Chebyshev's method for number fields

José Felipe Voloch

Chebyshev [C] proved that the number of primes up to x was between multiples of $x/\log x$. In the simplified form of Chebyshev's method worked out by Erdös this can be obtained by looking at the prime factorization of the binomial coefficients $\binom{2n}{n}$. Note that $\binom{2n}{n} = (-4)^n\binom{-1/2}{n}$ is the n-th coefficient of the Taylor expansion of $(1-4x)^{-1/2}$. In the course of their work on Grothendieck's conjecture on differential equations, by considering Padé approximations to $(1+x)^{i\alpha}, i = 1, 2, \ldots$ where α is an irrational algebraic integer, D. and G. Chudnovsky proved that there are infinitely many primes which do not split in $\mathbf{Q}(\alpha)$ (which is of course a special case of Chebotarev's density theorem). The proof requires estimating complicated expressions in various binomial coefficients $\binom{i\alpha+j}{n}$. In this paper we show that, at least for $\mathbf{Q}(\alpha)/\mathbf{Q}$ Galois, we can very easily obtain not only that there are infinitely many primes which split completely in $\mathbf{Q}(\alpha)$, but that there are at least a multiple of $x^{1/d}/\log x$ ($d = [\mathbf{Q}(\alpha) : \mathbf{Q}]$) such primes up to x, by simple estimates of $\binom{\alpha}{n}$. We will also study the prime factorization of $\binom{\alpha}{n}$ in the light of the present knowledge about the distribution of primes to analyse the scope of this method.

Lemma 1. *Let $\alpha \in \mathbf{C}, \alpha \notin \mathbf{N}$, then $\log|\binom{\alpha}{n}| = o(n)$.*

Proof: This is of course elementary and well-known. The function $(1+x)^\alpha$ is holomorphic on the unit disk and has a singularity at $x = 1$ so its Taylor expansion about zero has radius of convergence 1, hence the result.

J. Vaaler pointed out that the $o(n)$ can be improved to $O(\log n)$ but we won't need this sharper result.

For now on let α be an irrational algebraic integer such that $\mathbf{Q}(\alpha)/\mathbf{Q}$ is Galois and denote by $\alpha = \alpha_1, \ldots, \alpha_d$ the conjugates of α, so $d = [\mathbf{Q}(\alpha) : \mathbf{Q}]$. Put $f(x) = \prod(x - \alpha_i)$, the minimal polynomial of α over $\mathbf{Q}$. Under our assumptions, $f(x) \in \mathbf{Z}[x]$. Let A_n be the absolute norm of $\binom{\alpha}{n}$, which is a non-zero rational number. We denote by v_p the p-adic valuation associated to the prime p. We will often use that, since $\mathbf{Q}(\alpha)/\mathbf{Q}$ is Galois, for a

prime p of $\mathbf{Q}$, p splits completely in $\mathbf{Q}(\alpha)$ if and only if there is a prime of $\mathbf{Q}(\alpha)$ above p which is split over $\mathbf{Q}$.

Lemma 2. *Assume that p does not ramify in $\mathbf{Q}(\alpha)$. Then $v_p(A_n) \geq 0$, if p splits completely in $\mathbf{Q}(\alpha)$ and $v_p(A_n) \leq 0$ otherwise.*

Proof: If $x \in \mathbf{Z}_p$ then $\binom{x}{n} \in \mathbf{Z}_p$, so $\binom{\alpha}{n} \in \mathbf{Z}_p$ if p splits, giving the first statement of the lemma. For the second statement we note that there is no root of the minimal polynomial of α in $\mathbf{Z}/p\mathbf{Z}$ in that case for that would force p to split, so p does not divide the numerator of A_n.

Lemma 3. *Assume p does not split in $\mathbf{Q}(\alpha)$. Then $v_p(A_n) \leq -cn + O(\log n)$, for some $c > 0$ depending on p.*

Proof: If v_p denotes an extension of the p-adic valuation to $\mathbf{Z}_p[\alpha]$ then, since p does not split, $\mathbf{Z}_p[\alpha]$ is strictly bigger than $\mathbf{Z}_p$ so there exists an integer $s \geq 0$ such that $v_p(\alpha_j - k) \leq s, j = 1, \ldots, d, k \in \mathbf{Z}$. Given $j, 1 \leq j \leq d$, there are at most $n/p + O(1)$ integers $k, 0 \leq k < n$ with $v_p(\alpha_j - k) > 0$. At most $n/p^2 + O(1)$ of those satisfy $v_p(\alpha_j - k) > 1$ and so on, until at most $n/p^s + O(1)$ of those satisfy $v_p(\alpha_j - k) > s - 1$ but none satisfy $v_p(\alpha_j - k) > s$. As

$$A_n = \frac{\prod_{k=0}^{n-1} \prod_{j=1}^{d} (\alpha_j - k)}{n!^d}$$

we get $v_p(A_n) \leq d \sum_{i=1}^{s} n/p^i + O(1) - dv_p(n!) = -dn \sum_{i>s} 1/p^i + O(\log n)$, as was to be proved.

