Striding Towards a New Subexponential Factoring Algorithm

Francesco Sica

11 September 2009

Outline

Motivation

- Understanding factorisation and especially why the Number Field Sieve is the best current factoring approach.

Motivation

- Understanding factorisation and especially why the Number Field Sieve is the best current factoring approach.
- Understand why a more "natural " approach using the Riemann ζ function fails. Are we doomed to bang into a wall through an *analytic* approach?

Motivation

- Understanding factorisation and especially why the Number Field Sieve is the best current factoring approach.
- Understand why a more "natural " approach using the Riemann ζ function fails. Are we doomed to bang into a wall through an *analytic* approach?
- Get some money from RSA challenges (no current income).

Fermat's Idea

Suppose we can find x, y integers with $x^2 \equiv y^2 \pmod{N}$ and $x \not\equiv \pm y \pmod{N}$. Then $1 < \gcd(x - y, N) < N$ and this can be computed quickly, giving rise to a nontrivial factor of N.

Fermat's Idea

Suppose we can find x, y integers with $x^2 \equiv y^2 \pmod{N}$ and $x \not\equiv \pm y \pmod{N}$. Then $1 < \gcd(x - y, N) < N$ and this can be computed quickly, giving rise to a nontrivial factor of N.
To find x and y, the most successful technique uses smooth numbers (divisible by "small" primes only). It is due to Morrison & Brillhart.
This idea is at the heart of the most successful factoring methods (QS and NFS), except ECM.

Running Times

ECM, QS, NFS all have subexponential running times.

Running Times

ECM, QS, NFS all have subexponential running times.

- QS: $\exp\left((c_1 + o(1))(\log N)^{1/2}(\log\log N)^{1/2}\right)$
- ECM: $\exp\left((c_2 + o(1))(\log p)^{1/2}(\log\log p)^{1/2}\right)$, (where p is smallest prime dividing N)
- NFS: $\exp\left((c_3 + o(1))(\log N)^{1/3}(\log\log N)^{2/3}\right)$

Presentation of Current Work

We present an approach which is likely to yield

- a subexponential general purpose factoring algorithm.

Presentation of Current Work

We present an approach which is likely to yield

- a subexponential general purpose factoring algorithm.
- It does not use the Morrison-Brillhart paradigm.

Presentation of Current Work

We present an approach which is likely to yield

- a subexponential general purpose factoring algorithm.
- It does not use the Morrison-Brillhart paradigm.
- It is in my view more natural, as it relates quantities known for their intrinsic arithmetical significance.

Presentation of Current Work

We present an approach which is likely to yield

- a subexponential general purpose factoring algorithm.
- It does not use the Morrison-Brillhart paradigm.
- It is in my view more natural, as it relates quantities known for their intrinsic arithmetical significance.
- Translates an arithmetic problem into an analytic one.

Presentation of Current Work

We present an approach which is likely to yield

- a subexponential general purpose factoring algorithm.
- It does not use the Morrison-Brillhart paradigm.
- It is in my view more natural, as it relates quantities known for their intrinsic arithmetical significance.
- Translates an arithmetic problem into an analytic one.
- All running times are proven, no assumptions!

Presentation of Current Work

We present an approach which is likely to yield

- a subexponential general purpose factoring algorithm.
- It does not use the Morrison-Brillhart paradigm.
- It is in my view more natural, as it relates quantities known for their intrinsic arithmetical significance.
- Translates an arithmetic problem into an analytic one.
- All running times are proven, no assumptions!
- Much room for future improvements.

Approaching Multiplicative Functions

Let $\phi(n)$ be the Euler phi function. Suppose that N factors as $N = pq$, so that $\phi(N) = N - p - \frac{N}{p} + 1 = f(p)$.

Approaching Multiplicative Functions

Let $\phi(n)$ be the Euler phi function. Suppose that N factors as $N = pq$, so that $\phi(N) = N - p - \frac{N}{p} + 1 = f(p)$.
Then using Newton's method, an approximation to $\phi(N)$ will yield an approximation to p, which is enough to recover it.

Approaching Multiplicative Functions

Let $\phi(n)$ be the Euler phi function. Suppose that N factors as $N = pq$, so that $\phi(N) = N - p - \frac{N}{p} + 1 = f(p)$.
Then using Newton's method, an approximation to $\phi(N)$ will yield an approximation to p, which is enough to recover it.

How do we find a good approximation to $\phi(N)$?

