PACIFIC JOURNAL OF MATHEMATICS
Vol. 50, No. 2, 1974

ON A PROBLEM OF HURWITZ

NORMAN P. HERZBERG

A. Hurwitz proposed the problem of finding all the positive integers z, $\boldsymbol{x} = (x_1, \cdots, x_n)$ satisfying the diophantine equation $x_1^2 + \cdots + x_n^2 = z \cdot x_1, \cdots, x_n$. This paper investigates the question of which values of z can occur, using only the most elementary techniques. An algorithm is given for determining all permissible values of (z, n) for all n below a given bound. As an application it is established that the only possible values in the range $z \geqq (n + 15)/4$ are $z = n$, $z = (n + 8)/3$ when n is odd, and $z = (n + 15)/4$. As another application the fifteen values of $n \leqq 131{,}020$ for which the only permissible value of z is n have been found.

2. The problem of finding all the integer solutions z, $\boldsymbol{x} = (x_1, \cdots, x_n)$ of the equation

$$x_1^2 + \cdots + x_n^2 = z \cdot x_1, \cdots, x_n \tag{1}$$

was raised by A. Hurwitz in [1]. In that paper he showed that for $n > z$ there are no solutions. This is an easy consequence of Theorem 1 (see §3) and will be replaced by the stronger result in Theorem 3. To keep this paper self-contained, let us recall the following facts from [1].

For $n = 2$, the only solutions are $z = 2$, $x_1 = x_2$; for upon setting $x_1 = dy_1$, $x_2 = dy_2$ with $(y_1, y_2) = 1$, $y_1^2 + y_2^2 = zy_1y_2$, and so $z = 2$, $x_1 = x_2 = d$.

If $z, x_1, \cdots, x_j, \cdots, x_n$ is a solution, then so is $z, x_1, \cdots, x_j', \cdots, x_n$ where x_j' satisfies

$$x_j + x_j' = z \prod_{i \neq j} x_i .$$

The n solutions derived in this way are called the *neighbors* of z, $\boldsymbol{x}$. Define the *height* of a solution to be simply $x_1 + \cdots + x_n$, and call a solution *fundamental* if its height is no greater than the height of any of its neighbors. If a solution is not fundamental, it has a neighbor of strictly smaller height, and since the heights are all positive integers, in a finite number of steps we arrive at a fundamental solution. So we see that *it suffices to study fundamental solutions*. Moreover, it obviously suffices to study solutions that satisfy

$$x_1 \geqq x_2 \geqq \cdots \geqq x_n \geqq 1 . \tag{2}$$

Also, as Hurwitz point out, it is easy to see that fundamental solutions satisfying (2) are characterized by

$$(3) \qquad 2x_1 \leqq z \prod_{i=2}^{n} x_1 \ .$$

We now propose to study the system of Equations (1), (2), (3), *and shall regard* n *as well as* z *and* $x_1, \cdots, x_n$ *as variables.* By the first remark in this section we may also assume

$$(4) \qquad n \geqq 3 \ .$$

3. In this section we state our basic theorem. First some notation.

The *trivial solution* of (1)–(4) is $x_1 = \cdots = x_n = 1$, $z = n$. Call a nontrivial solution of (1)–(4) a SOL. For any SOL, we define

$$\chi(\boldsymbol{x}) = \text{the largest index } i \text{ for which } x_i > 1 \ .$$

THEOREM 1. *Let* $n, z, \boldsymbol{x}$ *be a SOL with* $k = \chi(\boldsymbol{x})$. *There is a chain of SOLs* $n^{(i)}, z, \boldsymbol{x}^{(i)}$, $i = 0, \cdots, t$ *such that*

(a) $\chi(\boldsymbol{x}) = \chi(\boldsymbol{x}^{(i)})$ *for all* i.

(b) *If* $z = 1$ *and* $k = 3$, *then* $n^{(0)}, \boldsymbol{x}^{(0)} = 3, (3, 3, 3)$. *Otherwise,* $n^{(0)} = 2^k \cdot z - 3k$, $x_1^{(0)} = \cdots = x_k^{(0)} = 2$.

(c) $n^{(i)} > n^{(i-1)}$ for $i = 1, \cdots, t$.

