

Integer Squares with Constant Second Difference

By Duncan A. Buell

Abstract. The problem addressed is this: Do there exist nonconsecutive integers n_0, n_1, n_2, \dots , such that the second differences of the squares of the n_i are constant? Specifically, can that constant be equal to 2? A complete characterization of sequences of length four can be given. The question of whether or not sequences of length five exist is still open but the existence or nonexistence of such sequences can be described in a more algorithmic way than the simple statement of the problem.

1. Introduction. It is trivial to note that the second differences of the squares of consecutive integers is always equal to 2. Problem A addressed in this paper is this: Does there exist a sequence of five squares of nonconsecutive integers, $n_0^2, n_1^2, n_2^2, n_3^2, n_4^2$, such that the second differences are constant? Can this constant be 2? It turns out that we can sensibly break Problem A into two subproblems. Subproblem B is: Does there exist a sequence of four squares of nonconsecutive integers, $n_0^2, n_1^2, n_2^2, n_3^2$, such that the second differences are constant (and equal to 2)? Subproblem C is: Does there exist a sequence of five integers, the first three and the last of which are squares, $n_0^2, n_1^2, n_2^2, N, n_4^2$, with n_0, n_1 , and n_2 not consecutive, such that the second differences are constant (and equal to 2)?

We have not been able to solve Problem A; the question of whether or not sequences of length five exist is still open. However, a complete characterization can be given both of the solutions of Subproblem B and of the solutions of Subproblem C, in the case where the constant second difference is 2. Our characterization of solutions to Subproblem B includes all sequences of four consecutive integers as well.

Solutions of Subproblem B are not difficult to obtain by direct search; examples are (6, 23, 32, 39), (39, 70, 91, 108), (16, 87, 122, 149), (108, 157, 194, 225), (51, 148, 203, 246), and (225, 296, 353, 402) [3]. Curiously enough, this is more than a problem in recreational mathematics; the late J. Richard Büchi observed that the nonexistence of sequences of nonconsecutive integer squares of length n with second differences 2, for any n , could be used to show that there is no algorithm for deciding when systems of diophantine equations of the form

$$\sum_{j=1}^N a_{ij}x_j^2 = b_i, \quad i = 1, \dots, M,$$

Received July 24, 1986; revised December 10, 1986.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 11B83.

©1987 American Mathematical Society
0025-5718/87 \$1.00 + \$.25 per page

with integral a_i, b_i , have integer solutions [4]. In addition, an unpublished paper of Hensley discusses this problem [3]. Similar problems have been considered by Barbeau [2] and by Allison [1].

2. A Characterization of Solutions. It is not difficult to see that we have a solution to Problem A, that is, integers $n_0^2, n_1^2, n_2^2, n_3^2, n_4^2$, with second differences equal to $2L$ if and only if, for some integer a , we have

$$(2.1) \quad \begin{aligned} n_1^2 &= (n_0 + a)^2, \\ n_2^2 &= (n_0 + 2a)^2 - 2(a^2 - L), \\ n_3^2 &= (n_0 + 3a)^2 - 6(a^2 - L), \\ n_4^2 &= (n_0 + 4a)^2 - 12(a^2 - L). \end{aligned}$$

We note that a must be congruent to L modulo 2. The case $a = 1$ is seen to be the case of consecutive integers, and is generally excluded from consideration.

Writing Eqs. (2.1) as differences of squares, letting u_1, u_2, v_1, v_2, w_1 , and w_2 be integers such that $u_1u_2 = 2, v_1v_2 = 6, w_1w_2 = 12$, and letting A and B be integers such that $AB = a^2 - L$, we see that it is necessary that we be able to rewrite the latter three equations as

$$\begin{aligned} [n_0 + 2a + n_2][n_0 + 2a - n_2] &= [u_1A][u_2B], \\ [n_0 + 3a + n_3][n_0 + 3a - n_3] &= [v_1lA/k][v_2kB/l], \\ [n_0 + 4a + n_4][n_0 + 4a - n_4] &= [w_1iA/j][w_2jB/i], \end{aligned}$$

where k, l, i , and j are nonzero integers, the bracketed right-hand expressions are integers, and the brackets indicate the symbolic factoring. It is clear that k and l can be chosen relatively prime, that i and j can be chosen relatively prime, and that each of these divides A or B and thus divides $a^2 - L$.

