AN ANALOG OF THE HARMONIC NUMBERS OVER THE SQUAREFREE INTEGERS

DICK BOLAND

Abstract. A nice, short result establishing an asymptotic equivalent of the harmonic numbers, $H_n$, in terms of the reciprocals of Euler’s totient of the squarefree integers. Designating the analog to $H_n$ by $F_n$, the limit of $F_n - H_n$ is shown to approach a constant that matches a known constant occurring in number theory to at least the first 6 significant digits. Proof that the constants are, or cannot be, equivalent, is an open problem and proffered as a challenge to the mathematical community. Even without the conjectured equivalence, the theorem allows a number theoretic construct utilizing $H_n$ to be reexamined over the squarefree integers in light of $F_n$. Mertens’ theorem is intimately involved and is used to give an example reformulation of the Nicolas criteria for the Riemann Hypothesis.

1. The Squarefree Totients Analog of the Harmonic Numbers

Where $\varphi(x)$ is Euler’s totient function, and letting $f$ be the squarefree positive integers, I prove that the sum of the reciprocals of Euler’s totient of the squarefree integers not exceeding $n$, is asymptotic to $H_n$, the $n^{th}$ Harmonic Number. Letting $F_n$ represent these analogs to $H_n$, there is a constant $\kappa = \lim_{n \to \infty} (F_n - H_n)$, which allows a number theoretic formula or construct utilizing the relation $H_n \simeq \log n + \gamma$, where gamma is Euler’s constant, to be reformulated into an equivalent statement over the squarefree integers in light of $F_n \simeq \log n + \gamma + \kappa$ and also Mertens’ Theorem [1] [p.351], which is intimately related. Examples using the Nicolas Criteria for the Riemann Hypothesis [2] are included.

Computing the first few significant digits of $\kappa$ is cause to conjecture that $\kappa = B_3 - \gamma$, where $B_3$ is the well known constant related to Mertens’ constant [3] [4] [5]. Proof of the conjecture that they are equivalent would ideally demonstrate whether the 2 derivations are ultimately by means of a rearrangement of the same set of summation terms. The notion that the 2 constructions of the same constant have distinct sets of summation terms, which is suggested by a cursory examination of the problem, is enticing and suggests a more profound number theoretic relation is lurking. The form of which might be viewed as a limiting balance between the sum of terms of the first type, characterized by integers less than some finite limit and those of the second type, characterized by an infinite set of integers greater than a related finite limit.

First, a simple lemma,
Lemma 1. Given the prime factorization of \( x = \prod_{i=1}^{r} p_i^\alpha_i \), The Euler totient of \( x \) is given by the formula \[ \varphi(x) = \prod_{i=1}^{r} (p_i^\alpha_i - p_i^{\alpha_i - 1}) \] thus, for squarefree \( x \),

\[ \varphi(x) = \prod_{i=1}^{r} (p_i - 1) \quad x \text{ is squarefree} \]

Then,

Theorem 1 (Asymptotic Equivalence of \( \sum_{k=1}^{n} \frac{\mu^2(k)}{\varphi(k)} \) to \( H_n \)). The \( n \)th Harmonic Number, \( H_n \), is asymptotically equivalent to the sum of the reciprocals of Euler’s totient of the squarefree integers not exceeding \( n \) and the difference between the 2 sums approaches the constant \( \kappa = 0.755366 \ldots \).

