The Irrationality of $\zeta(3)$

Manuel Eberl

December 14, 2021

Abstract

This article provides a formalisation of Beukers's straightforward analytic proof [2] that $\zeta(3)$ is irrational. This was first proven by Apéry [1] (which is why this result is also often called 'Apéry's Theorem') using a more algebraic approach. This formalisation follows Filaseta's presentation of Beukers's proof [5].

Contents

1 The Irrationality of $\zeta(3)$

theory *Zeta-3-Irrational*
imports
E-Transcendental.E-Transcendental
Prime-Number-Theorem.Prime-Number-Theorem
Prime-Distribution-Elementary.PNT-Consequences
begin

Apéry's original proof of the irrationality of $\zeta(3)$ contained several tricky identities of sums involving binomial coefficients that are difficult to prove. I instead follow Beukers's proof, which consists of several fairly straightforward manipulations of integrals with absolutely no caveats.

Both Apéry and Beukers make use of an asymptotic upper bound on $\text{lcm}\{1 \ldots n\}$ – namely $\text{lcm}\{1 \ldots n\} \in o(c^n)$ for any $c > e$, which is a consequence of the Prime Number Theorem (which, fortunately, is available in the *Archive of Formal Proofs* [4, 3]).

I follow the concise presentation of Beukers's proof in Filaseta's lecture notes [5], which turned out to be very amenable to formalisation.

There is another earlier formalisation of the irrationality of $\zeta(3)$ by Mahboubi *et al.* [6], who followed Apéry's original proof, but were ultimately forced to find a more elementary way to prove the asymptotics of $\text{lcm}\{1 \ldots n\}$ than the Prime Number Theorem.

First, we will require some auxiliary material before we get started with the actual proof.

1.1 Auxiliary facts about polynomials

lemma *degree-higher-pderiv*: *degree* ((*pderiv* ^^ *n*) *p*) = *degree p* − *n*
for *p* :: *'a*::{*comm-semiring-1*,*semiring-no-zero-divisors*,*semiring-char-0*} *poly*
⟨*proof*⟩

lemma *pcompose-power-left*: *pcompose* (*p* ^ *n*) *q* = *pcompose p q* ^ *n*
⟨*proof*⟩

lemma *pderiv-sum*: *pderiv* (∑*x*∈*A*. *f x*) = (∑*x*∈*A*. *pderiv* (*f x*))
⟨*proof*⟩

lemma *higher-pderiv-minus*: (*pderiv* ^^ *n*) (−*p* :: *'a* :: *idom poly*) = −(*pderiv* ^^ *n*) *p*
⟨*proof*⟩

lemma *pderiv-power*: *pderiv* (*p* ^ *n*) = *smult* (*of-nat n*) (*p* ^ (*n* − *1*)) ∗ *pderiv p*
⟨*proof*⟩

lemma *pderiv-monom*: *pderiv* (*monom c n*) = *monom* (*of-nat n* ∗ *c*) (*n* − *1*)
⟨*proof*⟩

lemma *higher-pderiv-monom*:
$k \le n \Longrightarrow$ (*pderiv* ^^ *k*) (*monom c n*) = *monom* (*of-nat* (*pochhammer* ($n - k$ + *1*) *k*) $*$ *c*) ($n - k$)
⟨*proof*⟩

lemma *higher-pderiv-mult*:
(*pderiv* ^^ *n*) ($p * q$) =
($\sum k \le n$. *Polynomial.smult* (*of-nat* (*n choose k*)) ((*pderiv* ^^ *k*) $p *$ (*pderiv* ^^ ($n - k$)) *q*))
⟨*proof*⟩

1.2 Auxiliary facts about integrals

theorem (**in** *pair-sigma-finite*) *Fubini-set-integrable*:
fixes *f* :: - ⇒ -::{*banach*, *second-countable-topology*}
assumes *f*[*measurable*]: *set-borel-measurable* ($M1 \bigotimes_M M2$) ($A \times B$) *f*
and *integ1*: *set-integrable M1 A* (λ*x*. ∫ *y*∈*B*. *norm* (*f* (*x*, *y*)) ∂*M2*)
and *integ2*: *AE x*∈*A in M1*. *set-integrable M2 B* (λ*y*. *f* (*x*, *y*))
shows *set-integrable* ($M1 \bigotimes_M M2$) ($A \times B$) *f*
⟨*proof*⟩

lemma (**in** *pair-sigma-finite*) *set-integral-fst′*:
fixes *f* :: - ⇒ ′*c* :: {*second-countable-topology*, *banach*}
assumes *set-integrable* ($M1 \bigotimes_M M2$) ($A \times B$) *f*
shows *set-lebesgue-integral* ($M1 \bigotimes_M M2$) ($A \times B$) *f* =
(∫ *x*∈*A*. (∫ *y*∈*B*. *f* (*x*, *y*) ∂*M2*) ∂*M1*)
⟨*proof*⟩

lemma (**in** *pair-sigma-finite*) *set-integral-snd*:
fixes *f* :: - ⇒ ′*c* :: {*second-countable-topology*, *banach*}
assumes *set-integrable* ($M1 \bigotimes_M M2$) ($A \times B$) *f*
shows *set-lebesgue-integral* ($M1 \bigotimes_M M2$) ($A \times B$) *f* =
(∫ *y*∈*B*. (∫ *x*∈*A*. *f* (*x*, *y*) ∂*M1*) ∂*M2*)
⟨*proof*⟩

