The sum of irreducible fractions with consecutive denominators is never an integer in a very weak arithmetic

Victor Pambuccian

Most problem solvers have encountered at some stage of their lives the problem asking for a proof that the sum

$$1 + \frac{1}{2} + \ldots + \frac{1}{n}$$

for $n \geq 2$ is never an integer. The proof one usually finds offered for this fact is based on Chebyshev's theorem (Bertrand's postulate). If one asks for a proof that, more generally, the sum

$$\frac{1}{n} + \frac{1}{n+1} + \ldots + \frac{1}{n+k}$$

with $k \geq 1$ can never be an integer, then the proof based on Chebyshev's theorem needs to be amended. One first notes that, if $k < n$, then the above sum must be less than 1, and thus cannot be an integer, and if $k \geq n$, then one applies the same proof based on Chebyshev's theorem (this fact seems to have been overlooked in [6], where the author wants to use Chebyshev's Theorem, but finds it necesary to make the proof dependent on another deep result, the Sylvester-Schur Theorem, as well). However, Kürschák [3] (see also [7]) found a much simpler proof, which relies on the very simple observation that among any number (≥ 2) of consecutive positive integers there is precisely one, which is divisible by the highest power of 2 from among all the given numbers. Aside from its didactical use, one may wonder whether Kürschák's proof is not in a very formal way much simpler, i. e. whether it does not require simpler methods of proof in the sense of formal logic.

When formalized, arithmetic is usually presented as Peano Arithmetic, which contains an induction axiom schema, stating, loosely speaking, that any set that can be defined by an elementary formula in the language of arithmetic (i.e. in terms of some undefined operation and predicate symbols, such as $+, \cdot, 1, 0, <$), which contains 1, and which contains $n+1$ whenever it contains n, is the set of all numbers. Several weak arithmetics have been studied, in which the types of elementary formulas allowed in the definitions of the sets used in induction are restricted by certain syntactic constraints (see [1]), and one might think that Kürschák's proof would make it in a weaker formal arithmetic than the one

dependent on Chebyshev's theorem. It turns out that, in fact, no amount of induction is needed at all!

To see this, let's first generalize the problem further, along the lines of the generalization in [5], so that there can be no proof based on Chebyshev's Theorem.

Theorem 1 *The sum*

$$\frac{m_0}{n} + \frac{m_1}{n+1} + \ldots + \frac{m_k}{n+k} \tag{1}$$

with $(m_i, n+i) = 1$, $m_i < n+i$, *and* $k \geq 1$ *is never an integer.*

Proof. (Kürschák [3]). Let $a = \max\{\alpha : 2^\alpha | (n+i) \text{ for some } 0 \leq i \leq k \}$. Then 2ª divides exactly one of the numbers n, $n+1$, ..., $n+k$. Let $l = \text{lcm } (n, n+1, \ldots, n+k)$. Suppose the sum in (1) is an integer b. Multiplying both (1) and b by l, we obtain on the one hand an odd number, and on the other an even number, which have to be equal. $\square$

Moreover, to make it a theorem of arithmetic, we will do away with the fractions appearing in it, and state it, for all positive $k \in \mathbb{N}$, as φ_k, the following statement (where we denote by $\overline{u}$ the term $((\ldots((1+1)+1)+\ldots)+1)$, in which there are u many 1's; the terms $\overline{u}$ will be referred to as *numerals*)

$$(\forall n)(\forall m_0)\ldots(\forall m_k)(\forall p) \bigvee_{i=0}^{k} ((\forall a)(\forall b)\, m_i a \neq (n+\overline{i})b + 1) \vee \bigvee_{i=0}^{k} n + \overline{i} < m_i$$

$$\vee \sum_{i=0}^{k}(m_i \prod_{0 \leq j \leq k, j \neq i} (n+\overline{j})) \neq p \prod_{j=0}^{k}(n+\overline{j}).$$

The arithmetic we will show it holds in is PA^-, which is expressed in a language containing as undefined operation and predicate symbols only $+$, $\cdot$, 1, 0, and $<$, and whose axioms A1-A15 were presented in [2, pp. 16-18]. We will repeat them here for the reader's convenience, and we will omit the universal quantifiers for all universal axioms.