Where \( \mu(x) \) is the Möbius function,

\[ \sum_{k=1}^{n} \frac{\mu^2(k)}{\varphi(k)} = H_n + O(1) \]

But, it is simpler to use \( f \) to indicate the squarefree integers as the square of the Möbius function serves no other purpose than to isolate them, and Eq.(3) is rewritten

\[ \sum_{f \leq n} \frac{1}{\varphi(f)} = H_n + O(1) \]

where, throughout, the term ‘\( F_n \)’ is an abbreviation of, and used interchangeably with, the lhs of Eq.(4). Another statement of the theorem is given by:

\[ \lim_{n \to \infty} \left( \sum_{f \leq n} \frac{1}{\varphi(f)} - H_n \right) = \kappa = 0.755366 \ldots \]

Proof. Where \( n\# \) is the \( \pi(n)^{th} \) primorial, and by Lemma 1, the sum of the reciprocals of the Euler totients of the divisors of \( n\# \) is given by reformulation of the partial product

\[ \prod_{p \leq n} \frac{1}{1 - \frac{1}{p}} \prod_{p \leq n} \left(1 + \frac{1}{p - 1}\right) = \sum_{d \mid n\#} \frac{1}{\varphi(d)} \]

The value of Eq.(6) is provided by Mertens’ Theorem and in relation to harmonic numbers in question, falls within the inequality \( (O(e^\gamma / \log n) \) [7]),(rhs [1] [p.341])

\[ \log n \sim H_n < e^\gamma (\log n + O(1 / \log n)) < H_{n\#} \sim \log n\# = \vartheta(n) \sim n \]

Expanding \( 1/(p - 1) \) in each term of Eq.(6) into its geometric series before taking the product, then yields the sum of the reciprocals of all positive integers formed exclusively by products of the primes \( \leq n \) and their powers. Thus, where \( \text{rad}(k) \) is the radical in the sense of the ‘squarefree root’ of positive integer \( k \) (ala the ABC conjecture),

\[ \prod_{p \leq n} \left(1 + \sum_{a=1}^{\infty} \frac{1}{p^a}\right) = \sum_{\text{rad}(k) \mid n\#} \frac{1}{k} \]
Eq.(8) is the partial parallel of the celebrated Euler relation between the product over all primes and the sum over all positive integers. Obviously, all \( k \leq n \) also meet the criteria \( \text{rad}(k) \mid n\# \). And also, for all \( f \leq n \), it must be true that \( f \mid n\# \). Thus the rhs’s of Eqs.(6 & 8) may each be split and equated as

\[
\sum_{f \leq n} \frac{1}{\varphi(f)} + \sum_{d \mid n \# \atop d > n} \frac{1}{\varphi(d)} = H_n + \sum_{k=n+1}^\infty \frac{1}{k}
\]

Giving the statement of the theorem in the form

\[
\sum_{f \leq n} \frac{1}{\varphi(f)} - H_n = \sum_{k=n+1}^\infty \frac{1}{k} - \sum_{d \mid n \# \atop d > n} \frac{1}{\varphi(d)} = O(1)
\]

Now, any divisor \( d \mid n\# \) where \( d > n \) must be a squarefree composite and the sum of the reciprocal totients in the middle term of Eq.(10) may be rewritten

\[
\sum_{d \mid n \# \atop d > n} \frac{1}{\varphi(d)} = \sum_{d \mid n \# \atop d > n} \prod_{p \mid d} \frac{1}{p-1} = \sum_{d \mid n \# \atop d > n} \prod_{p \mid d} \frac{1}{p} = \sum_{d \mid n \# \atop d > n} \prod_{k=n+1}^\infty \frac{1}{k}
\]

Substituting the far rhs of Eq.(11) into Eq.(10) yields

\[
\sum_{f \leq n} \frac{1}{\varphi(f)} - H_n = \sum_{k=n+1}^\infty \frac{1}{k} - O(1) = O(1)
\]

Because all \( f \leq n \) must divide \( n\# \), and all \( \text{rad}(k) \) are squarefree, and \( k > n \) in Eq.(12), a simple restatement of the theorem is

\[
\lim_{n \to \infty} \sum_{k \mid n \# \atop \text{rad}(k) \leq n} \frac{1}{k} = O(1) = \kappa
\]

