Computing Knopfmacher’s limit, or: My first foray into computational mathematics, reprise

Daniel Lichtblau
danl@wolfram.com
Wolfram Research, Inc.
100 Trade Center Dr.
Champaign IL 61820

ACA 2009
Montreal, Canada
Session: Applications of Math Software to Mathematical Research

ABSTRACT

I will discuss a problem I encountered over a decade ago, and worked on via internet with someone I (alas) never met. It involves a mix of number theory, real analysis, hard−core computation, and some slightly perplexing results.

In brief, we begin with a function expressed as a certain infinite product; Arnold Knopfmacher encountered it in an attempt to approximate the number of irreducible factors of univariate polynomials over Galois fields and raised the question of how to obtain a certain limit to this function. We derive and execute an effective algorithm for the task at hand. We’ll also indicate why the most "obvious" approach does not work well in practice, or at all in theory.

The problem

As posed by Arnold Knopfmacher to the Usenet group comp.soft−sys.math.mathematica in January 1999

Let $d(k)$ denote the smallest prime factor of k.

Define $m(k) = k - k/d(k)$

Define

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

We wish to compute numerically, to at least eight decimal places, the following limit.

$$\lim_{x\to 1^-} p(x)$$

That is to say, we compute the limit as x approaches 1 from the left (lesser) side.

Why do we care?

- There are similar formulas in a paper from 1995 by Knopfmacher and Warlimont, analyzing probabilities related to numbers of irreducible factors of distinct degrees in univariate polynomials over Galois fields

- It's an interesting computation

My history with this problem

I worked on it off and on for several days. Then someone else reading the forum contributed a similar result, but much more precise. His name was Jürgen Tischer, a math department faculty member of Universidad del Valle, Columbia. We corresponded a bit over a period of months, and I wrote up the results. I lost contact with him a year or two later. I had wanted to invite him to this conference. After getting nowhere with an internet search for a current contact address, I learned he had passed away in January of 2008. I felt it fitting to talk a bit about this problem, since it used ideas of his and also was one of my first forays into computational mathematics.

Candidate results

There were four responses, with two (mine and Tischer's) giving roughly the same results. The proposed values were

- 1.3397
- 2 (exactly)
- 2.292

So which is correct?

More important: How do we even know the limit exists?

Easy to show...

- A *lim sup* and *lim inf* both are readily computed.
- A *lim inf* is given by e^{γ}(the exponential of the Euler gamma constant). This is around 1.871, so...1.3397 will exit stage left.

◄ | ►

Definitions in *Mathematica*

```
d[k_] := Divisors[k][[2]]
```

$$\mathrm{m}[k\_] := k - \frac{k}{\mathrm{d}[k]}$$

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

General idea: Start with

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

Now take logs to get summations. Expand logs at 1 as power series, obtaining double summations. Switch order of summation (requires justification), and we find that log of a *lim sup* is $\gamma + \log(2)$. Finding a *lim inf* is similar though a bit more work.

What has changed in the past decade?

Ten years ago this computing took manual intervention. I had to do things like split sums, and do further contortions to take limits. Today some can be done directly. Here is one such that arose in the process.

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

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

◄ | ►

A start at approximating the actual limit

- Truncate the series for the logarithm
- Evaluate using exact or high precision arithmetic at $x = 1$
- Exponentiate

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

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

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

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

(The series that get truncated are the summations in *j*). From this tactic I was able to get 2.292 (so, as you may have guessed, exact 2 was not the correct result either).

Troublesome aspects

There are serious problems with this approach.

- Difficult to get good precision
- (Related, but more serious) It is quite difficult to bound the error. Indeed, it is not easy to show we have convergence.

◄ | ►

Tischer’s idea

Figure out exact forms for some of the infinite sums, so as to avoid truncation. In parts we cannot compute exactly, show that error is much better than what we have from above approach.

Start by writing $\log(p(x))$ (after a bit of algebra) as

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

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

Proposition: This approaches $\gamma + \log(2) + \sum_{k=2}^{\infty} \left(\frac{x^{k+1}}{k+1} - \frac{x^{m(k)}}{k+1} \right)$ as $x \to 1$ from below.