proposition (**in** *pair-sigma-finite*) *Fubini-set-integral*:
fixes *f* :: - ⇒ - ⇒ - :: {*banach*, *second-countable-topology*}
assumes *f*: *set-integrable* ($M1 \bigotimes_M M2$) ($A \times B$) (*case-prod f*)
shows (∫ *y*∈*B*. (∫ *x*∈*A*. *f x y* ∂*M1*) ∂*M2*) = (∫ *x*∈*A*. (∫ *y*∈*B*. *f x y* ∂*M2*) ∂*M1*)
⟨*proof*⟩

lemma (**in** *pair-sigma-finite*) *nn-integral-swap*:
assumes [*measurable*]: *f* ∈ *borel-measurable* ($M1 \bigotimes_M M2$)
shows ($\int^+ x$. *f x* ∂($M1 \bigotimes_M M2$)) = ($\int^+ (y,x)$. *f* (*x*,*y*) ∂($M2 \bigotimes_M M1$))
⟨*proof*⟩

lemma *set-integrable-bound*:
fixes *f* :: ′*a* ⇒ ′*b*::{*banach*, *second-countable-topology*}

and g :: $'a \Rightarrow 'c$::{*banach*, *second-countable-topology*}
shows *set-integrable* M A $f \Longrightarrow$ *set-borel-measurable* M A $g \Longrightarrow$
(AE x *in* M. $x \in A \longrightarrow$ *norm* (g x) $\leq$ *norm* (f x)) $\Longrightarrow$ *set-integrable* M A g
⟨*proof*⟩

lemma *set-integrableI-nonneg*:
fixes f :: $'a \Rightarrow$ *real*
assumes *set-borel-measurable* M A f
assumes AE x *in* M. $x \in A \longrightarrow 0 \leq f\ x$ ($\int^+ x{\in}A.\ f\ x\ \partial M$) $< \infty$
shows *set-integrable* M A f
⟨*proof*⟩

lemma *set-integral-eq-nn-integral*:
assumes *set-borel-measurable* M A f
assumes *set-nn-integral* M A f = *ennreal* x $x \geq 0$
assumes AE x *in* M. $x \in A \longrightarrow f\ x \geq 0$
shows *set-integrable* M A f
and *set-lebesgue-integral* M A f = x
⟨*proof*⟩

lemma *set-integral-0* [*simp*, *intro*]: *set-integrable* M A ($\lambda y.\ 0$)
⟨*proof*⟩

lemma *set-integrable-sum*:
fixes f :: - $\Rightarrow$ - $\Rightarrow$ - :: {*banach*, *second-countable-topology*}
assumes *finite* B
assumes $\bigwedge x.\ x \in B \Longrightarrow$ *set-integrable* M A (f x)
shows *set-integrable* M A ($\lambda y.\ \sum x{\in}B.\ f\ x\ y$)
⟨*proof*⟩

lemma *set-integral-sum*:
fixes f :: - $\Rightarrow$ - $\Rightarrow$ - :: {*banach*, *second-countable-topology*}
assumes *finite* B
assumes $\bigwedge x.\ x \in B \Longrightarrow$ *set-integrable* M A (f x)
shows *set-lebesgue-integral* M A ($\lambda y.\ \sum x{\in}B.\ f\ x\ y$) = ($\sum x{\in}B.$ *set-lebesgue-integral* M A (f x))
⟨*proof*⟩

lemma *set-nn-integral-cong*:
assumes $M = M'$ $A = B$ $\bigwedge x.\ x \in$ *space* $M \cap A \Longrightarrow f\ x = g\ x$
shows *set-nn-integral* M A f = *set-nn-integral* M' B g
⟨*proof*⟩

lemma *set-nn-integral-mono*:
assumes $\bigwedge x.\ x \in$ *space* $M \cap A \Longrightarrow f\ x \leq g\ x$
shows *set-nn-integral* M A f $\leq$ *set-nn-integral* M A g
⟨*proof*⟩

lemma
fixes *f* :: *real* ⇒ *real*
assumes $a \leq b$
assumes *deriv*: $\bigwedge x.\ a \leq x \Longrightarrow x \leq b \Longrightarrow$ (*F has-field-derivative f x*) (*at x within* $\{a..b\}$)
assumes *cont*: *continuous-on* $\{a..b\}$ *f*
shows *has-bochner-integral-FTC-Icc-real*:
has-bochner-integral lborel ($\lambda x.\ f\ x * indicator\ \{a\ ..\ b\}\ x$) ($F\ b - F\ a$) (**is** *?has*)
and *integral-FTC-Icc-real*: ($\int x.\ f\ x * indicator\ \{a\ ..\ b\}\ x\ \partial lborel$) = $F\ b - F\ a$ (**is** *?eq*)
⟨*proof*⟩

lemma *integral-by-parts-integrable*:
fixes *f g F G*::*real* ⇒ *real*
assumes $a \leq b$
assumes *cont-f*[*intro*]: *continuous-on* $\{a..b\}$ *f*
assumes *cont-g*[*intro*]: *continuous-on* $\{a..b\}$ *g*
assumes [*intro*]: $\bigwedge x.\ x \in \{a..b\} \Longrightarrow$ (*F has-field-derivative f x*) (*at x within* $\{a..b\}$)
assumes [*intro*]: $\bigwedge x.\ x \in \{a..b\} \Longrightarrow$ (*G has-field-derivative g x*) (*at x within* $\{a..b\}$)
shows *integrable lborel* ($\lambda x.((F\ x) * (g\ x) + (f\ x) * (G\ x)) * indicator\ \{a\ ..\ b\}\ x$)
⟨*proof*⟩