Where, in this form, summing terms to produce each successive convergent to \( \kappa \), call them \( \kappa_n \) with \( \kappa_1 = 0 \) and \( \kappa_n = F_n - H_n \), in order, involves adding or subtracting depending on whether or not the given \( n \) is squarefree. When squarefree \( n \) is encountered, the sum is increased via

\[
\kappa_n = \kappa_{n-1} + \frac{1}{\prod_{i=1}^r (p_i - 1)} = \kappa_{n-1} + \frac{1}{\varphi(f)} - \frac{1}{f} \quad \text{(squarefree n)}
\]

Given a squarefree \( n \) with \( r \) distinct prime factors \( (r = \omega(n)) \), Eq.(8) makes it clear that the term being added to \( \kappa_{n-1} \) in Eq.(14) represents the sum of the reciprocals of all integers of form \( \prod_{i=1}^r p_i^{\alpha_i} \) as \( \alpha_i \) runs over the positive integers for each \( p_i \), in all combinations except the single instance in which \( \alpha_i = 1 \forall i \), this latter quantity being \( 1/f \).

In the case where \( n \) is squareful, a term corresponding to \( 1/k = 1/n \) has already been added to the sum by some previous term \( n' = \text{rad}(n) \) by virtue of \( 1/\varphi(\text{rad}(n)) - 1/\text{rad}(n) \) having been added to the sum. But now that \( k \neq n \), the term \( 1/k \) must be subtracted from the sum, thus,

\[
\kappa_n = \kappa_{n-1} - \frac{1}{k} \quad \text{(squareful n)}
\]
Convergence of the sum is readily demonstrated by computation to sufficiently large \( n \), bearing in mind that the order of \( \varphi(n) \) is always 'nearly \( n \)' and the largest \( 1/\varphi(f) \) terms are bounded in accordance with \( \lim \varphi(x) = xe^{-\gamma}/\log \log x \) [1] [p.267]. Without computing the sum, convergence can be established crudely by a comparison test as follows. Where \( P \) is any perfect power,

\[
\sum_{k>n \atop \text{rad}(k) \leq n} \frac{1}{k} < \sum_{j=2}^{\infty} \frac{1}{j(j-1)} + \sum_{k>n \atop \text{rad}(k) \leq n \atop k \neq P} \frac{1}{k} = 1 + O(?)
\]

Now, the sum of \( 1/(j(j-1)) \) over all integers \( j > 1 \) is equivalently the sum of the reciprocals of all perfect powers > 1, with multiplicity (which sums to unity), so must include the reciprocals of all powers of integers \( k > n \) such that \( \text{rad}(k) < n \). This is a generously loose comparison test as the sum of the reciprocals of all powers without multiplicity converges to .87446... [8] [9] [p.66] and the largest terms, the reciprocals corresponding to the perfect powers \( \leq n \), are not included in the lhs sum. The \( O(?) \) term acknowledges the contribution of \( 1/k \) from each of the non-perfect power integers \( k > n, k \neq P \) such that \( \text{rad}(k) < n \) and convergence of their sum is inferred by the root test. One can surmise that \( O(?) \) is too small to allow the constant to actually approach 1 because each contributing term, or more appropriately a sum of such terms, such as \( (1/p^2q + 1/p^3q + 1/p^4q\ldots) = 1/(qp(p-1)) \), can be compared as being smaller than some other, unused \( P \) term or directly to a term in the \( 1/(j(j-1)) \) sum, i.e., terms not already included in the \( \kappa \) sum or which occur multiple times, such that the overall sum to \( \kappa \) is expected to be, relatively speaking, much less than 1.

Another way to visualize it is to formulate \( 1/\varphi(f) \) by

\[
\frac{1}{\varphi(f)} = \prod_{p|f} \frac{1}{p-1} = \prod_{p|f} \left(1 + \frac{1}{p^2} + \frac{1}{p^3} \ldots \right) = \sum_{\alpha_1=1}^{\infty} \sum_{\alpha_2=1}^{\infty} \cdots \sum_{\alpha_j=1}^{\infty} \prod_{i=1}^{r} \frac{1}{p_i^{\alpha_i}}
\]

which, summed over all \( f \leq n \) obviously must include each \( 1/k \) term of the \( H_n \) sum except \( 1/k = 1 \) which is included in the sum over \( 1/\varphi(f) \) by \( 1/\varphi(1) \) by definition, so subtracting out the \( H_n \) terms leaves the terms that now sum to \( \kappa_n \), the \( n^{th} \) convergent to \( \kappa \).