(d) $n^{(t)}, \boldsymbol{x}^{(t)} = n, \boldsymbol{x}$.

The proof is in the next section. Below we give some immediate corollaries of the theorem, using the same notation.

COROLLARY 1. *k must satisfy* $2^k - 3k \leqq n$. [By (b), since $z \geqq 1$.]

COROLLARY 2. *z must satisfy* $z \leqq (n+3)/2$. [By (b), since $k \geqq 1$.]

COROLLARY 3. *The only fundamental solution to Equation* (1) *with* $z \geqq n$ *is the trivial solution.*

4. In this section, we prove Theorem 1. First we state and prove some simple lemmas.

LEMMA 1. *Let* $n, z, \boldsymbol{x}$ *be a SOL.*
If $z = 3$, *then* $\chi(\boldsymbol{x}) \geqq 2$.
If $z \leqq 2$, *then* $\chi(\boldsymbol{x}) \geqq 3$.
If $z = 1$, *and* $\chi(x) = 3$, *then* $x_3 \geqq 3$.

Proof. If $z \leqq 3$ and $\chi(\boldsymbol{x}) = 1$, then by (3) $2x_1 \leqq 3$ which contradicts $x_1 > 1$. Hence $\chi(\boldsymbol{x}) \geqq 2$. If $z \leqq 2$ and $\chi(\boldsymbol{x}) \leqq 2$, then by (1) $x_1^2 + x_2^2 + (n-2) \cdot 1^2 \leqq 2 \cdot x_1 \cdot x_2 \cdot 1$. Thus $(x_1 - x_2)^2 \leqq 2 - n$. This contradicts (4). Finally suppose $z = 1$, $\chi(\boldsymbol{x}) = 3$ and $x_3 = 2$. Then by (1)

$x_1^2 + x_2^2 + (n+1) = 2x_1x_2$, a contradiction.

LEMMA 2. *Let n, z, x be a SOL with $k = \chi(x)$. When $z = 1$ and $k = 3$, $n^{(0)}, z, \boldsymbol{x}^{(0)} = 3, 1, (3, 3, 3)$ is a SOL. Otherwise if $n^{(0)} = z \cdot 2^k - 3k$ and $x_1^{(0)} = \cdots = x_k^{(0)} = 2$, $x_i^{(0)} = 1$ for $i = k+1, \cdots, n^{(0)}$, then $n^{(0)}$, z, x is a SOL with $\chi(x^{(0)}) = k$.*

Proof. Obviously $n^{(0)}, z, \boldsymbol{x}^{(0)} = 1, 3, (3, 3, 3)$ is a SOL. As for the other cases: $\sum x_i^2 = 4k + n^{(0)} - k$ while $z \prod x_i = z \cdot 2^k$, thus the definition of $n^{(0)}$ guarantees (1). (2) and (4) are trivial while to verify (3) we must check that

$$4 \leqq z \prod_{i>1} x_i$$

which is obvious when $z \geqq 4$ and true for $z \leqq 3$ by the constraints imposed by Lemma 1.

LEMMA 3. *Let $n, z, \boldsymbol{x}$ and $N, z, \boldsymbol{X}$ be two SOLs such that*
(a) $\chi(\boldsymbol{x}) = \chi(\boldsymbol{X}) = k$
(b) $X_1 > x_1$
(c) $X_j \geqq x_j$ *for* $j = 2, \cdots, k$.
Let r be the last index j for which $X_j > x_j$. Let s' be the first index j for which $x_j < x_1$, and define $s = s'$ if $s' \leqq r$, $s = 1$ if $s' > r$. Then $m, z, \boldsymbol{w}$ is a SOL if

$$\begin{aligned} m &= n - 2x_s - 1 + z \prod_{i \neq s} x_i \\ w_i &= x_i \text{ for } i \leqq k,\ i \neq s \\ w_s &= x_s + 1 \\ w_i &= 1 \text{ for } i > k\,. \end{aligned}$$

Moreover $m > n$.

Proof. We use the notation $\sum$ and $\prod$ to denote sums and products for which the index i runs from 1 to k, and append a prime to mean that $i \neq s$.

