

On an almost pure sieve

by

C. HOOLEY (Cardiff)

Combinatorial sieve methods can be classified according to the extent they depart from the exact exclusion principle of Legendre. Such a procedure, for instance, was adopted in the introductory chapter of our tract [5], in which a discussion of the basic elimination process was followed successively by descriptions of our simple asymptotic sieve, the first Brun sieve introduced in [1], and the second and more powerful sieve initiated by Brun in his seminal paper of 1920 [2]. But the more picturesque and elegant description of this classification in terms of purity or impurity of sieves is suggested to us by the name *Brun's Pure Sieve* that Halberstam and Richert assign to the earlier Brun sieve in their monograph [3]. Thus the purer the sieve the more closely it approximates to the fundamental exclusion process, while the more powerful it is the impurer and more complicated it is likely to be; in particular, sieves of substantial impurity applied to problems of difficulty normally yield only upper and lower positive bounds for what is sought. Naturally, under such an arrangement the so-called pure sieve of Brun is not actually completely unmixed because only the exclusion process of Legendre can have this status, albeit this qualification is merely a trifling matter of semantics that need not here detain us.

All versions or derivatives of Brun's method normally used are considerably more complex than his pure sieve. It is therefore not a little surprising that it has not been previously noticed that the addition of only a minor ingredient serves to convert the pure sieve into a much stronger instrument. So effective is the outcome that the resulting slightly impure sieve achieves upper and lower bounds of the same general power as those attained by much impurer versions of Brun's method, while retaining a simplicity that enables it to be used in circumstances where other methods of comparable strength are unworkable. Such a situation, for example, occurs when the constituents in upper or lower sieving bounds $\sum_{d|n} \varrho(d)$ for $\sum_{d|n} \mu(d)$ happen to appear in sums of the form

$$\sum_{n \leq x} \psi(n) \left(\sum_{d|n} \frac{\varrho(d)}{2^{\omega(d)}} \right)^2,$$

which possibly somewhat unexpectedly are also untreatable in terms of the elegant Selberg methods that fall outside our present classification of sieves. Estimates of optimal orders of magnitude for sums of this general type being needed in our forthcoming work [6] on integers in small intervals that are sums of two squares, we devote the present paper to the definition of our almost pure sieve and to an explanation of its mechanism by an application to a problem of prime twins type; in so doing, we supply an easy demonstration of the infinitude of pairs of integers differing by 2 that do not have more than fourteen prime factors—a result that is not markedly inferior to that first obtained by Brun in [2]. To the array of sieve methods we thus add a simple, versatile, and transparent tool that has much of the power enjoyed by more complicated procedures.

The almost pure sieve method used in our present application to almost prime twins has been deliberately described in such a way that only very minor modifications are needed for its indispensable operation in the above mentioned paper [6], in which the following new theorem is proved.

Let $M(m, h)$ be the number of integers n in the interval $m \leq n < m + h$ that are expressible as the sum of two squares. Then, if $h \leq x$ and $h/\sqrt{\log x} \rightarrow \infty$ as $x \rightarrow \infty$, we have

$$\frac{A_1 h}{\sqrt{\log x}} < M(m, h) < \frac{A_2 h}{\sqrt{\log x}}$$

for all m not exceeding x save for at most $o(x)$ exceptional values, where A_1, A_2 denote positive absolute constants.

Essentially best possible as a result involving almost all values of x , this theorem improves upon previous work of Harman's [4].

We follow the precedent set in [5] and describe the method in a general context, since the applications are not necessarily always confined to the divisibility properties of integers. Here, appropriately modifying our notation in [5] to suit the present occasion, we are presented with a finite set \mathcal{A} of N elements denoted generically by m , each of which may or may not possess one or more of the properties $\alpha_1, \dots, \alpha_n$ in a family \mathcal{S} indexed by $1, \dots, n$. This family is partitioned into h mutually exclusive sub-families \mathcal{S}_j of cardinality n_j in order that we may set up the present machinery for estimating the number $N_{1, \dots, n}$ of elements in \mathcal{A} that possess none of the properties α_i . A typical choice (possibly vacuous) of s_j elements from the indexing set of \mathcal{S}_j is then denoted by ι_{j, s_j} so that a sub-set of $1, \dots, n$ of cardinality s can be expressed as $\iota_s = (\iota_{1, s_1}, \dots, \iota_{h, s_h})$, where $s = s_1 + \dots + s_h$;

also $N(\iota_s) = N(\iota_{1,s_1}, \dots, \iota_{h,s_h})$ is to indicate the number of elements in \mathcal{A} having all the properties indexed by ι_s .

