Notes on Complements of Analysis

Jacopo Tissino

December 21, 2016

Chapter 1

Set theory

1.1 The ZFC axioms

Extensionality

$$\forall x : \forall y : \forall a : x = y \iff (a \in x \iff a \in y) \tag{1.1.1}$$

Existence of the null set

$$\exists x : \forall y : y \notin x \tag{1.1.2}$$

Foundation Every nonempty set contains an $\in$-minimal element:

$$\forall A : \exists x \in A : \forall y \in A : y \notin x \tag{1.1.3}$$

This means that there cannot be an infinite $\in$ chain like $A_1 \ni A_2 \ni A_3 \ni \dots$.
We can also say that $\forall A : \exists x \in A : x \cap A = \emptyset$.
This also exclude the existence of the set of all sets: $\nexists x : \forall y : y \in x$.

Separation Given a well-defined property $P(x)$, there exists a set such that

$$\forall y : \forall x : x \in y \wedge P(x) \tag{1.1.4}$$

This implies the existence of the empty set, and excludes Russel's paradox.

Pair sets

$$\forall a : \forall b : \exists x : \forall y : y \in x \iff (y = a \vee y = b) \tag{1.1.5}$$

This implies the existence of singlets, and of ordered pairs, defined as: $(a, b) := \{\{a\}, \{a, b\}\}$ ($a = \cap(a, b)$, $b = \cup(a, b) \setminus \cap(a, b)$).
Of course,

$$(a, b) = (c, d) \iff (a = c) \wedge (b = d) \tag{1.1.6}$$

Union set axiom

$$\forall x : \exists u : \forall z : \exists y : z \in u \iff (z \in y \wedge y \in x) \tag{1.1.7}$$

The usual notation is $u = \cup x$, or $A \cup B$. This also enables us to define intersections:

$$A \cap B = \big\{x \in \{A \cup B\} : x \in A \wedge x \in B\big\} \tag{1.1.8}$$

Power set axiom

$$\forall x : \exists p : \forall y : y \in p \iff y \subseteq x \tag{1.1.9}$$

The usual notation is: $p = \mathcal{P}(x)$.

Infinity

$$\exists x : \forall y : \emptyset \in x \wedge (y \in x \implies y \cup \{y\} \in x) \tag{1.1.10}$$

Replacement Given the set A, we can construct the set $\{x \in A : R(x)$. This allows us to construct infinite unions: given the sets A_i, $i \in \mathbb{N}$,

$$W = \mathbb{N} \to \{A_0, A_1, A_2 ...\} = I \tag{1.1.11}$$

then $\exists \cup I = \cup_{n\in\mathbb{N}} A_n$.

Choice Given a set A of nonempty sets, such that any two are disjoint, we can always find a set B containing exactly one element for any element of A.

(Reformulate as: $A \times B = \emptyset$ iff $A = \emptyset$ or $B = \emptyset$.)

1.1.1 Goedel

In 1938 Goedel proved that ZFC is coherent. In 1931 he proved that any coherent axiom set contains undecidable propositions. One example for ZFC is the continuum hypothesis.

1.2 The Von Neumann Integers

We define $0_{VN} = \emptyset$, and $S(n_{VN}) = n \cup \{n_{VN}\}$. So $1_{VN} = \{\emptyset\}$, $2_{VN} = \{\emptyset, \{\emptyset\}\}$, $3_{VN} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$...

The $\leq$ relation is thus replaced by $\subseteq$, and $<$ by $\in$.

The axiom of infinity seems to define the VN integers, but many sets could have those properties. So, we define the property $P(x) = \emptyset \in x \wedge (y \in x \implies y \cup \{y\} \in x)$

We'd like to intersect all the sets satisfying $P(x)$ (HOW?)

$$\omega := \{k \in x : k \in Y \iff \forall Y \in \mathcal{P}(x) : P(y) \tag{1.2.1}$$

to get the actual set $\mathbb{N}$.

1.3 Cardinality

Given the sets A and B, we say they have the same cardinality if there exists a bijective $f : A \to B$.

Having the same cardinality is an equivalence relation, but the set of all the sets with the same cardinality is not a set.