First Attempt with Riemann

Riemann zeta function is

$$\zeta(s) = \sum_{n \geq 1} \frac{1}{n^s} \quad \Re s > 1$$

It can be continued to a meromorphic function in $\mathbb{C}$ with simple pole with residue 1 at $s = 1$. Also

$$\frac{\zeta(s-1)}{\zeta(s)} = \sum_{n \geq 1} \frac{\phi(n)}{n^s} \quad \Re s > 2$$

Isolating $\phi(N)$

Classical: Compute $\Phi(x) = \sum_{n<x} \phi(n)$ by

$$\Phi(x) = \frac{1}{2\pi i} \int_{3-i\infty}^{3+i\infty} \frac{\zeta(s-1)}{\zeta(s)} \frac{x^s}{s} \, ds$$

and move line of integration "to the left".

Isolating $\phi(N)$

Classical: Compute $\Phi(x) = \sum_{n<x} \phi(n)$ by

$$\Phi(x) = \frac{1}{2\pi i} \int_{3-i\infty}^{3+i\infty} \frac{\zeta(s-1)}{\zeta(s)} \frac{x^s}{s}\, ds$$

and move line of integration "to the left". Problem: we hit the Riemann zeros, spooky beings! Can we avoid them?

Second Attempt with Riemann

We now consider $\sigma(N) = N + 1 + p + q$. As before, a close approximation to $\sigma(N)$ will reveal p. Here

$$\zeta(s)\zeta(s-1) = \sum_{n \geq 1} \frac{\sigma(n)}{n^s} \quad \Re s > 2$$

and hence if $S(x) = \sum_{n<x} \sigma(n)$ we get

$$S(x) = \frac{1}{2\pi i} \int_{3-i\infty}^{3+i\infty} \zeta(s-1)\zeta(s) \frac{x^s}{s} \, ds$$

Second Attempt with Riemann

We now consider $\sigma(N) = N + 1 + p + q$. As before, a close approximation to $\sigma(N)$ will reveal p. Here

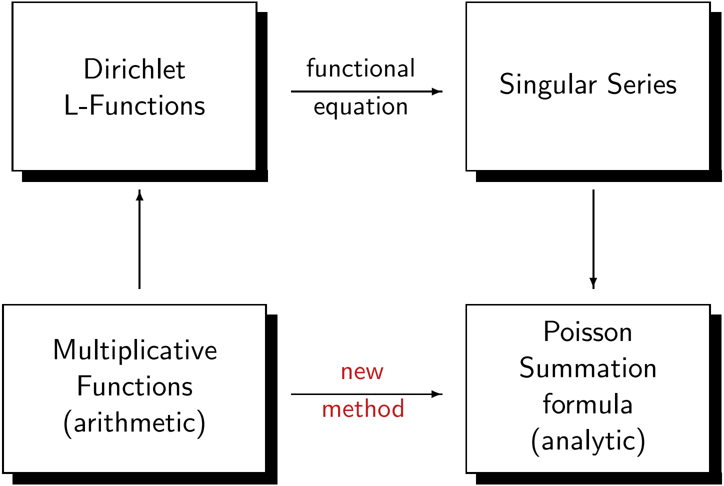
$$\zeta(s)\zeta(s-1) = \sum_{n \geq 1} \frac{\sigma(n)}{n^s} \quad \Re s > 2$$

and hence if $S(x) = \sum_{n<x} \sigma(n)$ we get

$$S(x) = \frac{1}{2\pi i} \int_{3-i\infty}^{3+i\infty} \zeta(s-1)\zeta(s)\, \frac{x^s}{s}\, ds$$

Problem: $|\zeta(s)| \approx |s|^{(1-\Re s)/2}$ as $|\Im s| \to \infty$ so cannot move the line of integration far enough to the left (to $\Re s \leq 0$)

The Mellin Transform Approach



Dirichlet L-functions

Let ℓ be prime and $\chi\colon (\mathbb{Z}/\ell)^* \to \mathbb{C}^*$ be a homomorphism, called Dirichlet character modulo ℓ. We define $\chi(n) = 0$ if ℓ divides $n \in \mathbb{Z}$. The Dirichlet L-function associated to χ is

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} \ , \quad \Re s > 1$$

Dirichlet L-functions

Let ℓ be prime and $\chi\colon (\mathbb{Z}/\ell)^* \to \mathbb{C}^*$ be a homomorphism, called Dirichlet character modulo ℓ. We define $\chi(n) = 0$ if ℓ divides $n \in \mathbb{Z}$. The Dirichlet L-function associated to χ is

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} \ , \quad \Re s > 1$$

It is an entire function if χ is not the trivial character, as we will henceforth suppose.

