ORDERS OF SIMPLE GROUPS AND THE BATEMAN–HORN CONJECTURE

GARETH A. JONES AND ALEXANDER K. ZVONKIN

Abstract. We use the Bateman–Horn Conjecture from number theory to give strong evidence of a positive answer to Peter Neumann's question, whether there are infinitely many simple groups of order a product of six primes. (Those with fewer than six were classified by Burnside, Frobenius and Hölder in the 1890s.) The groups satisfying this condition are $\mathrm{PSL}_2(8)$, $\mathrm{PSL}_2(9)$ and $\mathrm{PSL}_2(p)$ for primes p such that p^2-1 is a product of six primes. The conjecture suggests that there are infinitely many such primes p, by providing heuristic estimates for their distribution which agree closely with evidence from computer searches. We also briefly discuss the applications of this conjecture to other problems in group theory, such as the classifications of permutation groups and of linear groups of prime degree, the structure of the power graph of a finite simple group, and the construction of highly symmetric block designs.

1. Introduction

In the 1890s, Burnside [5, 6], Frobenius [11] and Hölder [13] classified those non-abelian finite simple groups G which have a number-theoretically small order, in the sense that $|G|$ is a product of relatively few primes, counting repetitions. Between them, they showed that there are only four such groups with at most five primes, namely the groups $\mathrm{PSL}_2(p)$ for $p = 5$, 7, 11 and 13, of orders

$$2^2 \cdot 3 \cdot 5, \quad 2^3 \cdot 3 \cdot 7, \quad 2^2 \cdot 3 \cdot 5 \cdot 11 \quad \text{and} \quad 2^2 \cdot 3 \cdot 5 \cdot 13.$$

(Recall that $\mathrm{PSL}_2(q)$ (or $L_2(q)$ in ATLAS [8] notation) has order $q(q^2-1)/d$ where $d = \gcd(q-1, 2)$, and that $\mathrm{PSL}_2(4) \cong \mathrm{PSL}_2(5) \cong \mathrm{A}_5$.)

Since the non-abelian finite simple groups are now classified, it is natural to ask which of them have order a product of six primes. It is straightforward to inspect the orders of the simple groups in such sources as [8] or [25], and to see that the only possibilities are the groups $\mathrm{PSL}_2(8)$ and $\mathrm{PSL}_2(9)$ ($\cong \mathrm{A}_6$) of orders

$$2^3 \cdot 3^2 \cdot 7 \quad \text{and} \quad 2^3 \cdot 3^2 \cdot 5,$$

and $\mathrm{PSL}_2(p)$, of order $p(p^2-1)/2$, for primes $p > 13$ such that p^2-1 is a product of six primes. For instance, see Example 6.3 in Section 6 for the elimination of small unitary groups.

Solomon, in his historical survey [24] of the classification of finite simple groups, notes that Peter Neumann, while editing Burnside's collected works, asked whether there are infinitely many primes p with this property. We do not have a definite answer to this question, but we have used the Bateman–Horn Conjecture to give what we feel is strong evidence that the answer is positive.

2020 *Mathematics Subject Classification.* Primary 20D05; secondary 11N32, 20B04, 20B10, 20D60, 20G40.

Key words and phrases. Finite simple group, group order, prime factor, prime degree, Bateman-Horn Conjecture.

To avoid confusion, we point out that the problem we consider here should be distinguished from the investigation of Kn groups, those simple groups of order divisible by n *distinct* primes. Nevertheless, we describe an application of our work to them in Section 9.

2. Factorisations

For a natural number $n \in \mathbb{N}$ with prime power factorisation

$$n = \prod_{i=1}^{k} p_i^{e_i} \quad (p_i \text{ prime},\ e_i \geq 1),$$

let

$$\Omega(n) := \sum_{i=1}^{k} e_i$$

denote the total number of prime factors of n, counting repetitions, and for a finite group G let

$$\Omega(G) := \Omega(|G|).$$

For each $m \in \mathbb{N}$ let $\mathcal{S}_m$ be the set of all non-abelian simple groups G such that $\Omega(G) = m$. As explained in the Introduction, $\mathcal{S}_6$ consists of the groups $\mathrm{PSL}_2(8)$ and $\mathrm{PSL}_2(9)$, together with the groups $\mathrm{PSL}_2(p)$ for those primes $p > 13$ such that $\Omega(p^2 - 1) = 6$.

For any integer p we have

$$\Omega(p^2 - 1) = \Omega(p - 1) + \Omega(p + 1).$$

Now assume that p is prime and $p > 13$. Then $p \pm 1$ are both even, one of them is divisible by 4, and one of them is divisible by 3. This gives us four prime factors contributing to $\Omega(p^2 - 1)$, and each of $p \pm 1$ must contribute at least one more prime factor since otherwise $p \leq 13$. Thus $\Omega(p^2 - 1) \geq 6$, and this lower bound is attained if and only if p satisfies one of the following conditions, depending on which of $p \pm 1$ is divisible by 3 or by 4:

a) $p - 1 = 4r$ and $p + 1 = 6s$,
b) $p - 1 = 6r$ and $p + 1 = 4s$,
c) $p - 1 = 2r$ and $p + 1 = 12s$,
d) $p - 1 = 12r$ and $p + 1 = 2s$,

where r and s are primes. All four of these cases are possible, for example with $p = 29$, 19, 23 and 37 respectively, but the question is whether any of them yields infinitely many primes p.