To check that $\boldsymbol{w}$ really is a SOL we must check (1) and (3). Now by (1) $\sum x_i^2 = z \prod x_i - (n - k)$. Thus $\sum w_i^2 = z \prod w_i = z \prod' x_i - (n - k) + 2x_s + 1$. So by the definition of m, (1) is satisfied.

If $s > 1$, then since $\boldsymbol{x}$ satisfies (3) so will $\boldsymbol{w}$. We may therefore assume $s = 1$. By the definition of s, $x_1 = \cdots = x_r$ and $x_{r+1} = X_{r+1}, \cdots, x_k = X_k$. Thus either (i) $r = 1$ or (ii) $r \geqq 2$ and $x_1 = x_2$. In case (i) we note that $N, z, \boldsymbol{X}$ satisfies (3), that $z \prod' w_i = z \prod' X_i$, and that $X_1 > x_1$ implies $2X_1 \geqq 2(x_1 + 1) = 2w_1$. Thus $\boldsymbol{w}$ satisfies (3). In case (ii) we must check that $z \prod' x_i \leqq 2x_1 + 2$. Dividing by $x_1 = x_2$ and recalling that $x_1 \geqq 2$ we see that it suffices to know that

$$z \prod_{i=3}^{k} x_i \geqq 3. \quad \text{(The empty product equals 1.)}$$

This is certainly true if $z \geqq 3$ and easily checked via the constraints of Lemma 1 when $z < 3$.

Finally we note that $m > n$ is equivalent to $z \prod' x_i \geqq 2(x_s + 1)$. Multiplying by x_s we see that it suffices to show $z \prod x_i \geqq 2(x_s^2 + x_s)$ and since $x_1 \geqq x_s$ it suffices to prove this when $s = 1$. Dividing by x_1 we obtain Equation (3) for $\boldsymbol{w}$, which was verified above.

Proof of Theorem 1. The $n^{(0)}, z, \boldsymbol{x}^{(0)}$ defined in (b) is a SOL by Lemma 2. If $(x_1, \cdots, x_k) \neq (x_1^{(0)}, \cdots, x_k^{(0)})$, we apply Lemma 3 (with $s = 1$) to obtain a SOL $n^{(1)}, z, \boldsymbol{x}^{(1)}$, with $n^{(1)} > n^{(0)}$. By induction: At step i, if $r > 1$, we will have either $x_1^{(i)} = \cdots = x_{s-1}^{(i)} > x_s^{(i)} = \cdots = x_r^{(i)}$ where $x_{s-1}^{(i)} = x_s^{(i)} + 1$, or $x_1^{(i)} = x_r^{(i)}$ and $s = 1$. Hence we will be able to apply Lemma 3. When $r = 1$, at $i = t$ say, we have $(x_1, \cdots, x_k) = (x_1^{(t)}, \cdots, x_k^{(t)})$ and by (1) both n and $n^{(t)}$ equal

$$k + z \prod_{i=1}^{k} x_i - \sum_{i=1}^{k} x_i^2 \,.$$

Hence, $n^{(t)}, z, \boldsymbol{x}^{(t)} = n, z, x$.

5. The following corollary is an easy consequence of the proof of Theorem 1.

COROLLARY 4. *Every SOL $n, z, \boldsymbol{x}$ satisfies $n \geqq x_1$.*

Proof. (We use the notation of Theorem 1.) To construct $\boldsymbol{x}^{(i+1)}$ from $\boldsymbol{x}^{(i)}$ we applied Lemma 3. Thus for $1 \leqq j \leqq k$

$$x_j^{(i+1)} - x_j^{(i)} = \begin{cases} 0 \text{ if } j \neq s \\ 1 \text{ if } j = s \,. \end{cases}$$

Since $n^{(i+1)} > n^{(i)}$,

$$\sum_{j=1}^{k} x_j^{(i+1)} - x_j^{(i)} \leqq n^{(i+1)} - n^{(i)} \,.$$

Summing these equations for $i = v, \cdots, t$ we get

$$(5) \qquad n^{(t)} = n \geqq n^{(v)} + \sum_{j=1}^{k} x_j - x_j^{(v)} \geqq n^{(v)} + x_1 - x_1^{(v)} \,.$$

