The exponential function

Paul Balister

In *Analysis II* (previewed in *Analysis I*) one encounters the following proof of

$$\exp(x + y) = \exp(x)\exp(y), \tag{1}$$

where $\exp(x)$ is defined by

$$\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}. \tag{2}$$

Proof. First we show that (2) has infinite radius of convergence (by, e.g., the *Ratio Test*). Then apply the *Differentiation Theorem for power series* to differentiate term by term:

$$\frac{\mathrm{d}}{\mathrm{d}x}\exp(x) = \sum_{n=0}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x}\frac{x^n}{n!} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \exp(x).$$

Now differentiate $F(x) := \exp(x)\exp(c - x)$ with respect to x using the *Product Rule* and *Chain Rule* to get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}F(x) &= \exp'(x)\exp(c - x) + \exp(x)\exp'(c - x)(-1)\\ &= \exp(x)\exp(c - x) - \exp(x)\exp(c - x) = 0.\end{aligned}$$

Now use the *Constancy Theorem* to deduce that the function $F(x)$ is constant on $\mathbb{R}$: $\exp(x)\exp(c - x) = F(x) = F(0) = \exp(c)$. Substituting $c = x + y$ now gives the result (1). $\square$

There are a couple of things that are rather unsatisfactory about this proof.

1. It uses a lot of machinery from *Analysis II* to prove something that is very simple, and certainly should be accessible using just the methods of *Analysis I*.

2. It only works for *real* x and y. To generalise to complex numbers needs significant extra effort.

To be fair, the conclusion of the *Constancy Theorem* also applies to complex functions, and indeed much more will be proved in the Part A course *Metric Spaces and Complex Analysis*. But in some sense that makes the 1st point above even worse — one needs to wait until the 2nd year to see a complete proof of this basic identity.

Here we present two proofs of (1) that rely only on results from *Analysis I* and also work for complex numbers. The first is in some sense more natural and elementary[1], but uses a different definition of the exponential. The second uses (2), but is slightly more opaque. We then show the two definitions give the same function.

For the first proof we try to use the following definition.[2]

$$\exp(x) = \lim_{n\to\infty} \big(1 + \tfrac{x}{n}\big)^n.$$

The hard part is showing that this converges. For *real* x it is not too difficult. A slight generalisation of question 4 on problem sheet 1 of *Analysis I* shows that the sequence is increasing in n for $n > |x|$ and bounded above. It is a bit more tricky for complex x however. To make life a little easier, we will just show convergence along a certain subsequence. This will be enough and makes the algebra a bit simpler. Hence we will define (using a different notation for exp as we haven't shown it matches (2) yet)

$$e(x) := \lim_{n\to\infty} e_n(x), \qquad \text{where} \qquad e_n(x) := \big(1 + \tfrac{x}{2^n}\big)^{2^n}. \tag{3}$$

We first observe that $e_n(x)$ satisfies the following easy identity:

$$e_{n+1}(x) = \Big(1 + \frac{x/2}{2^n}\Big)^{2^n\cdot 2} = e_n\big(\tfrac{x}{2}\big)^2. \tag{4}$$

The following lemma shows that $e_n(x) \approx 1 + x$ for x small with an explicit error bound.

Lemma 1. *For $|x| \leq \frac{1}{2}$ and any $n \geq 0$, $|e_n(x) - (1 + x)| \leq |x|^2$.*

