The evaluation of Knopfmacher's curious limit

Daniel Lichtblau
August, 2000

Wolfram Research, Inc.
100 Trade Centre Dr.
Champaign IL 61820
danl@wolfram.com

Joint work with: Jurgen Tischer
Universidad del Valle, Columbia

Abstract

We begin with a function expressed as a certain infinite product. It is a twice−mutated variation of another product that has its origins in counting irreducible factors of univariate polynomials over Galois fields. Knopfmacher's limit is taken as we approach `1` from below in this product. We derive and execute an algorithm that finds a good approxi−mation to this limit using moderate computational resources. We also investigate the coefficients of the power series for the logarithm of this product; these will be shown to exhibit size fluctuations that render the straightforward power series inadequate for the purpose of estimating the limit in question.

Statement of the problem

The problem was posed by Arnold Knopfmacher in a note to the Usenet news group comp.soft−sys.math.mathematica, January 1999 [Knopfmacher 1999a] (herein we make small changes in notation). We wish to obtain a numerical estimate (say *8* decimal digits) of the limit as *x* tends to *1* from below of the function

$$p[x] = \frac{\prod_{k=2}^{\infty}\left(1-\frac{x^{m[k]}}{k+1}\right)}{1-x}$$

where $m[k] = k - \frac{k}{d[k]}$ and $d[k]$ is the smallest prime factor of k.

There are similar formulas in [Knopfmacher and Warlimont 1995] analyzing probabilities related to numbers of irreducible factors of distinct degrees in univariate polynomials over Galois fields. According to [Knopfmacher 1999b] his coauthor studied a limiting case of such a formula. It gave rise to a limit seemingly harder than the one above. Knopfmacher extracted the present problem as a simpler case to try first, and subsequently posted it to the news group. Thus this example might be described as a double mutation of a problem that is of independent interest elsewhere in the realm of number theory. It later turned out that the original problem was the more tractable but by then the mutation had acquired a life of its own. In this report we show how to compute a good approxima−tion to the above limit using a blend of theory and the computational capabilities of *Mathematica* [Wolfram 1999] (*Mathematica* (TM) is a registered trademark of Wolfram Research, Incorporated).

In *Mathematica* this function may be written as

```
d[k_] := Divisors[k][[2]]
m[k_] := k - k/d[k]
p[x_] := Product[1 - x^m[k]/(k+1), {k, 2, Infinity}]/(1 - x)
```

Crude bounds

First we deduce crude bounds for the limit. Among other things this will demonstrate the existence of a *lim inf* and *lim sup*.

Proposition 1:

(a) A *lim inf* for $p[x]$ as $x \to 1$ is given by $e^{EulerGamma}$.

(b) A *lim sup* is given by $2\, e^{EulerGamma}$.

Proof: For $0 < x < 1$ note that

$$\frac{\prod_{k=2}^{\infty}\left(1-\frac{x^{\left\lceil\frac{k}{2}\right\rceil}}{k+1}\right)}{1-x} < p[x] < \frac{\prod_{k=2}^{\infty}\left(1-\frac{x^{k-1}}{k+1}\right)}{1-x}$$

Everything is positive so the inequalities are preserved on taking logarithms:

$$-\mathrm{Log}[1-x]+\sum_{k=2}^{\infty}\mathrm{Log}\left[1-\frac{x^{\left\lceil\frac{k}{2}\right\rceil}}{k+1}\right]<\ \mathrm{Log}[p[x]]<$$

$$-\mathrm{Log}[1-x]+\sum_{k=2}^{\infty}\mathrm{Log}\left[1-\frac{x^{k-1}}{k+1}\right]$$

We tackle the second inequality in order to prove (b).

$$\mathrm{Log}[p[x]]<-\mathrm{Log}[1-x]+\sum_{k=2}^{\infty}\mathrm{Log}\left[1-\frac{x^{k-1}}{k+1}\right]=\sum_{k=1}^{\infty}\frac{x^k}{k}-\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\frac{\left(\frac{x^k}{k+2}\right)^j}{j}=\sum_{k=1}^{\infty}\frac{2\,x^k}{k^2+2\,k}-\sum_{k=1}^{\infty}\sum_{j=2}^{\infty}\frac{x^{j\,k}}{j\,(k+2)^j}$$

The first sum in the limit as $x\to 1$ is readily computed to be $3/2$.

We evaluate this second sum at $x=1$ (justification of the interchange of sum and limit will be deferred to the next section). We first reverse the order of summation.

```
innersum = Expand[Sum[1 / (j * k^j), {k, 3, Infinity}]]
```

$$-\frac{1}{j}-\frac{2^{-j}}{j}+\frac{\mathrm{Zeta}[j]}{j}$$

We now sum this.

```
fullsum = Sum[innersum , {j, 2, Infinity}]
approx = N[fullsum]
```

$$\frac{1}{2}\,(3-2\,\mathrm{EulerGamma}-2\,\mathrm{Log}[2])$$

0.229637

```
logupbnd = 3 / 2 - fullsum
upbnd = N[Exp[logupbnd]]
```

EulerGamma + Log[2]

3.56214

So we see that an upper bound for $p[x]$ as $x\to 1$ is about 3.56.