If $z \neq 1$ or $\chi(\boldsymbol{x}) \neq 3$, then $x_1^{(0)} = 2$ and $n^{(0)} \geqq 4$. Thus by (5), $n \geqq x_1 + 2$. If $z = 1$ and $\chi(\boldsymbol{x}) = 3$, then $n^{(0)}, \boldsymbol{x}^{(0)} = 3, (3, 3, 3)$; $n^{(1)}, \boldsymbol{x}^{(1)} = 5, (4, 3, 3, 1, 1)$; and $n^{(2)}, \boldsymbol{x}^{(2)} = 10, (4, 4, 3, 1, \cdots, 1)$. Thus the corollary is true for $x = \boldsymbol{x}^{(0)}$ or $\boldsymbol{x}^{(1)}$. Setting $v = 2$ in (5), we have $n \geqq x_1 + 6$ otherwise.

6. Lemma 3 and Theorem 1 yield an algorithm that produces only SOLs, and each only once.

THEOREM 2. *The following seven step algorithm constructs all SOLs $n, z, \boldsymbol{x}$ with $n \leqq M$.*

Let A be a list of SOLs, initially empty. The set of SOLs put into A will be the SOLs sought.

(1) *Set $k = 1$ and $z = 4$.*

(2) *Using the current values of z and k, put the SOL constructed in Lemma 2 on the bottom of the list A.*

(3) *If A is empty, go to Step 6, otherwise remove the top SOL $n, z, \boldsymbol{x}$ from A.*

(4) *Define $w_1 = x_1 + 1$, $w_i = x_i$ for $i \geqq 2$, $k = \chi(\boldsymbol{x})$ and*

$$\nu = z \prod_{i=2}^{k} w_i - 2w_1 + 1 .$$

Let $m = n + \nu$. If $n < m < M$ define $w_i = 1$ for $i = n + 1, \cdots, m$. $m, z, \boldsymbol{w}$ is a new SOL. Put it on the bottom of A. (If m is not between n and M we do nothing.)

(5) *Find the smallest index $s \leqq k$ satisfying $x_1 - x_s = 1$. If no such s exists, go to Step 3; otherwise define $w_s = x_s + 1$, $w_i = x_i$ for $i \neq s$, $k = \chi(x)$ and*

$$\nu = z \prod_{i \neq s}^{k} w_i - 2w_s + 1 .$$

Let $m = n + \nu$. If $m > M$ go to Step 3. If $m \leqq M$ define $w_i = 1$ for $i = n + 1, \cdots, m$. $m, z, \boldsymbol{w}$ is a new SOL (since $n < m$ is always true). Put it on the bottom of A and go to Step 3.

(6) *Increase z by 1 and set $\nu = z \cdot 2^k - 3k$. If $\nu \leqq M$ go to Step 3, otherwise go to Step 7.*

(7) *Increase k by 1. If $k = 2$, set $z = 3$, otherwise set $z = 1$. Set $\nu = z \cdot 2^k - 3k$. If $\nu \leqq M$ go to Step 2, otherwise stop.*

Proof. Every SOL $n, z, \boldsymbol{x}$ satisfying $n \leqq M$ eventually is put on A because the algorithm produces a unique sequence of SOLs passing through the $\chi(\boldsymbol{x}) = k$ SOLs of the form $m, z, \boldsymbol{w}^{(j)}$ where $\chi(\boldsymbol{w}^{(j)}) = k$ and

$$\boldsymbol{w}^{(j)} = (x_j, \cdots, x_j, x_{j+1}, x_{j+2}, \cdots, x_k, 1, \cdots, 1) .$$

(Uniqueness is guaranteed by Step 5.)

Theorem 2 is extremely powerful, and it is no trouble to produce a table of SOLs by hand for moderately large n. The Appendix lists all solutions of (1)-(4) with $n \leqq 45$ except the trivial solution (when $z = n$). We have omitted those x_i which equal 1.

