Research Statement
William F. Galway

My research centers on computational number theory. It is the main concern of my thesis, and has also been the basis for several fruitful collaborations.

My thesis [Gal00a] deals with an analytic algorithm for computing $\pi(x)$, the number of primes below x. This algorithm, first described by Lagarias and Odlyzko, is based on numerical integration of a function related to the Riemann zeta function, $\zeta(s)$, in combination with summation of a "kernel function" evaluated at prime powers "near" x [LO84]. Later, with the development by Odlyzko and Schönhage of a fast method for computing $\zeta(s)$ [OS88], Lagarias and Odlyzko showed their algorithm could find $\pi(x)$ in $O(x^{1/2+\epsilon})$ time [LO87]. Although asymptotically this is the fastest known algorithm for computing $\pi(x)$, several unresolved technical issues have prevented any practical implementation until now. My thesis addresses these issues. Results include:

- Design of a kernel function superior to the kernel proposed in [LO87];
- Choice of a quadrature algorithm, with a careful analysis of quadrature error;
- Development of a new sieving algorithm requiring $O(x^{1/3+\epsilon})$ bits of memory for efficient operation, in contrast to previously known sieving methods, which require $O(x^{1/2+\epsilon})$ bits;
- An improved method for computing the Riemann zeta function to arbitrarily high accuracy.

Further details on my thesis are given in Sections 1 and 3.

I have performed several computations in collaboration with others. These include:

- Computation of 2^{27} coefficients of a power series, modulo 625, using a result due to Dennis Eichhorn to verify a conjecture of James Sellers on congruences for 2–colored Frobenius partitions.
- Computation of $5,400,000$ coefficients, modulo 9, of two modular forms, allowing Ken Ono to prove a congruence conjectured by J. Nekovář [Ono98]. In a separate project, I found all $n \leq 2 \cdot 10^{10}$ of the form $n = x^2+y^2+10z^2$ for integer x, y, z, allowing Ono and K. Soundararajan to give a complete list of numbers not of this form, conditionally on a GRH [OS97]. This software has since been used by Thomas Reinke, at Westfälische Wilhelms Universität, for research towards his Diplomarbeit.
- Finding all $n \leq 10^9$ with $n!+1 = m^2$. This is described in a paper with Bruce Berndt [BG00].

To extend results from my thesis and from earlier collaborations I propose research on:

- computing zeta and L-functions,
- the distribution of pseudoprimes.

1. Computing zeta and L-functions

The Riemann-Siegel formula for $\zeta(s)$ is derived by applying the saddle point method to find an asymptotic expansion of a certain integral, and is the preferred method for computing $\zeta(1/2 + it)$ when t is large [Edw74]. I have observed that this integral is well-suited to numerical quadrature, which offers advantages over the asymptotic expansion [Gal]. One advantage is that quadrature may be used to compute the integral to arbitrarily high accuracy, while the asymptotic expansion allows only limited accuracy. Another advantage is that analysis of the error appears to be much simpler than for the asymptotic expansion, for which good error bounds are only available in the case $\mathrm{Re}(s) = 1/2$ [Gab79]. I propose to explore similar techniques for other zeta and L-functions.

The Riemann-Siegel formula has a "main" sum of the form $\sum_{n=1}^{N} n^{-s}$. Berry and Keating have studied expansions in which this sum is replaced by a "smoothed" sum: $\sum_n a_n n^{-s}$, where the

a_n drop from 1 to 0 [BK92]. Their expansions are more accurate than the classical formula. My quadrature-based method also benefits from smoothing, and I hope to further explore this technique.

2. Distribution of pseudoprimes

Following the usual notation, "psp(n)" denotes that n is *pseudoprime*, i.e., $2^n \equiv 2 \bmod n$, n composite. Exploration of efficient primality tests for my thesis led me to study pseudoprimes with no small prime factors. I observed that many pseudoprimes of the form $n = pq$, with p, q prime, satisfy the relation $(p-1)/(q-1) = r/s$, with small integer r, s. This led to the following conjecture:

Let p, q, ℓ be odd primes, and let $P_2(x)$ be the counting function for odd pseudoprimes with two distinct prime factors, $P_2(x) := \#\{n \le x : n = pq,\ p < q,\ \mathrm{psp}(n)\}$. Then $P_2(x) \sim C\sqrt{x}/\ln^2(x)$ as $x \to \infty$, where

$$C = 4T \sum_{s \ge 1} \sum_{\substack{r > s \\ \gcd(r,s)=1}} \frac{\delta(rs)\,\rho(rs(r-s))}{(rs)^{3/2}} \approx 30.03\ldots \tag{1}$$

$$T = 2 \prod_{\ell} \frac{1 - 2/\ell}{(1-1/\ell)^2} \approx 1.320323632\ldots$$

$$\rho(m) = \prod_{\ell \mid m} \frac{\ell - 1}{\ell - 2}, \quad \textit{and} \quad \delta(m) = \begin{cases} 2 & \textit{if } 4 \mid m \\ 1 & \textit{otherwise.} \end{cases}$$

The argument supporting this conjecture depends on a longstanding conjecture of Hardy and Littlewood concerning the density of prime pairs (p, q) with $(p-1)/(q-1) = r/s$ [HL23, §5.33, Conjecture D] . In addition, I assume that prime pairs for which pq is pseudoprime are distributed as implied by the Chebotarëv density theorem. The value $C \approx 30.03\ldots$ was found by summing roughly 10^9 of the most significant terms from the defining sum (1).