We now prove (a). To get a lower bound we now minorize the exponents.

$$\mathrm{Log}[p[x]]>-\mathrm{Log}[1-x]+\sum_{k=2}^{\infty}\mathrm{Log}\left[1-\frac{x^{\left\lceil\frac{k}{2}\right\rceil}}{k+1}\right]=\sum_{k=1}^{\infty}\frac{x^k}{k}-\sum_{k=2}^{\infty}\sum_{j=1}^{\infty}\frac{\left(\frac{x^{\left\lceil\frac{k}{2}\right\rceil}}{k+2}\right)^j}{j}=$$

$$\frac{2\,x}{3}+\sum_{k=2}^{\infty}\left(\frac{1}{k}-\frac{1}{2\,k}-\frac{1}{2\,k+1}\right)x^k-\sum_{k=2}^{\infty}\sum_{j=2}^{\infty}\frac{x^{j\left\lceil\frac{k}{2}\right\rceil}}{j\,(k+1)^j}=$$

$$\frac{2\,x}{3}+\sum_{k=2}^{\infty}\frac{x^k}{4\,k^2+2\,k}-\sum_{k=2}^{\infty}\sum_{j=2}^{\infty}\frac{x^{j\left\lceil\frac{k}{2}\right\rceil}}{j\,(k+1)^j}$$

We will evaluate this at $x=1$ to obtain *EulerGamma*. Note that the single summation can readily be done by hand using basic calculus methods (one obtains $5/6-Log[2]$; we omit the details). So the limiting value of the first two terms is $3/2-Log[2]$. As it is important for the actual limit estimate later we indicate how to do the double summation without recourse to software; we will also show how it may be evaluated using *Mathematica*. We evaluate at $x=1$ so we want to find

$$\textbf{Limit}\left[\sum_{k=3}^{M}\sum_{j=2}^{\infty}\frac{1}{j\,k^j},\ M\to\infty\right]$$

$$\frac{3}{2}-\text{EulerGamma}-\text{Log}[2]$$

$$\text{Limit}\left[\sum_{k=3}^{M}\sum_{j=2}^{\infty}\frac{1}{j\,k^j},\ M\to\infty\right]$$

Now rewrite this as

$$\text{Limit}\left[\sum_{k=3}^{M}\sum_{j=1}^{\infty}\left(\frac{1}{j\,k^j}-\frac{1}{k}\right),\ M\to\infty\right]$$

Evaluating the inner sum gives

$$\text{Limit}\left[\sum_{k=3}^{M}\left(-\text{Log}\left[1-\frac{1}{k}\right]-\frac{1}{k}\right),\ M\to\infty\right]$$

which in turn is

$$\text{Limit}\left[\left((\text{Log}[3]-\text{Log}[2])+\ldots+(\text{Log}[M]-\text{Log}[M-1])-\sum_{k=3}^{M}\frac{1}{k}\right),\ M\to\infty\right]$$

Telescoping and simplification gives

$$\text{Limit}\left[\text{Log}[M]-\sum_{k=1}^{M}\frac{1}{k},\ M\to\infty\right]-\text{Log}[2]+\frac{3}{2}=-\text{EulerGamma}-\text{Log}[2]+3/2$$

Negating this and adding the result for the other terms gives a lower bound of *EulerGamma* for our sum. We exponentiate to obtain a lower bound for $p[1]$ that is near *1.78*, or half of our upper bound.

$$\textbf{N}\left[e^{\textbf{EulerGamma}}\right]$$

1.78107

Approximating the actual limit

Here I will discuss a nice analysis and resulting estimate worked out by Jurgen Tischer of Universidad del Valle, Columbia [Tischer 1999]. First I will mention a method I had previously tried that is not terribly promising. In brief we truncate the series for the logarithm, evaluate using exact or high precision arithmetic at $x = 1$, and that gives an estimate for the limit. Using this approach it is not difficult to approximate the result to four places without excessive computational effort. One learns that $p[1] \approx 2.292$. There are some problems with this method. First, it appears to require significant work to get even one more decimal place of precision. Moreover for truncation after N terms it is not clear how we might get a bound on the error. For odd k we get terms that are roughly $\frac{1}{2k}$ in the series and this alone makes an error bound for this approximation appear to be less than promising. We will return to these matters in the next section.

To proceed with the approximation we will look at the logarithm in more detail. It may be written as

$$\text{Log}[p[x]]=-\text{Log}[1-x]+\sum_{2|k}\text{Log}\left[1-\frac{x^{k/2}}{k+1}\right]+\sum_{2\nmid k,3|k}\text{Log}\left[1-\frac{x^{2k/3}}{k+1}\right]+\ldots=$$

$$-\sum_{k=1}^{\infty}\frac{x^k}{k}+\sum_{2|k}\frac{x^{k/2}}{k+1}+\sum_{2|k}\sum_{j=2}^{\infty}\frac{x^{j\,k/2}}{j\,(k+1)^j}+\sum_{2\nmid k,3|k}\frac{x^{2k/3}}{k+1}+\sum_{2\nmid k,3|k}\sum_{j=2}^{\infty}\frac{x^{(2/3)\,j\,k}}{j\,(k+1)^j}\ +\ldots$$

If we truncate the power series at degree N, say, it can be shown that error from dropping terms from the double sums (where $j \geq 2$) is $O\left(\frac{1}{N}\right)$. But it turns out that we need not rely on this error bound; Tischer derived a rather nice exact formula for these terms, summed to infinity. After showing his derivation we will then approximate the sum of the "main" terms, which are the single summations in the formula above.

We begin by rewriting *Log*[*p*[*x*]] (we will call this *lp*[*x*] from now on) as

$$\text{lp}[x] = x + \frac{x^2}{2} + \sum_{k=2}^{\infty}\left(\frac{x^{k+1}}{k+1} - \frac{x^{m[k]}}{k+1}\right) - \sum_{k=2}^{\infty}\sum_{j=2}^{\infty}\left(\frac{1}{j}\right)\left(\frac{x^{m[k]}}{k+1}\right)^j \tag{1}$$

Proposition 2: $lp[x] \to EulerGamma + Log[2] + \sum_{k=2}^{\infty}\left(\frac{x^{k+1}}{k+1} - \frac{x^{m[k]}}{k+1}\right)$ as $x \to 1$ from below.