Ere we delineate the present sieve it is helpful to digress momentarily by remarking that traditionally there have been two ways of approaching such a task. Some sieves lend themselves to an explanation via bounds for $N_{1,\dots,n}$ in terms of the sums $N(\iota_s)$, while others are better elucidated by the bounds provided for the characteristic function $f(m)$ of the non-excluded or sifted set. Formulae of the first type, however, are normally substantiated by considering their applicability to the special case $N = 1$ and hence by establishing results of the second type. Thus the difference in presentations is merely one of emphasis, corresponding in fact to a reversal of the order of summations in a double sum.

Although the more elementary sieve methods are usually expressed through the former method, the nature of our slightly impure sieve is more easily unveiled in terms of the characteristic function of the sifted set. This is bounded by means of the inequalities

$$(1) \quad f(m) \leq \prod_{1 \leq j \leq h} \left(\sum_{s_j \leq r_j} (-1)^{s_j} \sum_{\iota_j, s_j \in m} 1 \right)$$

and

$$(2) \quad f(m) \geq \prod_{1 \leq j \leq h} \left(\sum_{s_j \leq r_j} (-1)^{s_j} \sum_{\iota_j, s_j \in m} 1 \right) - \sum_{1 \leq k \leq h} \left(\sum_{\iota_k, r_{k+1} \in m} 1 \right) \prod_{j \neq k} \left(\sum_{s_j \leq r_j} (-1)^{s_j} \sum_{\iota_j, s_j \in m} 1 \right),$$

in which r_1, \dots, r_h are suitably chosen non-negative *even* integers and in which the symbolism $\iota \in m$ means that m is to possess all the properties indexed by a sub-set ι of $1, \dots, n$. The upper inequality has a right-hand side that is a product of expressions occurring in Brun's pure sieve and therefore follows, for example, from equation (5A) in [5]. The lower inequality is perforce a little more complicated, since the total number of negative factors on the right of (1) might well be even when the numbers r_j were taken to be odd in accordance with the lower pure Brun sieve for each set of properties \mathcal{S}_j . To prove (2) we first note that its right-hand side is unchanged when the variables j, k of multiplication and summation are restricted to values l for which a positive number u_l of properties in \mathcal{S}_l are possessed by m . Hence, also now using the Legendre formula, we see that (2) is certainly true either when m has none of the properties in \mathcal{S} or when there is an exponent l such that $1 \leq u_l \leq r_l$. In the remaining case, if j', k' denote indices for which $u_l \neq 0$ and for which therefore

$$(3) \quad u_l > r_l,$$

the substitution of j', k' for j, k in the right of (2) does not affect its value and gives rise to a multiplicand and summand that are seen to be equal, respectively, to

$$\sum_{s_{j'} \leq r_{j'}} (-1)^{s_{j'}} \binom{u_{j'}}{s_{j'}} = \binom{u_{j'} - 1}{r_{j'}}$$

and

$$\binom{u_{k'}}{r_{k'} + 1} \prod_{j' \neq k'} \binom{u_{j'} - 1}{r_{j'}}$$

in virtue of the identity

$$(1 - y)^{u-1} = (1 - y)^u (1 - y)^{-1} = \sum_r \left(\sum_{s=0}^r (-1)^s \binom{u}{s} \right) y^r.$$

The right-hand side of (2) is thus

$$\left\{ 1 - \sum_{k'} \binom{u_{k'}}{r_{k'} + 1} / \binom{u_{k'} - 1}{r_{k'}} \right\} \prod_{j'} \binom{u_{j'} - 1}{r_{j'}}$$

in which the first factor equals

$$1 - \sum_{k'} \frac{u_{k'}}{r_{k'} + 1} \leq 0$$

by (3); the proof of (2) is therefore complete.