Proof. We use induction on n, $n = 0$ is trivial as $e_0(x) = 1 + x$. Now for $|x| \leq \frac{1}{2}$ and $n \geq 0$ write $e_n(\frac{x}{2}) = 1 + \frac{x}{2} + \eta$ where, by induction, $|\eta| \leq |\frac{x}{2}|^2 = \frac{|x|^2}{4}$. Then

$$\begin{aligned}
\big|e_{n+1}(x) - (1 + x)\big| &= \big|e_n(\tfrac{x}{2})^2 - (1 + x)\big| && \text{from (4)}\\
&= \big|(1 + \tfrac{x}{2} + \eta)^2 - (1 + x)\big| && e_n(\tfrac{x}{2}) = 1 + \tfrac{x}{2} + \eta\\
&= \big|\tfrac{x^2}{4} + 2(1 + \tfrac{x}{2})\eta + \eta^2\big| && \text{expand and simplify}\\
&\leq \tfrac{|x|^2}{4} + |\eta|(2 + |x| + |\eta|) && \text{triangle inequality}\\
&\leq \tfrac{|x|^2}{4} + \tfrac{|x|^2}{4}(2 + |x| + \tfrac{|x|^2}{4}) && |\eta| \leq \tfrac{|x|^2}{4}\\
&\leq |x|^2(\tfrac{1}{4} + \tfrac{1}{4}(2 + \tfrac{1}{2} + \tfrac{1}{16})) && |x| \leq \tfrac{1}{2}\\
&\leq |x|^2. && \square
\end{aligned}$$

The next step is the crucial bit: showing $e_n(x)$ converges.

[1]'Elementary' is a technical term here. It means the proof does not use advanced concepts and theorems, but instead uses more basic techniques. It does *not* however mean the proof is simple or easy!

[2]Think 'compound interest'.

Lemma 2. *For any fixed $x \in \mathbb{C}$, $e_n(x)$ converges as $n \to \infty$.*

Proof. We first note that for any $n \geq 0$ and $|x| \leq \frac{1}{2}$, $|e_n(x)| \leq 1 + |x| + |x|^2 \leq \frac{7}{4}$ by Lemma 1 and the triangle inequality. Now assume $|x| \leq \frac{1}{2}$. We show that for all $n > m \geq 0$,

$$|e_n(x) - e_m(x)| \leq \left(\tfrac{7}{8}\right)^m |x|^2 \tag{5}$$

by induction on m (simultaneously for all $n > m$). Lemma 1 is just the case $m = 0$ and, assuming the result for m,

$$\begin{aligned} |e_{n+1}(x) - e_{m+1}(x)| &= \left|e_n(\tfrac{x}{2})^2 - e_m(\tfrac{x}{2})^2\right| \\ &= \left|e_n(\tfrac{x}{2}) + e_m(\tfrac{x}{2})\right|\left|e_n(\tfrac{x}{2}) - e_m(\tfrac{x}{2})\right| \\ &\leq \left(\tfrac{7}{4} + \tfrac{7}{4}\right) \cdot (\tfrac{7}{8})^m \left|\tfrac{x}{2}\right|^2 = (\tfrac{7}{8})^{m+1}|x|^2 \end{aligned}$$

for all $n + 1 > m + 1$. Hence (5) holds for all $n > m \geq 0$.

Now $(\frac{7}{8})^m |x|^2 \to 0$ as $m \to \infty$, so $(e_n(x))$ is a Cauchy sequence. Thus by the *Cauchy Convergence Criterion*, $(e_n(x))$ converges as $n \to \infty$ for any x with $|x| \leq \frac{1}{2}$.

For larger $|x|$ we note that if $e_n(x)$ converges as $n \to \infty$ for any $|x| < K$ then $e_n(x) = e_{n-1}(\frac{x}{2})^2$ converges for any $|x| < 2K$ by AOL. Hence we can inductively show $(e_n(x))$ converges as $n \to \infty$ for $|x| < 2^t$ for any $t \geq -1$. Thus $(e_n(x))$ converges as $n \to \infty$ for all $x \in \mathbb{C}$ and so definition (3) of $e(x)$ makes sense. □

Having shown $e(x)$ is well defined, the main result is relatively straightforward.