We can now solve for the n_m :

$$(2.2) \quad \begin{aligned} n_0 &= \frac{u_1A + u_2B}{2} - 2a, \\ n_1 &= \frac{u_1A + u_2B}{2} - a, \\ n_2 &= \frac{u_1A - u_2B}{2}, \\ n_3 &= \frac{v_1lA/k - v_2kB/l}{2}, \\ n_4 &= \frac{w_1iA/j - w_2jB/i}{2}. \end{aligned}$$

We will also need

$$(2.3) \quad n_0 + 3a = \frac{v_1lA/k + v_2kB/l}{2}.$$

These equations, together with the assumptions on the variables, are clearly necessary, and also sufficient, for the existence of integral squares with second differences $2L$.

Combining (2.2) and (2.3) and reducing, we obtain

$$A(u_1 - v_1l/k) + 2a + B(u_2 - v_2k/l) = 0.$$

We note that $u_1 - v_1l/k = 0$ leads to a finite number of possible (small) values a , which can be checked and shown to lead to no solutions to Subproblem B. We may thus assume that $u_1 - v_1l/k$ is not zero, multiply by A and then solve for A to obtain as a necessary condition for solutions to Problem A or Subproblem B that

$$(2.4) \quad A = \frac{-a \pm \sqrt{a^2 - (u_1 - v_1l/k)(u_2 - v_2k/l)(a^2 - L)}}{u_1 - v_1l/k}.$$

To have rational values n_m , we must have a rational value for A , which requires that the radicand be a rational square. Since k and l are relatively prime and divide $a^2 - L$, the radicand is actually an integer and so must be an integral square. If we denote this by t^2 and rearrange terms, we have as a necessary and sufficient condition for the existence of solutions to Problem A or to Subproblem B that we have solutions, with the previously-mentioned conditions on the variables, to

$$(2.5) \quad L \cdot (ku_1 - v_1l)(lu_2 - v_2k) = klt^2 - a^2\{kl - (ku_1 - v_1l)(lu_2 - v_2k)\}.$$

In the notation of binary quadratic forms, we are concerned with the representation of

$$(2.6) \quad L \cdot (ku_1 - v_1l)(lu_2 - v_2k)$$

by the binary quadratic form

$$(kl, 0, kl - (ku_1 - v_1l)(lu_2 - v_2k))$$

of discriminant

$$4 \cdot k \cdot l \cdot (kl - (ku_1 - v_1l)(lu_2 - v_2k)).$$

The above reduction applies to Subproblem B; a similar treatment shows that solutions to Subproblem C exist if and only if integers s and a exist so that the equation

$$(2.7) \quad L \cdot (ju_1 - w_1i)(iu_2 - w_2j) = jis^2 - a^2\{4ji - (ju_1 - w_1i)(iu_2 - w_2j)\}$$

holds, with the additional proviso that the derived values of A and B must be integral.

3. The Solutions to Subproblem B, for $L = 1$. In the special case $L = 1$, we know that one representation (2.5) exists, with $t^2 = a^2 = 1$ (this is the solution that yields consecutive integers). Thus, it is not the existence but merely the number of representations and the requirement that A and B be integral which are in question.

Considering symmetry and coupled sign changes, we may reduce (2.5) to the case $u_1 = 1$ and $u_2 = 2$, with v_1 and v_2 both positive.

Now consider the case $v_1 = 1, v_2 = 6$. We have, for Subproblem B,

$$(3.1) \quad n_0 = A/2 + B - 2a,$$

$$(3.2) \quad n_1 = A/2 + B - a,$$

$$(3.3) \quad n_2 = A/2 - B,$$

$$(3.4) \quad n_3 = \frac{lA}{2k} - \frac{3kB}{l},$$

$$(3.5) \quad A = \frac{k(a-t)}{l-k}, \quad B = \frac{-l(a+t)}{2(l-3k)}, \quad 1 = \frac{(-kl)(t^2-a^2)}{2(l-k)(l-3k)} + a^2.$$

We note that the choice of factorings for A and B follows from the convention of (2.4). Elementary but tedious analysis and the requirement that A and B be integral allow us to conclude that a and t are odd, that A is even, and $(a-t)/(l-k)$ and $(a+t)/(l-3k)$ are even integers. Working further, it can also be seen that k and l must be of opposite parity. A tedious but direct examination of all the other cases for v_1 and v_2 then shows that each reduces to this one. We have the following theorem.