Dirichlet L-functions (cont'd)

We let $r \geq 2$ and $\beta_m \in \mathbb{C}$ "fixed". Define

$$\sigma_r(n) = \sum_{d_1 d_2 \cdots d_{r-1} | n} d_1^{\beta_1} d_2^{\beta_2} \cdots d_{r-1}^{\beta_{r-1}}$$

Then

$$L(s,\chi)L(s-\beta_1,\chi)\cdots L(s-\beta_{r-1},\chi) = \sum_{n=1}^{\infty} \frac{\sigma_r(n)\chi(n)}{n^s}$$

Mellin Transform

We let $\nu \in \mathbb{C}$ with $\Re\nu > 1$. Define

$$f_\nu(t) = \begin{cases} (1-t)^{\nu-1} & 0 \le t \le 1 \\ 0 & t \ge 1 \end{cases}$$

The Mellin transform of f_ν is

$$\frac{\Gamma(\nu)\Gamma(s)}{\Gamma(\nu+s)} = \int_0^\infty f_\nu(t) t^{s-1}\, dt$$

Inverse Mellin Transform

We have

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} L(s,\chi)L(s-\beta_1,\chi)\cdots L(s-\beta_{r-1},\chi)\frac{\Gamma(\nu)\Gamma(s)}{\Gamma(\nu+s)}\,x^s\,ds$$
$$=\sum_{n\leq x}\sigma_r(n)\chi(n)f_\nu\left(\frac{n}{x}\right)$$

Call the right-hand side

$$F(\nu)=\sum_{n\leq x}\sigma_r(n)\chi(n)\left(1-\frac{n}{x}\right)^{\nu-1}$$

Isolating p dividing N

We estimate

$$F^{(k)}(\nu) = \sum_{n \le x} \sigma_r(n)\chi(n)\left(1 - \frac{n}{x}\right)^{\nu-1} \log^k\left(1 - \frac{n}{x}\right)$$

If $x = N + \dfrac{1}{N^2}$ we get

Isolating p dividing N

We estimate

$$F^{(k)}(\nu) = \sum_{n \le x} \sigma_r(n)\chi(n)\left(1 - \frac{n}{x}\right)^{\nu-1} \log^k\left(1 - \frac{n}{x}\right)$$

If $x = N + \dfrac{1}{N^2}$ we get

$$F^{(k)}(\nu) = (-3\log N)^k \sigma_r(N)\chi(N)N^{-3(\nu-1)} + O(N^3(2\log N)^{k+r})$$

Isolating p dividing N

We estimate

$$F^{(k)}(\nu) = \sum_{n \le x} \sigma_r(n)\chi(n)\left(1 - \frac{n}{x}\right)^{\nu-1} \log^k\left(1 - \frac{n}{x}\right)$$

If $x = N + \frac{1}{N^2}$ we get

$$F^{(k)}(\nu) = (-3\log N)^k \sigma_r(N)\chi(N)N^{-3(\nu-1)} + O(N^3(2\log N)^{k+r})$$

Choosing $k > c_1 r \log N$ and supposing we can compute $F^{(k)}(\nu)$ with good precision we get a value for $\sigma_r(N)$ up to an error $O(N^{-c_2})$, where $c_2 \to \infty$ as $c_1 \to \infty$. If $N = pq$, then as before this is sufficient to obtain p.

The Functional Equation

The Dirichlet L-function of a primitive character χ of modulus $\ell > 1$ is an entire function satisfying the functional equation (given here in asymmetric form)

$$L(s,\chi) = \frac{1}{2\pi i}\left(\frac{2\pi}{\ell}\right)^s \tau(\chi)\Gamma(1-s)L(1-s,\bar{\chi}) \cdot \left(e^{i\pi s/2} - \chi(-1)e^{-i\pi s/2}\right)$$

where $\tau(\chi)$ is the Gauss sum

$$\tau(\chi) = \sum_{m=1}^{\ell} \chi(m)\exp(2\pi i m/\ell)$$

New Identities

Moving the line of integration to the left and using the functional equation shows

$$F(\nu) \approx R + \frac{\tau(\chi)^r \left(\frac{2\pi i}{\ell}\right)^{r-\beta_1-\cdots-\beta_{r-1}} \Gamma(\nu)}{(2\pi i)^{r+1}} x(\cos \pi\nu - \sin \pi\nu)$$

$$\times \int_{(1+1/r)} \left\{ \left(\frac{2\pi i}{\ell}\right)^r x \right\}^{-s} \Gamma(s-\nu)\Gamma(s+\beta_1)\cdots\Gamma(s+\beta_{r-1})$$

$$L(s,\bar\chi)L(s+\beta_1,\bar\chi)\cdots L(s+\beta_{r-1},\bar\chi)\, ds$$

where R is some residue, independent of x and ν. In view of the previous expression, it is appropriate to choose $\ell \approx x^{1/r}$ so that the integral does not depend on x (hence N).