In case (a) we have $p \equiv 1 \bmod (4)$ and $p \equiv -1 \bmod (6)$, which are equivalent to $p \equiv 5 \bmod (12)$. If we put $p = 12t + 5$ where $t \in \mathbb{N}$ then $r = 3t + 1$ and $s = 2t + 1$. The question is now whether these three polynomials

$$f_i(t) = 12t + 5, \quad 3t + 1 \quad \text{and} \quad 2t + 1 \quad (i = 1, 2, 3)$$

can simultaneously take prime values for infinitely many $t \in \mathbb{N}$. The situation is similar in the other three cases, the only difference being that the triples of polynomials $f_i(t)$ are $12t + 7$, $2t + 1$ and $3t + 2$ in case (b), $12t - 1$, $6t - 1$ and t in case (c), and $12t + 1$, t and $6t + 1$ in case (d).

3. Number-theoretic conjectures

These four cases are instances of a much more general problem in number theory, namely whether a given set of polynomials $f_1(t), \ldots, f_k(t) \in \mathbb{Z}[t]$ can simultaneously take prime values for infinitely many $t \in \mathbb{N}$. Bunyakovsky [4] considered the case $k = 1$ in 1857. The following conditions are obviously necessary for a single polynomial $f(t) \in \mathbb{Z}[t]$ to take prime values for infinitely many $t \in \mathbb{N}$:

(1) it has a positive leading coefficient,
(2) it is irreducible in $\mathbb{Z}[t]$, and
(3) it is not identically zero modulo any prime.

Bunyakovsky conjectured that these conditions are also sufficient. For example, they are satisfied by the polynomial $t^2 + 1$, the subject of the Euler–Landau Conjecture on primes of this form. The Bunyakovsky Conjecture has been proved only in the case where $\deg f = 1$: this is simply a reformulation of Dirichlet's Theorem on primes in an arithmetic progression.

Our cases (a) to (d) are instances of Dickson's Conjecture, which is that polynomials $f_i(t)$ $(i = 1, \ldots, k)$ of degree $\deg f_i = 1$ simultaneously take prime values for infinitely many $t \in \mathbb{N}$ if and only if they all satisfy the first two Bunyakovsky conditions and their product satisfies the third. Particular cases include the twin primes and Sophie German primes conjectures, with $f_i(t) = t, t+2$ and $t, 2t + 1$ respectively. Again, this conjecture has been proved only in the case $k = 1$, whereas our six primes problem has $k = 3$.

In 1957 Schinzel's Hypothesis [23] generalised both the Bunyakovsky and Dickson Conjectures by removing the condition $\deg f_i = 1$ from the latter. In 1962 Bateman and Horn [3], extending earlier work of Hardy and Littlewood [12] on the twin primes and other related conjectures, proposed an asymptotic estimate $E(x)$ for the number $Q(x)$ of $t \in \mathbb{N}$ with $t \leq x$ such that $f_i(t)$ is prime for $i = 1, \ldots, k$. The Bateman–Horn Conjecture (BHC) asserts that

$$Q(x) \sim E(x) := C \int_a^x \frac{dt}{\prod_{i=}^k \ln f_i(t)} \quad \text{as} \quad x \to +\infty \tag{1}$$

where C, known as a *Hardy–Littlewood constant*, is given by

$$C = C(f_1, \ldots, f_k) := \prod_{r \text{ prime}} \left(1 - \frac{1}{r}\right)^{-k} \left(1 - \frac{\omega_f(r)}{r}\right), \tag{2}$$

with $\omega_f(r)$ denoting the number of roots of $f := f_1 \ldots f_k \bmod (r)$, and a in (1) chosen to avoid singularities of the integral, where some $f_i(t) = 1$. In the next section we will give a short heuristic argument to explain these formulae, but as yet there is no proof (again, apart from the quantified version of Dirichlet's Theorem, due to de la Vallée Poussin). If Schinzel's conditions are satisfied, the infinite product in (2) converges to a limit $C > 0$. (See [1] for a proof, and for an interesting account of the background to the BHC.) Now the definite integral in (1) diverges to $+\infty$ as $x \to +\infty$, so $E(x) \to +\infty$ and hence, if the BHC is true, $Q(x) \to +\infty$ also, proving that there are infinitely many $t \in \mathbb{N}$ such that each $f_i(t)$ is prime.