Using tables of pseudoprimes computed by Richard Pinch [Pin00, Pin], we see that the conjecture $P_2(x)/(\sqrt(x)/\ln^2(x)) \to 30.03\ldots$, $(x \to \infty)$, is only roughly supported by the data.

x	$P_2(x)$	$P_2(x)/(\sqrt{x}/\ln^2(x))$
10^{10}	6501	34.468
10^{11}	17207	34.908
10^{12}	46080	35.181
10^{13}	123877	35.100

I propose:

- to make the computation of C rigorous, with explicit error bounds;
- to extend Pinch's tables of pseudoprimes, which would further test the conjecture, would clarify the error term in our estimates, and would be of independent interest (Pinch and I hope to collaborate on this work);
- to treat pseudoprimes with k prime factors, relating this to similar work of Granville and Pomerance [GP];
- to generalize my analysis to cover pseudoprimes to an arbitrary base b, i.e., to composite numbers n with $b^n \equiv b \bmod n$.

3. Thesis Details

The Lagarias-Odlyzko algorithm is based on the identity which, for $\sigma > 1$, may be written as

$$\pi^*(x) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} \frac{x^s}{s} \ln \zeta(s)\, ds, \tag{2}$$

where $\pi^*(x) = \sum_{m \ge 1} \pi(x^{1/m})/m$.

Recall that a "kernel function" $\phi(u)$ and its Mellin transform $\widehat{\phi}(s)$ are related by

$$\widehat{\phi}(s) = \int_0^\infty \phi(u) u^{s-1}\, du, \qquad\qquad \phi(u) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} \widehat{\phi}(s) u^{-s}\, ds.$$

Lagarias and Odlyzko noted that although (2) is unsuited for computation because the integral converges very slowly, it may be generalized to

$$\pi^*(x) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} \widehat{\phi}(s)\ln\zeta(s)\, ds + \sum_{p^m} \frac{1}{m}(\chi(p^m;x) - \phi(p^m)), \tag{3}$$

where $\chi(u;x)$ denotes the characteristic function of the interval $[0,x]$ and where the sum is evaluated over powers of primes, p^m, $m \geq 1$. If $\phi(u)$ closely approximates $\chi(u;x)$, and $\widehat{\phi}(s)$ damps out quickly as $\mathrm{Im}(s) \to \pm\infty$, then we may closely approximate (3) by truncating both the integral and sum:

$$\pi^*(x) \approx \frac{1}{2\pi i}\int_{\sigma-iT}^{\sigma+iT} \widehat{\phi}(s)\ln\zeta(s)\, ds + \sum_{p^m\in[u_0,u_1]} \frac{1}{m}(\chi(p^m;x) - \phi(p^m)). \tag{4}$$

A fast method of approximating $\pi^*(x)$ gives a fast algorithm for $\pi(x)$, because finding

$$\pi(x) = \pi^*(x) - \sum_{m\geq 2} \frac{1}{m}\pi(x^{1/m})$$

requires computation of a sum with roughly $\log_2(x)$ non-zero terms, and because all of $\pi(x^{1/2})$, $\pi(x^{1/3})$, ... , may be computed in $O(x^{1/2+\epsilon})$ operations. Given a sufficiently accurate approximation of $\pi(x)$, we may round to the nearest integer to find $\pi(x)$ exactly.

Lagarias and Odlyzko suggested a kernel function which allows both a short quadrature interval and a short summation interval, and allows fast computation of $\widehat{\phi}(s)$ and $\phi(u)$. Their $O(x^{1/2+\epsilon})$ time bound is based on order of magnitude estimates of the time required to perform quadrature and to locate primes [LO87]. They did not deal with issues such as precise estimates of error terms, which would be required for an implementation.

My thesis proposes use of the Mellin transform pair

$$\phi(u) = \phi(u;x,\lambda) = \frac{1}{2}\operatorname{erfc}\left(\frac{\ln(u/x)}{\sqrt{2}\lambda}\right), \qquad\qquad \widehat{\phi}(s) = \widehat{\phi}(s;x,\lambda) = e^{\lambda^2 s^2/2}\,\frac{x^s}{s}.$$

where $\operatorname{erfc}(z)$ is the complementary error function, and $\lambda > 0$ controls the kernel's cutoff rate. I give precise bounds on error terms, and discuss optimal choices for parameters such as λ, σ, and T. Letting $G(s)$ denote the integrand appearing in (3), I also present a "number theoretical" quadrature formula, of the form

$$\frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} G(s)\, ds \approx \frac{h}{2\pi}\sum_{|kh|\leq T} G(\sigma+ikh) - \sum_{p^m\leq x^{1/2+\epsilon}} \ldots$$

where the sum over p^m serves as a correction term, allowing larger h for a given error bound.