We say that $|a| \leq |B|$ if $\exists f : A \to B$ injective. This is an order relation: it is

- reflexive: $|A| \leq |A|$;
- transitive: $|A| \leq |B| \wedge |B| \leq |C| \implies |A| \leq |C|$;
- antisymmetric: $|A| \leq |B| \wedge |B| \leq |A| \implies |A| = |B|$;
- connected if there is a good ordering: $\forall A, B : |A| \leq |B| \vee |B| \leq |A|$.

Saying that there exists a surjective $g : B \to A$ is equivalent to saying there exists an injective $f : A \to B$.

Proof. If there exists an injective $f : A \to B$, we define $g(y) := f^{-1}(y)$ if $y \in f(A)$, and any a_0 otherwise. g is surjective.

If there exists a surjective $g : B \to A$, we define $f(x)$ as any element in $g^{-1}\{x\}$. We need the axiom of choice for this. □

Proof of point 1.3, by Cantor-Bernstein-Schroeder. We have the two bijective functions $f : A \to B$ and $g : B \to A$. Then □

Chapter 2

Inequalities

2.1 Means

We shall treat sequences (n-uples) in the form $a = (a_1, a_2, \ldots, a_n)$ where $\forall i : a_i \geq 0$.

Definition 2.1.1. The r-mean of a, denoted $M_r(a)$ or just M_r, is:

$$M_r(a) := \left(\frac{1}{n} \sum_{i=1}^{n} a_i^r \right)^{\frac{1}{r}} \tag{2.1.1}$$

If $r < 0$ and $\exists i : a_i = 0$, we take $M_r = 0$.

Some notable means are:

- the arithmetic mean $A := M_1$; any mean can be written as $M_r = A(a_i^r)^{1/r}$
- the harmonic mean $H := M_{-1}$
- the geometric mean $G(a) := \sqrt[n]{\prod_i a_i} = \exp[A(\log a_i)]$

Definition 2.1.2. Given the set of weights $p = (p_1, p_2, \ldots, p_n)$, where $\forall i : p_i > 0$, the weighted sum M_r is:

$$M_r = M(a, p) := \left(\frac{\sum_i p_i a_i^r}{\sum_i p_i} \right)^{1/r} \tag{2.1.2}$$

The weighted geometric mean, from the definition, is

$$G := \left(\prod_i a_i^{p_i} \right)^{1/\sum_i p_i} \tag{2.1.3}$$

We notice that means are 1-homogeneous, that is, $\forall\lambda \in \mathbb{R}^+ : M_r(\lambda a) = \lambda M_r(a)$.

We can always suppose that the weights all add up to 1. If this is true, we call then q_i.

We notice that $\min a_i \leq M_r(a) \leq \max a_i$, with equality iff all the a_i are equal, or if $r < 0$ and $\exists a_i = 0$. The same is true for the geometric mean.

Theorem 2.1.1.

$$\lim_{r\to 0} M_r(a) = G(a) \tag{2.1.4}$$

Proof. Suppose that $\forall i : a_i > 0$. We look at the log of M_r:

$$\lim_{r\to 0} \log M_r = \lim_{r\to 0} \frac{1}{r} \log \sum_i q_i a_i^r \tag{2.1.5}$$

Now, $\sum_i q_i a_i^r$ goes to 1 under our hypotheses.

So, adding and subtracting 1 to the sum (since the q_i add up to 1), and multiplying and dividing, we get:

$$\lim_{r\to 0} \frac{1 + \log\left(\sum_i q_i(a_i^r - 1)\right)}{\sum_i q_i(a_i^r - 1)} \frac{\sum_i q_i(a_i^r - 1)}{r} \tag{2.1.6}$$

but by the limit $\lim_{x\to 0} \log(1+x)/x = 1$ the first fraction goes to 1, so we get:

$$\lim_{r\to 0} \frac{\sum_i q_i(a_i^r)}{r} \tag{2.1.7}$$

which by the linearity of the limit and the limit $\lim_{x\to 0}(a^x - 1)/x = \log a$ equals:

$$\sum_i q_i \log a_i = \log\left(\prod_i a_i^{q_i}\right) = \log G(a) \tag{2.1.8}$$

