Bull. London Math. Soc. 34 (2002) 279–283
DOI: 10.1112/S002460930100889X

HUA'S LEMMA AND SIMULTANEOUS DIAGONAL EQUATIONS

JÖRG BRÜDERN AND TREVOR D. WOOLEY

ABSTRACT

This paper concerns systems of r homogeneous diagonal equations of degree k in s variables, with integer coefficients. Subject to a suitable non-singularity condition, it is shown that the expected asymptotic formula holds for the number of such systems inside a box $[-P,P]^s$, provided only that $s > (3r+1)2^{k-2}$. By way of comparison, classical methods based on the use of Hua's lemma would establish a similar conclusion, provided instead that $s > r2^k$.

In the study of the number of solutions of additive equations of smaller degree, Hua's lemma continues to play a prominent role in establishing asymptotic formulae. A generalization of this lemma due to Cook [**2**] provides a bound of similar strength for a system of equations, the number of variables required increasing in proportion to the number of equations. In order to be precise, we introduce some notation. Consider an $r \times rt$ integral matrix $A = (a_{ij})$ that contains t disjoint $r \times r$ non-singular submatrices. With this matrix A we associate the linear forms

$$\gamma_j = \sum_{i=1}^{r} a_{ij}\alpha_i \qquad (1 \leqslant j \leqslant rt).$$

Also, when $k \geqslant 3$ is an integer, we define the exponential sum

$$f(\alpha) = \sum_{|x| \leqslant P} e(\alpha x^k),$$

where $e(z)$ denotes $\exp(2\pi iz)$. Cook's lemma asserts that whenever $t = 2^{l-1}$, with $1 \leqslant l \leqslant k$, then for each positive number ε one has

$$\int_{[0,1]^r} |f(\gamma_1)f(\gamma_2)\dots f(\gamma_{rt})|^2 \, d\boldsymbol{\alpha} \ll P^{r(2^l-l)+\varepsilon}. \tag{1}$$

When $r = 1$, the bound (1) is tantamount to Hua's lemma (see Vaughan [**8**, Lemma 2.5]). Moreover, when the coefficient matrix A contains t disjoint $r \times r$ diagonal submatrices, then the integral on the left-hand side of (1) factorises into one-dimensional mean values. Hence, in the absence of sharper estimates for $r = 1$, and without further constraints on the coefficient matrix A, the estimate (1) is the best that can be attained. The purpose of this note is to show that for most coefficient matrices, sharper bounds are nonetheless accessible.

It is convenient to refer to an $r \times s$ matrix A that contains no singular $r \times r$ submatrix as being *highly non-singular*. In the sense usually adopted in analytic number theory, it is apparent that almost all coefficient matrices are highly non-singular. Also, when $r, s \in \mathbb{N}$, we define the exponent $\Lambda(s,r)$ to be the least real

Received 14 October 2000; revised 14 April 2001.

2000 *Mathematics Subject Classification* 11D72, 11P55.

The second author is a Packard Fellow and is supported in part by NSF grant DMS-9970440.

number with the property that whenever A is a highly non-singular $r \times s$ matrix, then for any positive number ε one has

$$\int_{[0,1]^r} |f(\gamma_1)f(\gamma_2)\dots f(\gamma_s)|^2\, d\boldsymbol{\alpha} \ll P^{\Lambda(s,r)+\varepsilon},$$

where the implicit constant depends at most on ε and A.

THEOREM. *Let r and s be natural numbers with $s \geqslant r$. Then $\Lambda(s,r) \leqslant \mathrm{M}(s,r)$, where*

$$\mathrm{M}(s,r) = \begin{cases} s, & \text{when } s \leqslant 3r-1, \\ 2s - rk, & \text{when } s \geqslant (3r+1)2^{k-3}, \end{cases}$$

and

$$\mathrm{M}(s,r) = \left(2 - 2^{1-l}\right)s - \left(l - \tfrac{1}{2}\right)r + \tfrac{1}{2},$$

when

$$2 \leqslant l \leqslant k-1 \quad \text{and} \quad \max\left\{3r-1\,,\, 2^{l-1}r\right\} < s \leqslant \min\left\{2^l r\,,\, (3r+1)2^{k-3}\right\}.$$

We note that the bound provided by our theorem is essentially dominated by the diagonal solutions of the underlying diophantine system so long as $s \leqslant 3r-1$, whereas in Cook's estimate (1), such is the case only for $s = rt \leqslant 2r$. Likewise, Cook's lemma provides an essentially best possible upper bound for $s = rt \geqslant 2^{k-1}r$, whereas our estimate already reaches this barrier for $s \geqslant 2^{k-3}(3r+1)$. When $k \geqslant 9$ or thereabouts, Vinogradov's methods may be applied to obtain estimates that are sharper than those presented in our theorem (but Heath-Brown's mean value estimate [**5**] lacks sufficient power to improve our estimates for $\Lambda(s,r)$ when $r \geqslant 2$).