THEOREM. *There exist nonconsecutive integers n_0, n_1, n_2, n_3 , whose squares have second differences $n_3^2 - 2n_2^2 + n_1^2 = n_2^2 - 2n_1^2 + n_0^2 = 2$ if and only if n_0, n_1, n_2 , and n_3 are given by*

$$(3.6) \quad \begin{aligned} n_0 &= xk - yl - 2a = x(-2l + 3k) + y(-3l + 6k), \\ n_1 &= xk - yl - a = x(-l + 2k) + y(-2l + 3k), \\ n_2 &= xk + yl, \\ n_3 &= xl + 3yk, \end{aligned}$$

where k, l, x , and y are integers, k and l are relatively prime and of opposite parity such that

$$(3.7) \quad 1 = x^2(l-k)^2 + xy(2l^2 - 6kl + 6k^2) + y^2(l-3k)^2.$$

At issue in the solution of Subproblem B is thus whether the binary quadratic form

$$(3.8) \quad ((l-k)^2, 2l^2 - 6kl + 6k^2, (l-3k)^2)$$

of discriminant

$$D = 4 \cdot k \cdot l \cdot (l-2k) \cdot (2l-3k)$$

represents 1, that is, is equivalent to the principal form of discriminant D . The form (3.8) is equivalent to the forms

$$(3.9) \quad (4k^2 - 2kl, +4k^2 - 2kl, (l-k)^2)$$

and

$$(3.10) \quad (4l^2 - 6kl, 4l^2 - 6kl, (l-k)^2),$$

both of which are ambiguous. Since the product $(2kl + 4k^2)(6kl - 4l^2)$ is the discriminant, exactly one of the two lead coefficients is smaller in magnitude than the square root of $|D|$.

It can be seen that for negative discriminants we have a representation of 1 if and only if $(l - k)^2 = 1$, and thus that all of the cases of negative discriminant are degenerate, providing only sequences of four consecutive integers.

If D is positive, though, exactly one of

$$(4k^2 - 2kl, K(4k^2 - 2kl), *)$$

and

$$(4l^2 - 6kl, K(4l^2 - 6kl), *)$$

is a reduced ambiguous form, for some suitable K and the third coefficient chosen to fit. We can actually be more specific. It is always true that

$$-4k^2 + 2kl < 4l^2 - 6kl.$$

Both expressions above, however, must be of like sign for the case of positive discriminant, since the product is the discriminant. Therefore, of the two forms listed above, that form is reduced which has a *negative* lead coefficient.

Remarks. 1. A form represents 1 if and only if it is equivalent to the principal form. There are, however, in any ambiguous cycle exactly two ambiguous forms. Thus, if the reduced form above is equivalent to the principal form, it must be the other ambiguous form (besides the principal form) in the principal cycle. Among other things, this implies that the form $(-1, *, **)$ is *not* in the principal cycle, that the negative Pellian equation $X^2 - DY^2 = -1$ or -4 is not solvable, and that the principal cycle has length equal to twice an odd number (since the signs of leading coefficients alternate in sign).

2. The existence of solutions to Subproblem B reduces to the known difficult problem of determining whether or not an indefinite binary quadratic form is equivalent to the principal form. The actual solutions, once existence is assured, come from the recurrence relationships which are Lucas sequences obtained from the coefficients in the fundamental solutions of $X^2 - DY^2 = 4$.

4. The Solutions to Subproblem C and to Problem A, for $L = 1$. The different possible cases of Subproblem B all reduce to the same case, as evidenced by the above theorem. This is not true of Subproblem C; we obtain seven similar but distinct possible cases, of which one is this:

Solutions for Subproblem C exist if we have integers $p, q, i,$ and j such that

$$4 = p^2(i - j)^2 + pq(16ij + 4(i - j)(i - 6j)) + 4q^2(i - 6j)^2$$

with i and j both odd, i relatively prime to 3, and

$$n_0 = pj - 2qi - 2a,$$

$$n_1 = pj - 2qi - a,$$

$$n_2 = pj + 2qi,$$

$$n_4 = pi + 12qj.$$

In this case, we have $A = 2pj, B = 2qi,$ and $a^2 = 1 - 4pqij$. In order to solve Problem A, the A and B values must match, up to exchange and signs, from Subproblem B to Subproblem C. That is, we must have $(xk, yl) = \pm(pj, 2qi)$ or $\pm(2qi, pj)$. It would seem improbable, then, that solutions to Problem A exist. Not only must both representations (2.5) and (2.7) exist, but the value of a in both representations must be the same.

