Towards a New Subexponential Factoring Algorithm

Francesco Sica

University of Calgary

28 April 2010

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 so far. Are we doomed to bang into a wall through an *analytic* approach?

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

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

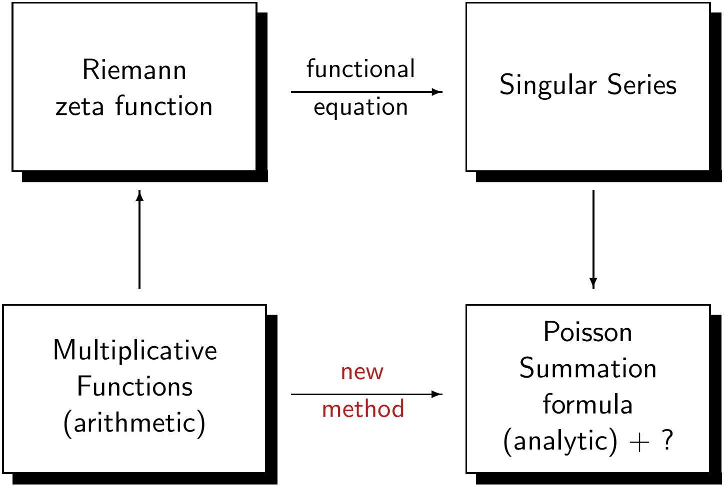
$$\zeta(s)\zeta(s-1) = \sum_{n \geq 1} \frac{\sigma(n)}{n^s} \quad \Re s > 2$$

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



The Multiplicative Function

We let $r \geq 2$ and $\beta_m \leq 0$ "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

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

The Test Function

We let $\nu \in \mathbb{R}$ with $\nu > 1$. 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

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

Call the right-hand side

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

Isolating p dividing N

We estimate

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

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

Isolating p dividing N

We estimate

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

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

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

Isolating p dividing N

We estimate

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

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

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

Choosing $k > c_1(\nu + r) \log N$ 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 Riemann zeta function is an meromorphic function with a single pole at 1 with residue 1 satisfying the functional equation (given here in asymmetric form)

$$\zeta(s) = \frac{(2\pi)^s}{\pi}\Gamma(1-s)\sin\left(\frac{\pi s}{2}\right)\zeta(1-s)$$

New Identities

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

$$F(\nu) \approx \rho + \frac{(2\pi i)^{r-\beta_1-\cdots-\beta_{r-1}}\,\Gamma(\nu)}{(2\pi i)^{r+1}}\, x(\cos \pi\nu - \sin \pi\nu)$$
$$\times \int_{(1+1/r)} \{(2\pi i)^r x\}^{-s}\,\Gamma(s-\nu)\Gamma(s+\beta_1)\cdots\Gamma(s+\beta_{r-1})$$
$$\zeta(s)\zeta(s+\beta_1)\cdots\zeta(s+\beta_{r-1})\,ds$$

where ρ is some easily expressible residue. In view of the previous expression, it is appropriate to choose r so that $x^{1/r} \approx e$ 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

$$\sum_{d_1,\ldots,d_r\geq 1} \frac{d_1^{-\beta_1} d_2^{-\beta_2}\cdots d_{r-1}^{-\beta_{r-1}} \log^j d_r}{(d_1\cdots d_r)^2}\cdot e^{2\pi i r (d_1\cdots d_r)^{1/r}}$$

with $j \leq k = O(\log^2 x)$. Here $r = [\log x]$.

Computing the Singular Series

How do we evaluate the singular series? Find an algorithm which computes it within e^{-t} in $O(t^c)$ binary operations for some fixed $c > 0$. Can allow dependence on r to be bad, for instance e^r. In general, need to evaluate

$$\sum_{d_1,\ldots,d_r \geq 1} \frac{\log^j d_r}{d_1^{c_1} d_2^{c_2} \cdots d_r^{c_r}} \cdot e^{iAd_1^{a_1} \cdots d_r^{a_r}}$$

with $A = O\big(\max\big(1/a_1, \ldots, 1/a_r\big)\big)$ and $a_1 + \cdots + a_r = 1$.