Theorem 3. *For all $x, y \in \mathbb{C}$, $e(x + y) = e(x)e(y)$.*

Proof. We observe that

$$\frac{e_n(x)e_n(y)}{e_n(x+y)} = \left(\frac{(1+2^{-n}x)(1+2^{-n}y)}{1+2^{-n}(x+y)}\right)^{2^n} = \left(1 + \frac{4^{-n}xy}{1+2^{-n}(x+y)}\right)^{2^n} = \left(1 + \frac{z_n}{2^n}\right)^{2^n} = e_n(z_n), \tag{6}$$

where $z_n = \frac{2^{-n}xy}{1+2^{-n}(x+y)}$ (and the denominator in (6) is nonzero for sufficiently large n by definition of $e_n(x)$). But $z_n \to 0$ as $n \to \infty$ by the fact that $2^{-n} \to 0$ and AOL so, for sufficiently large n, $|z_n| \leq \frac{1}{2}$. But then $|e_n(z_n) - (1 + z_n)| \leq |z_n|^2$ by Lemma 1. But $z_n \to 0$, so $e_n(z_n) \to 1$ by sandwiching.

Now take (6) in the form $e_n(x + y)e_n(z_n) = e_n(x)e_n(y)$ and apply AOL and Lemma 2 to obtain $e(x + y) = e(x)e(y)$. □

We remark that this proof uses the *Cauchy Convergence Criterion*, which was pretty much inevitable as we wanted to show a sequence of complex numbers converged to something we could not previously describe. (In the other proofs it is hidden away in the various results that are used.) Other than that it just uses AOL, sandwiching, and simple bounding techniques (triangle inequality and standard algebra).

We also remark that once differentiation is introduced, Lemma 1 tells us precisely that $e'(0) = 1$. Then $e'(x) = e(x)$ can be deduced from the *Chain Rule* and Theorem 3: $e'(x+c) = e'(x+c)\frac{\mathrm{d}}{\mathrm{d}x}(x+c) = \frac{\mathrm{d}}{\mathrm{d}x}e(x+c) = \frac{\mathrm{d}}{\mathrm{d}x}e(c)e(x) = e(c)e'(x)$, and set $x = 0$.

We now come to the second proof of (1). This uses more properties of series from *Analysis I*. We start with a general result that could have been proved in *Analysis I* and is of independent interest (and was question 3 on the 2013 Prelims MII exam!).

Lemma 4 (Multiplication of absolutely convergent series)**.** *If $\sum a_k$ and $\sum b_k$ converge absolutely and we define $c_k := a_0b_k + a_1b_{k-1} + \cdots + a_kb_0 = \sum_{j=0}^{k} a_jb_{k-j}$, then $\sum c_k$ converges absolutely and*

$$\sum_{k=0}^{\infty} c_k = \sum_{k=0}^{\infty} a_k \sum_{k=0}^{\infty} b_k.$$

We remark that this is basically a theorem about rearranging terms in a series, summing them up in a different order. As a result it should not come as much of a surprise that the condition of *absolute* convergence is needed. The lemma would be false in general if the sequences converged, but not absolutely.

Proof. By absolute convergence of $\sum a_k$ and $\sum b_k$ we can define $A := \sum_{k=0}^{\infty} |a_k|$ and $B := \sum_{k=0}^{\infty} |b_k|$. Now fix $\varepsilon > 0$. By convergence of the sums defining A and B, there exists an N such that $\sum_{k=N+1}^{\infty} |a_k| < \frac{\varepsilon}{A+B}$ and $\sum_{k=N+1}^{\infty} |b_k| < \frac{\varepsilon}{A+B}$. Now[3]

$$\begin{aligned}\sum_{i=0}^{n} a_i \sum_{j=0}^{n} b_j - \sum_{k=0}^{n} c_k &= \sum_{i=0}^{n}\sum_{j=0}^{n} a_ib_j - \sum_{k=0}^{n}\sum_{i+j=k} a_ib_j \\ &= \sum_{i=0}^{n}\sum_{j=0}^{n} a_ib_j - \sum_{\substack{i=0\\ i+j=k\le n}}^{n}\sum_{j=0}^{n} a_ib_j = \sum_{\substack{i=0\\ i+j>n}}^{n}\sum_{j=0}^{n} a_ib_j.\end{aligned}$$