A 1 $(x + y) + z = x + (y + z)$

A 2 $x + y = y + x$

A 3 $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

A 4 $x \cdot y = y \cdot x$

A 5 $x \cdot (y + z) = x \cdot y + x \cdot z$

A 6 $x + 0 = x \wedge x \cdot 0 = 0$

A 7 $x \cdot 1 = x$

A 8 $(x < y \wedge y < z) \rightarrow x < z$

A 9 $\neg x < x$

A 10 $x < y \vee x = y \vee y < x$

A 11 $x < y \rightarrow x + z < y + z$

A 12 $(0 < z \wedge x < y) \rightarrow x \cdot z < y \cdot z$

A 13 $(\forall x)(\forall y)(\exists z)\, x < y \rightarrow x + z = y$

A 14 $0 < 1 \wedge (x > 0 \rightarrow (x > 1 \vee x = 1))$

A 15 $x > 0 \vee x = 0$

What is missing from PA^-, and makes it so weak (indeed, the positive cone of every discretely ordered ring is a model of PA^-), is the absence of any form of induction.

The proof that φ_k holds in PA^- will be carried out in an arbitrary model $\mathfrak{M}$ of PA^-. The idea of proof will be to show that all variables that appear in φ_k must be numerals. An essential ingredient of the proof is the following fact, which holds in PA^- (see [2, Lemma 2.7, p. 22]), for all positive $k \in \mathbb{N}$

$$x < \overline{k} \rightarrow x = 0 \vee x = 1 \vee \ldots \vee x = \overline{k-1},$$

and which allows us to deduce that any element that is bounded from above by a numeral must be a numeral.

Suppose that, for some positive $k \in \mathbb{N}$, φ_k does not hold in $\mathfrak{M}$. Then, for all $i = 0, \ldots, k$, there are m_i, p, a_i and b_i with $m_i a_i = (n+i)b_i + 1$ and such that

$$\sum_{i=0}^{k}\left(m_i \prod_{0 \le j \le k, j \ne i}(n + \overline{j})\right) = p \prod_{j=0}^{k}(n + \overline{j}). \tag{2}$$

This can be rewritten, by leaving only the first term of the sum on the left-hand side, and sending all others to the right-hand side with changed sign, as $m_0(n + \overline{1})\ldots(n + \overline{k}) = nq$, where by q we have denoted $p \prod_{j=1}^{k}(n + \overline{j}) - (\sum_{i=1}^{k} m_i \prod_{0 \le j \le k, j \ne i}(n + \overline{j}))$. The product $(n+\overline{1})\ldots(n+\overline{k})$ can also be written as a polynomial in n, whose free term is $\overline{k!}$, i. e. as $nr + \overline{k!}$, thus $m_0(nr + \overline{k!}) = nq$. Given that there are a_0 and b_0 such that $m_0 a_0 = nb_0 + 1$, if we multiply both sides of the equality $m_0(nr + \overline{k!}) = nq$ by a_0 we obtain $(nb_0 + 1)\overline{k!} = n(a_0 q - a_0 m_0 r)$, thus $\overline{k!} = n(a_0 q - a_0 m_0 r - b_0 \overline{k!})$. We know that $\mathfrak{M}$ must contain a copy of $\mathbb{N}$, and it may contain other elements as well, called *nonstandard* numbers. Could n be in $\mathfrak{M}$ but not of the form $\overline{m}$ for some $m \in \mathbb{N}$? If it were such an element of $\mathfrak{M}$, then it would be greater than all $\overline{m}$ with $m \in \mathbb{N}$, and thus so would $n(a_0 q - a_0 m_0 r - b_0 \overline{k!})$, unless $a_0 q - a_0 m_0 r - b_0 \overline{k!} = 0$, which cannot be the case, as $\overline{k!}$ is not zero. However, $n(a_0 q - a_0 m_0 r - b_0 \overline{k!})$ cannot be greater than all

$\overline{m}$ with $m \in \mathbb{N}$, for it is equal to such a number, namely to $\overline{k!}$. Thus n must be an $\overline{m}$ for some $m \in \mathbb{N}$. This means that in (2) all variables are numerals, i. e. $\overline{i}$'s for some $i \in \mathbb{N}$. However, we know, from Kürschák's proof, that such an equation cannot exist, so, for all $k \in \mathbb{N}$, φ_k holds in PA$^-$.