Poisson Summation Approach

Compute

$$\sum_{d_1,\ldots,d_r \geq 1} e^{-\frac{1}{d_1^R}} \cdots e^{-\frac{1}{d_r^R}} \frac{\log^j d_r}{d_1^{c_1} d_2^{c_2} \cdots d_r^{c_r}} \cdot e^{iAd_1^{a_1} \cdots d_r^{a_r}}$$

Poisson Summation Approach

Compute

$$\sum_{d_1,\ldots,d_r \geq 1} e^{-\frac{1}{d_1^R}} \cdots e^{-\frac{1}{d_r^R}} \frac{\log^j d_r}{d_1^{c_1} d_2^{c_2} \cdots d_r^{c_r}} \cdot e^{iAd_1^{a_1} \cdots d_r^{a_r}}$$
$$= \sum_{d_1,\ldots,d_r \geq 1} f(d_1, \ldots, d_r)$$

Poisson Summation Approach

Compute

$$
\begin{aligned}
&\sum_{d_1,\ldots,d_r\geq 1} e^{-\frac{1}{d_1^R}} \cdots e^{-\frac{1}{d_r^R}} \frac{\log^j d_r}{d_1^{c_1} d_2^{c_2} \cdots d_r^{c_r}} \cdot e^{iAd_1^{a_1}\cdots d_r^{a_r}} \\
&= \sum_{d_1,\ldots,d_r\geq 1} f(d_1,\ldots,d_r) \\
&= \sum_{x_1,\ldots,x_r\in\mathbb{Z}} \hat{f}(x_1,\ldots,x_r)
\end{aligned}
$$

Poisson Summation Approach

Compute

$$\sum_{d_1,\ldots,d_r\geq 1} e^{-\frac{1}{d_1^R}}\cdots e^{-\frac{1}{d_r^R}}\frac{\log^j d_r}{d_1^{c_1}d_2^{c_2}\cdots d_r^{c_r}}\cdot e^{iAd_1^{a_1}\cdots d_r^{a_r}}$$

$$= \sum_{d_1,\ldots,d_r\geq 1} f(d_1,\ldots,d_r)$$

$$= \sum_{x_1,\ldots,x_r\in\mathbb{Z}} \hat{f}(x_1,\ldots,x_r)$$

Estimate the rate of decrease of $\hat{f}$.

Poisson Summation Approach

Compute

$$\begin{aligned}\sum_{d_1,\dots,d_r\geq 1} e^{-\frac{1}{d_1^R}}\cdots e^{-\frac{1}{d_r^R}}\frac{\log^j d_r}{d_1^{c_1}d_2^{c_2}\cdots d_r^{c_r}}\cdot e^{iAd_1^{a_1}\cdots d_r^{a_r}} &= \sum_{d_1,\dots,d_r\geq 1} f(d_1,\dots,d_r)\\ &= \sum_{x_1,\dots,x_r\in\mathbb{Z}}\hat{f}(x_1,\dots,x_r)\end{aligned}$$

Estimate the rate of decrease of $\hat{f}$. Evaluate single coefficients.

Rate of Decrease of $\hat{f}$

$$\hat{f}(x_1,\ldots,x_r) = \int_{\mathbb{R}^r} e^{-\frac{1}{u_1^R}} \cdots e^{-\frac{1}{u_r^R}} \frac{\log^j u_r}{u_1^{c_1} u_2^{c_2} \cdots u_r^{c_r}} e^{iAu_1^{a_1}\cdots u_r^{a_r}} e^{-2\pi i(x_1u_1+\cdots+x_ru_r)}\, du_1\cdots du_r$$

We can modify contours ($u_m \to e^{\pm i\theta} u_m$ for $\theta > 0$) to get

$$\hat{f}(x_1,\ldots,x_r) = e^{\pm i\theta\star} \int_{\mathbb{R}^r} e^{-\frac{e^{\pm iR\theta}}{u_1^R}} \cdots e^{-\frac{e^{\pm iR\theta}}{u_r^R}} \frac{\log^j e^{\pm i\theta} u_r}{u_1^{c_1} u_2^{c_2} \cdots u_r^{c_r}} e^{ie^{\pm i\theta}Au_1^{a_1}\cdots u_r^{a_r}} e^{2\pi i e^{i\theta}(|x_1|u_1+\cdots+|x_r|u_r)}\, du_1\cdots du_r$$