5. Computations and Examples. We have computed all discriminants $D = 4 \cdot k \cdot l(l - 2k)(2l - 3k)$ with $1 \leq k, l \leq 1000$, produced the relevant ambiguous forms, and determined whether or not the forms represent 1. It is easy to show that these limits provide for a complete list of the relevant positive discriminants less than 1341348008. There are only 5816 such discriminants, 42 of which appear twice. (For example, 14560 appears with (k, l) equal to $(4, 13)$ and to $(7, 4)$.) Of these 5858 instances, 2668 of the forms represent 1, and 3190 do not. Since in most instances, the actual sequences of n_m are rather large, we have not computed these.

In computing the discriminants and determining representability, it is clear that the cases $l - k = \pm 1$ and $l - 3k = \pm 1$ are special, as we can trivially guarantee the existence of representations. It is easy to show that only $l - k = -1$ need be considered, and then that the solutions and sequences can be explicitly given by recurrences. Using the standard theory of representation by forms, or equivalently of Lucas sequences, one can show that the following are true.

Example 1. If $l - k = -1$, then $D = 4\Delta = 4(k - 1)(k)(k + 1)(k + 2) = 4(k^4 + 2k^3 - k^2 - 2k)$. The representation (3.7) is

$$1 = x^2 + (2k^2 + 2k + 2)xy + (2k + 1)^2y^2$$

which is equivalent to the representation

$$1 = (x + 3y)^2 + (2k^2 + 2k - 4)(x + 3y)y + (-2k^2 - 2k + 4)y^2$$

by the principal reduced form $(1, 2k^2 + 2k - 4, -2k^2 - 2k + 4)$ of discriminant D . The negative Pell equation $X^2 - DY^2 = -4$ is not solvable, and all solutions of $X^2 - DY^2 = 4$ are

$$\frac{X + Y\sqrt{D}}{2} = \left(\frac{2T + U\sqrt{D}}{2} \right)^n,$$

or the conjugates, for positive or negative exponents n , with $T = k^2 + k - 1$ and $U = 1$. Up to sign, all values of $x_i, y_i, a_i, n_{0,i}, n_{1,i}, n_{2,i}, n_{3,i}$, involved in or derived from the representation (3.7) are obtained from

$x_0 = -1$	$x_1 = -2k^2 - 2k$	
$y_0 = 0$	$y_1 = 1$	
$a_0 = 1$	$a_1 = 2k^2 - 1$	$a_i = -x_i - (2k + 1)y_i$
$n_{0,0} = -k - 2$	$n_{0,1} = -2k^3 - 6k^2 - k + 3$	$n_{0,i} = (k + 2)x_i + (3k + 3)y_i$
$n_{1,0} = -k - 1$	$n_{1,1} = -2k^3 - 4k^2 - k + 2$	$n_{1,i} = (k + 1)x_i + (k + 2)y_i$
$n_{2,0} = -k$	$n_{2,1} = -2k^3 - 2k^2 + k - 1$	$n_{2,i} = kx_i + (k - 1)y_i$
$n_{3,0} = -k + 1$	$n_{3,1} = -2k^3 + 5k$	$n_{3,i} = (k - 1)x_i + 3ky_i$

with each of these satisfying the recurrence $\alpha_{i+2} = 2T\alpha_{i+1} - \alpha_i$ for $i \geq 0$.

We present in Tables 1 and 2 examples of this case.