◄ | ►

Sketch of proof

Clearly we only need focus on the double summation. Switch summation order and use

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

Middle is squeezed to

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

So we can find:

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

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

Several steps require justification! We interchanged a summation order, then a sum with a limit...

◄ | ►

That remaining summation

We now need to estimate the remaining part. We split by smallest divisors.

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

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

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

The reordering is fine: for $0 < x < 1$ each term is negative so we can do this.

Now we need to compute

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

◄ | ►

Remaining summation...

We need some functions.

```
q[j_] := Product[Prime[k], {k, 1, j}]
r[j_] := Product[(Prime[k] - 1), {k, 1, j-1}]
frac[j_] := r[j]/q[j]
```

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

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

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

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.

Remaining summation...

In general we have the following lemmas.

- The step size of congruence classes for prime (j) is q(j) as defined above
- The number of congruence classes is $r(j)$
- When we partition in this way, the limit for each subsum depends only on the prime and is independent of congruence class

Upshot:

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

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

◄ | ►

Remaining summation...

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

$$\sum_{j=1}^{M}\mathrm{frac}[j]\,\mathrm{Log}\left[1-\frac{1}{\mathrm{Prime}[j]}\right]+$$

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

We can readily bound that tail sum.

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

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

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

$$\mathrm{Log}\left[1-\frac{1}{\mathrm{Prime}[M]}\right]$$

◄ | ►

Computing our estimate

We can now put all this together.

```
estimate[n_] := Sum[frac[j] Log[1 - 1/Prime[j]], {j, 1, n}]

error[n_] := Sum[frac[j] Log[1 - 1/Prime[j]], {j, n+1, ∞}]
```

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

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

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

```
Timing[{mfact, est} =
  numestimate[2 * 10 ^ 8 + 1, 25]]
```

```
{17524.4, {0.02533245484926073 9,
  -0.4408622662133543648819837}}
```

Computing...

Upper bound on error. Set $n = 2 \times 10\text{^}8$.

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

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

$$57/2000 \int_n^{\infty} \frac{1}{(j\,\log[j])^2}\, dj$$

```
errormax =
  (57 / 2000) Integrate[1 / (j Log[j]) ^2,
    {j, 2*10^8, Infinity}]
N[errormax]
```

$$\frac{1}{2000}\, 57 \left(\text{ExpIntegralEi}[-\text{Log}[200\,000\,000]] + \frac{1}{200\,000\,000\,\text{Log}[200\,000\,000]}\right)$$

3.54571×10^{-13}

◄ | ►

Computing...

Finally we get our estimate and error bound.

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

```
2.29217369524804969041O395
```

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

8.12817×10^{-13}

◄ | ►

Further items of interest

We can investigate the error term of the "naive" summation approach by looking at the series of the log of the product.

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

$$\text{logp}[x\_, n\_] := \text{Normal}\left[\text{Series}\left[-\text{Log}[1-x] + \sum_{k=2}^{2*n} \text{Log}\left[1 - \frac{x^{m[k]}}{k+1}\right], \{x, 0, n\}\right]\right]$$

We can use this to get the signs of the terms. I show them in a run−length form.

```
{{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}}
```

They seem to alternate, with sporadic runs of three positive terms. Strange...

Further items...

But stranger is the magnitudes of these coefficients. They are not even bounded by $O\left(\frac{1}{n}\right)$.

Can show:

- They are bounded by $O\left(\frac{\log[\log[n]]}{n}\right)$
- This bound is tight (we can show that infinitely many coefficients will approach it closely).
- The ones that approach closely have interesting factorization patterns (which is why they approach closely).
- Figuring this out was more math than computation (our jobs are not going to the machines just yet).

Upshot: a naive summation will clearly give very poor convergence.

Some open problems

- Understand the sign patterns of error approximants.
- Find a more efficient way to compute the estimate to high precision.
- Find an exact closed form for the limit.
- Automate more of the symbolic analysis: some still requires manual intervention.
- Determine whether the error bound/estimate is tight. If not, improve it (this would be a "cheap" way of getting more digits).

References

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

DL (1999). The evaluation of Knopfmacher's curious limit. Manuscript.

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