Proof: The limiting value of first two terms in (*1*) is obvious so we focus on the last sum. We showed above how to evaluate it by hand at $x = 1$, now we will do so using *Mathematica*. It is not hard to see that for $x < 1$ everything in sight converges nicely and so we can switch summation order to obtain

$$\sum_{j=2}^{\infty}\left(\frac{1}{j}\right)\sum_{k=2}^{\infty}\left(\frac{x^{m[k]}}{k+1}\right)^j$$

Fix j and look at the inner sum. Since $\frac{k}{2} \leq m[k] \leq k$ and we approach from $x < 1$ we have

$$\sum_{k=2}^{\infty}\left(\frac{x^k}{k+1}\right)^j \leq \sum_{k=2}^{\infty}\left(\frac{x^{m[k]}}{k+1}\right)^j \leq \sum_{k=2}^{\infty}\left(\frac{x^{k/2}}{k+1}\right)^j$$

Simple limit arguments and Abel's theorem [Rudin 1976 p. 174] demonstrate that as $x \to 1$ the middle sum is squeezed to

$$\sum_{k=2}^{\infty}\left(\frac{1}{k+1}\right)^j$$

We must justify the interchange of limit with the outer sum. This may be done using the Weierstrass M−test [Rudin 1976 p. 148]. First note that for all $x \in [0, 1]$ we have

$$\left(\frac{1}{j}\right)\sum_{k=2}^{\infty}\left(\frac{x^{m[k]}}{k+1}\right)^j \leq \left(\frac{1}{j}\right)\sum_{k=2}^{\infty}\left(\frac{1}{k+1}\right)^j$$

and all the summands are nonnegative. Therefore if we know that the right hand side summed over j converges then we have also justified the interchange of limit and outer sum. Alternatively we might overestimate the tail of the outer sum with an integral and show that it goes to zero as the lower bound increases.

Our sum at $x = 1$ is

$$\sum_{j=2}^{\infty}\left(\frac{1}{j}\right)\sum_{k=2}^{\infty}\left(\frac{1}{k+1}\right)^j = \frac{3}{2} - \text{EulerGamma} - \text{Log}[2]$$

as was demonstrated in the previous section. This finishes the proof of proposition 2.

Now we must work with that remaining sum. As all summands are negative we may reorder them; we have

$$\sum_{k=2}^{\infty}\left(\frac{x^{k+1}}{k+1} - \frac{x^{m[k]}}{k+1}\right) =$$

$$\sum_{d[k]=2}^{\infty}\left(\frac{x^{k+1}}{k+1} - \frac{x^{\left(\frac{k}{2}\right)}}{k+1}\right) + \sum_{d[k]=3}^{\infty}\left(\frac{x^{k+1}}{k+1} - \frac{x^{\left(\frac{2k}{3}\right)}}{k+1}\right) + \sum_{d[k]=5}^{\infty}\left(\frac{x^{k+1}}{k+1} - \frac{x^{\left(\frac{4k}{5}\right)}}{k+1}\right) + \ldots \tag{2}$$

We are in effect sieving our sum over smallest prime divisors. It will turn out that we can evaluate these subsums and eventually truncate in a way that gives a very accurate result. To start out, we want to evaluate

$$\sum_{d[k]=\text{Prime}[j]}^{\infty}\left(\frac{x^{k+1}}{k+1}-\frac{x^{\left(\frac{(\text{Prime}[j]-1)\,k}{\text{Prime}[j]}\right)}}{k+1}\right)$$

We will need some simple functions below.

$$\mathbf{q[j\_] := \prod_{k=1}^{j} Prime[k]}$$

$$\mathbf{r[j\_] := \prod_{k=1}^{j-1} (Prime[k] - 1)}$$

$$\mathbf{frac[j\_] := \frac{r[j]}{q[j]}}$$

The terms *k* for which *Prime*[*j*] is the smallest divisor larger than *1* fall into finitely many congruence sets.

For example, when the prime in question is *5*, the applicable values for *k* are {*5*, *25*, *35*, *55*, *65*, *85*, *95*, ...}. This may be partitioned as {*5*, *35*, *65*, *95*, ...} and {*25*, *55*, *85*, ...}. In each case, the step size is *30*.

Proposition 3: For *Prime*[*j*] the step size in each congruence class will be

$$q[j]=\prod_{m=1}^{j}\text{Prime}[m]$$

Proof: This is perhaps well known but we provide a proof for sake of completeness. Clearly if *Prime*[*j*] is the smallest divisor of *t* then it is also the smallest divisor of $t+q[j]$, $t+2\,q[j]$, ... Suppose there is some smaller $\tilde{q}[j]$ with this property. Then for some $k<j$ we have $\text{Prime}[k]\nmid\tilde{q}[j]$. Hence *Prime*[*k*] is relatively prime to both *t* and $\tilde{q}[j]$.

Suppose *y* and *z* are distinct members of the set

$$\left\{t,\ t+\tilde{q}[j],\ t+2\,\tilde{q}[j],\ \ldots,\ t+(\text{Prime}[k]-1)\,\tilde{q}[j]\right\}$$

with $z>y$. Then

$$z-y=(k_1-k_2)\,\tilde{q}[j]$$

where

$$k_1-k_2<\text{Prime}[k]$$

Hence

$$\text{Prime}[k]\nmid z-y$$

so *y* and *z* occupy distinct congruence classes modulo *Prime*[*k*]. Thus a pigeonhole argument shows that *Prime*[*k*] must divide one of these, contradicting the assumption that *Prime*[*j*] is the smallest prime divisor.