Taking the single instance of Eq.(17) corresponding to \( f = p_j \# \) and multiplying both sides by \( p_j \# \) yields,

\[
\frac{p_j \#}{\varphi(p_j \#)} = \sum_{\alpha_1=0}^{\infty} \sum_{\alpha_2=0}^{\infty} \cdots \sum_{\alpha_j=0}^{\infty} \prod_{i=1}^{j} \frac{1}{p_i^{\alpha_i}}
\]

which is identically the products and sum given in Eq.(6) as well as the critical ratio of the Nicolas criteria for the Riemann Hypothesis.

Barring the conjectured equivalence, the fastest convergence to \( \kappa \) I could find comes by computing \( \kappa_n \) via Eqs.(14-15), from which it is clear from computations to \( n > 4 \times 10^9 \) that

\[
\lim_{n \to \infty} \left( \sum_{f \leq n} \frac{1}{\varphi(f)} - H_n \right) = \kappa = 0.755366\ldots
\]
are significant digits, and the values of the sum near this value of $n$ meander nearby significant digits 0.75536661 . . . , further suggesting an apparent agreement with the following conjecture.

**Conjecture 1.** $\kappa = \lim_{n \to \infty} (F_n - H_n)$ is the same constant given by

$$\kappa^* = \sum_{j=1}^{\infty} \frac{\log p_j}{p_j(p_j - 1)} = \sum_{j=2}^{\infty} \frac{\mu(j)\zeta'(j)}{\zeta(j)} = 0.755366610831688021159 \ldots$$

Which is also given by the Mertens’ related constant $B_3$ [3] [4] [5] less gamma, $\kappa^* = B_3 - \gamma$, where

$$B_3 = \lim_{n \to \infty} \left( \log n - \sum_{p \leq n} \frac{\log p}{p} \right)$$

$\kappa^*$ also arises in the variance of the average order of $\omega(n)$

$$\text{var}(\omega(n)) \sim \log \log n + B_1 - \zeta'(2) - \zeta(2) + \frac{c_1}{\log n} + \frac{c_2}{(\log n)^2} + \ldots$$

where $B_1$ is Mertens’ constant and $\zeta_p(s)$ is the prime zeta function [12] [1] [p.356]. The presumed to be $\kappa$ constant occurs as a main term in the formulation of the coefficients $c_1$ and $c_2$.

Assuming the conjecture, one has

$$\lim_{n \to \infty} \left( \sum_{f \leq n} \frac{1}{\varphi(f)} - \sum_{p \leq n} \frac{\log p}{p} \right) = 2(\gamma + \kappa^*)$$

from which, by using

$$\kappa^* = \lim_{n \to \infty} \sum_{p \leq n} \frac{\log p}{p(p-1)}$$

one can directly derive

$$\gamma = \lim_{n \to \infty} \left( \log n - \sum_{p \leq n} \frac{\log p}{p-1} \right)$$

without necessarily needing to employ the method of taking the limit of $\zeta(s)$ as $s \to 1$ as is commonly used to derive Eq.(25) [4] [1] [p.356].

Even without the equivalence, using $F_n$ to denote the $n^{th}$ squarefree totients analog of the harmonic number, allows a number theoretic relation or construct utilizing the relation

$$H_n \simeq \log n + \gamma$$

to be examined in terms of the squarefree integers in light of

$$F_n \simeq \log n + \gamma + \kappa$$

And, again assuming the conjecture, analogous equivalents to other, more exotic constructs can be derived, such as

$$\gamma^2 = \lim_{n \to \infty} \left( \log^2(n) - H_n \sum_{p \leq n} \frac{\log p}{p-1} \right)$$

$$\gamma^2 = \lim_{n \to \infty} \left( \log^2(n) - F_n \sum_{p \leq n} \frac{\log p}{p} \right)$$
for example.