4. Heuristic argument for the BHC

According to the Prime Number Theorem, a good asymptotic estimate for the number $\pi(x)$ of primes $p \le x$ is obtained by integrating the probability of t being prime, to give

$$\pi(x) \sim \int_2^x \frac{dt}{\ln t} \quad \text{as} \quad x \to +\infty. \tag{3}$$

(This is significantly more accurate than the widely-used estimate $x/\ln x$ for $\pi(x)$. Thus, for example, for the value of $\pi(x)$ for $x = 10^{28}$, which may be found in [22], entry A006880, the relative error of the estimate $x/\ln x$ is $-1.576\,\%$, while the relative error of the estimate (3) is, approximately, $10^{-12}\,\%$.) This suggests that if $f(t)$ is a polynomial in $\mathbb{Z}[t]$ satisfying Bunyakovsky's conditions then the expected number of prime values of $f(t)$ for $t \le x$ might be estimated by

$$\int_a^x \frac{dt}{\ln f(t)} \quad \text{as} \quad x \to +\infty, \tag{4}$$

where a is chosen so that $f(t) > 1$ for $t \ge a$. For a finite set of polynomials $f_1, \dots, f_k$ satisfying Schinzel's conditions we can therefore make a first attempt at an estimate

$$\int_a^x \frac{dt}{\prod_i \ln f_i(t)} \tag{5}$$

for the expected number of $t \le x$ with all $f_i(t)$ prime, multiplying probabilities on the assumption that these polynomials behave independently of each other. However, they are not independent, and the Hardy–Littlewood constant C is a product, over all primes r, of correction factors

$$\left(1 - \frac{1}{r}\right)^{-k} \left(1 - \frac{\omega_f(r)}{r}\right)$$

which replace the probability $\left(1 - \frac{1}{r}\right)^k$ that k randomly and independently chosen elements of $\mathbb{Z}_r$ are all non-zero with the probability that the product $f(t) = \prod_i f_i(t)$ is non-zero mod (r), so that no $f_i(t)$ is divisible by r.

Remark 4.1. Note that the factor $C(f)$ modifies not only formula (5), but also (4). As to (3), we have $f(t) = t$, and it is easy to see that in this case the Hardy–Littlewood constant is $C(f) = 1$.

Of course, the above is only a very brief outline of more persuasive heuristic arguments which can be used to support the BHC. See, however, the following remark.

Remark 4.2. To the best of our knowledge, there is no completely adequate probabilistic model for the distribution of primes which would imply, even informally, the Bateman–Horn Conjecture (BHC). Maybe, we must simply change direction and construct a model which would be based on the BHC. At least, this latter conjecture has such strong supporting experimental data that a model based on it just has no other way than to be an adequate description of a statistical behaviour of primes. This task is not yet accomplished.

However, we must keep in mind that the *density*, taken as a probability measure on $\mathbb{N}$ or on $\mathbb{Z}$, is only finitely additive and not countable additive. Therefore, it is impossible just to take a theorem of Probability Theory and apply it directly to the distribution of primes. The only particular case

of the BHC which is proved up to now (the quantified version of Dirichlet's theorem due to de la Vallée Poussin, mentioned above) was proved by analytic methods. As to Probability Theory, it is used as a heuristic tool, as a source of conjectures to be proved by other, non-probabilistic methods.

In the absence of a proof, we have to rely on the fact that, time after time, the BHC produces estimates which agree remarkably closely with experimental data (see subsequent sections of this paper). Indeed, such evidence would regarded as convincing proof in other areas, such as Physics or Law.

5. Applying the BHC

We applied the BHC to the four triples of polynomials $f_i(t)$ in cases (a) to (d) of the six primes problem. In each case Maple can evaluate the definite integral in (1) almost instantly, using numerical integration. (In the early1960s, when systems like Maple were unavailable and such programming had to be done 'from scratch', Bateman and Horn simplified the integration by ignoring all coefficients of the polynomials f_i, replacing the definite integral in (1) with

$$\frac{1}{\prod_{i=1}^{k} \deg f_i} \int_2^x \frac{dt}{(\ln t)^k}.$$

The resulting estimates are asymptotically equivalent, but a little less accurate.

In case (a) we have

$$f(t) = (12t + 5)(3t + 1)(2t + 1).$$

This has three roots modulo each prime except 2 and 3, where it has one, so this gives us the terms in the infinite product (2). The product converges slowly, but by taking the product over all primes $r \leq 10^9$ one can get a good approximation to the limit, namely $C(f) = 5.71649719$. If we take $x = 10^9$, for example, then the actual number of $t \leq 10^9$ such that each $f_i(t)$ is prime is $Q_a(10^9) = 614\,423$, whereas the estimate is $E(x) = E_a(10^9) = 615\,580.70$, representing a relative error of +0.188 %. After the example $p = 29$ given above, arising when $t = 2$, the next primes in case (a) appear when $t = 14$, giving $p = 173$, and then $t = 26$, giving $p = 317$.

The other three cases all give the same Hardy–Littlewood constant C as in (a), leading to very similar BHC estimates. For instance, in case (b) we have $Q_b(10^9) = 615\,369$, while the estimate is $E_b(10^9) = 615\,580.614$, an error of +0.034%. In the other two cases $Q_c(10^9) = 616\,509$ and $Q_d(10^9) = 616\,289$, whereas the estimates are $E_c(10^9) = 616\,720.62$ and $E_d(10^9) = 616\,720.51$, giving errors of +0.034 % and +0.070 %.