Rate of Decrease of $\hat{f}$ (cont'd)

$$\hat{f}(x_1,\ldots,x_r) = e^{\pm i\theta\star} \int_{\mathbb{R}^r} e^{-\frac{e^{\pm iR\theta}}{u_1^R}} \cdots e^{-\frac{e^{\pm iR\theta}}{u_r^R}} \frac{\log^j e^{\pm i\theta} u_r}{u_1^{c_1} u_2^{c_2} \cdots u_r^{c_r}} e^{ie^{\pm i\theta} A u_1^{a_1} \cdots u_r^{a_r}} e^{2\pi i e^{i\theta}(|x_1|u_1+\cdots+|x_r|u_r)} \, du_1 \cdots du_r$$

therefore

$$\left|\hat{f}(x_1,\ldots,x_r)\right| \ll \int_{\mathbb{R}^r} e^{-\frac{\cos(R\theta)}{u_1^R}} \cdots e^{-\frac{\cos(R\theta)}{u_r^R}} \frac{\log^j u_r}{u_1^{c_1} u_2^{c_2} \cdots u_r^{c_r}} e^{\pm A(\sin\theta) u_1^{a_1} \cdots u_r^{a_r}} e^{-2\pi(\sin\theta)(|x_1|u_1+\cdots+|x_r|u_r)} \, du_1 \cdots du_r$$

Rate of Decrease of $\hat{f}$ (end)

A New Factoring Algorithm

Francesco Sica

Outline

Introduction

Heuristics

Factoring with the Riemann zeta function

Conclusion

The cool case:

$$\begin{aligned} A(\sin\theta)u_1^{a_1}\cdots u_r^{a_r} &\le A(\sin\theta)(a_1u_1+\cdots+a_ru_r) \\ &\le (2\pi-\delta)(\sin\theta)(u_1+\cdots+u_r) \end{aligned}$$

i.e. if $A \le (2\pi-\delta)\min(1/a_1,\ldots,1/a_r)$ for some $\delta > 0$.

Rate of Decrease of $\hat{f}$ (end)

The cool case:

$$\begin{aligned} A(\sin\theta)u_1^{a_1}\cdots u_r^{a_r} &\le A(\sin\theta)(a_1u_1+\cdots+a_ru_r) \\ &\le (2\pi-\delta)(\sin\theta)(u_1+\cdots+u_r) \end{aligned}$$

i.e. if $A \le (2\pi-\delta)\min(1/a_1,\ldots,1/a_r)$ for some $\delta > 0$.
Problem: In our case

$$A = (2\pi + O(\log r/r))\min(1/a_1,\ldots,1/a_r)$$

Final Considerations

- Poisson Summation or purely analytic approach seems difficult to carry out for singular series.

Final Considerations

- Poisson Summation or purely analytic approach seems difficult to carry out for singular series.
- Hope one can find a method to treat the seemingly innocuous but revealing

$$\sum_{n,m\geq 1} \frac{\cos(An^a m^{1-a})}{n^2 m^2}$$

within e^{-t^3} when $a = 1/t$ and $A = 3\pi$. In this case the best algorithm I found runs in $O(C^t)$ whereas if $A = \pi$ this is essentially $O(t^2)$.

Final Considerations

- Poisson Summation or purely analytic approach seems difficult to carry out for singular series.
- Hope one can find a method to treat the seemingly innocuous but revealing

$$\sum_{n,m\geq 1} \frac{\cos(An^a m^{1-a})}{n^2 m^2}$$

within e^{-t^3} when $a = 1/t$ and $A = 3\pi$. In this case the best algorithm I found runs in $O(C^t)$ whereas if $A = \pi$ this is essentially $O(t^2)$.

- Ultimately, this would show that factoring could be done in $O(e^r)$ for $r = \log^c N$ (subexponential).

Final Considerations

- Poisson Summation or purely analytic approach seems difficult to carry out for singular series.
- Hope one can find a method to treat the seemingly innocuous but revealing

$$\sum_{n,m\geq 1} \frac{\cos(An^a m^{1-a})}{n^2 m^2}$$

within e^{-t^3} when $a = 1/t$ and $A = 3\pi$. In this case the best algorithm I found runs in $O(C^t)$ whereas if $A = \pi$ this is essentially $O(t^2)$.

- Ultimately, this would show that factoring could be done in $O(e^r)$ for $r = \log^c N$ (subexponential).
- Need to perform extensive numerical calculations.

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 subexponential factoring algorithm with proven running time faster than what is currently known.
- Preprint available on ArXiv.

THANK YOU! ☺