Integer sequences related to \( \kappa \) and \( F_n \) have been entered into Sloane’s On Line Encyclopedia of Integer Sequences as given in [13][14][15][16][17] and [18]

1.1. **Equivalents to Mertens’ Theorem.** Mertens’ Theorem can take many forms

\[
\prod_{p \leq n} \left( 1 + \frac{1}{p-1} \right) = \sum_{d \mid n \#} \frac{1}{\varphi(d)} = \prod_{p \leq n} \frac{p}{\varphi(p\#)} = \prod_{p \leq n} \frac{1}{1 - \frac{1}{p}} = e^\gamma \log n + \delta_n
\]

Where \( \delta_n = O(e^{\gamma} / \log n) \) [7]. Also then, substituting from Eqs.(6 & 9),

\[
H_n + \kappa_n + \sum_{d \mid n \#; d > n} \frac{1}{\varphi(d)} = e^\gamma \log n + \delta_n
\]

Which leads to the following constant approached by the ratio:

\[
\lim_{n \to \infty} \left( \frac{\sum_{d \mid n \#; d > n} \frac{1}{\varphi(d)}}{\sum_{d \mid n \#} \frac{1}{\varphi(d)}} \right) = \lim_{n \to \infty} \left( \frac{(e^\gamma - 1) \log n - \kappa_n - \gamma_n + \delta_n}{e^\gamma \log n + \delta_n} \right) = \frac{e^\gamma - 1}{e^\gamma} = 0.4385405 \ldots
\]

Another interesting observation comes by letting \( n \) be the \( j \)th prime, for which it is obvious that

\[
F_{p_j \#} - F_{p_j} = \log p_{j-1 \#} + (\gamma_{p_j \#} + \kappa_{p_j \#}) - (\gamma_{p_j} + \kappa_{p_j})
\]

such that subtracting Eq.(32) yields

\[
\sum_{f \mid p_j \#; f < p_j \#} \frac{1}{\varphi(f)} = \log p_{j-1 \#} - (e^\gamma - 1) \log p_j - \delta_{p_j} + \gamma_{p_j \#} + \kappa_{p_j \#}
\]

\[
\sum_{f \mid p_j \#; f < p_j \#} \frac{1}{\varphi(f)} = \log p_{j \#} - e^\gamma \log p_j - \delta_{p_j} + \gamma_{p_j \#} + \kappa_{p_j \#}
\]

\[
\sum_{f \mid p_j \#; f < p_j \#} \frac{1}{\varphi(f)} = H_{p_j \#} - (e^\gamma \log p_j + \delta_{p_j})
\]

Which, in very broad strokes, requires ‘quite a few’ primes greater than \( p_j \) in order to form ‘a lot’ of squarefree integers between \( p_j \) and \( p_j \# \) that possess at least 1 prime factor > \( p_j \). This is because the \( \text{lhs} \) will not take on any magnitude until subsequent primes appear, while the \( \text{rhs} \) requires the given magnitude of the \( \text{lhs} \) using a formulation from the primes not exceeding \( p_j \), where it is noted that \( H_{p_j \#} \) can be formulated without knowing the status of the integers exceeding \( p_j \). Of course, even for modest numbers, the magnitude of \( p_j \# \) greatly exceeds \( p_j \), suggesting a dubious, at best, value of the relation in establishing any kind of
meaningful bounds, particularly considering that the Prime Number Theorem is already in play via Mertens’ Theorem.