In the case where $\exists a_i = 0$, we take the sets $b = \{a_i \neq 0\}$, $s = \{\text{the corresponding } q_i\}$. Now, in the limit

$$\lim_{r\to 0^+} M_r(a, q) = \lim_{r\to 0^+} \left(\sum_i q_i a_i^r\right)^{1/r} \tag{2.1.9}$$

we would like to swap the as for the bs and the qs for the ss, but we need to account for the fact that $\sum_i s_i < 1$. So

$$\lim_{r\to 0^+} \left(\sum_i s_i\right)^{1/r} M_r(b, s) = 0 = G(a) = \lim_{r\to 0^-} M_r(a, q) \tag{2.1.10}$$

by definition. So we define $M_0 := G$. □

Theorem 2.1.2.

$$\lim_{r\to+\infty} M_r(a_i) = \max(a_i) \tag{2.1.11}$$

$$\lim_{r\to-\infty} M_r(a_i) = \min(a_i) \tag{2.1.12}$$

Proof. We take a_k to be the maximum a_i. Then, since $(q_i a_i^r)^{1/r} \leq \left(\sum_i q_i a_i^r\right)^{1/r}$, we can write

$$q_k^{1/r} a_k \leq M_r(a_i) \leq \max(a_i) \tag{2.1.13}$$

which, by the squeeze theorem, implies the thesis.

For $r \to -\infty$, we just need to notice that

$$M_{-r} a_i = \frac{1}{M_r\left(\frac{1}{a_i}\right)} \tag{2.1.14}$$

and that the maximum of the $1/a_i$ corresponds to the minimum of the a_i. □

Cauchy's Inequality

Theorem 2.1.3. Given two sequences of numbers a_i and b_i with the usual properties:

$$\left(\sum_i a_i b_i\right)^2 \leq \left(\sum_i a_i^2\right)\left(\sum_i b_i^2\right) \tag{2.1.15}$$

The equality holds iff the vectors a and b are linearly dependent.

Proof. We can rearrange the inequality like:

$$\left(\sum_i a_i^2\right)\left(\sum_j b_j^2\right) - \left(\sum_k a_k b_k\right)^2 \geq 0 \tag{2.1.16}$$

$$\left(\sum_i a_i^2\right)\left(\sum_j b_j^2\right) - \left(\sum_i a_i b_i\right)\left(\sum_j a_j b_j\right) \geq 0 \tag{2.1.17}$$

$$\sum_{i,j} a_i^2 b_j^2 - \sum_{i,j} a_i b_i a_j b_j \geq 0 \tag{2.1.18}$$

$$\frac{1}{2}\sum_{i,j} 2\left(a_i^2 b_j^2 - a_i b_i a_j b_j\right) \geq 0 \tag{2.1.19}$$

$$\frac{1}{2}\sum_{i,j} \left(a_i b_j - b_i a_j\right)^2 \geq 0 \tag{2.1.20}$$

Where, in the last passage, we have swapped some indices which would have been summed over in another iteration anyway. Now, this is clearly true.

a and b are proportional iff $\forall i, j : a_i b_j - b_i a_j = 0$, that is, the matrix they span has rank 1. □

This implies that $\forall r > 0 : M_r \leq M_{2r}$, with equality iff all the a_i are equal. This can be easily proven by setting $a_i := \sqrt{p_i}$ and $b_i := \sqrt{p_i} a_i^r$ and applying the theorem.

Theorem 2.1.4. $G \leq A$.

Proof. $A = M_1 \geq M_{1/2} \geq M_{1/4} \geq M_{1/8} \geq \cdots \geq \lim_{r \to 0} M_r = G$ □

Theorem 2.1.5 (Young's Inequality).

$$\forall a, b \geq 0 : \forall p > 1 : ab \leq \frac{a^p}{p} + \frac{b^{p'}}{p'} \tag{2.1.21}$$

where $p^{-1} + p'^{-1} = 1$. Equality holds iff $b = a^{p-1}$.

Proof. We just need to use $G \leq A$ with the set $(a^p, b^{p'})$ and as weights (p^{-1}, p'^{-1}). □

Integral version of the proof. Suppose WLOG that $b \leq a^{p-1}$. Now, graph the function $y = x^{p-1}$. Now, in $[0; +\infty] \times [0; +\infty]$, consider the area between $x = a$ and $y = b$, it is ab and surely less than the sum of these integrals:

$$ab \leq \int_0^a x^{p-1} \, \mathrm{d}x + \int_0^b y^{\frac{1}{p-1}} \, \mathrm{d}y = \frac{a^p}{p} + \frac{b^{\frac{1}{p-1}+1}}{\frac{1}{p-1}+1} = \frac{a^p}{p} + \frac{b^{p'}}{p'} \tag{2.1.22}$$

□

Hoelder

Theorem 2.1.6. Given some n-uples, (a_{ji}) (j is the index of the tuple number, and i is the element in the tuple) and some weights α_i such that $\sum_i \alpha_i = 1$, the following holds:

$$\sum_i \left(\prod_j a_{ji}^{\alpha_j} \right) \leq \prod_j \left(\sum_i a_{ji} \right)^{\alpha_j} \tag{2.1.23}$$

with equality iff all the n-uples are proportional.