We fix $n > 2N$. Then if $i + j > n$ either $i > N$ or $j > N$. Thus by the triangle inequality

$$\begin{aligned}\Big|\sum_{i=0}^{n} a_i \sum_{j=0}^{n} b_j - \sum_{k=0}^{n} c_k\Big| \le \sum_{\substack{i=0\\ i+j>n}}^{n}\sum_{j=0}^{n} |a_ib_j| &\le \sum_{i=N+1}^{n}\sum_{j=0}^{n} |a_ib_j| + \sum_{i=0}^{n}\sum_{j=N+1}^{n} |a_ib_j| \\ &= \sum_{i=N+1}^{n} |a_i| \sum_{j=0}^{n} |b_j| + \sum_{i=0}^{n} |a_i| \sum_{j=N+1}^{n} |b_j| \\ &< \tfrac{\varepsilon}{A+B} \cdot B + A \cdot \tfrac{\varepsilon}{A+B} = \varepsilon.\end{aligned}$$

Hence $|\sum_{k=0}^{n} a_k \sum_{k=0}^{n} b_k - \sum_{k=0}^{n} c_k| \to 0$ as $n \to \infty$. By AOL and sandwiching, $\sum c_k$ converges and $\sum_{k=0}^{\infty} c_k = \sum_{k=0}^{\infty} a_k \sum_{k=0}^{\infty} b_k$.

For absolute convergence, replace a_k and b_k by $a'_k := |a_k|$ and $b'_k := |b_k|$ and note that $\sum c'_k$ converges where $c'_k := \sum |a_jb_{k-j}| \ge |c_k|$. Now $\sum |c_k|$ converges by the *Comparison Test*. $\square$

[3] It might help here to arrange the terms a_ib_j in a grid and think about which regions of the grid are included in each double sum.

Theorem 5. *For all $x, y \in \mathbb{C}$, $\exp(x+y) = \exp(x)\exp(y)$.*

Proof. For any $x, y \in \mathbb{C}$ the power series for $\exp(x)$ and $\exp(y)$ converge absolutely as the series has infinite radius of convergence (by e.g., the *Ratio Test*). Hence by Lemma 4

$$\begin{aligned}\exp(x)\exp(y) &= \sum_{n=0}^{\infty}\frac{x^n}{n!}\sum_{n=0}^{\infty}\frac{y^n}{n!} = \sum_{n=0}^{\infty}\sum_{j=0}^{n}\frac{x^j}{j!}\cdot\frac{y^{n-j}}{(n-j)!} \\ &= \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{j=0}^{n}\binom{n}{j}x^j y^{n-j} = \sum_{n=0}^{\infty}\frac{(x+y)^n}{n!} = \exp(x+y),\end{aligned}$$

where in the second line we have used the definition $\binom{n}{j} = \frac{n!}{j!(n-j)!}$ and the *Binomial Theorem.* □

Theorem 6. *For all $x \in \mathbb{C}$, $e(x) = \exp(x)$.*

Direct proof without assuming Theorems 3 and 5. Set $N = 2^m$ and use the *Binomial Theorem*

$$e_m(x) = \Big(1 + \frac{x}{N}\Big)^N = 1 + \binom{N}{1}\Big(\frac{x}{N}\Big) + \binom{N}{2}\Big(\frac{x}{N}\Big)^2 + \cdots + \binom{N}{N}\Big(\frac{x}{N}\Big)^N.$$

