Linear Algebra and its Applications 485 (2015) 72–102

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Montgomery's method of polynomial selection for the number field sieve

Nicholas Coxon

LORIA, Campus Scientifique, BP 239, 54506 Vandœuvre-lès-Nancy Cedex, France

ARTICLE INFO

Article history:
Received 18 December 2014
Accepted 21 July 2015
Available online 30 July 2015
Submitted by V. Mehrmann

MSC:
11Y05
11Y16

Keywords:
Integer factorisation
Number field sieve
Polynomial selection
Montgomery's method

ABSTRACT

The number field sieve is the most efficient known algorithm for factoring large integers that are free of small prime factors. For the polynomial selection stage of the algorithm, Montgomery proposed a method of generating polynomials which relies on the construction of small modular geometric progressions. Montgomery's method is analysed in this paper and the existence of suitable geometric progressions is considered.

1. Introduction

In this paper, N denotes a positive integer that is destined to be factored. When N is large and free of small factors, the most efficient publicly known algorithm for determining its factors is the number field sieve [20]. Such N include RSA [28] moduli, for which numerous record factorisations have been achieved with the number field sieve, including the current 768-bit record [17].

E-mail address: nicholas.coxon@inria.fr.

http://dx.doi.org/10.1016/j.laa.2015.07.025

The number field sieve is comprised of several stages, commonly referred to as polynomial selection, sieving, filtering, linear algebra and square root computation. The polynomial selection stage requires the selection of coprime irreducible polynomials $f_1, f_2 \in \mathbb{Z}[x]$ that have a common root modulo N. After polynomial selection, sieving is used to identify coprime integer pairs (a, b) such that the prime factors of $f_i(a/b)b^{\deg f_i}$ are below some bound y_i for $i = 1, 2$. Obtaining sufficiently many pairs with this property, called *relations*, is the most time consuming stage of the number field sieve, with the time taken greatly influenced by the choice of polynomials [24,25].

Let $\Psi(x, y)$ denote the number of positive integers less than x that are free of prime factors greater than y. Canfield, Erdős and Pomerance [6] showed that for any $\varepsilon > 0$, $\Psi(x, x^{1/u}) = xu^{-u(1+o(1))}$ for $u \to \infty$, uniformly in the region $x \geq u^{u(1+\varepsilon)}$. It follows, heuristically, that in the polynomial selection stage of the number field sieve, the polynomials f_1 and f_2 should be chosen to minimise the size of the values $f_1(a/b)b^{\deg f_1}$ and $f_2(a/b)b^{\deg f_2}$ over the pairs (a, b) considered in the sieve stage. Thus, it is necessary for the polynomials to have small coefficients. As a result, the degrees of f_1 and f_2 should not be too small. However, the degrees should not be too large either, since $f_i(a/b)b^{\deg f_i}$ is a homogeneous polynomial of degree $\deg f_i$ in a and b. In practice, low-degree polynomials are used. For example, the two largest factorisations of RSA moduli [17,4] both used a sextic polynomial together with a linear polynomial. To quantify the coefficient size of a polynomial, the skewed 2-norm $\|.\|_{2,s}$ is used. The norm is defined as follows: if $f = \sum_{i=0}^{d} a_i x^i$ is a degree d polynomial with real coefficients, then

$$\|f\|_{2,s} = \sqrt{\sum_{i=0}^{d} \left(a_i s^{i-d/2}\right)^2} \quad \text{for all } s > 0.$$

The parameter s captures the shape of the sieve region, which is modelled by a rectangular region $[-A, A] \times (0, B]$ or an elliptic region

$$\left\{(x, y) \in \mathbb{R}^2 \mid 0 < y \leq B\sqrt{1 - (x/A)^2}\right\}$$

such that $A/B = s$. In practice, the polynomial selection stage proceeds by first generating many “raw” polynomial pairs with small coefficients. Then various methods of optimisation [25,3,2] are used to improve the quality of the raw pairs by taking into account additional factors that influence a pair’s yield of relations, such as the presence of real roots and roots modulo small primes [24,25].

The methods of polynomial selection used in all recent record factorisations [24,25,15,16] produce polynomials f_1 and f_2 such that one polynomial is linear. However, it is expected that a significant advantage is gained by using two nonlinear polynomials [9, Section 6.2.7] (see also [27, Section 4] for practical considerations relating to sieving). Montgomery [22,23] provided a method for generating two nonlinear polynomials with small coefficients. This paper extends and sharpens Montgomery’s original analysis of the method.