Having concluded our discussion of the sieve in a general context, we illustrate its relevance to more familiar situations in the theory of numbers by considering its application to problems of twin primes type. We shall therefore be involved with the production of upper and lower bounds for the number $\pi_2(x, \xi)$ of positive integers m not exceeding x which are to have the feature that $m(m + 2)$ is to be indivisible by any prime p not exceeding a certain limit ξ , concentrating almost entirely on the lower bound because only a small amount of the reasoning used for this is needed for the other. The primes p not exceeding ξ now describe the unwanted properties, wherefore the integers m not exceeding x with properties appertaining to p_1, \dots, p_s are just those for which $m(m + 2)$ is divisible by $d = p_1 \dots p_s$ and are thus $N(x, d)$ in number when the previous notation $N(\iota)$ is appropriately adapted. Consequently,

$$(4) \quad N(x, d) = \frac{x\nu(d)}{d} + O(\nu(d)) = \frac{x\nu(d)}{d} + O(d^\epsilon),$$

where $\nu(d)$ is the multiplicative function of square-free numbers d defined by $\nu(p) = 1$ if $p = 2$ and $\nu(p) = 2$ otherwise.

Considerable latitude in the apportionment of the properties to sub-families is permissible provided that it be made within a framework having cer-

tain prescribed features. Yet some care should be taken in the specification of the numerical parameters defining the structure lest we stray too far from the theoretical limits of the method. First, having written $\xi = x^{1/u}$, we let

$$(5) \quad \eta = \log \log x,$$

which function tends to infinity so slowly as $x \rightarrow \infty$ that $\pi_2(x, \xi)$ can be directly evaluated asymptotically by Legendre's principle whenever $\xi \leq \eta$. In the contrary situation $\xi > \eta$ to which we may now confine ourselves, we bring in parameters a, a_1 that are selected here to be 3.99, 4, respectively, and use the former to define the sequence $\xi_0, \dots, \xi_R, \xi_{R+1}$ by

$$(6) \quad \xi_j = \begin{cases} \xi^{1/a^j} & \text{for } j \leq R - 1, \\ \eta & \text{for } j = R, \\ 1 & \text{for } j = R + 1, \end{cases}$$

where R is the least exponent j for which $\xi^{1/a^j} \leq \eta$. For $j = 1, \dots, R + 1$ the family \mathcal{S}_j of unwanted properties is then to appertain to the primes p_j satisfying $\xi_j < p_j \leq \xi_{j-1}$, a typical square-free product (possibly empty) of which is to be denoted by d_j . The restricting agent r_j on the number s_j of properties in this family \mathcal{S}_j —in other words, the number $\omega(d_j)$ of distinct prime factors of d_j —to be used in (2) is determined by those of the equations

$$(7) \quad r_1 = r_2 = r_3 = 10; \quad r_j = 14(j - 3) \quad (j \geq 4)$$

that appertain to subscripts not exceeding R , while r_{R+1} is formally taken to be ∞ in reflection of the fact that s_{R+1} is only circumscribed by the cardinality of \mathcal{S}_{R+1} .

Since, in the current circumstances, formula (2) becomes

$$(8) \quad f(m) \geq \prod_{1 \leq j \leq R+1} \sum_{\substack{d_j | m(m+2) \\ \omega(d_j) \leq r_j}} \mu(d_j) - \sum_{1 \leq k \leq R+1} \left(\sum_{\substack{d_k | m(m+2) \\ \omega(d_k) = r_k + 1}} 1 \right) \prod_{j \neq k} \sum_{\substack{d_j | m(m+2) \\ \omega(d_j) \leq r_j}} \mu(d_j)$$