Computing the sum over p^m in (4) requires efficient location of primes near x. To achieve this, I present a modification of an algorithm due to Atkin and Bernstein [AB99], using a technique similar to that used by Voronoï to treat the "Dirichlet divisor problem" [Vor03]. This reduces the memory requirements to $O(x^{1/3+\epsilon})$ bits from $O(x^{1/2+\epsilon})$ bits for the Atkin-Bernstein algorithm [Gal00b].

References

[AB99] A. O. L. Atkin and D. J. Bernstein, *Prime sieves using binary quadratic forms*, Dept. of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, 60607-7045. Preprint available at `http://pobox.com/~djb/papers/primesieves.dvi`, 1999.

[BG00] Bruce C. Berndt and William F. Galway, *On the Brocard-Ramanujan Diophantine equation* $n! + 1 = m^2$, Ramanujan J. (2000), MR 1 754 629.

[BK92] Michael V. Berry and Jonathan P. Keating, *A new asymptotic representation for* $\zeta(\frac{1}{2} + it)$ *and quantum spectral determinants*, Proc. Roy. Soc. London Ser. A **437** (1992), no. 1899, 151–173, MR **93j**:11057.

[Edw74] H. M. Edwards, *Riemann's zeta function*, Pure and Applied Mathematics, Vol. 58, Academic Press, New York, 1974, MR **57** #5922.

[Gab79] Wolfgang Gabcke, *Neue Herleitung und Explizite Restabschätzung der Riemann-Siegel-Formel*, Ph.D. thesis, Georg-August-Universität zu Göttingen, 1979.

[Gal] William F. Galway, *Computing the Riemann zeta function by numerical quadrature*, to be submitted to the Proceedings of the AMS Conference on Dynamical, Spectral and Arithmetic Zeta-Functions.

[Gal00a] William F. Galway, *Analytic computation of the prime-counting function*, Ph.D. thesis, University of Illinois at Urbana-Champaign, 2000, (expected).

[Gal00b] William F. Galway, *Dissecting a sieve to cut its need for space*, Algorithmic Number Theory (ANTS-IV) (Wieb Bosma, ed.), Lecture Notes in Computer Science, vol. 1838, Springer, Berlin, July 2000, pp. 297–312.

[GP] Andrew Granville and Carl Pomerance, *Two contradictory conjectures concerning Carmichael numbers*, (to appear).

[HL23] G. H. Hardy and J. E. Littlewood, *Some problems of 'partitio numerorum'; III: on the expression of a number as a sum of primes*, Acta Mathematica **44** (1923), 1–70.

[LO84] J. C. Lagarias and A. M. Odlyzko, *New algorithms for computing* $\pi(x)$, Number theory (New York, 1982), Lecture Notes in Mathematics, vol. 1052, Springer-Verlag, New York, 1984, pp. 176–193, MR **85j**:11182.

[LO87] J. C. Lagarias and A. M. Odlyzko, *Computing* $\pi(x)$*: an analytic method*, Journal of Algorithms **8** (1987), no. 2, 173–191, MR **88k**:11095.

[Ono98] Ken Ono, *A note on a question of J. Nekovář and the Birch and Swinnerton-Dyer conjecture*, Proc. Amer. Math. Soc. **126** (1998), no. 10, 2849–2853, MR **99a**:11081.

[OS88] Andrew M. Odlyzko and A. Schönhage, *Fast algorithms for multiple evaluations of the Riemann zeta function*, Trans. Amer. Math. Soc. **309** (1988), no. 2, 797–809, MR **89j**:11083.

[OS97] Ken Ono and K. Soundararajan, *Ramanujan's ternary quadratic form*, Invent. Math. **130** (1997), no. 3, 415–454, MR **99b**:11036.

[Pin] R. G. E. Pinch, *Lists of pseudoprimes below* 10^{13}, `ftp://ftp.dpmms.cam.ac.uk/pub/rgep/PSP/`.

[Pin00] R. G. E. Pinch, *The pseudoprimes up to* 10^{13}, Algorithmic Number Theory (ANTS-IV) (Wieb Bosma, ed.), Lecture Notes in Computer Science, vol. 1838, Springer, Berlin, July 2000, pp. 459–473.

[Vor03] Georges Voronoï, *Sur un problème du calcul des fonctions asymptotiques*, J. Reine Angew. Math. **126** (1903), 241–282.

Dept. of Mathematics, U. of Illinois at Urbana-Champaign, 1409 West Green St., Urbana, IL 61801

E-mail address: `galway@math.uiuc.edu`

URL: `http://www.math.uiuc.edu/~galway`