1+\left(\frac{i}{n}\right)^2}$, so the sum above looks like a Riemann sum. Actually, $\sum_{i=1}^{n} \frac{n}{N+i^2}$ is a Riemann sum that approximates the integral $\int_0^1 \frac{1}{1+x^2} dx$ on the finite interval $[0,1]$, hence

$$\lim_{n\to\infty} \sum_{i=1}^{n} \frac{n}{N+i^2} = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{1}{n} \cdot \frac{1}{1+\left(\frac{i}{n}\right)^2} = \int_0^1 \frac{1}{1+x^2} dx = \arctan(1) - \arctan(0) = \frac{\pi}{4} - 0 = \frac{\pi}{4}.$$

As i ranges from 1 to n^2, $\frac{i}{n}$ ranges from $\frac{1}{n}$ to n, so $\sum_{i=1}^{N} \frac{n}{N+i^2}$ is a Riemann sum on the infinite interval $[0,\infty)$ which converges, as $n \to \infty$, to $\int_0^\infty \frac{1}{1+x^2} dx$. Finally,

$$\begin{aligned}
\int_0^\infty \frac{1}{1+x^2} dx &= \lim_{t\to\infty} \int_0^t \frac{1}{1+x^2} dx \\
&= \lim_{t\to\infty} \arctan(x)\Big|_0^t \\
&= \lim_{t\to\infty} \arctan(t) - \arctan(0) = \frac{\pi}{2} - 0 = \frac{\pi}{2},
\end{aligned}$$

which is the value of the limit that we had to evaluate. □