in the notation described above, we have

$$\begin{aligned} \pi_2(x, \xi) \geq & \sum_{\omega(d_j) \leq r_j (j=1, \dots, R+1)} \mu(d_1) \dots \mu(d_{R+1}) N(x, d_1 \dots d_{R+1}) \\ & - \sum_{1 \leq k \leq R+1} \sum_{\substack{\omega(d_j) \leq r_j (j \neq k) \\ \omega(d_k) = r_k + 1}} \frac{\mu(d_1) \dots \mu(d_{R+1})}{\mu(d_k)} N(x, d_1 \dots d_{R+1}) \end{aligned}$$

by substitution in

$$\pi_2(x, \xi) = \sum_{m \leq x} f(m)$$

and a change in the orders of summation in the resulting multiple sums. Hence, recalling our convention about $\omega(d_{R+1})$ and letting θ_{d_1, \dots, d_R} denote the condition that there be $R-1$ indices i from $1, \dots, R$ for which $\omega(d_i) \leq r_i$ and that the remaining one satisfy $\omega(d_i) \leq r_i + 1$, we infer from (4) that

$$\begin{aligned} (9) \quad & \pi_2(x, \xi) \\ & \geq x \sum_{\omega(d_j) \leq r_j (j=1, \dots, R+1)} \frac{\mu(d_1) \dots \mu(d_{R+1}) \nu(d_1 \dots d_{R+1})}{d_1 \dots d_{R+1}} \\ & \quad - x \sum_{1 \leq k \leq R} \sum_{\substack{d_1, \dots, d_{R+1} \\ \omega(d_j) \leq r_j (j \neq k) \\ \omega(d_k) = r_k + 1}} \frac{\mu(d_1) \dots \mu(d_{R+1}) \nu(d_1, \dots, d_{R+1})}{\mu(d_k) d_1 \dots d_{R+1}} \\ & \quad + O\left(\sum_{\substack{d_1, \dots, d_{R+1} \\ \theta_{d_1, \dots, d_R}}} (d_1 \dots d_{R+1})^\varepsilon \right) \\ & = x \prod_{1 \leq j \leq R+1} \sum_{\omega(d_j) \leq r_j} \frac{\mu(d_j) \nu(d_j)}{d_j} \\ & \quad - x \sum_{1 \leq k \leq R} \sum_{\omega(d_k) = r_k + 1} \frac{\nu(d_k)}{d_k} \prod_{j \neq k} \sum_{\omega(d_j) \leq r_j} \frac{\mu(d_j) \nu(d_j)}{d_j} \\ & \quad + O\left(\sum_{\substack{d_1, \dots, d_{R+1} \\ \theta_{d_1, \dots, d_R}}} (d_1 \dots d_{R+1})^\varepsilon \right) \\ & = x \left\{ 1 - \sum_{1 \leq k \leq R} \left(\sum_{\omega(d_k) = r_k + 1} \frac{\nu(d_k)}{d_k} / \sum_{\omega(d_k) \leq r_k} \frac{\mu(d_k) \nu(d_k)}{d_k} \right) \right\} \\ & \quad \times \prod_{1 \leq j \leq R+1} \sum_{\omega(d_j) \leq r_j} \frac{\mu(d_j) \nu(d_j)}{d_j} + O\left(\sum_{\substack{d_1, \dots, d_{R+1} \\ \theta_{d_1, \dots, d_R}}} (d_1 \dots d_{R+1})^\varepsilon \right) \\ & = x \left(1 - \sum_{1 \leq j \leq R} \left(\sum_2^{(j)} / \sum_1^{(j)} \right) \right) \prod_{1 \leq j \leq R+1} \sum_1^{(j)} + O\left(\sum_3 \right), \quad \text{say,} \end{aligned}$$

since we shall see that $\sum_1^{(j)}$ is certainly non-zero as soon as we begin to develop this initial inequality.

Our exploitation of (9) is speeded by using the following lemma, whose proof is not unconnected with the principles of Brun's pure sieve but whose existence is usually ignored in the treatment thereof.

LEMMA. Let σ_r denote the r -th elementary symmetric function of numbers a_1, \dots, a_n lying between 0 and 1. Then

$$\sum_{0 \leq r \leq s} (-1)^r \sigma_r - \prod_{1 \leq i \leq n} (1 - a_i)$$

is non-negative or non-positive according as s is even or odd.