The accuracy of these estimates is comparable to that obtained in many other applications of the BHC, such as to twin or Sophie Germain primes (see also subsequent sections of this paper). Based on this evidence, we make the following conjecture:

Conjecture 5.1. *There are infinitely many primes p satisfying the conditions in each of cases (a) to (d), and in particular there are infinitely many groups $\mathrm{PSL}_2(p)$ of order a product of six primes.*

6. More than six primes

One can easily modify the preceding arguments to obtain similar evidence for the existence of infinitely many groups $\mathrm{PSL}_2(p)$ in $\mathcal{S}_m$ for any $m \geq 6$. For instance, this can be done by replacing the conditions in case (a) with

$$p - 1 = 4r \quad \text{and} \quad p + 1 = 2as \quad (a := 3^{m-5}),$$

where p, r and s are primes. Then $p \equiv 1 \bmod (4)$ and $p \equiv -1 \bmod (2a)$, so $p \equiv 2a - 1 \bmod (4a)$. Writing $p = 4at + 2a - 1$ we have

$$r = \frac{p-1}{4} = at + \frac{a-1}{2} \quad \text{and} \quad s = \frac{p+1}{2a} = 2t + 1,$$

so we have an instance of the BHC with

$$f_i(t) = 4at + 2a - 1, \quad at + \frac{a-1}{2} \quad \text{and} \quad 2t + 1 \quad (i = 1, 2, 3). \tag{6}$$

Example 6.1. Let us take $m = 7$, so that $a = 3^{m-2} = 9$. The polynomials we need to consider are

$$f_1(t) = 36t + 17, \quad f_2(t) = 9t + 4 \quad \text{and} \quad f_3(t) = 2t + 1.$$

For $r = 2$ and 3 the product $(36t + 17)(9t + 4)(2t + 1)$ has a unique root modulo r, while for all the other prime r it has three roots. Therefore, the constant factor $C(f)$ remains the same as in Section 5. Then, taking, for example, $x = 10^9$, we find that the number of $t \leq x$ for which the values of all the three polynomials are prime is $Q(x) = 556\,373$, while the BHC estimate gives $E(x) = 556\,520.2$. The relative error of this estimate is 0.026 %.

We make the following conjecture:

Conjecture 6.2. *Given any $a = 3^e$ ($e \geq 1$), the three polynomials in* (6) *simultaneously take prime values for infinitely many $t \in \mathbb{N}$. In particular, given any $m \geq 6$ there are infinitely many groups* $\mathrm{PSL}_2(p)$ *of order a product of m primes.*

This raises the question of whether one can *prove* that there is an infinite set of primes p such that $\Omega(p^2 - 1)$ is bounded above, or more generally that there is an infinite set $\mathcal{S}$ of non-abelian finite simple groups G with $\Omega(G)$ bounded above. Any such set $\mathcal{S}$ must contain only finitely many alternating groups, and hence any groups $G \in \mathcal{S}$ of Lie type must have bounded Lie rank, for otherwise their Weyl groups would involve alternating groups of unbounded degree. Likewise, the field of definition $\mathbb{F}_q$ ($q = p^e$) of G must have bounded degree e over $\mathbb{F}_p$, for otherwise the order of the Sylow p-subgroups of G, which contain copies of the additive group of $\mathbb{F}_q$, would be an unbounded power of p. This suggests that the simplest way of constructing candidates for $\mathcal{S}$ is to use the groups $\mathrm{PSL}_2(p)$ as we have done here, with the Lie rank and the degree e both equal to 1.

Example 6.3. As another example where the rank and degree are both 1, consider the simple unitary groups $G = \mathrm{PSU}_3(p)$ ($= U_3(p)$ in ATLAS notation) of order $p^3(p^3 + 1)(p - 1)/d$, $p > 2$ prime, where $d = \gcd(3, p+1)$. Since $p^3 + 1 = (p+1)(p^2 - p + 1)$ it is not hard to see that $\Omega(G) \geq 9$, attained if and only if $\Omega(p^2 - 1) = 6$ (as in the case of $\mathrm{PSL}_2(p)$) and $p^2 - p + 1$ or $(p^2 - p + 1)/3$ is prime. (The latter case arises when $p \equiv -1 \bmod (3)$, so that $d = 3$.) When we considered $\mathrm{PSL}_2(p)$

in Sections 2 and 5 the prime p was represented by $f_1(t)$ in each of cases (a) to (d), so one can apply the BHC to $\mathrm{PSU}_3(p)$ by adding the polynomial $f_4(t) = f_1(t)^2 - f_1(t) + 1$ to the triples $f_i(t)$ $(i = 1, 2, 3)$ used earlier in cases (b) and (d). In cases (a) and (c) we have $f_1(t) \equiv -1 \bmod (3)$ so $f_1(t)^2 - f_1(t) + 1$ is divisible by 3 and we can take $f_4(t) = (f_1(t)^2 - f_1(t) + 1)/3$. The following example gives strong evidence that there are infinitely many groups $\mathrm{PSU}_3(p) \in \mathcal{S}_9$.