Proof. If one of the tuples is 0 in every position, then the theorem is automatically proven.

Otherwise, we can divide the left side of the inequality by the right to get:

$$\frac{\sum_i \left(\prod_j a_{ji}^{\alpha_j} \right)}{\prod_j \left(\sum_i a_{ji} \right)^{\alpha_j}} = \sum_i \left(\prod_j \left(\frac{a_{ji}}{\sum_k a_{jk}} \right)^{\alpha_j} \right) \leq 1 \tag{2.1.24}$$

but

$$\sum_i \left(\prod_j \left(\frac{a_{ji}}{\sum_k a_{jk}} \right)^{\alpha_j} \right) \leq \sum_i \left(\sum_j \alpha_j \left(\frac{a_{ji}}{\sum_k a_{jk}} \right) \right) = \sum_j \alpha_j \frac{\sum_i a_{ji}}{\sum_k a_{jk}} = 1 \tag{2.1.25}$$

by $G \leq A$, and since $\sum_k \alpha_k = 1$. $\square$

Chapter 3

Convex functions

3.1 Definition

Definition 3.1.1. Given an interval $I \subseteq \mathbb{R}$, $f : I \to \mathbb{R}$ is convex if, $\forall x_1, x_2 \in I$ and $\forall \lambda \in [0, 1]$:

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \tag{3.1.1}$$

This is read as "f applied to a convex combination of x_1 and x_2 is less than or equal to a convex combination of $f(x_1)$ and $f(x_2)$."

$x = \lambda x_1 + (1 - \lambda)x_2$ is a convex combination (a kind of weighted average) of x_1 and x_2; we clearly have $x_1 \leq x \leq x_2$, reaching equality on one side or the other for $\lambda = 1$ or $\lambda = 0$ respectively.

An alternative definition is: f is convex if $\forall x_1, x, x_2 : x_1 \leq x \leq x_2$:

$$f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) \geq f(x) \tag{3.1.2}$$

Proof. □

Theorem 3.1.1. If f is convex, then it is Lipschitz-continuous, which implies it is continuous, at least in the interior of an interval. Also, if the derivative is defined, then it is bounded.

Proof. □

The left and right derivatives of a convex function are defined everywhere. The following hold:

1. $f'_-(x) \leq f'_+(x)$;
2. both are non decreasing;
3. $x < y \implies f'_+(x) \leq f'_-(y)$.

Theorem 3.1.2.

$$\forall x_0 \in I : \exists m \in \mathbb{R} : \forall x \in I : f(x) \geq f(x_0) + m(x - x_0) \tag{3.1.3}$$

Clearly, m is between the left and right derivative.

Theorem 3.1.3. The set of points in which $f : I \to \mathbb{R}$ is not differentiable (that is, the left and right derivatives are not continuous) is at most countable.

Corollary 3.1.3.1. If f is differentiable in I, then it is convex if f' is increasing, and if it is two times differentiable, it is convex if $f'' \geq 0$.

The right derivative is right-continuous, the left derivative is left-continuous. The points where f is differentiable coincide with those where f'_+ and f'_- are continuous.

Theorem 3.1.4 (Jensen's inequality). If F is convex, then $\forall g : [0,1] \to \mathbb{R}$:

$$\int_0^1 F(g(x))\,\mathrm{d}x \geq F\left(\int_0^1 g(x)\,\mathrm{d}x\right) \tag{3.1.4}$$

Chapter 4

Discrete dynamical systems

These are systems in which time is quantized, and things change over time. We are given a function, and we analyze its behaviour when we iterate it. An example is the logistic equation:

$$P_{n+1} = \lambda P_n(1 - P_n) \tag{4.0.1}$$

Another one is the algorithm to calculate $\sqrt{5}$: choose any P_0 and apply

$$P_{n+1} = \frac{1}{2}\left(P_n + \frac{5}{P_n}\right) \tag{4.0.2}$$

4.1 Basic notation

The notation we will use for iteration is:

$$f^n(x) = \overbrace{f \circ f \circ \cdots \circ f \circ f}^{n}(x) \tag{4.1.1}$$

Definition 4.1.1. Given a point $x_0 \in \mathbb{R}$ and a function $f : \mathbb{R} \to \mathbb{R}$, the orbit of x_0 is the sequence of the $x_n := f^n(x_0)$ with $n \in \mathbb{N}$.