Immediately, we have

$$(10) \quad \sum_1^{(j)} \geq \prod_{\xi_j < p \leq \xi_{j-1}} \left(1 - \frac{\nu(p)}{p}\right) > 0,$$

thus vindicating the assertion made after (9). Indeed, since

$$\prod_{p \leq y} \left(1 - \frac{\nu(p)}{p}\right) \sim \frac{c}{\log^2 y} \left\{1 + O\left(\frac{1}{\log y}\right)\right\},$$

we even have

$$(11) \quad \begin{aligned} \sum_1^{(j)} &\geq \frac{\log^2 \xi_j}{\log^2 \xi_{j-1}} \left\{1 + O\left(\frac{1}{\log \xi_j}\right)\right\} \\ &\geq \frac{1}{a^2} + o(1) \geq \frac{1}{a_1^2} \quad (j = 1, \dots, R) \end{aligned}$$

by (6). Also, in like manner,

$$\begin{aligned} \sum_{\xi_j < p \leq \xi_{j-1}} \frac{1}{p} &= \log \log \xi_{j-1} - \log \log \xi_j + O\left(\frac{1}{\log \xi_j}\right) \\ &\leq \log a + o(1) \leq \log a_1 \quad (j = 1, \dots, R) \end{aligned}$$

so that, by (9) and the elementary inequality $N! \geq (N/e)^N$,

$$(12) \quad \begin{aligned} \sum_2^{(j)} &\leq \frac{1}{(r_j + 1)!} \left(2 \sum_{\xi_j < p \leq \xi_{j-1}} \frac{1}{p}\right)^{r_j + 1} \\ &\leq \left(\frac{2e \log a_1}{r_j + 1}\right)^{r_j + 1} \leq \left(\frac{2e \log a_1}{t_j}\right)^{t_j} \quad (j = 1, \dots, R) \end{aligned}$$

under the sufficient condition

$$(13) \quad 2e \log a_1 < t_j < r_j + 1.$$

The indices r_j were chosen in (7) so that favourable estimates for the sums $\sum_2^{(j)}$ are consistent with a satisfactory value of \sum_3 when u is a small

constant. To attend to the former sums, we set

$$t_j = 2e\alpha_j \log a_1$$

in conformity with (13), where $\alpha_j > 1$ is a suitably close lower approximation to the number $(r_j + 1)/2e \log a_1$ that can be verified, for example, by the use of mathematical tables. Then, considering the values of r_j in (7), we can display the appropriate values of α_j and lower bounds β_j for $2e\alpha_j \log \alpha_j$ in the following table:

$$\begin{array}{ll} j = 1, 2, 3 & j \geq 4 \\ \alpha_j = 1.455 & (1.85)(j - 3) \\ \beta_j = 2.95 & 6(j - 3). \end{array}$$

Therefore, by (11) and (12),

$$\begin{aligned} \sum_2^{(j)} / \sum_1^{(j)} &< a_1^2 \left(\frac{1}{\alpha_j} \right)^{2e\alpha_j \log a_1} \\ &= \frac{1}{a_1^{2e\alpha_j \log \alpha_j - 2}} < \frac{1}{a_1^{\beta_j - 2}} \quad (j = 1, \dots, R), \end{aligned}$$

whence

$$\begin{aligned} (14) \quad 1 - \sum_{1 \leq j \leq R} \left(\sum_2^{(j)} / \sum_1^{(j)} \right) &> 1 - \sum_{1 \leq j \leq R} \frac{1}{a_1^{\beta_j - 2}} \\ &> 1 - \frac{3}{4^{19/20}} - \sum_{k=1}^{\infty} \frac{1}{4^{4k}} \\ &= 1 - \frac{3}{4^{19/20}} - \frac{1}{255} > \frac{3}{20} \end{aligned}$$

by a relatively crude calculation.