Example 6.4. Let us consider one particular example coming from case (c): we have

$$f_1(t) = 12t - 1, \quad f_2(t) = 6t - 1, \quad f_3(t) = t, \quad \text{and} \quad f_4(t) = \frac{f_1(t)^2 - f_1(t) + 1}{3} = 48t^2 - 12t + 1.$$

In order to compute the constant $C(f)$ we need to know the value $\omega_f(r)$ which is the number of roots of the product $f(t) = f_1 f_2 f_3 f_4$ in $\mathbb{Z}_r$, with r prime. For $r = 2$ and 3 this product has a single root. For $r > 3$ the factors f_1, f_2, f_3 provide us with three roots, while the quadratic polynomial f_4 may have two roots or no roots. Let us look at this case in more detail.

The discriminant of f_4 is $-48 = 4^2 \cdot (-3)$. Therefore, f_4 has two roots in $\mathbb{Z}_r$ if -3 is a quadratic residue modulo (r), and no roots if not. Recall the notation for the Legendre symbol: for p prime and $q \in \mathbb{Z}$

$$\left(\frac{q}{p}\right) = \begin{cases} 0 & \text{if } q \equiv 0 \bmod (p); \\ 1 & \text{if } q \text{ is a quadratic residue mod } (p); \\ -1 & \text{otherwise.} \end{cases}$$

(See [14, Chapter 7] for quadratic residues and the Legendre symbol.) This symbol is multiplicative: for all $q, q' \in \mathbb{Z}$ we have

$$\left(\frac{qq'}{p}\right) = \left(\frac{q}{p}\right)\left(\frac{q'}{p}\right), \quad \text{therefore} \quad \left(\frac{-3}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{3}{p}\right).$$

The value of $\left(\frac{-1}{p}\right)$ is known: it is 1 if $p = 4k + 1$, and -1 if $p = 4k - 1$. As to $\left(\frac{3}{p}\right)$, we may use Gauss' Law of Quadratic Reciprocity: for p, q prime we have $\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right)$ if one or both p and q are of the form $4k + 1$, and $\left(\frac{q}{p}\right) = -\left(\frac{p}{q}\right)$ if both p and q are of the form $4k - 1$. Collecting all these data, we find out that

$$\left(\frac{-3}{p}\right) = \left(\frac{p}{3}\right) = \begin{cases} 1 & \text{if } p \equiv 1 \bmod 3, \\ -1 & \text{if } p \equiv -1 \bmod 3. \end{cases}$$

We may now conclude: for a prime $r > 3$ the product $f(t) = f_1(t) f_2(t) f_3(t) f_4(t)$ has five roots in $\mathbb{Z}_r$ if $r \equiv 1 \bmod (3)$, and has three roots if $r \equiv -1 \bmod (3)$. This permits us to compute the constant $C(f)$: taking the product over the primes $r \leq 10^9$ we get $C(f) = 12.10128533$. The number of $t \leq 10^9$ such that all $f_i(t)$ are prime is $Q(10^9) = 30\,452$; the estimate of this number given by the BHC is 30 504.71; the relative error is 0.173 %.

7. Permutation groups of prime degree

Although we have concentrated in this paper on the six primes problem, our involvement with the BHC began with a different problem in group theory. This was motivated by our study of Klein's papers [19, 20] on equations (and hence coverings and permutation groups) of degree 7

and 11, and our desire to extend his results to all primes (see [15] for details). In doing this we encountered an important but rarely discussed gap in the classification of transitive permutation groups of prime degree, a problem going back two and a half centuries to Lagrange. Building on early work by Galois and Burnside, the classification of finite simple groups implies that these groups are as follows:

(a) subgroups of $\mathrm{AGL}_1(p)$ containing the translation group, for primes p;
(b) alternating and symmetric groups A_p and S_p, for primes $p \geq 5$;
(c) $\mathrm{PSL}_2(11)$ and Mathieu groups M_{11} and M_{23}, of degrees 11, 11 and 23;
(d) subgroups G of $\mathrm{P\Gamma L}_n(q)$ containing $\mathrm{PSL}_n(q)$, in those cases when the natural degree $d = (q^n - 1)/(q - 1)$ is prime.

In (c), $\mathrm{PSL}_2(11)$ has two actions of degree 11, on the cosets of two conjugacy classes of subgroups isomorphic to A_5; the Mathieu groups act on the points of the Steiner systems $S(4, 5, 11)$ and $S(4, 7, 23)$. In (d), G acts on the d points (and also the d hyperplanes if $n \geq 3$) of the projective geometry $\mathbb{P}^{n-1}(\mathbb{F}_q)$ of dimension $n - 1$ over $\mathbb{F}_q$.

It is an open problem whether the degree d in case (d) is prime for infinitely many pairs (n, q). Such *projective primes*, as we call them, include the Fermat primes, of the form $q + 1 = 2^{2^f} + 1$, for $n = 2$, and the Mersenne primes, of the form $2^n - 1$, for $q = 2$. Five Fermat primes are known, and it is conjectured that there are no more; at the time of writing 51 Mersenne primes are known, and it is conjectured that there are infinitely many of them. In investigating this problem (see [16] for details), having nothing new to contribute to the extensive research on those very difficult topics, we restricted our attention to the cases $n, q \geq 3$.

If we write $q = p^e$, we are asking whether p and

$$d := p^{(n-1)e} + p^{(n-2)e} + \cdots + p^e + 1 \tag{7}$$

can both be prime for infinitely many p. Clearly, this requires n to be prime, and a simple argument involving cyclotomic polynomials shows that e must be a power of n, possibly equal to $n^0 = 1$.