TABLE 1. $(l, k, T, \Delta) = (1, 2, 5, 24)$

x	y	a	n_0	n_1	n_2	n_3
-1	0	1	-4	-3	-2	-1
-12	1	7	-39	-32	-23	-6
-119	10	69	-386	-317	-228	-59
-1178	99	683	-3821	-3138	-2257	-584
-11661	980	6761	-37824	-31063	-22342	-5781
-115432	9701	66927	-374419	-307492	-221163	-57226

TABLE 2. $(l, k, T, \Delta) = (2, 3, 11, 120)$

x	y	a	n_0	n_1	n_2	n_3
-1	0	1	-5	-4	-3	-2
-24	1	17	-108	-91	-70	-39
-527	22	373	-2371	-1998	-1537	-856
-11570	483	8189	-52054	-43865	-33744	-18793
-254013	10604	179785	-1142817	-963032	-740831	-412590
-5576716	232805	3947081	-25089920	-21142839	-16264538	-9058187

Example 2. If $l - 3k = \epsilon = \pm 1$, then

$$D = 4\Delta = 4(9k^4 + 18\epsilon k^3 + 11k^2 + 2\epsilon k);$$

the negative Pell equation $X^2 - DY^2 = -4$ is not solvable; the fundamental values of T and U are $9k^2 + 9\epsilon k + 1$ and 3, respectively; up to sign, all values of $x_i, y_i, a_i, n_{0,i}, n_{1,i}, n_{2,i}, n_{3,i}$, involved in or derived from the representation (3.7) are obtained from

$$\begin{aligned} x_0 &= 0 & x_1 &= 3 \\ y_0 &= 1 & y_1 &= -2 \\ a_0 &= \epsilon & a_1 &= 6k + \epsilon & a_i &= (2k + \epsilon)x_i + \epsilon y_i \\ n_{0,0} &= -3k - 3\epsilon & n_{0,1} &= -3k & n_{0,i} &= (-3k - 2\epsilon)x_i + (-3k - 3\epsilon)y_i \\ n_{1,0} &= -3k - 2\epsilon & n_{1,1} &= 3k + \epsilon & n_{1,i} &= (-k - \epsilon)x_i + (-3k - 2\epsilon)y_i \\ n_{2,0} &= 3k + \epsilon & n_{2,1} &= -3k - 2\epsilon & n_{2,i} &= kx_i + (3k + \epsilon)y_i \\ n_{3,0} &= 3k & n_{3,1} &= 3k + 3\epsilon & n_{3,i} &= (3k + \epsilon)x_i + 3ky_i \end{aligned}$$

with each of these satisfying the recurrence $\alpha_{i+2} = 2T\alpha_{i+1} - \alpha_i$ for $i \geq 0$. We note that for $\epsilon = +1$ we have

$$\begin{aligned} n_{0,2} &= -2(3k + 1)^3 + 5(3k + 1), \\ n_{1,2} &= 2(3k + 1)^3 + 2(3k + 1)^2 - (3k + 1) + 1, \\ n_{2,2} &= -2(3k + 1)^3 - 4(3k + 1)^2 - (3k + 1) + 2, \\ n_{3,2} &= 2(3k + 1)^3 + 6(3k + 1)^2 + (3k + 1) - 3, \end{aligned}$$

and for $\epsilon = -1$ we have

$$\begin{aligned} n_{0,2} &= -2(3k - 2)^3 - 6(3k - 2)^2 - (3k - 2) + 3, \\ n_{1,2} &= 2(3k - 2)^3 + 4(3k - 2)^2 + (3k - 2) - 2, \\ n_{2,2} &= -2(3k - 2)^3 - 2(3k - 2)^2 + (3k - 2) - 1, \\ n_{3,2} &= 2(3k - 2)^3 - 5(3k - 2), \end{aligned}$$

so that these examples have already appeared in Special Case 1. We present in Tables 3 and 4 examples of this case.

TABLE 3a. $(l, k, T, \Delta, \epsilon) = (5, 2, 19, 40, -1)$

x	y	a	n_0	n_1	n_2	n_3
0	1	-1	-3	-4	5	6
3	-2	11	-6	5	-4	3
114	-77	419	-225	194	-157	108
4329	-2924	15911	-8544	7367	-5962	4101
164388	-111035	604199	-324447	279752	-226399	155730

TABLE 3b. $(l, k, T, \Delta, \epsilon) = (7, 2, 55, 336, 1)$

x	y	a	n_0	n_1	n_2	n_3
0	1	1	-9	-8	7	6
3	-2	13	-6	7	-8	9
330	-221	1429	-651	778	-887	984
36297	-24308	157177	-71604	85573	-97562	108231