We will also need to know the number of congruence classes for given *Prime*[*j*].

Proposition 4: The number of congruence classes for *Prime*[*j*] is

$$r[j]=\prod_{k=1}^{j-1}(\text{Prime}[k]-1)$$

Proof: This too is known but will likewise be proven for sake of completeness. First a check: For $j = 1$, $Prime[j]$ is 2, the number of congruence classes is 1. For $j = 2$, $Prime[j]$ is 3, the number of congruence classes is again 1. For $j = 3$, $Prime[j]$ is 5 and there are two congruence classes. For $j = 4$, $Prime[j]$ is 7 and there are 8 congruence classes. The general assertion is that in the set $1 \leq n \leq q[j]$ there are $r[j]$ elements whose smallest prime divisor is $Prime[j]$. We checked this explicitly above for the first few primes. A general proof will be obtained working directly with $frac[j] = r[j]/q[j]$. It suffices to show that

$$\text{frac}[j] = \frac{(\text{Prime}[j-1]-1)}{\text{Prime}[j]}\,\text{frac}[j-1]$$

This in essence defines the sieving process. At step j the fraction of integers with smallest prime divisor equal to $Prime[j]$ will be $\frac{1}{Prime[j]}$ times the fraction not already removed for having smaller divisors. In other words, we have

$$\text{frac}[j] = \frac{1 - \sum_{k=1}^{j-1} \text{frac}[k]}{\text{Prime}[j]}$$

Thus

$$\text{frac}[j] = \frac{\left(1 - \sum_{k=1}^{j-2} \text{frac}[k]\right) - \text{frac}[j-1]}{\text{Prime}[j]} = \frac{\frac{\left(1-\sum_{k=1}^{j-2}\text{frac}[k]\right)\text{Prime}[j-1]}{\text{Prime}[j-1]} - \text{frac}[j-1]}{\text{Prime}[j]} = \frac{\text{Prime}[j-1]\,\text{frac}[j-1] - \text{frac}[j-1]}{\text{Prime}[j]}$$

which completes the proof.

Now that we have determined the number of congruence classes for each prime we will need to show that the relevant limits for each prime are independent of these classes. So we must find the limits of these sums for all congruence class members as $x \to 1$. First we form such a sub−subsum as below.

$$\text{Sum}\left[\frac{x^{k+1}}{k+1} - \frac{x^{\frac{(\text{Prime}[j]-1)\,k}{\text{Prime}[j]}}}{k+1}, \{k, \text{init}, \infty, \text{step}\}\right]$$

Next we find the limit of this sum as $x \to 1$.

$$\mathbf{s1 = Sum}\left[\frac{x^{k+1}}{k+1} - \frac{x^{\frac{(\text{Prime}[j]-1)\,k}{\text{Prime}[j]}}}{k+1}, \{k, \text{init}, \infty, \text{step}\}\right]$$

$$\frac{x^{1+\text{init}}\,\text{Hypergeometric2F1}\left[\frac{1+\text{init}}{\text{step}}, 1, 1+\frac{1+\text{init}}{\text{step}}, x^{\text{step}}\right]}{1+\text{init}} - \frac{1}{1+\text{init}}\left(x^{\frac{\text{init}\,(-1+\text{Prime}[j])}{\text{Prime}[j]}}\,\text{Hypergeometric2F1}\left[\frac{1+\text{init}}{\text{step}}, 1, 1+\frac{1+\text{init}}{\text{step}}, x^{\frac{\text{step}\,(-1+\text{Prime}[j])}{\text{Prime}[j]}}\right]\right)$$

While this may look awkward we see that it has a tractable limit.

$$\mathbf{lims1 = Limit[s1,\ x \to 1]}$$

$$\left(-\text{Gamma}\left[1+\frac{1+\text{init}}{\text{step}}\right]\text{Log}[-\text{step}] + \text{Gamma}\left[1+\frac{1+\text{init}}{\text{step}}\right]\text{Log}\left[-\frac{\text{step}\,(-1+\text{Prime}[j])}{\text{Prime}[j]}\right]\right) \Big/ \left((1+\text{init})\,\text{Gamma}\left[\frac{1+\text{init}}{\text{step}}\right]\right)$$

Actually this can be simplified considerably.

```
lims1 = FullSimplify[lims1]
```

$$\frac{-\text{Log}[-\text{step}] + \text{Log}\left[\text{step}\left(-1 + \frac{1}{\text{Prime}[j]}\right)\right]}{\text{step}}$$

We next use an improper high−school simplification because we know it holds on our domain:

```
lims1 = lims1 /. Log[r_] - Log[s_] → Log[r / s]
```

$$\frac{\text{Log}\left[1 - \frac{1}{\text{Prime}[j]}\right]}{\text{step}}$$

In addition to simplicity this limit has the virtue of being independent of congruence class. That is, it is does not depend on the parameter *init*. This significantly simplifies the problem of evaluating the full sums.