lemma *integral-by-parts*:
fixes *f g F G*::*real* ⇒ *real*
assumes [*arith*]: $a \leq b$
assumes *cont-f*[*intro*]: *continuous-on* $\{a..b\}$ *f*
assumes *cont-g*[*intro*]: *continuous-on* $\{a..b\}$ *g*
assumes [*intro*]: $\bigwedge x.\ x \in \{a..b\} \Longrightarrow$ (*F has-field-derivative f x*) (*at x within* $\{a..b\}$)
assumes [*intro*]: $\bigwedge x.\ x \in \{a..b\} \Longrightarrow$ (*G has-field-derivative g x*) (*at x within* $\{a..b\}$)
shows ($\int x.\ (F\ x * g\ x) * indicator\ \{a\ ..\ b\}\ x\ \partial lborel$)
$= F\ b * G\ b - F\ a * G\ a - \int x.\ (f\ x * G\ x) * indicator\ \{a\ ..\ b\}\ x\ \partial lborel$
⟨*proof*⟩

lemma *interval-lebesgue-integral-by-parts*:
assumes $a \leq b$
assumes *cont-f*[*intro*]: *continuous-on* $\{a..b\}$ *f*
assumes *cont-g*[*intro*]: *continuous-on* $\{a..b\}$ *g*
assumes [*intro*]: $\bigwedge x.\ x \in \{a..b\} \Longrightarrow$ (*F has-field-derivative f x*) (*at x within* $\{a..b\}$)
assumes [*intro*]: $\bigwedge x.\ x \in \{a..b\} \Longrightarrow$ (*G has-field-derivative g x*) (*at x within* $\{a..b\}$)
shows (*LBINT x=a..b.* $F\ x * g\ x$) = $F\ b * G\ b - F\ a * G\ a$ − (*LBINT x=a..b.* $f\ x * G\ x$)

⟨*proof*⟩

lemma *interval-lebesgue-integral-by-parts-01*:
assumes *cont-f*[*intro*]: *continuous-on* {*0..1*} *f*
assumes *cont-g*[*intro*]: *continuous-on* {*0..1*} *g*
assumes [*intro*]: $\bigwedge x.\ x \in \{0..1\} \Longrightarrow$ (*F has-field-derivative f x*) (*at x within* {*0..1*})
assumes [*intro*]: $\bigwedge x.\ x \in \{0..1\} \Longrightarrow$ (*G has-field-derivative g x*) (*at x within* {*0..1*})
shows (*LBINT x=0..1. F x* ∗ *g x*) = *F 1* ∗ *G 1* − *F 0* ∗ *G 0* − (*LBINT x=0..1. f x* ∗ *G x*)
⟨*proof*⟩

lemma *continuous-on-imp-set-integrable-cbox*:
fixes *h* :: *'a* :: *euclidean-space* ⇒ *real*
assumes *continuous-on* (*cbox a b*) *h*
shows *set-integrable lborel* (*cbox a b*) *h*
⟨*proof*⟩

1.3 Shifted Legendre polynomials

The first ingredient we need to show Apéry's theorem is the *shifted Legendre polynomials*

$$P_n(X) = \frac{1}{n!}\frac{\partial^n}{\partial X^n}(X^n(1-X)^n)$$

and the auxiliary polynomials

$$Q_{n,k}(X) = \frac{\partial^k}{\partial X^k}(X^n(1-X)^n)\ .$$

Note that P_n is in fact an *integer* polynomial.

Only some very basic properties of these will be proven, since that is all we will need.

context
fixes *n* :: *nat*
begin

definition *gen-shleg-poly* :: *nat* ⇒ *int poly* **where**
gen-shleg-poly k = (*pderiv* ^^ *k*) ([:*0*, *1*, −*1*:] ^ *n*)

definition *shleg-poly* **where** *shleg-poly* = *gen-shleg-poly n div* [:*fact n*:]

We can easily prove the following more explicit formula for $Q_{n,k}$:

$$Q_{n,k}(X) = \sum_{i=0}^{k}(-1)^{k-1}\binom{k}{i}n^{\underline{i}}\,n^{\underline{k-i}}\,X^{n-i}\,(1-X)^{n-k+i}$$

lemma *gen-shleg-poly-altdef*:

assumes *k* ≤ *n*
shows *gen-shleg-poly k* =
(∑ *i*≤*k*. *smult* ((−*1*)^(*k*−*i*) ∗ *of-nat* (*k choose i*) ∗
pochhammer (*n*−*i*+*1*) *i* ∗ *pochhammer* (*n*−*k*+*i*+*1*) (*k*−*i*))
([:*0*, *1*:] ^ (*n* − *i*) ∗ [:*1*, −*1*:] ^ (*n*−*k*+*i*)))
⟨*proof*⟩

lemma *degree-gen-shleg-poly* [*simp*]: *degree* (*gen-shleg-poly k*) = *2* ∗ *n* − *k*
⟨*proof*⟩

lemma *gen-shleg-poly-n*: *gen-shleg-poly n* = *smult* (*fact n*) *shleg-poly*
⟨*proof*⟩

lemma *degree-shleg-poly* [*simp*]: *degree shleg-poly* = *n*
⟨*proof*⟩

lemma *pderiv-gen-shleg-poly* [*simp*]: *pderiv* (*gen-shleg-poly k*) = *gen-shleg-poly* (*Suc k*)
⟨*proof*⟩