Definition 4.1.2. x_0 is a fixed point if $f(x_0) = x_0$.

Definition 4.1.3. x_0 is periodic with period k if $f^k(x_0) = x_0$. The least k for which this is true is the minimal period.

Definition 4.1.4. A point is definitively fixed (or periodic) if in its orbit there is a fixed (or periodic) point.

4.2 Graphical analysis

4.3 Fixed points

Theorem 4.3.1. Given the continuous function $f : [a, b] \to [a, b]$, f has at least one fixed point.

Proof. We can define the auxiliary function $g(x) := f(x) - x$. We then see that $g(a) \geq 0$ and $g(b) \leq 0$, then $\exists c \in [a, b] : g(c) = 0$, so $f(c) = c$. $\square$

Definition 4.3.1. The fixed point x_0 of the function f is said to be:

1. attractive if $\left|f'(x_0)\right| < 1$;
2. neutral if $\left|f'(x_0)\right| = 1$;
3. repulsive if $\left|f'(x_0)\right| > 1$.

For example, x^2 has two fixed points: 0 is attractive, 1 is repulsive.
The reason for this definition is the following theorem:

Theorem 4.3.2. If x_0 is attractive, then $\exists I(x_0, \delta)$ (an interval) such that $\forall x \in I : f^n(x) \to x_0$.
If x_0 is repulsive, then $\exists I(x_0, \delta)$ such that $\forall x \in I : \exists n \in \mathbb{N} : f^n(x) \notin I$.

Attractive points.

$$\left|f'(x_0)\right| = \lim_{x \to x_0} \left| \frac{f(x) - f(x_0)}{x - x_0} \right| \tag{4.3.1}$$

If $\left|f'(x_0)\right| < 1$, then $\exists \delta > 0, \lambda > 0$ such that

$$\forall x \in I(x_0, \delta) : \left| \frac{f(x) - f(x_0)}{x - x_0} \right| < \lambda < 1 \tag{4.3.2}$$

where we can think of λ as $f'(x_0) + \varepsilon$.
Then, if $x \in I$, we can show that the orbit converges to x_0:

$$\left|f(x) - f(x_0)\right| = \left|f(x) - x_0\right| < \lambda |x - x_0| < \lambda \delta < \delta \tag{4.3.3}$$

So the distance from x_0 has diminished:

$$\left|f^n(x) - x_0\right| < \lambda^n |x - x_0| < \lambda^n \delta \to 0 \tag{4.3.4}$$

$\square$

Repulsive points. If $\left|f'(x_0) > 1\right|$, $\exists \delta > 0, \lambda > 0$ such that $\forall x \in I(x_0, \delta)$:

$$\left| \frac{f(x) - f(x_0)}{x - x_0} \right| > \lambda > 1 \tag{4.3.5}$$

where we can think of λ as being $f'(x_0) - \varepsilon$. Like before, then, we can write

$$\left|f^n(x_0) - x_0\right| > \lambda^n |x - x_0| \to +\infty \tag{4.3.6}$$

so for some $n \in \mathbb{N}$ the point will escape the interval. $\square$

Now, we should point out that while a point will surely escape a repulsive point, there is no guarantee that it will stay outside of it; it might even come back to x_0 itself.

4.4 Periodic points

If x_0 is $n-$periodic for f, then it is fixed for f^n.

An $n-$periodic point is said to be attractive or repulsive if it is for f^n. Then, we call the orbit of x_0 attractive or repulsive.

Theorem 4.4.1. If $x_n \to L \in \mathbb{R}$, and f is continuous in L, then L is a fixed point for f.