In terms of the function *frac*, we have shown

Proposition 5:

$$\text{Limit}\left[\sum_{d[k]=\text{Prime}[j]}^{\infty}\left(\frac{x^{k+1}}{k+1} - \frac{x^{\left(\frac{(\text{Prime}[j]-1)\,k}{\text{Prime}[j]}\right)}}{k+1}\right), x \to 1\right] = \text{frac}[j]\,\text{Log}\left[1 - \frac{1}{\text{Prime}[j]}\right] \tag{3}$$

We now return to the full sum in (*2*). Pulling the limit inside the first M terms of the right hand side gives

$$\begin{gathered}\text{Limit}\left[\sum_{k=2}^{\infty}\left(\frac{x^{k+1}}{k+1} - \frac{x^{m[k]}}{k+1}\right), x \to 1\right] = \\ \sum_{j=1}^{M}\text{frac}[j]\,\text{Log}\left[1 - \frac{1}{\text{Prime}[j]}\right] + \text{Limit}\left[\sum_{d[k]\geq\text{Prime}[M+1]}^{\infty}\left(\frac{x^{k+1}}{k+1} - \frac{x^{\left(\frac{(d[k]-1)\,k}{d[k]}\right)}}{k+1}\right), x \to 1\right]\end{gathered} \tag{4}$$

We now want to show the tail gets arbitrarily small as $M \to \infty$ in order to interchange limit with infinite sum, as this will yield

Theorem 1:

$$\text{Limit}\left[\sum_{k=2}^{\infty}\left(\frac{x^{k+1}}{k+1} - \frac{x^{m[k]}}{k+1}\right), x \to 1\right] = \sum_{j=1}^{\infty}\text{frac}[j]\,\text{Log}\left[1 - \frac{1}{\text{Prime}[j]}\right]$$

Proof: The tail is bounded in absolute value by

$$\text{Limit}\left[\sum_{d[k]\geq\text{Prime}[M+1]}^{\infty}\left(\frac{x^{\left(\frac{(d[k]-1)\,k}{d[k]}\right)}}{k+1} - \frac{x^{k+1}}{k+1}\right), x \to 1\right]$$

and we have

$$\sum_{d[k]\geq\text{Prime}[M+1]}^{\infty}\left(\frac{x^{\left(\frac{(d[k]-1)\,k}{d[k]}\right)}}{k+1} - \frac{x^{k+1}}{k+1}\right) < \sum_{d[k]\geq\text{Prime}[M+1]}^{\infty}\left(\frac{x^{\left(\frac{(\text{Prime}[M]-1)\,k}{\text{Prime}[M]}\right)}}{k+1} - \frac{x^{k+1}}{k+1}\right) < \sum_{k=0}^{\infty}\left(\frac{x^{\left(\frac{(\text{Prime}[M]-1)\,k}{\text{Prime}[M]}\right)}}{k+1} - \frac{x^{k+1}}{k+1}\right)$$

The sum on the right is a special case of (*3*) above. It has a fairly simple closed form.

$$\textbf{tailsum} = \textbf{Sum}\left[\frac{x^{k+1}}{k+1} - \frac{x^{\frac{(\text{Prime}[M]-1)\,k}{\text{Prime}[M]}}}{k+1}, \{k, 0, \infty\}\right]$$

$$-\frac{x\,\text{Log}[1-x] - x^{\frac{1}{\text{Prime}[M]}}\,\text{Log}\left[1 - x^{\frac{-1+\text{Prime}[M]}{\text{Prime}[M]}}\right]}{x}$$

We compute the limit as $x \to 1$ from below.

```
Limit[tailsum, x→1, Direction→1]
```

```
Log[1 - 1/Prime[M]]
```

As this goes to 0 for $M \to \infty$ we have shown that the error term in (4) can be made arbitrarity small. This suffices to prove theorem 1.

Now we need to estimate our sum. We have

Corollary: $p[x] \to 2.2921736953$ as $x \to 1$ from below.

Proof: We will do a careful computation of our estimate and bounded error. Define

$$\text{estimate[n\_]} := \sum_{j=1}^{n} \text{frac[j] Log}\left[1 - \frac{1}{\text{Prime[j]}}\right]$$

$$\text{error[n\_]} := \sum_{j=n+1}^{\infty} \text{frac[j] Log}\left[1 - \frac{1}{\text{Prime[j]}}\right]$$

We will bound the magnitude of the error term and explicitly compute the estimate for some large n. To bound this error, first observe that the products *frac*[*j*] are conveniently bounded by $\frac{constant}{Prime[j]}$; this follows from a simple telescoping argument since successive factors are of the form $\frac{Prime[k-1]-1}{Prime[k]}$. We will get a reasonable value for *constant* as part of the computation that estimates the value. For this we require a faster numeric version of the algorithm for *frac*[*j*] and *estimate*[*j*]. This will allow us to avoid the incessant recomputation implicit in the definition of *estimate*[*n*]; moreover in using numeric approximation we avoid generating large rationals. This will introduce some error but as we use *Mathematica*'s significance arithmetic [Keiper 1992] at sufficiently high precision it will be small compared to the bounds that we also compute. We will overestimate those bounds so arithmetic correct to *20* digits will more than suffice for our purposes.

```
numestimate[n_, prec_] :=
 Module[{sum = -1/2*Log[2], frac = 1/2, p1 = Prime[1], p2},
  Do[
   p2 = Prime[j];
   frac *= (N[p1 - 1, prec] / p2);
   sum += frac*Log[N[1 - 1/p2, prec]];
   p1 = p2,
   {j, 2, n}];
  {frac*p2, sum}]
```

```
Timing[{mfact, est} = numestimate[2*10^7+1, 25]]
```

```
{14470.91 Second, {0.0284446218701220421, -0.4408622662100870964715674}}
```