Lemma 4. *If p splits in $\mathbf{Q}(\alpha)$ then $v_p(A_n) \ll \log n / \log p$.*

Proof: Consider A_n, α as p-adic integers. We have that

$$A_n = \frac{\prod_{k=0}^{n-1} \prod_{j=1}^{d} (\alpha_j - k)}{n!^d}.$$

Given $j, 1 \leq j \leq d$, there are $n/p + O(1)$ integers $k, 0 \leq k < n$ with $k \equiv \alpha_j (\mathrm{mod}\, p)$ and $n/p^2 + O(1)$ of those satisfy $k \equiv \alpha_j (\mathrm{mod}\, p^2)$ and so on. However, if $p^r | (\alpha_j - k)$ then $p^r \ll k^d \leq n^d$, so $r \ll \log n / \log p$. Therefore the p-adic valuation of the numerator of the

above expression for A_n is at most $dn/(p-1) + O(\log n/\log p)$. But the last expression is also a lower bound for the p-adic valuation of the denominator of the above expression for A_n, namely $n!^d$. The lemma follows.

Theorem 1. *If S is the set of primes splitting in $\mathbf{Q}(\alpha)$ then $\#\{p \leq x \mid p \in S\} \gg x^{1/d}/\log x$.*

Proof: For x sufficiently large, let y be the unique positive solution to $f(y) = x$ and put $n = [y]$. By lemma 1, $\log|A_n| = o(n)$. On the other hand, $\log|A_n| = \sum_{p\in S} v_p(A_n)\log p + \sum_{p\notin S} v_p(A_n)\log p$. Clearly, if $v_p(A_n) > 0$ with $p \in S$, then p divides $f(k)$ for some $k, 0 \leq k < n$, so $p \leq f(n) \leq x$ and by lemma 4, $v_p(A_n)\log p \ll \log x$, so

$$\sum_{p\in S} v_p(A_n)\log p \ll \#\{p \leq x \mid p \in S\}\log x.$$

By lemma 2, for all but finitely many primes p not in S, $v_p(A_n) \leq 0$ and for any $p \notin S$, $v_p(A_n) \leq -cn$ eventually. Therefore, provided there exists at least one prime $p \notin S$, $-\sum_{p\notin S} v_p(A_n)\log p \gg n \gg x^{1/d}$, and the result follows. If S is the set of all primes then the theorem follows from Chebyshev's original argument.

Let's stop pretending we don't know anything about the distribution of primes. The estimate $\#\{p \leq x \mid p \in S\} \sim x/d\log x$ is equivalent to the prime ideal theorem for $\mathbf{Q}(\alpha)$. We will now discuss the limitations of the above method. The contribution of a prime p that doesn't split or ramify in $\mathbf{Q}(\alpha)$ to the logarithm of the denominator of A_n is $d\log p v_p(n!) = d\log p\sum_{i=1}^{\infty}[n/p^i]$, so the logarithm of the denominator of A_n is

$$dn\sum_{p\leq n, p\notin S}\frac{\log p}{p-1} + O(n)$$

and this is the same as $(d-1)n\log n + O(n)$. Using lemmas 1 and 4 we get that, for any $c > 1$, $\sum_{p\in S, p\leq cn} v_p(A_n)\log p = O(n)$. Therefore large primes must be contributing to the numerator of A_n. Note that $v_p(A_n) > 0$ for $p > n$ if and only if there exists $a, 1 \leq a < n, f(a) \equiv 0 (\mathrm{mod}\, p)$. Thus, large primes p which contribute to the numerator of A_n are those for which there is a small solution to $f(x) \equiv 0(\mathrm{mod}\, p)$. In the case that

α is imaginary quadratic, Duke et al. [DFI] have shown that the values of a/p, where $0 \le a < p, f(a) \equiv 0 (\mathrm{mod}\, p)$, are uniformly distributed in $[0,1]$ as p varies. No such result is known for general α, but a weak statement about small values of a/p can be deduced from the above and, replacing α by 2α we get values of a/p near $1/2$ but is unclear how far this can be pushed.

Here is a numerical example. Let, as usual, $i^2 = -1$ and let A be the norm of $\binom{i}{50}$ ($A = A_{50}$ in the above notation). Its value is approximately $A = 0.001441\ldots$. The numerator of A is

8542180816279975513293380143686677865317970703903412860607361376623932520681407 77,

which factors as $13^2 17^3 29^3$ 37 $53^2$61 $73^2$89 97 101 109 113 137 149 157 181 197 257 313 353 401 421 461 577 613 677 761 1013 1201 1297 1601 and we can see that indeed primes much larger than 50 occur. The denominator of A is

592784436219772736387514437405254968475156860253773651024470394041201404156839985152,

which factors as $2^{69} 3^{44} 7^{16} 11^8 19^4 23^4 31^2 43^2 47^2$.

Acknowledgements: The author would like to thank J. Vaaler and F. Rodríguez Villegas for helpful conversations. We also acknowledge the use of the software PARI for the numerical calculations. The author would also like to thank the TARP (grant #ARP-006) and the the NSA (grant MDA904-97-1-0037) for financial support.

References.

[C] P. L. Chebyshev, *Memoire sur les nombres premiers*, J. Math. pures et appl. **17**(1852) 366-390.

[CC] D. Chudnovsky and G. Chudnovsky, *Applications of Padé approximations to the Grothendieck conjecture on linear differential equations*,in Number theory (New York, 1983–84), 52–100, Lecture Notes in Math., 1135, Springer, Berlin-New York, 1985.

[DFI] W. Duke, J. B. Friedlander, H. Iwaniec, *Equidistribution of roots of a quadratic congruence to prime moduli*, Ann. of Math. **141**(1995), 423–441.

Dept. of Mathematics, Univ. of Texas, Austin, TX 78712, USA

e-mail: voloch@math.utexas.edu