Proof.

$$L = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} f(x_n) = f(L) \tag{4.4.1}$$

□

We can use the same proof for the case in which $x_{kn} \to L$, with $k \in \mathbb{N}$:

$$L = \lim_{n\to\infty} x_{k(n+1)} = \lim_{n\to\infty} f^k(x_{kn}) = f^k(L) \tag{4.4.2}$$

4.4.1 Chain rule

To check whether a point is fixed for f^n we need to look at its derivative: we can use the chain rule to get

$$(f^n)'(x_0) = (f^{n-1})'(f(x_0))f'(x_0) = \prod_{i=0}^{n-1} f'(f^i(x_0)) = \prod_{i=0}^{n-1} f'(x_i) \tag{4.4.3}$$

Where we use the notation in which $x_{n+1} = f(x_n)$. We then notice that $\forall i \in \{1 \dots n\} : (f^n)'(x_0) = (f^n)'(f^i(x_0))$: we can swap the order of the points.

Chapter 5

Miscellaneous

5.1 Wallis

Theorem 5.1.1 (Wallis).

$$\lim_{n\to+\infty} \frac{\prod_{i=1}^{n}(2i)^2}{(2n+1)\prod_{i=1}^{n-1}(2i+1)^2} = \frac{\pi}{2} \tag{5.1.1}$$

Proof. We define the succession I_n as:

$$I_n = \int_0^{\frac{\pi}{2}} (\sin x)^n \,\mathrm{d}x \tag{5.1.2}$$

Clearly a first property is $I_{n+1} \leq I_n$. Then, if $n \geq 2$, we can calculate:

$$\begin{aligned}
I_n &= \int_0^{\frac{\pi}{2}} \sin^2(x) \sin^{n-2} x \,\mathrm{d}x \\
&= \int_0^{\frac{\pi}{2}} (1-\cos^2 x) \sin^{n-2} x \,\mathrm{d}x \\
&= I_{n-2} - \int_0^{\frac{\pi}{2}} (\cos^2 x) \left(\frac{\sin^{n-1} x}{n-1}\right)' \mathrm{d}x \\
&= I_{n-2} - \frac{1}{n-1}\left(\cos x \sin^{n-1} x\Big]_0^{\frac{\pi}{2}} + \int_0^{\frac{\pi}{2}} \sin^n x \,\mathrm{d}x\right) \\
&= I_{n-2} - \frac{1}{n-1} I_n
\end{aligned}$$

therefore $I_n = I_{n-2}((n-1)/n)$.
Now, consider:

$$\frac{I_{2n+1}}{I_{2n}} = \frac{2n}{2n+1} \leq \frac{I_2 n + 1}{I_{2n}} \leq 1 \tag{5.1.3}$$

this, by the squeeze theorem, implies

$$\lim_{n \to +\infty} \frac{I_{2n+1}}{I_{2n}} = 1 \tag{5.1.4}$$

But we can expand the I_n into products of the even or odd I_n preceding them, so we get:

$$\lim_{n \to +\infty} \frac{I_{2n+1}}{I_{2n}} = \frac{I_1 \prod_{i=1}^{n} \frac{2i}{2i+1}}{I_0 \prod_{i=1}^{n} \frac{2i-1}{2i}} = \frac{\prod_{i=1}^{n} (2i)^2}{(2n+1) \prod_{i=1}^{n-1} (2i+1)^2} \frac{I_1}{I_0} \tag{5.1.5}$$

and $I_0 = \pi/2$, $I_1 = 1$. □

5.2 Stirling

Theorem 5.2.1. For large enough numbers, the factorial can be approximated as $n! \sim n^n e^{-n} \sqrt{2\pi n}$; that is:

$$\lim_{n \to +\infty} \frac{n!}{n^n e^{-n} \sqrt{2\pi n}} = 1 \tag{5.2.1}$$

Proof. We define the succession a_n as:

$$a_n := \log \left(\frac{n!}{n^n e^{-n} \sqrt{n}} \right) = \log n! - \log \left(\left(n + \frac{1}{2} \right) \log n - n \right) \tag{5.2.2}$$

So we just need to prove that $\lim_{n \to +\infty} a_n = \log(\sqrt{2\pi})$. First, we will show that it is strictly decreasing:

$$\begin{aligned} a_n - a_{n+1} &= \left(n + \frac{1}{2} \right) \log \left(\frac{n+1}{n} \right) - 1 \\ &= (2n+1) \left(\frac{1}{2} \log \left(\frac{1 + \frac{1}{2n+1}}{1 - \frac{1}{2n+1}} \right) - \frac{1}{2n+1} \right) \end{aligned} \tag{5.2.3}$$