The following functions are the interpretation of the shifted Legendre polynomials and the auxiliary polynomials as a function from reals to reals.

definition *Gen-Shleg* :: *nat* ⇒ *real* ⇒ *real*
where *Gen-Shleg k x* = *poly* (*of-int-poly* (*gen-shleg-poly k*)) *x*

definition *Shleg* :: *real* ⇒ *real* **where** *Shleg* = *poly* (*of-int-poly shleg-poly*)

lemma *Gen-Shleg-altdef*:
assumes *k* ≤ *n*
shows *Gen-Shleg k x* = (∑ *i*≤*k*. (−*1*)^(*k*−*i*) ∗ *of-nat* (*k choose i*) ∗
of-int (*pochhammer* (*n*−*i*+*1*) *i* ∗ *pochhammer* (*n*−*k*+*i*+*1*) (*k*−*i*)) ∗
x ^ (*n* − *i*) ∗ (*1* − *x*) ^ (*n*−*k*+*i*))
⟨*proof*⟩

lemma *Gen-Shleg-0* [*simp*]: *k* < *n* ⟹ *Gen-Shleg k 0* = *0*
⟨*proof*⟩

lemma *Gen-Shleg-1* [*simp*]: *k* < *n* ⟹ *Gen-Shleg k 1* = *0*
⟨*proof*⟩

lemma *Gen-Shleg-n-0* [*simp*]: *Gen-Shleg n 0* = *fact n*
⟨*proof*⟩

lemma *Gen-Shleg-n-1* [*simp*]: *Gen-Shleg n 1* = (−*1*) ^ *n* ∗ *fact n*
⟨*proof*⟩

lemma *Shleg-altdef*: *Shleg x* = *Gen-Shleg n x* / *fact n*
⟨*proof*⟩

lemma *Shleg-0* [*simp*]: *Shleg 0* = *1* **and** *Shleg-1* [*simp*]: *Shleg 1* = (−*1*) ^ *n*
⟨*proof*⟩

lemma *Gen-Shleg-0-left*: *Gen-Shleg 0 x* = *x* ^ *n* ∗ (*1* − *x*) ^ *n*
⟨*proof*⟩

lemma *has-field-derivative-Gen-Shleg*:
(*Gen-Shleg k has-field-derivative Gen-Shleg* (*Suc k*) *x*) (*at x*)
⟨*proof*⟩

lemma *continuous-on-Gen-Shleg*: *continuous-on A* (*Gen-Shleg k*)
⟨*proof*⟩

lemma *continuous-on-Gen-Shleg′* [*continuous-intros*]:
continuous-on A f ⟹ *continuous-on A* (λ*x*. *Gen-Shleg k* (*f x*))
⟨*proof*⟩

lemma *continuous-on-Shleg*: *continuous-on A Shleg*
⟨*proof*⟩

lemma *continuous-on-Shleg′* [*continuous-intros*]:
continuous-on A f ⟹ *continuous-on A* (λ*x*. *Shleg* (*f x*))
⟨*proof*⟩

lemma *measurable-Gen-Shleg* [*measurable*]: *Gen-Shleg n* ∈ *borel-measurable borel*
⟨*proof*⟩

lemma *measurable-Shleg* [*measurable*]: *Shleg* ∈ *borel-measurable borel*
⟨*proof*⟩

end

1.4 Auxiliary facts about the ζ function

lemma *Re-zeta-ge-1*:
assumes *x* > *1*
shows *Re* (*zeta* (*of-real x*)) ≥ *1*
⟨*proof*⟩

lemma *sums-zeta-of-nat-offset*:
fixes *r* :: *nat*
assumes *n*: *n* > *1*
shows (λ*k*. *1* / (*r* + *k* + *1*) ^ *n*) *sums* (*zeta* (*of-nat n*) − (∑*k*=*1*..*r*. *1* / *k* ^ *n*))
⟨*proof*⟩

lemma *sums-Re-zeta-of-nat-offset*:
fixes *r* :: *nat*

assumes n: $n > 1$
shows $(\lambda k.\ 1\ /\ (r + k + 1)\ \hat{}\ n)$ *sums* $(Re\ (zeta\ (of\text{-}nat\ n)) - (\sum k{=}1..r.\ 1\ /\ k\ \hat{}\ n))$
⟨*proof*⟩

1.5 Divisor of a sum of rationals

A finite sum of rationals of the form $\frac{a_1}{b_1} + \ldots + \frac{a_n}{b_n}$ can be brought into the form $\frac{c}{d}$, where d is the LCM of the b_i (or some integer multiple thereof).

lemma *sum-rationals-common-divisor*:
fixes $f\ g :: {'a} \Rightarrow int$
assumes *finite A*
assumes $\bigwedge x.\ x \in A \Longrightarrow g\ x \neq 0$
shows $\exists c.\ (\sum x{\in}A.\ f\ x\ /\ g\ x) = real\text{-}of\text{-}int\ c\ /\ (LCM\ x{\in}A.\ g\ x)$
⟨*proof*⟩

lemma *sum-rationals-common-divisor'*:
fixes $f\ g :: {'a} \Rightarrow int$
assumes *finite A*
assumes $\bigwedge x.\ x \in A \Longrightarrow g\ x \neq 0$ **and** $(\bigwedge x.\ x \in A \Longrightarrow g\ x\ dvd\ d)$ **and** $d \neq 0$
shows $\exists c.\ (\sum x{\in}A.\ f\ x\ /\ g\ x) = real\text{-}of\text{-}int\ c\ /\ real\text{-}of\text{-}int\ d$
⟨*proof*⟩