We concentrated on the simplest and most frequently-arising case $n = 3$, $e = 1$, applying the BHC to the polynomials

$$f_1(t) = t, \quad f_2(t) = t^2 + t + 1.$$

Typical results obtained were

- $E(10^{10}) = 1.579642126 \times 10^7$ and $Q(10^{10}) = 15\,801\,827$, an error of $-0.03420956\,\%$,
- $E(10^{11}) = 1.292974079 \times 10^8$ and $Q(10^{11}) = 129\,294\,308$, an error of $+0.00239757\,\%$.

For other pairs (n, e), such as $(5, 1)$ and $(3, 3)$, the results were good but much less convincing, since the primes increase so rapidly that relatively few of them were within our computing range. Nevertheless, on the basis of this evidence we make the following conjecture:

Conjecture 7.1. *For each prime $n \geq 3$ there are infinitely many prime powers q such that $\mathrm{PSL}_n(q)$ is a transitive permutation group of prime degree $d = (q^n - 1)/(q - 1)$.*

8. Linear groups of prime degree

In [9] Dixon and Zalesskii classified the irreducible finite subgroups $G \leq \mathrm{SL}_d(\mathbb{C})$ of prime degree d, in the case where the socle S of their image $\overline{G}$ in $\mathrm{PSL}_d(\mathbb{C})$ is non-abelian and acts primitively, that is, preserving no non-trivial direct sum decomposition of $\mathbb{C}^d$. (In this section 'degree' and 'primitive' always refer to the dimension and structure of the vector space on which G acts, and not to any permutation representation of G.) In this case S is simple and $\overline{G} \leq \mathrm{Aut}\, S$. Theorem 1.2 of [9] gives a finite list of families of simple groups S which can arise, with necessary and sufficient conditions on d and their parameters for such groups G to exist. It is unknown whether some of these families are finite or infinite. Here we give a brief outline of how we have used the BHC to give strong evidence that they are infinite (see [16] for full details).

A typical example has $S \cong \mathrm{PSU}_n(q)$ for $n \geq 3$, where q is a prime power p^e and the degree

$$d = \frac{q^n + 1}{q + 1} = q^{n-1} - q^{n-2} + \cdots - q + 1$$

is prime. The pair of polynomials

$$f_1(t) = t \quad \text{and} \quad f_2(t) = t^{(n-1)e} - t^{(n-2)e} + \cdots - t^e + 1 \tag{8}$$

satisfy the conditions of Schinzel's Hypothesis if n is prime and e is a power of n. This is the same as the condition in Section 7 for $(q^n-1)/(q-1)$ to be a projective prime, and indeed the substitution $t \mapsto -t$ shows that the Hardy–Littlewood constants $C(f)$ are the same in both cases: they are equal to 1.521730.

We may compare the results, for example, for $x = 10^{10}$. The definite integrals in (1) are very similar: in fact, they differ by 0.456 (and the estimates $E(x)$ thus differ by $C(f) \cdot 0.456 \approx 0.7$). Therefore, we may suppose that the actual numbers $Q(x)$ will also be close to each other. And, indeed, we have seen above that the value of $Q(x)$ for $f_1(t) = t$, $f_2(t) = t^2 + t + 1$ was 15 801 827, while for the polynomials $f_1(t) = t$, $f_2(t) = t^2 - t + 1$ we have 15 801 414 (the relative error in this case is $-0.0316\,\%$). It is interesting to note that, again for $x = 10^{10}$, the computation of the estimate $E(x)$ on a particular (modest) laptop we have used took 0.002 seconds, while the computation of the exact value of $Q(x)$ took 54 hours. The computation of $E(x)$ for $x = 10^{30}$ *with* 30 *digits of accuracy* takes 2.4 seconds, while the computation of $Q(x)$ for this value of x is far beyond our reach. Regrettably, we don't have a (conjectural) upper bound for the error term.

Another family appearing in [9] consists of groups $S \cong \mathrm{PSL}_2(q)$ where q and the degree $d = (q - 1)/2$ are both prime. This is equivalent to d being a Sophie Germain prime, a case where the BHC has already provided strong evidence of infinitely many examples.

A more difficult family in [9] has $S \cong \mathrm{PSL}_2(q)$ for prime degrees $d = \frac{q+1}{2}$, where $q = p^{2^k} \geq 5$ for an odd prime p and integer $k \geq 0$. To apply the BHC we took

$$f_1(t) = 2t + 1 \ (= p) \quad \text{and} \quad f_2(t) = \frac{(2t+1)^{2^k} + 1}{2} = \sum_{i=1}^{2^k} \binom{2^k}{i} 2^{i-1} t^i + 1 \ (= d)$$

for some fixed $k \geq 0$. In this case the calculation of the Hardy–Littlewood constants $C(f)$ for $k \geq 1$ is less straightforward (the case $k = 0$ is similar to the Sophie Germain primes problem). We have

$$\omega_f(r) = \begin{cases} 0 & \text{if } r = 2, \\ 2^k + 1 & \text{if } r \equiv 1 \bmod (2^{k+1}), \\ 1 & \text{otherwise,} \end{cases}$$

leading to the results for small k shown in Table 1:

k	$C(f)$	$Q(10^9)$	$E(10^9)$	relative error
1	4.426783	5 448 994	5 448 648.05	−0.006 %
2	10.433814	6 373 197	6 365 668.39	−0.118 %
3	7.885346	2 394 012	2 395 075.38	0.044 %
4	14.642571	2 219 445	2 218 975.66	−0.021 %

TABLE 1. $Q(10^9)$ is the number of $t \leq 10^9$ such that both $f_1(t)$ and $f_2(t)$ are prime; $E(10^9)$ is the BHC estimate for $Q(10^9)$.