7. In this section, we will apply Theorem 2 to get a better bound on z than that given by Corollary 2.

Suppose $n, z, \boldsymbol{x}$ is a SOL with $k = \chi(\boldsymbol{x})$, and suppose $n \neq 2z - 3$. In particular, if $k = 1$, then $n \neq n^{(0)}$. Hence either (i) $k \geqq 2$ or (ii) $k = 1$ and $n \geqq n^{(1)}$. In case (i) by Theorem 1 (b)

$$z \leqq (n + 3k)/2^k \leqq (n + 6)/4 .$$

In case (ii) since $n^{(0)} = 2z - 3$ and $n^{(1)} = n^{(0)} + z - 5$, we see that $z \leqq (n + 8)/3$. Now if $n \geqq 14$, $(n + 8)/3 \leqq (n + 6)/4$, while for $n \leqq 14$, $z \leqq (n + 8)/3$ by inspection.

THEOREM 3. *The only SOLs $n, z, \boldsymbol{x}$ with $z > (n + 8)/3$ are the SOLs with n odd, $z = (n + 3)/2$, $\boldsymbol{x} = (2, 1, \cdots, 1)$.*

Proof. Since n even implies $n \neq 2z - 3$, there are no SOLs with $z > (n + 8)/3$. If n is odd and $n = 2z - 3$, then $\chi(\boldsymbol{x}) = 1$ and $n = n^{(0)}$, $\boldsymbol{x} = \boldsymbol{x}^{(0)}$ of Theorem 1 (b).

Theorem 3 is hardly the best possible. For any n, each SOL $n, z, \boldsymbol{x}$ is the end point of one of the chains described in Theorem 1, and in general, the longer the chain, the larger n must be compared to z. So for example if $n \geqq n^{(2)}$, $z \leqq (n + 15)/4$ when $\chi(\boldsymbol{x}) = 1$ and if $\chi(\boldsymbol{x}) \geqq 2$ and $z \geqq 3$, then $z \leqq (n + 10)/8$. *Thus there are no solutions to* (1) *when* $(n + 8)/3 > z > (n + 15)/4$, *etc.*.

8. Hurwitz asked if there exists n for which the only solutions to (1) have $z = n$. There are.

PROPOSITION. *There are* 15 *values of* $n \leqq 301020$ *for which* (1)-(4) *has no nontrivial solutions. They occur when* $n = 12, 24, 32, 48, 60, 108, 240, 384, 480, 608, 972, 984, 1020,$ *and* 2688.

This is the result of a computer program implementing Theorem 2. Suppose a computer has $\boldsymbol{b}$ binary bits per word. Since one only wants to remember which n have at least one SOL, this information can be stored in a single bit. Hence at most $[n/b] + 1$ words are needed to keep track of which n have a SOL. Suppose $\chi(\boldsymbol{x}) = k \geqq 17$, then $2^k - 3k \geqq 301, 021$. Thus all SOLs for which $n \leqq 301, 020$ have $k \leqq 16$. By Theorem 3, $z < 2^{16}$. It is possible to show that for $n \geqq 55$, $x_1 < \sqrt{2n}$. Hence $x_1 < 2^9$. Thus, if $b \geqq 25$, n, z, and k can be packed into one computer word, and $x_1, \cdots, x_k$ can be packed $[b/8]$ to a computer word. So e.g., if $b = 25$, no more than six computer words are needed for the x_i. The list A of active solutions will not grow too large if the solutions are packed in this way. Finally let me comment that removing SOLs from the end of A, rather than the beginning (see Step 3) will save considerable computing time, since the stack A need not be "pushed down" after a SOL is removed. Moreover, if the last entry for each SOL is the word containing (n, z, k), then upon removing the last word of A one knows how many words were needed to store $x_1, \cdots, x_k$.

It is tempting to conjecture that there is at least one SOL for all $n > 2688$.

PROPOSITION. *There are nontrivial solutions to* (1) *whenever* $n \equiv 1 \bmod u$ *and* $n > u^2$, *or* $n \equiv 2 \bmod u^2$ *for any integer* $u > 1$.

Proof. If $n, z, \boldsymbol{x}$ is a SOL with $\chi(\boldsymbol{x}) = k$, then so is $n' = n + d \prod x_i$, $z' = z + d$, $\boldsymbol{x}' = (x_1, \cdots, x_k, 1, \cdots, 1)$ for any $d \geqq 0$. Apply this fact to the SOLs, $n = u^2 + 2$, $z = 3$, $z = 2u$, $\boldsymbol{x} = (u, 1, \cdots, 1)$ and the SOLs, $n = u^2 + 2$, $z = 3$, $\boldsymbol{x} = (u, u, 1, \cdots, 1)$.