We note that for each *fixed* k

$$\binom{N}{k}\Big(\frac{x}{N}\Big)^k = \frac{N(N-1)\cdots(N-k+1)}{N^k\cdot k!}x^k = \frac{x^k}{k!}\prod_{j=0}^{k-1}\Big(1-\frac{j}{N}\Big) \to \frac{x^k}{k!},$$

as $m \to \infty$ (and hence $N \to \infty$) by AOL. We also note that $|\binom{N}{k}(\frac{x}{N})^k| \le |\frac{x^k}{k!}|$ so each term in $e_m(x)$ is no larger than the corresponding term in $\exp(x)$. So fix $\varepsilon > 0$ and pick n large enough such that $\sum_{k=n+1}^{\infty}|\frac{x^k}{k!}| < \frac{\varepsilon}{4}$. We can do this as the series for $\exp(x)$ converges absolutely. Now pick m_0 large enough so that for each $k = 2, \dots, n$ and each $N \ge 2^{m_0}$

$$\Big|\binom{N}{k}\Big(\frac{x}{N}\Big)^k - \frac{x^k}{k!}\Big| < \frac{\varepsilon}{2n}.$$

We can do this for each $k = 2, \dots, n$ and just take the largest m_0 that is needed (n here is fixed, so there are only finitely many ks). Then for $m \ge m_0$,

$$\begin{aligned}|\exp(x) - e_m(x)| &\le \sum_{k=2}^{n}\Big|\binom{N}{k}\Big(\frac{x}{N}\Big)^k - \frac{x^k}{k!}\Big| + \sum_{k=n+1}^{N}\binom{N}{k}\Big|\frac{x}{N}\Big|^k + \sum_{k=n+1}^{\infty}\Big|\frac{x^k}{k!}\Big| \\ &\le \sum_{k=2}^{n}\Big|\binom{N}{k}\Big(\frac{x}{N}\Big)^k - \frac{x^k}{k!}\Big| + \sum_{k=n+1}^{N}\Big|\frac{x^k}{k!}\Big| + \sum_{k=n+1}^{\infty}\Big|\frac{x^k}{k!}\Big| \\ &\le \sum_{k=2}^{n}\frac{\varepsilon}{2n} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} < \varepsilon.\end{aligned}$$

As ε was arbitrary, $e_m(x) \to \exp(x)$ as $m \to \infty$ and so $e(x) = \exp(x)$. □

Note that we could have used this as an alternative proof of Lemma 2. However it uses results on convergence of power series and the *Binomial Theorem*, whereas the proof of Lemma 2 did not.

Once we have (1) for all complex x and y, it is easy to quickly deduce standard results about both exp and the trigonometric functions.

For example, we can use the power series definitions of $\cos x$ and $\sin x$ to prove for all *complex* x,

$$\begin{aligned} e^{ix} &= \cos x + i \sin x, \\ e^{-ix} &= \cos x - i \sin x, \end{aligned}$$

or equivalently *define* (for any *complex* x)

$$\cos x = \frac{e^{ix} + e^{-ix}}{2} \qquad \sin x = \frac{e^{ix} - e^{-ix}}{2i}.$$

Addition formulae follow immediately:

$$\begin{aligned} \cos(x+y) + i\sin(x+y) &= e^{i(x+y)} = e^{ix}e^{iy} = (\cos x + i \sin x)(\cos y + i \sin y) \\ &= (\cos x \cos y - \sin x \sin y) + i(\cos x \sin y + \sin x \cos y), \\ \cos(x+y) - i\sin(x+y) &= e^{-i(x+y)} = e^{-ix}e^{-iy} = (\cos x - i \sin x)(\cos y - i \sin y) \\ &= (\cos x \cos y - \sin x \sin y) - i(\cos x \sin y + \sin x \cos y). \end{aligned}$$

Solving these simultaneous equations (add and subtract) then gives that for all $x, y \in \mathbb{C}$

$$\begin{aligned} \cos(x+y) &= \cos x \cos y - \sin x \sin y, \\ \sin(x+y) &= \cos x \sin y + \sin x \cos y. \end{aligned}$$

Other trigonometric formulae can be derived in a similar manner, valid for all complex numbers.