An examination of the conditions of summation in \sum_3 shews that the size of any number $d_1 \dots d_{R+1}$ corresponding to a contributing set d_1, \dots, d_{R+1} is circumscribed by $x^\gamma \eta^\eta$, where

$$\begin{aligned} \gamma &= \frac{1}{u} \left(1 + \sum_{1 \leq j \leq R} \frac{r_j}{a^{j-1}} \right) < \frac{1}{u} \left(1 + 10 + \frac{10}{a} + \frac{10}{a^2} + \frac{14}{a^3} \sum_{k=0}^{\infty} \frac{k+1}{a^k} \right) \\ &= \frac{1}{u} \left(1 + \frac{10a}{a-1} + \frac{14}{a(a-1)^2} \right) < \frac{74}{5u}. \end{aligned}$$

Hence

$$(15) \quad \sum_3 = O\left(\frac{x}{\log^3 x} \right)$$

whenever $u \geq 74/5$.

The first result attained by the method follows from (9), (10), (14), and (15), which give

$$(16) \quad \pi_2(x, \xi) > \frac{3x}{20} \prod_{p \leq \xi} \left(1 - \frac{\nu(p)}{p}\right) = \frac{3x}{40} \prod_{2 < p \leq \xi} \left(1 - \frac{2}{p}\right)$$

when $\xi \leq x^{5/74}$. In particular, we have established the existence of infinitely many pairs of natural numbers differing by 2 neither of which has more than fourteen prime factors.

A slightly shorter account would have been possible if we had been merely content to derive an inequality of type (16) for values of ξ not exceeding a smaller limit of the type x^{A_1} . On the other hand, some further fine tuning of the procedures might have resulted in some minor improvements, which, however, are not worth seeking through this avenue because stronger and more complicated methods are available for this particular purpose. In like but much simpler manner, the formula (1) yields the upper bound $O(x/\log^2 x)$ for the number of prime twins not exceeding x .

The development adopted above was chosen to suit the needs of a theorem about pseudo-prime twins. Yet, if we look for an improvement in (16) that is suitable for large values of u , then we must vary the previous process by selecting the parameters in terms of u . This is done easily by setting b to be the least even integer exceeding $u/2$ and then redefining r_j and α_j as $r_j = bj$ and $\alpha_j = (r_j + 1)/(2e \log a_1)$, it being easily verified that (15) continues to hold. Since an examination of the consequential change to (14) reveals that now

$$1 - \sum_{1 \leq j \leq R} \left(\frac{\sum_2^{(j)}}{\sum_1^{(j)}} \right) > 1 - e^{-A_2 u \log u},$$

we obtain the lower bound that is latent in the formula

$$\pi_2(x, x^{1/u}) = x \left(1 + O(e^{-Au \log u}) + O\left(\frac{1}{\log x}\right) \right) \prod_{p \leq x^{1/u}} \left(1 - \frac{\nu(p)}{p} \right),$$

the upper bound aspect of which is obtainable via (1). Our method thus produces *fundamental lemmata*, keener versions of which have been derived by Halberstam, Richert [3], and others by the more usual forms of Brun's method.

References

- [1] V. Brun, *Über das Goldbachsche Gesetz und die Anzahl der Primzahlpaare*, Arch. Math. Naturvidenskab 34 (1915), no. 8.
- [2] —, *Le crible d'Erathosthène et le théorème de Goldbach*, Videnskapselskapets Skrifter Mat.-nat. Kl. Kristiania 1920, no. 3.

- [3] H. Halberstam and H. E. Richert, *Sieve Methods*, Academic Press, 1975.
- [4] G. Harman, *Sums of two squares in short intervals*, Proc. London Math. Soc. (3) 62 (1991), 225–241.
- [5] C. Hooley, *Applications of Sieve Methods to the Theory of Numbers*, Cambridge Univ. Press, 1976.
- [6] —, *On the intervals between numbers that are sums of two squares: IV*, J. Reine Angew. Math., to appear.

SCHOOL OF MATHEMATICS
UNIVERSITY OF WALES COLLEGE OF CARDIFF
CARDIFF, U.K.

Received on 26.1.1993
and in revised form on 17.11.1993

(2369)