1.6 The first double integral

We shall now investigate the double integral

$$I_1 := \int_0^1 \int_0^1 -\frac{\ln(xy)}{1-xy}\, x^r y^s \,\mathrm{d}x\,\mathrm{d}y\ .$$

Since everything is non-negative for now, we can work over the extended non-negative real numbers and the issues of integrability or summability do not arise at all.

definition *beukers-nn-integral1* :: $nat \Rightarrow nat \Rightarrow ennreal$ **where**
$beukers\text{-}nn\text{-}integral1\ r\ s =$
$(\int^{+}(x,y){\in}\{0{<}..{<}1\}{\times}\{0{<}..{<}1\}.\ ennreal\ (-ln\ (x{*}y)\ /\ (1 - x{*}y) * x\hat{}r * y\hat{}s)\ \partial lborel)$

definition *beukers-integral1* :: $nat \Rightarrow nat \Rightarrow real$ **where**
$beukers\text{-}integral1\ r\ s = (\int (x,y){\in}\{0{<}..{<}1\}{\times}\{0{<}..{<}1\}.\ (-ln\ (x{*}y)\ /\ (1 - x{*}y) * x\hat{}r * y\hat{}s)\ \partial lborel)$

lemma
fixes $x\ y\ z :: real$
assumes xyz: $x \in \{0{<}..{<}1\}\ y \in \{0{<}..{<}1\}\ z \in \{0..1\}$
shows *beukers-denom-ineq*: $(1 - x * y) * z < 1$ **and** *beukers-denom-neq*: $(1 - x * y) * z \neq 1$
⟨*proof*⟩

We first evaluate the improper integral

$$\int_0^1 -\ln x \cdot x^e \, \mathrm{d}x = \frac{1}{(e+1)^2} \ .$$

for any $e > -1$.

lemma *integral-0-1-ln-times-powr*:
  assumes *e* > −*1*
  shows (*LBINT x*=*0*..*1*. −*ln x* ∗ *x powr e*) = *1* / (*e* + *1*)^2
    and *interval-lebesgue-integrable lborel 0 1* (λ*x*. −*ln x* ∗ *x powr e*)
⟨*proof*⟩

lemma *interval-lebesgue-integral-lborel-01-cong*:
  assumes ⋀*x*. *x* ∈ {*0*<..<*1*} ⟹ *f x* = *g x*
  shows *interval-lebesgue-integral lborel 0 1 f* =
          interval-lebesgue-integral lborel 0 1 g
  ⟨*proof*⟩

lemma *nn-integral-0-1-ln-times-powr*:
  assumes *e* > −*1*
  shows (∫$^+$*y*∈{*0*<..<*1*}. *ennreal* (−*ln y* ∗ *y powr e*) ∂*lborel*) = *ennreal* (*1* / (*e* + *1*)^2)
⟨*proof*⟩

lemma *nn-integral-0-1-ln-times-power*:
  (∫$^+$*y*∈{*0*<..<*1*}. *ennreal* (−*ln y* ∗ *y* ^ *n*) ∂*lborel*) = *ennreal* (*1* / (*n* + *1*)^2)
⟨*proof*⟩

Next, we also evaluate the more trivial integral

$$\int_0^1 x^n \ \mathrm{d}x \ .$$

lemma *nn-integral-0-1-power*:
  (∫$^+$*y*∈{*0*<..<*1*}. *ennreal* (*y* ^ *n*) ∂*lborel*) = *ennreal* (*1* / (*n* + *1*))
⟨*proof*⟩

I_1 can alternatively be written as the triple integral

$$\int_0^1 \int_0^1 \int_0^1 \frac{x^r y^s}{1-(1-xy)w} \ \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}w \ .$$

lemma *beukers-nn-integral1-altdef*:
  beukers-nn-integral1 r s =
    (∫$^+$(*w*,*x*,*y*)∈{*0*<..<*1*}×{*0*<..<*1*}×{*0*<..<*1*}.
      ennreal (*1* / (*1*−(*1*−*x*∗*y*)∗*w*) ∗ *x*^*r* ∗ *y*^*s*) ∂*lborel*)
⟨*proof*⟩

context

fixes *r s* :: *nat* **and** *I1 I2'* :: *real* **and** *I2* :: *ennreal* **and** *D* :: (*real* × *real* × *real*) *set*
assumes *rs*: *s* ≤ *r*
defines *D* ≡ {*0*<..<*1*}×{*0*<..<*1*}×{*0*<..<*1*}
begin

By unfolding the geometric series, pulling the summation out and evaluating the integrals, we find that

$$I_1 = \sum_{k=0}^{\infty} \frac{1}{(k+r+1)^2(k+s+1)} + \frac{1}{(k+r+1)(k+s+1)^2}\ .$$

lemma *beukers-nn-integral1-series*:
beukers-nn-integral1 r s = (∑ *k*. *ennreal* (*1*/((*k*+*r*+*1*)^*2*∗(*k*+*s*+*1*)) + *1*/((*k*+*r*+*1*)∗(*k*+*s*+*1*)^*2*)))
⟨*proof*⟩