Setting n to be $2*10^7+1$, an upper bound on the magnitude of *error*[*n*] is given by

$$\text{mfact} \sum_{j=n+1}^{\infty} \frac{1}{\text{Prime[j]}} \left| \text{Log}\left[1 - \frac{1}{\text{Prime[j]}}\right] \right|$$

This in turn is bounded by

$$\text{mfact}\left(1 + \frac{1}{\text{Prime[n]}}\right) \sum_{j=n+1}^{\infty} \frac{1}{j^2\,\text{Log[j]}^2}$$

Here we have overestimated the error by using the asymptotic term $j\,Log[j]$ from the prime number theorem to underestimate $Prime[j]$ [Ribenboim 1996 p. 249], and we use the basic calculus result that $\left(1+\frac{1}{Prime[n]}\right)$ overestimates the ratio of $-Log\left[1-\frac{1}{Prime[j]}\right]$ to $\frac{1}{Prime[j]}$ for $j>n$. We will now ignore this factor because we more than compensate when we round upward the value of *mfact* in the fourth decimal place. Thus we have as a bound:

$$(.0285)\sum_{j=n+1}^{\infty}\frac{1}{j^2\,Log[j]^2}$$

Using simple calculus we overestimate this sum by an integral:

```
errormax = (57 / 2000) ∫_{2×10^8}^{∞} 1/(j Log[j])^2 ⅆj
```

$$\frac{57\left(\text{ExpIntegralEi}[-\text{Log}[20\,000\,000]]+\frac{1}{20\,000\,000\,\text{Log}[20\,000\,000]}\right)}{2000}$$

```
N[errormax]
```

4.52951×10^{-12}

Now we add $estimate\left[2\times10^{7}+1\right]$ to the exact terms given in proposition 2. Then we will exponentiate the sum to get an approximation of $p[x]$ as $x\to1$ from below. We also multiply this by $e^{errormax}-1$ to get a bound on the error in our estimate of the limit.

```
p1 = Exp[EulerGamma + Log[2] + est]
```

2.29217369525553883711607

```
errbound = N[Exp[errormax] - 1] * p1
```

1.03824×10^{-11}

So we have obtained a result that is correct to about the eleventh decimal place, demonstrating the corollary. In [Tischer 1999] the computations are done with around 10^8 terms. This can achieve a result with approximately one more decimal place of accuracy. In actual fact it was presented as being slightly less accurate because the error bound was looser.

Further remarks about the truncated power series

We will use the *Mathematica Series* function to work with the truncated power series. First we note a curiousity in the sign pattern of the coefficients.

```
logp[x_, n_] := Normal[Series[-Log[1 - x] + Sum[Log[1 - x^m[k]/(k + 1)], {k, 2, 2*n}], {x, 0, n}]]

terms100 = logp[x, 100];
coeffs100 = CoefficientList[terms100 , x];
signs100 = Map[Sign, coeffs100];
Map[{#[[1]], Length[#]} &, Split[signs100]]

{{0, 1}, {1, 1}, {-1, 1}, {1, 1}, {-1, 1}, {1, 1}, {-1, 1}, {1, 3}, {-1, 1},
 {1, 1}, {-1, 1}, {1, 1}, {-1, 1}, {1, 1}, {-1, 1}, {1, 1}, {-1, 1}, {1, 1},
 {-1, 1}, {1, 1}, {-1, 1}, {1, 3}, {-1, 1}, {1, 1}, {-1, 1}, {1, 1}, {-1, 1},
 {1, 3}, {-1, 1}, {1, 1}, {-1, 1}, {1, 1}, {-1, 1}, {1, 1}, {-1, 1}, {1, 1},
 {-1, 1}, {1, 1}, {-1, 1}, {1, 1}, {-1, 1}, {1, 3}, {-1, 1}, {1, 1}, {-1, 1},
 {1, 1}, {-1, 1}, {1, 3}, {-1, 1}, {1, 1}, {-1, 1}, {1, 1}, {-1, 1},
 {1, 3}, {-1, 1}, {1, 1}, {-1, 1}, {1, 1}, {-1, 1}, {1, 1}, {-1, 1},
 {1, 1}, {-1, 1}, {1, 1}, {-1, 1}, {1, 1}, {-1, 1}, {1, 3}, {-1, 1},
 {1, 3}, {-1, 1}, {1, 1}, {-1, 1}, {1, 1}, {-1, 1}, {1, 1}, {-1, 1},
 {1, 1}, {-1, 1}, {1, 1}, {-1, 1}, {1, 1}, {-1, 1}, {1, 1}, {-1, 1}}
```

We see an unusual phenomenon wherein signs by-and-large alternate except for sporadic runs of three consecutive positive coefficients. This holds at least for the first *2000* coefficients. It is an open question as to whether or why this persists.

In attempting to get a numerical approximation from this not−quite−alternating series, at first glance it appeared that one could sum n terms and hope to obtain O(n) convergence. Further work makes it clear that this cannot be the case, at least not for all values of n. Convergence is slow (though it does in fact seem to converge) for good reason; some of the terms are asymptotically larger than $O\left(\frac{1}{n}\right)$. As we will see, one can bound from above the magnitudes of the coefficients that arise in the power series by $O\left(\frac{\log[\log[n]]}{n}\right)$ for n bigger than 3. Further work will reveal that this bound is the best we can do, and indeed the coefficients that approach it will be not too scarce.

In order to obtain results more readily we will now work with only the main terms. Specifically, we discard higher−order terms from all but the main logarithmic term (the one that is singular at $x = 1$). The series in question is now

$$\text{logpmainterms[x\_]} := \sum_{k=1}^{\infty} \frac{x^k}{k} - \sum_{k=2}^{\infty} \frac{x^{m[k]}}{k+1}$$

where the first sum is for $-Log[1 - x]$ and the second is for the main terms from the remaining logarithms.