1.2. Regarding the Nicolas Criteria for the Riemann Hypothesis. The Nicolas Criteria establishes that the truth of the Riemann Hypothesis is equivalent to the statement that

\[ \frac{p_j^\#}{\varphi(p_j^\#)} > e^\gamma \log \log p_j^\# \]  

for all but finitely many values of \( j \) [2]. An equivalent statement of the Nicolas criteria is that for only finitely many \( j \)

\[ \sum_{\substack{f \mid p_j^\# \atop f < p_j^\#}} \frac{1}{\varphi(f)} < \log p_j^\# - e^\gamma \log \log p_j^\# + \kappa_{p_j^\#} + \gamma_{p_j^\#} \]  

**Proof.** The lhs of the inequality of Eq.(38) is precisely \( \sum_{d \mid p_j^\#} \frac{1}{\varphi(d)} \), thus

\[ \sum_{d \mid p_j^\#} \frac{1}{\varphi(d)} > e^\gamma \log \log p_j^\# \]  

The lhs of Eq.(40) may be substituted with \( F_{p_j^\#} = H_{p_j^\#} + \kappa_{p_j^\#} \) less the reciprocals of the totients of the squarefree numbers that do not divide and are \( < p_j^\# \) yielding

\[ H_{p_j^\#} + \kappa_{p_j^\#} - \sum_{\substack{f \mid p_j^\# \atop f < p_j^\#}} \frac{1}{\varphi(f)} > e^\gamma \log \log p_j^\# \]

\[ \log p_j^\# + \kappa_{p_j^\#} + \gamma_{p_j^\#} > e^\gamma \log \log p_j^\# \]

which is rearranged to

\[ \sum_{\substack{f \mid p_j^\# \atop f < p_j^\#}} \frac{1}{\varphi(f)} < \log p_j^\# - e^\gamma \log \log p_j^\# + \kappa_{p_j^\#} + \gamma_{p_j^\#} \]  

Another interesting observation of this criteria comes by substituting into Eq.(38) from Eq.(30) to obtain

\[ e^\gamma \log p_j + \delta_n > e^\gamma \log \log p_j^\# \]

The error term, \( \delta_n \) in Mertens’ Theorem is positive for \( n < 10^8 \) and a discussion attributed to Rosser and Schoenfeld in [10], indicates that, as of 1962 anyway, the question of whether it ever reverses sign had apparently not been investigated, but they suggested that their methods might apply to such an investigation and seemed to intimate that sign changes should be likely. If, on the other hand, \( \delta_n \) were provably positive for all \( n \), then it is apparent from Eq.(44) that proof that

\[ p_j > \log p_j^\# \]

for all but finitely many primes, would be sufficient proof of the Riemann Hypothesis. If the criteria of Eq.(45) were provably true, an infinitude of negative values
for $\delta_n$ would not necessarily be fatal if the magnitude could be shown to have no impact on the larger proposition.

References

http://physics.open.ac.uk/~dbroadhu/cert/cohenb3.ps
http://www.research.att.com/~njas/sequences/A083343
http://www.research.att.com/~njas/sequences/A072102
[10] MathWorld - MertensTheorem
http://algo.inria.fr/csolve/asym.pdf
http://www.research.att.com/~njas/sequences/A138312
[14] D. Boland - Decimal expansion of constant ‘kappa’ = (lim $n \to \infty$) $F_n - H_n$, where $H_n$ are harmonic numbers, $F_n$ are squarefree totient analogs of $H_n$. The On-Line Encyclopedia of Integer Sequences, Sequence A138313
http://www.research.att.com/~njas/sequences/A138313
http://www.research.att.com/~njas/sequences/A138316
http://www.research.att.com/~njas/sequences/A138317
[17] D. Boland - Numerators of the difference between the squarefree totient analogs of the harmonic numbers and the harmonic numbers: $F_n - H_n$. The On-Line Encyclopedia of Integer Sequences, Sequence A138320
http://www.research.att.com/~njas/sequences/A138320
[18] D. Boland - Denominators of the difference between the squarefree totient analogs of the harmonic numbers and the harmonic numbers: $F_n - H_n$. The On-Line Encyclopedia of Integer Sequences, Sequence A138321
http://www.research.att.com/~njas/sequences/A138321