We will now use the fact that, $\forall t \in [0; 1]$:

$$\frac{1}{2} \log \left(\frac{1+t}{1-t} \right) - t < \frac{t^3}{3(1-t^2)} \tag{5.2.4}$$

this can be proven with derivatives,[1] but a deeper reason is the fact that $\forall t \in [-1; 1]$ we can write the following MacLaurin series:

$$\log(1+t) = \sum_{i=1}^{+\infty} \frac{(-1)^{i+1} t^i}{i} \tag{5.2.6}$$

$$\log(1-t) = -\sum_{i=1}^{+\infty} \frac{t^i}{i} \tag{5.2.7}$$

Then we can expand:

$$\log\left(\frac{1+t}{1-t}\right) = \log(1+t) - \log(1-t) = 2\sum_{i=0}^{+\infty} \frac{t^{2i+1}}{2i+1} \tag{5.2.8}$$

so

$$\frac{1}{2}\log\left(\frac{1+t}{1-t}\right) < t + \frac{t^3}{3}\left(\sum_{i=1}^{+\infty} t^{2i}\right) = t + \frac{t^3}{3(1-t^2)} \tag{5.2.9}$$

Now we can take $t = (2n+1)^{-1}$ in equation (5.2.3), to get the following formula for $a_n - a_{n+1}$:

$$f(t) = \frac{1}{t}\left(\frac{1}{2}\log\left(\frac{1+t}{1-t}\right) - t\right) < \frac{t^2}{3(1-t^2)} \tag{5.2.10}$$

$$= \frac{(2n+1)^{-2}}{3(1-(2n+1)^{-2}} \tag{5.2.11}$$

$$= \frac{1}{12n^2+12n} \tag{5.2.12}$$

$$= 12\left(\frac{1}{n} - \frac{1}{n+1}\right) \tag{5.2.13}$$

So, we have proven that the function $a_n - (12n)^{-1}$ is strictly increasing; that is, its limit is either real or $+\infty$. Now, we shall prove that a_n is stricty decreasing, that is, $0 < a_n - a_{n+1}$.

It is enough for $f(t)$ to satisfy this condition:

$$f'(t) = \frac{1}{2}\frac{1}{1+t} + \frac{1}{2}\frac{1}{1-t} > 0 \quad \forall t \in [0,1] \tag{5.2.14}$$

---

[1]

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\left(\log\left(\frac{1+t}{1-t}\right) - t\right) - \frac{t^3}{3(1-t^2)}\right) = \frac{-2t^4-3t^2+3}{3(t^2-1)^2} < 0 \tag{5.2.5}$$

and we have equality for $t = 0$ in (5.2.4).

and $f(0) = 0$, so $\forall t \in [0,1] : f(t) > 0$. So a_n is strictly decreasing, thus its limit is either real or $-\infty$. But the limits of a_n and $a_n - (12n)^{-1}$ must be the same since their difference has limit 0: so $\lim_{n\to+\infty} a_n = c \in \mathbb{R}$.

This means that

$$\lim_{n\to+\infty} \frac{n!}{n^n e^{-n}\sqrt{2\pi n}} = e^c \tag{5.2.15}$$

Now we can apply Wallis:

$$\begin{aligned}
\lim_{n\to+\infty} \frac{\displaystyle\prod_{i=1}^{n}(2i)^2}{\displaystyle(2n+1)\prod_{i=1}^{n-1}(2i+1)^2} &= \lim_{n\to\infty} \frac{(2^n n!)^2}{(2n+1)\left(\frac{(2n)!}{2^n n!}\right)^2} \\
&= \lim_{n\to\infty} \frac{2^{4n}(n!)^4}{(2n!)^2(2n+1)} \\
&= \lim_{n\to\infty} \frac{2^{4n}\left(e^c \frac{n^n}{e^n}\sqrt{n}\right)^4}{\left(e^c \frac{(2n)^{2n}}{e^{2n}}\sqrt{2n}\right)^2 (2n+1)} \\
&= \lim_{n\to\infty} (e^c)^2 \frac{n^2}{2n(2n+1)} = \frac{(e^c)^2}{4} = \frac{\pi}{2}
\end{aligned}$$

$$e^c = \sqrt{2\pi} \implies c = \log(\sqrt{2\pi}) \tag{5.2.16}$$

□

Contents