Thus the remarks below actually apply to a slightly altered problem, but one that is not different in significant aspects from the original (and in particular, with a bit of work one can show that the derivations of coefficient size bounds will still apply). We write the series for *logpmainterms* in the symbolic form $\sum_{k=1}^{\infty}\left(c[k]\, x^k\right)$, thus implicitly defining a "coefficient function" $c[k]$. Let us see explicitly how one would describe a general coefficient. Contributions to the term in x^k arise from the $\frac{1}{k}$ coefficient in the first sum, as well as from coefficients of the form $\frac{1}{j+1}$ for all j for which $m[j] = k$.

Proposition 6: For odd k we have

$$c[k] = \frac{1}{k} - \frac{1}{2k+1} = \frac{k+1}{2k^2 + k}$$

Proof: This follows from the following simple observations.

(i) $d[2k] = 2$, hence $m[2k] = k$.

(ii) For j even and $j \neq 2k$, $m[j] = \frac{j}{2} \neq k$.

(iii) For j odd, $\frac{j}{d[j]}$ is odd, hence $m[j]$ is even. So $m[j] \neq k$.

We conclude that for odd k there are exactly two contributions to $c[k]$ and they are as in the statement of the proposition.

Corollary: Proposition 6 shows (as previously noted) that the coefficients are in general no smaller than $O\left(\frac{1}{k}\right)$. But they can in fact be bigger. The next result gives an indication of how large they can be.

Theorem 2:

(a) An upper bound on size of $c[n]$ is $\frac{Log_2[Log_2[n+1]]+1}{n}$.

(b) There are coefficients that approach bound to within a factor of $3/4$. We construct one such explicitly. From the method one will see how to approach the upper bound within any factor less than 1.

Proof: We now investigate coefficients of those terms of even degree. Let us suppose that $m[j] = k$ for some j. Let $d[j] = p$. Then we know that $j - \frac{j}{p} = \frac{(p-1)\,j}{p} = k$ and thus $p\,k = (p-1)\,j$. Immediately we see that

(i) $p - 1 \mid k$

(ii) $p \mid j$

(this is in any case obvious because p is given as the smallest prime divisor of j)

(iii) $\frac{k}{p-1} = \frac{j}{p}$

and hence $\frac{k}{p-1}$ has no prime divisor smaller than p (again because p is the smallest prime divisor of j).

Now we suppose we have all possible j for which $m[j] = k$. Specifically, suppose

$$k = \frac{p_1 - 1}{j_1} = \frac{p_2 - 1}{j_2} = ... = \frac{p_s - 1}{j_s}$$

where $d[j_r] = p_r$ for all $1 \le r \le s$ and we moreover have ordered these so that

$$2 = p_1 < p_2 < ... < p_s \le k + 1$$

and

$$2\,k = j_1 > j_2 > ... > j_s \ge k + 1$$

As we saw, $p_r - 1 \mid k$ for each $1 \le r \le s$.

Claim 1: $p_r - 1$ contains all divisors of k that are smaller than p_r .

Proof: If not then the equality $p_r\,k = (p_r - 1)\,j_r$ implies that j_r has a divisor (and hence a prime divisor) strictly smaller than p_r . This would contradict the fact that $p_r = d[j_r]$.

Claim 2: For each $1 \le r \le s - 1$ we have $p_r - 1 \mid p_{r+1} - 1$.

Proof: $p_r - 1$ contains all divisors of k that are smaller than p_r. It cannot contain any other factors; there is no room left over as it is itself smaller than p_r. A similar statement holds for the divisors of $p_{r+1} - 1$ with respect to p_{r+1}. Thus the divisors of $p_r - 1$ are a proper subset of those of $p_{r+1} - 1$.

We obtain the chain $p_2 - 1 \mid p_3 - 1 \mid ... \mid p_s - 1$ of length $s - 1$ (we do not bother with p_1 because it is 2, so $p_1 - 1$ is 1), and more− over we have inequalities concerning the quotients. Specifically, for $1 \le r \le s - 1$ we may write

$$p_{r+1} - 1 = q_r\,(p_r - 1)$$

where $q_r \ge p_r$. Hence

$$p_{r+1} \ \ge\ p_r\,(p_r - 1) + 1$$

Now using the fact that $p_r - 1 \ge \frac{2}{3}\,p_r$ we obtain

$$k + 1 \ge p_s \ge \left(\left(\frac{2}{3}\right)^{2^{s-2}}\right) p_2^{2^{s-2}} \ge 2^{2^{s-2}}$$

We now take logarithms to base 2 to see $Log_2[k + 1] > 2^{s-2}$. Again taking logarithms to base 2, we get

$$\mathrm{Log}_2\left[\mathrm{Log}_2[k + 1]\right] + 2 \ge s$$

Thus

$$|\,c[k]\,| = \left|\frac{1}{k} - \sum_{m[j]=k} \frac{1}{j + 1}\right| < \frac{\mathrm{Log}_2\left[\mathrm{Log}_2[k + 1]\right] + 1}{k}$$

where we use the fact that the denominators in the sum all lie between $k + 1$ and $2\,k$. This completes the proof of part (a).