TABLE 4a. $(l, k, T, \Delta, \epsilon) = (11, 4, 109, 1320, -1)$

x	y	a	n_0	n_1	n_2	n_3
0	1	-1	-9	-10	11	12
3	-2	23	-12	11	-10	9
654	-437	5015	-2607	2408	-2191	1950
142569	-95264	1093247	-568314	524933	-477628	425091

TABLE 4b. $(l, k, T, \Delta, \epsilon) = (13, 4, 181, 3640, 1)$

x	y	a	n_0	n_1	n_2	n_3
0	1	1	-15	-14	13	12
3	-2	25	-12	13	-14	15
1086	-725	9049	-4329	4720	-5081	5418
393129	-262448	3275713	-1567086	1708627	-1839308	1961301

Example 3. The careful reader will probably notice that some of the sequences from the introduction, (6, 23, 32, 39), (39, 70, 91, 108), (108, 157, 194, 225), (225, 296, 353, 402), appear in our tables, and that these have the rather curious feature that the last member of the previous sequence is the first member of the next. Indeed, if $l - k = -1$, as in Special Case 1 above, it is easy to see that as functions of k , $n_{0,1}(k) = n_{3,1}(k + 1)$, and the rest of the relations follow.

Example 4. The next simplest case after those of Examples 1 and 2 is the case $(l - k)^2 = 9$. We can choose $l - k = -3$, and the representation problem (3.7) is to find x and y such that

$$1 = 9x^2 + (2k^2 + 6k + 18)xy + (2k + 3)^2y^2$$

$$= 9(x + 3y)^2 + (2k^2 + 6k - 36)(x + 3y)y + (-2k^2 - 6k + 36)y^2.$$

For the case $l = 1, k = 4$ we present in Table 5 the first few solutions. In this case, $D = 1320$ and the fundamental solution of $T^2 - 280U^2 = 1$ has $T = 251$ and $U = 15$. The previous recurrences $\alpha_{n+2} = 2T\alpha_{n+1} - \alpha_n$ again hold for all seven variables.

TABLE 5

x	y	a	n_0	n_1	n_2	n_3
6	-1	-7	39	32	23	-6
3021	-506	-3497	19584	16087	11578	-3051
1516536	-254011	-1755487	9831129	8075642	5812133	-1531596

6. A Final Remark. We have not been able to determine a way to obtain distinct sequences as solutions to Subproblem B. We make, however, the following observation. The polynomial in the representation of (3.7),

$$(l - k)^2x^2 + (2l^2 - 6kl + 6k^2)xy + (l - 3k)^2y^2,$$

is identical to the polynomial

$$(x + y)^2l^2 + (-2x^2 - 6xy - 6y^2)lk + (x + 3y)^2k^2,$$

and the values of a and the n_i are unchanged (up to sign) by the exchange of l for x and k for $-y$. Considering this as a binary quadratic form in variables x and y with coefficients that are functions of l and k , we have discriminant $D = 4kl(l - 2k)(2l - 3k)$. Exchanging the roles of coefficients and variables, we have discriminant $D' = -4xy(x + 2y)(2x + 3y)$. We note that $(2l^2 - 6kl + 6k^2)$ is always positive, so that x and y are of opposite sign and both discriminants are positive. In enumerating solutions to Subproblem B, we need only consider the smaller discriminant, which forces, among other things, the values of the variables x and y to be not both larger in magnitude than the square root of D . Since the values of x and y become too large after only one iteration of the Lucas recurrence, we see that *only* the fundamental representation of (3.7) can produce distinct sequences. Among other things, this implies that all the sequences in Tables 1 through 4 can be discarded, as they will appear as fundamental solutions later in an enumeration.

Supercomputing Research Center
4380 Forbes Boulevard
Lanham, Maryland 20706

1. D. ALLISON, "On square values of quadratics," *Math. Proc. Cambridge Philos. Soc.*, v. 99, 1986, pp. 381–383.
2. E. J. BARBEAU, "Numbers differing from consecutive squares by squares," *Canad. Math. Bull.*, v. 28, 1985, pp. 337–342.
3. DOUGLAS C. HENSLEY, "Sequences of squares with second difference of 2 and a problem of logic," unpublished.
4. LEONARD LIPSCHITZ, personal correspondence.