On the basis of this we conjecture that for each $k \geq 0$ there are infinitely many primes p such that d is prime.

The above example is, in fact, the particular case $n = 1$ of a more general family involving the symplectic groups $S = \mathrm{PSp}_{2n}(q)$ where $d = (q^n + 1)/2$ is prime, $n = 2^j$ for some integer $j \geq 0$, and q is as before, but without the restriction $q \geq 5$ when $n > 1$. (Note that $\mathrm{PSp}_2(q) \cong \mathrm{PSL}_2(q)$.) For a fixed pair j, k we applied the BHC to the polynomials

$$f_1(t) = 2t + 1 \ (= p) \quad \text{and} \quad f_2(t) = \frac{(2t+1)^{2^{j+k}} + 1}{2} \ (= d).$$

These are the same as the preceding pair f_1, f_2, but with $j + k$ replacing k, so the same estimates $E(x)$ and search results $Q(x)$ apply in this case.

There are several other potentially infinite families of groups in [9, Theorem 1.2], but since they involve exponential functions rather than polynomials they are beyond the scope of the BHC.

In a corrigendum to [9] Dixon and Zalesskii showed that if G is primitive, and the socle S is non-abelian and imprimitive, then G has an imprimitive commutator subgroup $G' \cong \mathrm{PSL}_n(q)$ where $d = (q^n - 1)/(q - 1)$, with q odd or $q = 2$. Our results on projective primes apply in this case for odd q, and they are also relevant to [10] where the same authors have considered imprimitive linear groups of prime degree, transitively permuting the one-dimensional subspaces in a direct sum decomposition.

9. Other group-theoretic applications of the BHC

A *Kn group* is a simple group of order divisible by n distinct primes. The K1 groups are the cyclic groups of prime order. By a well-known theorem of Burnside there are no K2 groups, and it is known that there are just eight K3 groups, namely $\mathrm{PSL}_2(q)$ for $q = 5$, 7, 8, 9 and 17, $\mathrm{PSL}_3(3)$, $\mathrm{PSU}_3(3)$ and $\mathrm{PSp}_4(3)$. In Problem 13.65 of the Kourovka Notebook [18] Shi asked whether the set

of K4 groups is finite or infinite. The corresponding question for Kn groups appears to be open for each $n \geq 4$. In [21, Theorem 2] Kondrat'ev has classified the K5 groups; of his 12 classes, (1) and (12) are clearly finite, while the groups in each of classes (2) to (11) are defined in terms of a prime p satisfying some number-theoretic conditions, and it is not clear whether any of these classes is infinite. His class (5) consists of the groups $\mathrm{PSL}_2(p)$ and $\mathrm{PGL}_2(p)$ for primes $p \geq 41$ such that $p^2 - 1$ is divisible by just four primes. If a prime p satisfies condition (a), (b), (c) or (d) in Section 2, then the primes dividing $p^2 - 1$ are 2, 3, q and r, so by our earlier results the BHC gives strong evidence that this set of groups is infinite.

Cameron, Manna and Mehatari [7] have recently studied the *power graph* $P(G)$ of a finite group G. The vertices of $P(G)$ are the elements of G, and a pair of them are joined by an edge if one of them is a power of the other. In their Theorem 1.3 they characterise those non-abelian finite simple groups G for which $P(G)$ is a *cograph*. (This is a graph with no induced subgraph isomorphic to the path P_4 with four vertices; cographs form the closure of the one-vertex graph K_1 under the operations of disjoint union and complement.) In their Problem 1.4 they ask whether there are infinitely many such groups G.

One family appearing in their characterisation consists of the groups $\mathrm{PSL}_2(q)$ for odd prime powers $q \geq 5$ such that $(q \pm 1)/2$ are each either a prime power or a product of two primes. The groups $\mathrm{PSL}_2(p)$ in cases (a) and (b) of the six primes problem (see Section 2) satisfy this condition, so the BHC gives strong evidence of a positive answer to their question.

In [2] Amarra, Devillers and Praeger have recently constructed families of block designs which have interesting symmetry properties (a group of automorphisms acting transitively on blocks, and transitively but imprimitively on points) and which maximise various parameters. Their constructions, based on finite fields, require certain quadratic polynomials to take prime power values. By using the BHC, together with some extensions from primes to prime powers, we have in [17] provided strong evidence that these families are infinite.