The Singular Series

Using the multiplication theorem

$$\Gamma(s)\Gamma\left(s+\frac{1}{r}\right)\Gamma\left(s+\frac{2}{r}\right)\cdots\Gamma\left(s+\frac{r-1}{r}\right) = (2\pi)^{(r-1)/2} r^{1/2-rs}\Gamma(rs)$$

we arrive at evaluating terms for $F^{(k)}(\nu)$ which consist of derivatives of $\Gamma(\nu)(\cos\pi\nu - \sin\pi\nu)$ times the following series

$$\frac{1}{\tau(\chi)}\sum_{m=1}^{\ell}\chi(m)\sum_{d_1,\ldots,d_r\geq 1}\frac{d_1^{-\beta_1}d_2^{-\beta_2}\cdots d_{r-1}^{-\beta_{r-1}}\log^j(d_1\cdots d_r)}{(d_1\cdots d_r)^{\frac{1}{2r}-\frac{\beta_1+\cdots+\beta_{r-1}-\nu}{r}}} \cdot e^{2\pi i\left(\frac{m}{\ell}d_1\cdots d_r+(d_1\cdots d_r)^{1/r}\right)}$$

Computing the Singular Series

The singular series can be written as

$$S_j = \sum_{d_1,\ldots,d_r \geq 1} \frac{\log^j d_r}{d_1^{a_1} d_2^{a_2} \cdots d_r^{a_r}} e^{2\pi i\left(\beta d_1 \cdots d_r + \gamma (d_1 \cdots d_r)^{1/r}\right)}$$

Approximate the singular series as

$$\sum_{(d_1,\ldots,d_r)\in\mathbb{Z}^r} f(d_1,\ldots,d_r) = \sum_{(\delta_1,\ldots,\delta_r)\in\mathbb{Z}^r} \hat{f}(\delta_1,\ldots,\delta_r)$$

by Poisson summation. Here f is a $C_c^\infty(\mathbb{R}^r)$ function interpolating the summands of the singular series so that its Fourier transform $\hat{f}$ is decreasing super-polynomially. We therefore need to compute only $O(N^\epsilon)$ terms with precision $O(N^{-c})$, which should be possible in polynomial time.

Work in Progress

- Write up this step

Work in Progress

- Write up this step
- Ultimately, this shows that factoring could be done in $O(N^{1/r})$ for any r (subexponential). But in this last step with need to sample at least 2^r points. Therefore, we can only negotiate $r \approx \sqrt{\log N}$ and runtime is $O\left(e^{\sqrt{\log N}}\right)$ at best (naive estimate).

Work in Progress

- Write up this step
- Ultimately, this shows that factoring could be done in $O(N^{1/r})$ for any r (subexponential). But in this last step with need to sample at least 2^r points. Therefore, we can only negotiate $r \approx \sqrt{\log N}$ and runtime is $O\left(e^{\sqrt{\log N}}\right)$ at best (naive estimate).
- Need to implement it in practice.

Conclusion

- Completely new approach to factoring.

Conclusion

- Completely new approach to factoring.
- Advantage is that it transforms the arithmetic problem of factoring N into a purely analytic one (evaluation of the singular series, "independent" of N).

Conclusion

A New Factoring Algorithm

Francesco Sica

- Outline
- Introduction
- Heuristics
- Factoring with L-functions
- Conclusion

- Completely new approach to factoring.
- Advantage is that it transforms the arithmetic problem of factoring N into a purely analytic one (evaluation of the singular series, "independent" of N).
- Should lead to a deterministic factoring algorithm with proven running time $O(\exp(c_1\sqrt{\log N \log\log N}))$

Conclusion

- Completely new approach to factoring.
- Advantage is that it transforms the arithmetic problem of factoring N into a purely analytic one (evaluation of the singular series, "independent" of N).
- Should lead to a deterministic factoring algorithm with proven running time $O(\exp(c_1\sqrt{\log N \log\log N}))$
- Hoping to extend this to $O(\exp(c_3(\log N)^{1/3}(\log\log N)^{2/3}))$

Thank You!