Remembering that $\zeta(3) = \sum k^{-3}$, it is easy to see that if $r = s$, this sum is simply

$$2\left(\zeta(3) - \sum_{k=1}^{r} \frac{1}{k^3}\right)\ .$$

lemma *beukers-nn-integral1-same*:
assumes *r* = *s*
shows *beukers-nn-integral1 r s* = *ennreal* (*2* ∗ (*Re* (*zeta 3*) − (∑ *k*=*1*..*r*. *1* / *k* ^ *3*)))
and *2* ∗ (*Re* (*zeta 3*) − (∑ *k*=*1*..*r*. *1* / *k* ^ *3*)) ≥ *0*
⟨*proof*⟩

lemma *beukers-integral1-same*:
assumes *r* = *s*
shows *beukers-integral1 r s* = *2* ∗ (*Re* (*zeta 3*) − (∑ *k*=*1*..*r*. *1* / *k* ^ *3*))
⟨*proof*⟩

In contrast, for $r > s$, we find that

$$I_1 = \frac{1}{r-s} \sum_{k=s+1}^{r} \frac{1}{k^2}\ .$$

lemma *beukers-nn-integral1-different*:
assumes *r* > *s*
shows *beukers-nn-integral1 r s* = *ennreal* ((∑ *k*∈{*s*<..*r*}. *1* / *k* ^ *2*) / (*r* − *s*))
⟨*proof*⟩

lemma *beukers-integral1-different*:
assumes *r* > *s*
shows *beukers-integral1 r s* = (∑ *k*∈{*s*<..*r*}. *1* / *k* ^ *2*) / (*r* − *s*)
⟨*proof*⟩

end

It is also easy to see that if we exchange r and s, nothing changes.

lemma *beukers-nn-integral1-swap*:
  beukers-nn-integral1 r s = *beukers-nn-integral1 s r*
  ⟨*proof*⟩

lemma *beukers-nn-integral1-finite*: *beukers-nn-integral1 r s* < ∞
  ⟨*proof*⟩

lemma *beukers-integral1-integrable*:
  set-integrable lborel ({*0*<..<*1*}×{*0*<..<*1*})
    (λ(*x*,*y*). (−*ln* (*x*∗*y*) / (*1* − *x*∗*y*) ∗ *x*^*r* ∗ *y*^*s* :: *real*))
⟨*proof*⟩

lemma *beukers-integral1-integrable′*:
  set-integrable lborel ({*0*<..<*1*}×{*0*<..<*1*}×{*0*<..<*1*})
    (λ(*z*,*x*,*y*). (*x*^*r* ∗ *y*^*s* / (*1* − (*1* − *x*∗*y*) ∗ *z*) :: *real*))
⟨*proof*⟩

lemma *beukers-integral1-conv-nn-integral*:
  beukers-integral1 r s = *enn2real* (*beukers-nn-integral1 r s*)
⟨*proof*⟩

lemma *beukers-integral1-swap*: *beukers-integral1 r s* = *beukers-integral1 s r*
  ⟨*proof*⟩

1.7 The second double integral

context
  fixes *n* :: *nat*
  fixes *D* :: (*real* × *real*) *set* **and** *D′* :: (*real* × *real* × *real*) *set*
  fixes *P* :: *real* ⇒ *real* **and** *Q* :: *nat* ⇒ *real* ⇒ *real*
  defines *D* ≡ {*0*<..<*1*} × {*0*<..<*1*} **and** *D′* ≡ {*0*<..<*1*} × {*0*<..<*1*} × {*0*<..<*1*}
  defines *Q* ≡ *Gen-Shleg n* **and** *P* ≡ *Shleg n*
begin

The next integral to consider is the following variant of I_1:

$$I_2 := \int_0^1 \int_0^1 -\frac{\ln(xy)}{1-xy} P_n(x) P_n(y)\,\mathrm{d}x\,\mathrm{d}y\ .$$

definition *beukers-integral2* :: *real* **where**
  beukers-integral2 = (∫ (*x*,*y*)∈*D*. (−*ln* (*x*∗*y*) / (*1*−*x*∗*y*) ∗ *P x* ∗ *P y*) ∂*lborel*)

I_2 is simply a sum of integrals of type I_1, so using our results for I_1, we can write I_2 in the form $A\zeta(3) + \frac{B}{\mathrm{lcm}\{1...n\}^3}$ where A and B are integers and $A > 0$:

lemma *beukers-integral2-conv-int-combination*:
obtains *A B* :: *int* **where** *A* > *0* **and**
beukers-integral2 = *of-int A* ∗ *Re* (*zeta 3*) + *of-int B* / *of-nat* (*Lcm* {*1..n*} ^ *3*)
⟨*proof*⟩

lemma *beukers-integral2-integrable*:
set-integrable lborel D (λ(*x*,*y*). −*ln* (*x*∗*y*) / (*1* − *x*∗*y*) ∗ *P x* ∗ *P y*)
⟨*proof*⟩