Another generalization of the original problem, proved by T. Nagell in [4], states that the sum

$$\frac{1}{m} + \frac{1}{m+n} + \frac{1}{m+2n} + \cdots + \frac{1}{m+kn}$$

is never an integer if n, m, k are positive integers. The proof is rather involved and uses both a Kürschák-style argument and Chebyshev's theorem. This statement turns out to be, with $\overline{k}$ instead of k, valid in PA^- as well. To see this, let, for all positive $k \in \mathbb{N}$, ν_k stand for

$$(\forall m)(\forall n)(\forall p)\, m > 0 \wedge n > 0 \rightarrow \sum_{i=0}^{k} \prod_{0 \le j \le k, j \neq i} (m + \overline{j}n) \neq p \prod_{0 \le j \le k} (m + \overline{j}n), \quad (3)$$

and let $\mathfrak{M}$ be again a model of PA^-. Notice that, if $m > \overline{k}$, then, then $\neg\nu_k$ cannot hold for any n and p. To see this, suppose that, for some n and p, we have equality in (3). Given that $(m+n)\ldots(m+\overline{k}n)$ is the largest of all the summands on the left-hand side, and there are $\overline{k}$ summands, the sum on the left-hand side is $\le \overline{k}(m+n)\ldots(m+\overline{k}n)$, and thus $< m(m+n)\ldots(m+\overline{k}n)$, thus equality cannot hold in (3). Thus $m \le \overline{k}$, and thus (see [2, p. 20]) m must be standard, i. e. it must be $\overline{u}$ for some $0 < u \le k$. It remains to be shown that n must be standard as well. To see this, suppose again that, for some $n > 0$ and p, we have equality in (3). Notice that, since $m(m+2n)\ldots(m+\overline{k}n)$ is the largest product among all $\prod_{0 \le j \le k, j \neq i}(m+\overline{j}n)$, for $i = 1, 2, \ldots k$, the sum on the left hand side of our equality is $\le (m+n)\ldots(m+\overline{k}n)+\overline{k}m(m+2n)\ldots(m+\overline{k}n)$ (with equality if and only if $k = 1$). Thus $pm(m+n)\ldots(m+\overline{k}n) \le (m+n+\overline{k}m)(m+2n)\ldots(m+\overline{k}n)$, which implies $pm(m+n) \le m+n+\overline{k}m$. If n were nonstandard, then this inequality were possible only if $pm = 1$, i. e. if $p = m = 1$, which is not possible, for in that case the first summand on the left hand side of (3) is equal to the right hand side, thus the left hand side must be larger than the right hand side, so equality could not have taken place in (3). Now that m, n, k have all been shown to be standard, Nagell's proof implies the truth of our statement, which thus holds in PA^-.

By Gödel's completeness theorem, there must exist syntactic proofs that φ_k and ν_k hold in PA$^-$, i. e. formal derivation of φ_k and ν_k from the axioms of PA$^-$. Such a formal proof for φ_k cannot use the idea behind Kürschák's proof, for it is not even true in PA$^-$ that among $n, n+1, \ldots n+k$, there is a multiple of 2 (see [2, p. 18]). Thus there must exist even simpler proofs for both φ_k and ν_k and they are worth finding. Such proofs would reveal the real reasons why these results hold, and the reason must be of an algebraic nature, for there is no traditional number theory to be found in PA^-.

References

[1] P. D'Aquino, Weak fragments of Peano Arithmetic, in: *The Notre Dame Lectures*, (ed. P. Cholak), Lecture Notes in Logic, 18, pp. 149–185, A. K. Peters, 2005.

[2] R. Kaye, *Models of Peano Arithmetic*, Oxford University Press, Oxford, 1991.

[3] J. Kürschák, On the harmonic series (Hungarian), Mat. Phys. Lapok 27 (1918), 299–300.

[4] T. Nagell, Eine Eigenschaft gewisser Summen, Norsk Mat. Forenings Skrifter 13 (1923), 10–15.

[5] R. Oblath, Some number-theoretical theorems (Hungarian), Math. és phys. lapok 27 (1918), 91–94.

[6] R. Oblath, Über einen arithmetischen Satz von Kürschák, Comment. Math. Helv. 8 (1935), 186–187.

[7] G. Pólya, G. Szegö, *Aufgaben und Lehrsätze der Analysis*, Springer-Verlag, Berlin, 1971, vol. II, VIII. Abschnitt, Aufgabe 251, S. 159.

Department of Integrative Studies, Arizona State University - West Campus, Phoenix, AZ 85069-7100, U.S.A
pamb@math.west.asu.edu