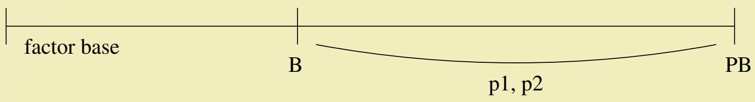
Theoretical Analysis of Relations in PPMPQS

Universiteit Leiden

Willemien Ekkelkamp
W.H.Ekkelkamp@cwi.nl | www.math.leidenuniv.nl/~ekkelkam

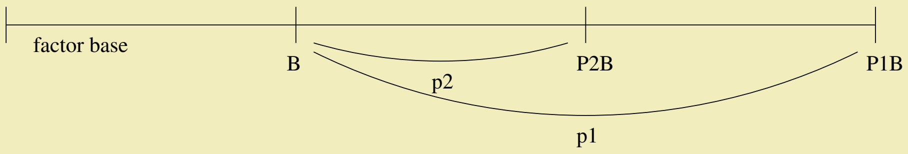
Basic PPMPQS

The multiple polynomial quadratic sieve (MPQS) is one of the algorithms used for factoring large numbers. One of its variations makes use of relations with two large primes (PPMPQS). These two large primes p_1 and p_2 have the same upper bound PB. Their lower bound is B, which is the bound of the factor base.



Variation

Instead of using the same upper bound for the two large primes, A.K. Lenstra and M.S. Manasse[a] mentioned the idea to choose two different bounds for these primes, and this variation has been implemented in the computer algebra package Magma. In a picture these relations can be seen as follows, with $B < p_2 < P2B$, $B < p_1 < P1B$, and $p_2 < p_1$.



Analysis of relations

In order to improve our understanding of this variation we made a theoretical analysis of the densities of the different types of relations occurring in PPMPQS with different bounds for the large primes. For complete, partial and partial-partial relations with the same bound for the two large primes this has already been done by R.J. Lambert[b].

To give a theoretical estimate for the number of partial-partial relations with different bounds for the two large primes, we derived the following generalization of a result of Lambert. Here $\Psi(x, y_1, y_2, y_3)$ denotes the number of positive integers $\leq x$ with greatest prime factor $\leq y_1$, one but greatest prime factor $\leq y_2$ and all other prime factors $\leq y_3$, and ρ is the Dickman ρ function.

Theorem *For* $0 < \alpha < \omega < \beta < 1/2$ *the limit* $\lim_{x\to\infty} \Psi(x, x^\beta, x^\omega, x^\alpha)/x$ *exists and equals*

$$\frac{1}{2}\int_\alpha^\omega \int_\alpha^\omega \rho\left(\frac{1-\lambda_1-\lambda_2}{\alpha}\right)\frac{d\lambda_1}{\lambda_1}\frac{d\lambda_2}{\lambda_2} +$$

$$\int_\alpha^\omega \int_\omega^\beta \rho\left(\frac{1-\lambda_1-\lambda_2}{\alpha}\right)\frac{d\lambda_1}{\lambda_1}\frac{d\lambda_2}{\lambda_2} \; .$$

Footnotes

[a]Factoring with Two Large Primes, Math.Comp. 63 (1994) 785-798

[b]Computational aspects of discrete logarithms, Ph.D. thesis, University of Waterloo (1996)

Experiments

We want to compare our theoretical analysis with practice, and the basic version of PPMPQS with the mentioned variation. To that end we have run PPMPQS for four different numbers and counted the number of partial-partial relations and compared them with the theoretical counts derived from our theorem. For each number, the sieving threshold was kept the same by keeping $P1B * P2B$ constant.

Results

Counts of partial-partial relations (pp) after sieving with 100 096 polynomials, both experimentally and theoretically.

#	digits	$y_3 = B$	$y_2 = P2B$	$y_1 = P1B$
1	80	249 797	24 979 712	24 979 712
2	80	249 797	12 489 854	49 959 435
3	91	300 007	30 000 690	30 000 690
4	91	300 007	15 000 354	60 001 434
5	101	482 231	48 223 067	48 223 067
6	101	482 231	24 111 564	96 446 162
7	110	747 853	149 570 695	149 570 695
8	110	747 853	74 785 318	299 141 028

#	pp (exp.)	ratio	pp (th.)	ratio	difference
1	93 556		85 607		9.29%
2	123 775	1.32	113 774	1.33	8.79%
3	6756		7200		−6.17%
4	8973	1.33	9618	1.34	−6.71%
5	1391		1477		−5.82%
6	1863	1.34	1985	1.34	−6.15%
7	557		544		2.39%
8	737	1.32	721	1.33	2.22%

Notice that the number of partial-partial relations for the case $y_2 < y_1$ is about one third larger than for $y_2 = y_1$. This will lead to more cycles and to a shorter running time of PPMPQS, since the CPU-times to generate them are comparable.

Conclusion

Our experiments show good agreement with the theoretical analysis. The study shows that it is advantageous to choose different upper bounds for the large primes. More experiments are necessary to find the optimal choice of y_1 and y_2.