1.8 The triple integral

Lastly, we turn to the triple integral

$$I_3 := \int_0^1 \int_0^1 \int_0^1 \frac{(x(1-x)y(1-y)w(1-w))^n}{(1-(1-xy)w)^{n+1}} \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}w \ .$$

definition *beukers-nn-integral3* :: *ennreal* **where**
beukers-nn-integral3 =
(∫⁺(*w*,*x*,*y*)∈*D*′. ((*x*∗(*1*−*x*)∗*y*∗(*1*−*y*)∗*w*∗(*1*−*w*))^*n* / (*1*−(*1*−*x*∗*y*)∗*w*)^(*n*+*1*)) ∂*lborel*)

definition *beukers-integral3* :: *real* **where**
beukers-integral3 =
(∫ (*w*,*x*,*y*)∈*D*′. ((*x*∗(*1*−*x*)∗*y*∗(*1*−*y*)∗*w*∗(*1*−*w*))^*n* / (*1*−(*1*−*x*∗*y*)∗*w*)^(*n*+*1*)) ∂*lborel*)

We first prove the following bound (which is a consequence of the arithmetic–geometric mean inequality) that will help us to bound the triple integral.

lemma *beukers-integral3-integrand-bound*:
fixes *x y z* :: *real*
assumes *xyz*: *x* ∈ {*0*<..<*1*} *y* ∈ {*0*<..<*1*} *z* ∈ {*0*<..<*1*}
shows (*x*∗(*1*−*x*)∗*y*∗(*1*−*y*)∗*z*∗(*1*−*z*)) / (*1*−(*1*−*x*∗*y*)∗*z*) ≤ *1* / *27* (**is** *?lhs* ≤ -)
⟨*proof*⟩

Connecting the above bound with our results of I_1, it is easy to see that $I_3 \leq 2 \cdot 27^{-n} \cdot \zeta(3)$:

lemma *beukers-nn-integral3-le*:
beukers-nn-integral3 ≤ *ennreal* (*2* ∗ (*1* / *27*) ^ *n* ∗ *Re* (*zeta 3*))
⟨*proof*⟩

lemma *beukers-nn-integral3-finite*: *beukers-nn-integral3* < ∞
⟨*proof*⟩

lemma *beukers-integral3-integrable*:
set-integrable lborel D′ (λ(*w*,*x*,*y*). (*x*∗(*1*−*x*)∗*y*∗(*1*−*y*)∗*w*∗(*1*−*w*))^*n* / (*1*−(*1*−*x*∗*y*)∗*w*)^(*n*+*1*))
⟨*proof*⟩

lemma *beukers-integral3-conv-nn-integral*:
beukers-integral3 = *enn2real beukers-nn-integral3*
⟨*proof*⟩

lemma *beukers-integral3-le*: *beukers-integral3* ≤ *2* ∗ (*1* / *27*) ^ *n* ∗ *Re* (*zeta 3*)
⟨*proof*⟩

It is also easy to see that $I_3 > 0$.

lemma *beukers-nn-integral3-pos*: *beukers-nn-integral3* > *0*
⟨*proof*⟩

lemma *beukers-integral3-pos*: *beukers-integral3* > *0*
⟨*proof*⟩

1.9 Connecting the double and triple integral

In this section, we will prove the most technically involved part, namely that $I_2 = I_3$. I will not go into detail about how this works – the reader is advised to simply look at Filaseta's presentation of the proof.

The basic idea is to integrate by parts n times with respect to y to eliminate the factor $P(y)$, then change variables $z = \frac{1-w}{1-(1-xy)w}$, and then apply the same integration by parts n times to x to eliminate $P(x)$.

The first expand

$$-\frac{\ln(xy)}{1-xy} = \int_0^1 \frac{1}{1-(1-xy)z}\,\mathrm{d}z\ .$$

lemma *beukers-aux-ln-conv-integral*:
fixes *x y* :: *real*
assumes *xy*: *x* ∈ {*0*<..<*1*} *y* ∈ {*0*<..<*1*}
shows −*ln* (*x*∗*y*) / (*1*−*x*∗*y*) = (*LBINT z*=*0*..*1*. *1* / (*1*−(*1*−*x*∗*y*)∗*z*))
⟨*proof*⟩

The first part we shall show is the integration by parts.

lemma *beukers-aux-by-parts-aux*:
assumes *xz*: *x* ∈ {*0*<..<*1*} *z* ∈ {*0*<..<*1*} **and** *k* ≤ *n*
shows (*LBINT y*=*0*..*1*. *Q n y* ∗ (*1*/(*1*−(*1*−*x*∗*y*)∗*z*))) =
(*LBINT y*=*0*..*1*. *Q* (*n*−*k*) *y* ∗ (*fact k* ∗ (*x*∗*z*)^*k* / (*1*−(*1*−*x*∗*y*)∗*z*) ^(*k*+*1*)))
⟨*proof*⟩

lemma *beukers-aux-by-parts*:
assumes *xz*: *x* ∈ {*0*<..<*1*} *z* ∈ {*0*<..<*1*}
shows (*LBINT y*=*0*..*1*. *P y* / (*1*−(*1*−*x*∗*y*)∗*z*)) =
(*LBINT y*=*0*..*1*. (*x*∗*y*∗*z*)^*n* ∗ (*1*−*y*)^*n* / (*1*−(*1*−*x*∗*y*)∗*z*) ^(*n*+*1*))
⟨*proof*⟩

We then get the following by applying the integration by parts to y:

lemma *beukers-aux-integral-transform1*:
fixes *z* :: *real*
assumes *z*: *z* ∈ {*0*<..<*1*}
shows (*LBINT* (*x*,*y*):*D*. *P x* ∗ *P y* / (*1*−(*1*−*x*∗*y*)∗*z*)) =
(*LBINT* (*x*,*y*):*D*. *P x* ∗ (*x*∗*y*∗*z*)^*n* ∗ (*1*−*y*)^*n* / (*1*−(*1*−*x*∗*y*)∗*z*)^(*n*+*1*))
⟨*proof*⟩