Proof. We prove the theorem by induction on r with a subinduction on s. Observe first that when $r = 1$, the statement of the theorem follows from Hua's lemma via Hölder's inequality. Also, when $r \in \mathbb{N}$ and $s \leqslant 2r$, the desired conclusion is contained in Cook's lemma. Finally, on making a trivial estimate for the implicit exponential sums, it is apparent that the conclusion of the theorem for $s > (3r+1)2^{k-3}$ is immediate from that when $s = (3r+1)2^{k-3}$.

In order to discuss the remaining cases, we introduce the numbers $\sigma(l,r)$, which we define by

$$\sigma(l,r) = \begin{cases} 2r, & \text{when } l = 0, \\ 3r-1, & \text{when } l = 1, \\ 2^l r, & \text{when } 2 \leqslant l \leqslant k-2, \\ (3r+1)2^{k-3}, & \text{when } l = k-1. \end{cases}$$

Consider natural numbers r and s with $2r < s \leqslant (3r+1)2^{k-3}$ and $r \geqslant 2$, and let A denote an integral highly non-singular $r \times s$ matrix. We may suppose that $\Lambda(s',r') \leqslant \mathrm{M}(s',r')$ whenever $r' < r$, and likewise whenever $r' = r$ and $s' < s$. We take l to be the unique natural number satisfying $\sigma(l-1,r) < s \leqslant \sigma(l,r)$, and put

$$u = \max\left\{\sigma(l-1,r)\,,\, s - 2^{l-1}\right\} \quad \text{and} \quad t = s - u.$$

Then, by Hölder's inequality, one has

$$\int_{[0,1]^r} |f(\gamma_1)\dots f(\gamma_s)|^2\, d\boldsymbol{\alpha} \leqslant I_0^{1-t2^{1-l}} \prod_{u<j\leqslant s} I_j^{2^{1-l}}, \tag{2}$$

where

$$I_0 = \int_{[0,1]^r} |f(\gamma_1)\dots f(\gamma_u)|^2\, d\boldsymbol{\alpha}$$

and

$$I_j = \int_{[0,1]^r} |f(\gamma_1)\dots f(\gamma_u)|^2 |f(\gamma_j)|^{2^l} d\boldsymbol{\alpha}.$$

For each fixed j with $u < j \leqslant s$, by considering the underlying diophantine system and performing the appropriate elementary row operations, one finds that the last integral is equal to

$$\int_{[0,1]^r} |f(\widetilde{\gamma}_1)\dots f(\widetilde{\gamma}_u)|^2 |f(b\alpha_r)|^{2^l} d\boldsymbol{\alpha}, \tag{3}$$

where $\widetilde{\gamma}_m = \widetilde{\gamma}_m(\boldsymbol{\alpha})$ is defined by

$$\widetilde{\gamma}_m = \sum_{i=1}^{r} \widetilde{a}_{im}\alpha_i \qquad (1 \leqslant m \leqslant u)$$

and the integers $\widetilde{a}_{im}$ and the positive integer b arise from the row operations alluded to above. It is therefore apparent that the matrix $(\widetilde{A}|\mathbf{b})$, where $\widetilde{A} = (\widetilde{a}_{im})$ and $\mathbf{b} = (0,\dots,0,b)^T$, remains highly non-singular. But, by Weyl differencing, one obtains

$$|f(b\alpha)|^{2^l} \ll P^{2^l-1} + P^{2^l-l-1} \sum_{0<|h|\leqslant bk!(2P)^k} c_h e(\alpha h), \tag{4}$$

where the integers c_h satisfy $c_h = O(|h|^{\varepsilon})$. On substituting (4) into (3), one finds that the mean value (3) is bounded above by

$$P^{2^l-1+\Lambda(u,r)+\varepsilon} + P^{2^l-l-1} \sum_{0<|h|\leqslant bk!(2P)^k} c_h T(h),$$

where

$$T(h) = \int_{[0,1]^r} |f(\widetilde{\gamma}_1)\dots f(\widetilde{\gamma}_u)|^2 e(\alpha_r h)\, d\boldsymbol{\alpha}.$$