Now that we have this upper bound for coefficients in the power series, it is natural to ask whether one can derive an asymptotically smaller bound. As stated in part (b), this is not possible. We outline why this is so. First we investigate numerically: we will search for coefficients that approach this bound in ratio. Within the first few thousand terms we can find several that come close.

```
approxlogp[x_, n_] := Normal[Series[Sum[x^k/k, {k, 1, n}] - Sum[x^m[k]/(k + 1), {k, 2, 2*n}], {x, 0, n}]]

Timing[terms4k = approxlogp[x, 2^12];]
coeffs4k = CoefficientList[terms4k, x];

{20.82 Second, Null}
```

Now we look at the sizes of the coefficients. While as we know none can be larger than $\frac{Log[2,Log[2,n]]+1}{n}$ (for $n > 3$, that is), it appears that some come near to this. For example, in the first 4096 coefficients one is about 0.66 of this maximum and it is roughly three times larger than $\frac{1}{n}$.

```
normedcoeffs = Drop[Abs[coeffs4k], 4] *
   Table[n / (Log[2, Log[2, N[n]]] + 1), {n, 5, Length[coeffs4k]}];
```

```
Max[normedcoeffs]
```

```
0.661706
```

Let us investigate more closely that large normalized coefficient.

```
Position[normedcoeffs , Max[normedcoeffs]]
```

```
{{3063}}
```

We dropped four coefficients (corresponding to terms of degree zero through three), so we want to look at the *3067*th coefficient in the list, which goes with the term of degree *3066*.

```
(Log[2, Log[2, 3067]] + 1) / 3067 // N
```

```
0.00126164
```

```
Abs[coeffs4k[[3067]]] // N
```

```
0.000978188
```

```
Abs[coeffs4k[[3067]] * 3067] // N
```

```
3.0001
```

It is claimed above that the best bound we can attain on these coefficients is in fact $O\left(\frac{\log[\log[n]]}{n}\right)$. To see how to approach this bound, it is instuctive to see how we obtain the coefficient for x^{3066}.

```
FactorInteger[3066]
```

```
{{2, 1}, {3, 1}, {7, 1}, {73, 1}}
```

We have a chain of primes

$$\{1+1,\ 2+1,\ 2*3+1,\ 2*3*7+1,\ 2*3*7*73+1\}$$

wherein all factors in the indicated products are themselves prime, and each last factor is at least as large as the previous element in the list (note that *73* is the smallest prime greater than or equal to *43* that works). In general we will look for values of *k* that have an increasing chain of prime factors q_r such that

$$p_{r+1} - 1 = q_r\,(p_r - 1)$$

for some ascending set of primes p_r. For example, to find a term after the *3066*th that makes the ratio larger, we look for the smallest prime *q* larger than *3066* such that $3066\,q + 1$ is also prime. This happens for $q = 3137$, giving $k = 9\,618\,042$. The ratio

$$\frac{\text{Log}[2,\ \text{Log}[2,\ n]] + 1}{n}$$

to the size of the coefficient of $x^{9\,618\,042}$ turns out to be near *0.72*. Carrying this construction yet another step gives the next *q* as $9\,618\,361$, so the value for *k* is $9\,618\,361 * 9\,618\,042$ or $92\,509\,800\,069\,162$. The ratio now is about *0.765*. This finishes the proof of part (b) and moreover demonstrates how one might search for coefficients closer in ratio to the upper bound of part (a).

Note that one can get more terms near the upper bound simply by taking for our next prime factor *q* any prime slightly larger than the smallest one that works, as some of these may also work. For example, if we use *79* rather than *73* as a factor above (noting that $2*3*7*79+1 = 3319$ is prime), we get a coefficient with ratio just under *.66*. One thus sees a general means by which to obtain coefficients not far in ratio below the upper bound given above.

As a final remark, let us recall that a straightforward computation with the truncated power series, taking $x = 1$, gave a result correct to about *4* digits of precision. We know they are correct because they agree with the result of Tischer's method, which we prove above to give a correct result to several more places. Also at the beginning of this report we found a *lim inf* and *lim sup* for the power series sum as $x \to 1$. That notwithstanding, we do not have a stand−alone proof that the power series itself converges to the limit. An indirect proof is that Tischer's method converges, and one observes that the truncated power series only omits from Tischer's truncated sum terms of the same sign, hence it too must converge. Of course the construction of terms with coefficients of size $O\left(\frac{log[log[n]]}{n}\right)$ forces one to conclude that estimation of the limit by truncation of the power series will be problematic, but nonetheless an independent proof of convergence would be of interest in its own right.

Acknowledgements

I thank Harold Diamond for several helpful discussions about this problem, and for inviting me to speak about it in the weekly departmental seminar in analytical number theory. I also thank the University of Illinois Math Department for extending its courtesy to me while I was a visitor during the academic year 1998−9.

References

J. Keiper (1992). Numerical computation with *Mathematica* (tutorial notes). An electronic version may be found at: http://mathsource.com/Content22/Enhancements/Numerical/0204−028

A. Knopfmacher (1999a). Usenet news group post to comp.soft−sys.math.mathematica. http://library.wolfram.com/mathgroup/archive/1999/Jan/msg00023.html

A. Knopfmacher (1999b). Private communication.

A. Knopfmacher and R. Warlimont (1995). Distinct degree factorizations for polynomials over a finite field. Transactions of the American Mathematical Society **347**(6):2235−2243.

P. Ribenboim (1996). The *New* Book of Prime Number Records. Springer.

W. Rudin (1976). Principles of mathematical Analysis (third edition). McGraw−Hill.

J. Tischer (1999). Usenet news group post to comp.soft−sys.math.mathematica. http://library.wolfram.com/mathgroup/archive/1999/Jan/msg00355.html

S. Wolfram (1999). The *Mathematica* Book (4th edition). Wolfram Media/Cambridge University Press.