10. Acknowledgements

The authors are grateful to Peter Cameron and Cheryl Praeger for helpful comments on applying the BHC to their work, and to Natalia Maslova for pointing out the connection with Kn groups. Alexander Zvonkin is supported by the French ANR project Combiné (ANR-19-CE48-0011).

References

[1] S. L. Aletheia-Zomlefer, L. Fukshansky and S. R. Garcia, The Bateman–Horn conjecture: heuristics, history, and applications, *Expo. Math.* 38 (2020), 430–479. Also available at `arXiv-math[NT]:1807.08899v4`.

[2] C. Amarra, A. Devillers and C. E. Praeger, Delandsheer–Doyen parameters for block-transitive point-imprimitive block designs, `arXiv-math[CO]:2009.00282`.

[3] P. T. Bateman and R. A. Horn, A heuristic asymptotic formula concerning the distribution of prime numbers, *Math. Comp.* 16 (1962), 220–228.

[4] V. Bouniakowsky, Sur les diviseurs numériques invariables des fonctions rationnelles entières, *Mém. Acad. Sci. St. Péteresbourg*, 6$^{\text{e}}$ série, vol. VI (1857), 305–329. Available at: `https://books.google.fr/books?hl=fr&id=wXIhAQAAMAAJ&pg=PA305#v=onepage&q&f=false`.

[5] W Burnside, Notes on the theory of groups of finite order. I: On the proof of Sylow's Theorem. II. On the possibility of simple groups whose orders are the products of four primes, *Proc. London Math. Soc.* 25 (1894), 9–18. Collected Papers, vol. I, 401–410, Note II.
[6] W Burnside, Notes on the theory of groups of finite order, *Proc. London Math. Soc.* 26 (1895), 191–214. Collected Papers, vol. I, 561–214, Note VII.
[7] P. J. Cameron, P. Manna and R. Mehatari, On finite groups whose power graph is a cograph, `arXiv-math[GR]: 2106.14217v3`.
[8] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker and R. A. Wilson, *ATLAS of Finite Groups*, Clarendon Press, Oxford, 1985. Available at https://epdf.pub/queue/atlas-of-finite-groups-maximal-subgroups-and-ordinary-characters-for-simple-grou.html.
[9] J. D. Dixon and A. E. Zalesskii, Finite primitive linear groups of prime degree, *J. London Math. Soc.* (2) 57 (1998), 126–134; corrigendum *ibid.* 77 (2008), 808–812.
[10] J. D. Dixon and A. E. Zalesskii, Finite imprimitive linear groups of prime degree, *J. Algebra* 276 (2004), 340–370.
[11] F. G. Frobenius, Über auflösbare Gruppen, *Sitzb. Akad. Wiss. Berlin* (1893), 337–345. Also in *Gesammelte Abhandlungen* (ed. J-P. Serre) II, Springer-Verlag, Berlin - Heidelberg - New York (1968), 565–573.
[12] G. H. Hardy and J. E. Littlewood, Some problems of 'Partitio numerorum'; III: On the expression of a number as a sum of primes, *Acta Math.* 114 (1923), 215–273.
[13] O. Hölder, Die einfachen Gruppen im ersten und zweiten Hundert der Ordnungszahlen, *Math. Ann.* 40 (1892), 55–88.
[14] G. A. Jones and J. M. Jones, *Elementary Number Theory*, Springer, 1998.
[15] G. A. Jones and A. K. Zvonkin, Klein's ten planar dessins of degree 11, and beyond, `arxiv.math:2104.12015`, to appear.
[16] G. A. Jones. and A. K. Zvonkin, Groups of prime degree and the Bateman–Horn Conjecture, `arxiv.math: 2106.00346`, to appear.
[17] G. A. Jones. and A. K. Zvonkin, Block designs and prime values of polynomials, `https://arxiv.org/abs/ 2105.03915`.
[18] E. Khukhro and V. Mazurov (eds), *Kourovka Notebook*, `https://arxiv.org/pdf/1401.0300.`
[19] F. Klein, Über die Transformationen siebenter Ordnung der elliptischen Funktionen, *Math. Ann.* 14 (3) (1878), 428–471.
[20] F. Klein, Über die Transformationen elfter Ordnung der elliptischen Funktionen, *Math. Ann.* 15 (3–4) (1879), 533–555.
[21] A. S. Kondrat'ev, Finite almost simple 5-primary groups and their Gruenberg–Kegel graphs, *Siberian Electronic Math. Reports* 11 (2014), 634–674.
[22] The Online Encyclopedia of Integer Sequences, `https://oeis.org`.
[23] A. Schinzel and W. Sierpiński, Sur certaines hypothèses concernant les nombres premiers, *Acta Arith.* 4 (1958), 185–298; erratum 5 (1958), 259.
[24] R. Solomon, A brief history of the classification of the finite simple groups, *Bull. Amer. Math. Soc. (N.S.)* 38 (2001), 315–352.
[25] R. A. Wilson, *Finite Simple Groups*, Springer, 2009.

School of Mathematical Sciences, University of Southampton, Southampton SO17 1BJ, UK
Email address: `G.A.Jones@maths.soton.ac.uk`

LaBRI, Université de Bordeaux, 351 Cours de la Libération, F-33405 Talence Cedex, France
Email address: `zvonkin@labri.fr`