On considering the underlying diophantine system, one finds that $T(h) \geqslant 0$, and that

$$\sum_h T(h) = \int_{[0,1]^{r-1}} |f(\widehat{\gamma}_1)\dots f(\widehat{\gamma}_u)|^2 d\alpha_1 \dots d\alpha_{r-1},$$

where $\widehat{\gamma}_m = \widetilde{\gamma}_m(\alpha_1,\dots,\alpha_{r-1},0)$, for $1 \leqslant m \leqslant u$. But the matrix obtained from $\widetilde{A}$ by deleting the rth row is a highly non-singular $(r-1)\times u$ matrix, and hence we deduce that

$$\sum_{0<|h|\leqslant bk!(2P)^k} c_h T(h) \ll P^{\Lambda(u,r-1)+2k\varepsilon}.$$

On noting that $I_0 \ll P^{\Lambda(u,r)+\varepsilon}$, we thus conclude from (2) that

$$\begin{aligned}\Lambda(s,r) \leqslant\ & t2^{1-l} \max\left\{2^l - 1 + \Lambda(u,r),\ 2^l - l - 1 + \Lambda(u,r-1)\right\} \\ & + \left(1 - t2^{1-l}\right)\Lambda(u,r).\end{aligned} \tag{5}$$

In order to extract the desired conclusion from this last bound, it is useful to observe that in all the situations under consideration, one has $\mathrm{M}(u,r-1) \leqslant \mathrm{M}(u,r)+l$.

To verify this assertion, we note first that when $l = 1$, one has $u \leqslant \sigma(1,r)-1 = 3r-2$, and further that $\mathrm{M}(u,r) = u$ and

$$\mathrm{M}(u,r-1) = \begin{cases} u, & \text{when } u \leqslant 3r-4, \\ \frac{3}{2}u - \frac{1}{2}(3r-4), & \text{when } 3r-4 < u \leqslant 3r-2. \end{cases}$$

Thus we find that when $l = 1$, one has

$$\mathrm{M}(u,r-1) - \mathrm{M}(u,r) \leqslant \max\{0, \tfrac{1}{2}(u-3r+4)\} \leqslant 1.$$

We observe next that for $2 \leqslant l \leqslant k-2$, one has $\sigma(l,r) - 2^{l-1} < \sigma(l+1,r-1)$. Then for the latter values of l, one has $u < \sigma(l+1,r-1)$. Consequently,

$$\mathrm{M}(u,r) = \left(2-2^{1-l}\right)u - \left(l-\tfrac{1}{2}\right)r + \tfrac{1}{2},$$

and

$$\mathrm{M}(u,r-1) = \begin{cases} \left(2-2^{1-l}\right)u - \left(l-\frac{1}{2}\right)(r-1) + \frac{1}{2}, & \text{when } u \leqslant \sigma(l,r-1), \\ \left(2-2^{-l}\right)u - \left(l+\frac{1}{2}\right)(r-1) + \frac{1}{2}, & \text{when } u > \sigma(l,r-1), \end{cases}$$

whence

$$\mathrm{M}(u,r-1) - \mathrm{M}(u,r) \leqslant \max\left\{l-\tfrac{1}{2}, 2^{-l}\left(2^l r - 2^{l-1}\right) + l - r + \tfrac{1}{2}\right\} = l.$$

Finally, when $l = k-1$, one finds that

$$\mathrm{M}(u,r) \leqslant \left(2-2^{2-k}\right)u - \left(k-\tfrac{3}{2}\right)r + \tfrac{1}{2},$$

and

$$\mathrm{M}(u,r-1) = \begin{cases} \left(2-2^{2-k}\right)u - \left(k-\frac{3}{2}\right)(r-1) + \frac{1}{2}, & \text{when } u \leqslant \sigma(k-1,r-1), \\ 2u - (r-1)k, & \text{when } u > \sigma(k-1,r-1), \end{cases}$$

and hence

$$\begin{aligned} \mathrm{M}(u,r-1) - \mathrm{M}(u,r) &\leqslant \max\left\{k-\tfrac{3}{2}, 2^{2-k}\left((3r+1)2^{k-3} - 2^{k-2}\right) + k - \tfrac{3}{2}r - \tfrac{1}{2}\right\} \\ &= \max\left\{k-\tfrac{3}{2}, k-1\right\} \\ &= k-1. \end{aligned}$$