COROLLARY. *If* (1) *has only trivial solutions, then* $n \equiv 0$ *or* $8 \bmod 12$.

[Set $u = 2, 3$.]

I take this opportunity to thank Ed Bender for many valuable discussions.

APPENDIX

(See the end of Section 6.)

N	Z	X1	X2	X3	X4	X5
3	1	3	3	3		
4	1	2	2	2	2	
5	1	4	3	3		
	4	2				
6	3	2	2			
7	1	3	2	2	2	
	2	2	2	2		
	3	3	2			
	5	2				
8	1	4	2	2	2	
9	6	2				
10	1	4	4	3		
	2	3	2	2		
	4	2	2			
	6	3				
11	2	4	2	2		
	3	3	3			
	7	2				
12		NONE				
13	1	5	4	3		
	3	4	3			
	4	3	2			
	7	3				
	8	2				
14	1	3	3	2	2	
	1	6	4	3		
	4	4	2			
	5	2	2			
15	3	2	2	2		
	9	2				
16	8	3				
17	1	2	2	2	2	2
	2	3	3	2		
	8	4				
	10	2				
18	3	4	4			
	6	2	2			
19	1	4	3	2	2	
	1	5	5	3		
	1	4	4	4		
	5	3	2			
	9	3				
	11	2				
20	2	2	2	2	2	
	4	3	3			
21	3	5	4			
	9	4				
	12	2				
22	1	5	3	2	2	
22	2	4	3	2		
	3	3	2	2		
	3	6	4			
	5	4	2			
	7	2	2			
	10	3				
23	1	6	3	2	2	
	1	6	5	3		
	4	2	2	2		
	5	5	2			
	13	2				
24		NONE				
25	1	7	5	3		
	2	5	3	2		
	4	4	3			
	6	3	2			
	10	4				
	11	3				
	14	2				
26	1	5	4	4		
	2	6	3	2		
	8	2	2			
	10	5				
27	1	3	3	3	2	
	3	4	2	2		
	3	5	5			
	15	2				
28	1	3	2	2	2	2
	1	4	4	2	2	
	4	5	3			
	12	3				
29	4	6	3			
	5	3	3			
	11	4				
	16	2				
30	1	6	6	3		
	2	3	3	3		
	3	5	2	2		
	6	4	2			
	9	2	2			
31	1	6	4	4		
	2	3	2	2	2	
	2	4	4	2		
	3	6	2	2		
	3	6	5			
	5	2	2	2		
	7	3	2			
	11	5				
	13	3				

N	Z	X1	X2	X3	X4	X5	N	Z	X1	X2	X3	X4	X5
	17	2					38	6	3	3			
32		NONE						7	4	2			
33	3	7	5					11	2	2			
	6	5	2				39	1	9	6	3		
	12	4						6	2	2	2		
	18	2						21	2				
34	1	7	4	4			40	1	6	4	2	2	
	4	3	2	2				2	4	2	2	2	
	4	4	4					16	3				
	6	6	2				41	2	4	3	3		
	10	2	2					4	5	4			
	14	3						13	5				
35	1	5	4	2	2			14	4				
	1	8	4	4				22	2				
	1	7	6	3			42	12	2	2			
	3	3	3	2			43	1	7	4	2	2	
	19	2						1	7	7	3		
36	3	2	2	2	2			2	6	4	2		
	12	5						3	7	6			
37	1	4	2	2	2	2		4	4	2	2		
	1	5	5	4				5	5	3			
	5	4	3					7	5	2			
	8	3	2					9	3	2			
	12	6						13	6				
	13	4						17	3				
	15	3						23	2				
	20	2					44	1	5	2	2	2	2
38	1	4	3	3	2			1	8	4	2	2	
	1	8	6	3			45	15	4				
	2	5	4	2				24	2				
	3	6	6										

REFERENCE

1. A. Hurwitz, *Über eine Aufgabe der unbestimmten Analysis*, Archiv. der Math. und Phys., III Reine, Bd **11** (1907), 185-196.

Received October 11, 1972.

INSTITUTE FOR DEFENSE ANALYSES