We then change variables for z:

lemma *beukers-aux-integral-transform2*:
assumes *xy*: *x* ∈ {*0*<..<*1*} *y* ∈ {*0*<..<*1*}
shows (*LBINT z*=*0*..*1*. (*x*∗*y*∗*z*)^*n* ∗ (*1*−*y*)^*n* / (*1*−(*1*−*x*∗*y*)∗*z*)^(*n*+*1*)) =
(*LBINT w*=*0*..*1*. (*1*−*w*)^*n* ∗ (*1*−*y*)^*n* / (*1*−(*1*−*x*∗*y*)∗*w*))
⟨*proof*⟩

Lastly, we apply the same integration by parts to x:

lemma *beukers-aux-integral-transform3*:
assumes *w*: *w* ∈ {*0*<..<*1*}
shows (*LBINT* (*x*,*y*):*D*. *P x* ∗ (*1*−*y*)^*n* / (*1*−(*1*−*x*∗*y*)∗*w*)) =
(*LBINT* (*x*,*y*):*D*. (*1*−*y*)^*n* ∗ (*x*∗*y*∗*w*)^*n* ∗ (*1*−*x*)^*n* / (*1*−(*1*−*x*∗*y*)∗*w*)^(*n*+*1*))
⟨*proof*⟩

We need to show the existence of some of these triple integrals.

lemma *beukers-aux-integrable1*:
set-integrable lborel (({*0*<..<*1*} × {*0*<..<*1*}) × {*0*<..<*1*})
(λ((*x*,*y*),*z*). *P x* ∗ *P y* / (*1*−(*1*−*x*∗*y*)∗*z*))
⟨*proof*⟩

lemma *beukers-aux-integrable2*:
set-integrable lborel D′ (λ(*z*,*x*,*y*). *P x* ∗ (*x*∗*y*∗*z*)^*n* ∗ (*1*−*y*)^*n* / (*1*−(*1*−*x*∗*y*)∗*z*)
^(*n*+*1*))
⟨*proof*⟩

lemma *beukers-aux-integrable3*:
set-integrable lborel D′ (λ(*w*,*x*,*y*). *P x* ∗ (*1*−*w*)^*n* ∗ (*1*−*y*)^*n* / (*1*−(*1*−*x*∗*y*)∗*w*))
⟨*proof*⟩

Now we only need to put all of these results together:

lemma *beukers-integral2-conv-3*: *beukers-integral2* = *beukers-integral3*
⟨*proof*⟩

1.10 The main result

Combining all of the results so far, we can derive the key inequalities

$$0 < A\zeta(3) + B < 2\zeta(3) \cdot 27^{-n} \cdot \mathrm{lcm}\{1 \ldots n\}^3$$

for integers A, B with $A > 0$.

lemma *zeta-3-linear-combination-bounds*:

obtains $A\ B$ :: *int*
where $A > 0$
$A * Re\ (zeta\ 3) + B \in \{0 <..\ 2 * Re\ (zeta\ 3) * Lcm\ \{1..n\}\ \hat{}\ 3\ /\ 27\ \hat{}\ n\}$
⟨*proof*⟩

If $\zeta(3) = \frac{a}{b}$ for some integers a and b, we can thus derive the inequality $2b\zeta(3) \cdot 27^{-n} \cdot \text{lcm}\{1 \ldots n\}^3 \geq 1$ for any natural number n.

lemma *beukers-key-inequality*:
fixes a :: *int* **and** b :: *nat*
assumes $b > 0$ **and** *ab*: $Re\ (zeta\ 3) = a\ /\ b$
shows $2 * b * Re\ (zeta\ 3) * Lcm\ \{1..n\}\ \hat{}\ 3\ /\ 27\ \hat{}\ n \geq 1$
⟨*proof*⟩

end

lemma *smallo-power*: $n > 0 \Longrightarrow f \in o[F](g) \Longrightarrow (\lambda x.\ f\ x\ \hat{}\ n) \in o[F](\lambda x.\ g\ x\ \hat{}\ n)$
⟨*proof*⟩

This is now a contradiction, since $\text{lcm}\{1 \ldots n\} \in o(3^n)$ by the Prime Number Theorem – hence the main result.

theorem *zeta-3-irrational*: $zeta\ 3 \notin \mathbb{Q}$
⟨*proof*⟩

end

References

[1] R. Apéry. Irrationalité de ζ2 et ζ3. In *Journées Arithmétiques de Luminy*, number 61 in Astérisque, pages 11–13. Société mathématique de France, 1979.

[2] F. Beukers. A note on the irrationality of $\zeta(2)$ and $\zeta(3)$. *Bulletin of the London Mathematical Society*, 11(3):268–272, 1979.

[3] M. Eberl. Elementary facts about the distribution of primes. *Archive of Formal Proofs*, Feb. 2019. http://isa-afp.org/entries/Prime_Distribution_Elementary.html, Formal proof development.

[4] M. Eberl and L. C. Paulson. The prime number theorem. *Archive of Formal Proofs*, Sept. 2018. http://isa-afp.org/entries/Prime_Number_Theorem.html, Formal proof development.

[5] M. Filaseta. Math 785: Transcendental number theory (lecture notes, part 4), 2011.

[6] A. Mahboubi and T. Sibut-Pinote. A formal proof of the irrationality of $\zeta(3)$, 2019.