Equipped with the discussion of the previous paragraph, and making use of the inductive hypothesis, we infer from (5) that

$$\begin{aligned} \Lambda(s,r) &\leqslant t2^{1-l}\max\left\{2^l - 1 + \mathrm{M}(u,r), 2^l - l - 1 + \mathrm{M}(u,r-1)\right\} + \left(1 - t2^{1-l}\right)\mathrm{M}(u,r) \\ &= 2t - t2^{1-l} + \mathrm{M}(u,r) \\ &= \mathrm{M}(s,r). \end{aligned}$$

In view of our earlier comments, the conclusion of the theorem now follows by induction in all cases.

We finish this note by discussing an application of our theorem to the solubility of simultaneous diagonal equations. Let r and s be natural numbers with $s > (3r+1)2^{k-2}$. Let $N(P)$ denote the number of integral solutions of the system of equations

$$\sum_{j=1}^{s} c_{ij}x_j^k = 0 \qquad (1 \leqslant i \leqslant r) \tag{6}$$

with $|x_j| \leqslant P$ $(1 \leqslant j \leqslant s)$, where (c_{ij}) is an $r \times s$ integral matrix with the property that

the first $(3r+1)2^{k-3}$ columns form a highly non-singular submatrix, and likewise also the final $(3r+1)2^{k-3}+1$ columns. Then, by standard methods originating in work of Davenport and Lewis [**3**, **4**] (a model for the relevant argument may be found in Brüdern and Cook [**1**]), one may apply our theorem to show that for large P one has

$$N(P) = v_\infty \left(\prod_p v_p \right) P^{s-rk} + o(P^{s-rk}), \tag{7}$$

where v_∞ is the area of the manifold defined by (6) in the box $[-1,1]^s$, and

$$v_p = \lim_{h\to\infty} p^{h(r-s)} \mathrm{card}\left\{ \mathbf{x} \in \left(\mathbb{Z}/p^h\mathbb{Z}\right)^s : \sum_{j=1}^{s} c_{ij} x_j^k \equiv 0 \pmod{p^h} \ (1 \leqslant i \leqslant r) \right\}.$$

We note that when $k=3$, the methods of this note permit the proof of the asymptotic formula (7) whenever $s \geqslant 6r+3$. Meanwhile, even the strongest speculative hypotheses concerning mean values of cubic Weyl sums (see Heath-Brown [**6**] and Hooley [**7**]) provide such a conclusion only for $s \geqslant 6r+1$. Indeed, the sharpest conclusions available in the literature hitherto (see Brüdern and Cook [**1**]) establish (7) for $s > 8r$ and the lower bound $N(P) \gg (\prod_p v_p) P^{s-rk}$ for $s > 7r$, though with weaker conditions on the coefficient matrix.

References

1. J. Brüdern and R. J. Cook, 'On simultaneous diagonal equations and inequalities', *Acta Arith.* 62 (1992) 125–149.

2. R. J. Cook, 'A note on a lemma of Hua', *Quart. J. Math. Oxford* (2) 23 (1972) 287–288.

3. H. Davenport and D. J. Lewis, 'Cubic equations of additive type', *Philos. Trans. Roy. Soc. London Ser. A* 261 (1966) 97–136.

4. H. Davenport and D. J. Lewis, 'Simultaneous equations of additive type', *Philos. Trans. Roy. Soc. London Ser. A* 264 (1969) 557–595.

5. D. R. Heath-Brown, 'Weyl's inequality, Hua's inequality, and Waring's problem', *J. London Math. Soc.* (2) 38 (1988) 216–230.

6. D. R. Heath-Brown, 'The circle method and diagonal cubic forms', *Philos. Trans. Roy. Soc. London Ser. A* 356 (1998) 673–699.

7. C. Hooley, 'On hypothesis K^* in Waring's problem', *Sieve methods, exponential sums and their applications in number theory* (*Cardiff, 1995*), London Math. Soc. Lecture Note Ser. 237 (Cambridge University Press, 1997) 175–185.

8. R. C. Vaughan, *The Hardy–Littlewood method*, 2nd edn (Cambridge University Press, 1997).

Mathematisches Institut A
Universität Stuttgart
D-70511 Stuttgart
Germany

bruedern@mathematik.uni-stuttgart.de

Department of Mathematics
University of Michigan
East Hall
525 East University Avenue
Ann Arbor, MI 48109-1109
USA

wooley@math.lsa.umich.edu