QUADRATIC FORMS

A mathematical vignette

Ed Barbeau, University of Toronto

1. Difference of squares. *Question: What positive integers can be written as the difference of two squares?* This is a good problem to ask students to explore. After a little experimentation, there is a good chance that students will be able to guess the situation, namely that as long as the integer is not twice an odd integer, *i.e.* not congruent to 2 modulo 4, such a representation is possible. Some students may discover that one can cover all the odd integers by taking the difference between two consecutive squares, since $(x+1)^2 - x^2 = 2x+1$, and that multiples of 4 can be obtained by doubling the integers used for the odds.

A key observation is that any difference of squares can be written as a product $x^2 - y^2 = (x-y)(x+y)$. A second key obervation is that $x+y$ and $x-y$ have the same parity. So if a representation of a number as the difference of two squares is possible, then that number can be written as the product of two numbers with the same parity. Is the converse true?

Suppose that $n = ab$, where a and b have the same parity. Then, with $a > b$, we can show that the system $x + y = a; x - y = b$ has nonnegative integer solutions x and y and so $n = x^2 - y^2$ has the desired representation. This suggests additional questions to be answered: given n, how many distinct representations are possible? Are there any numbers for which such a representation is unique? Which multiples of 4 can be written as the difference of two odd squares?

Another tool to tackle the same question is modular arithmetic. While some older students find this difficult, it may be that the difficulty arises in part from its unfamiliarity with respect to what they have learned and been working with for years. Younger students may actually be more comfortable with it. In the present situation, modulo 4, we find that any square is congruent to 0 or 1, so that it is impossible for two squares to differ by a numer that is congruent to 2 modulo 4. A finer analysis is possible if we work modulo 8, for odd squares are always congruent to 1 modulo 8, while even squares are congruent to 0 or 4. An odd multiple of 4 cannot be the difference of two odd squares, but every multiple of 8 can. In fact, for any integer x,

$$(2x+1)^2 - (2x-1)^2 = 8x.$$

We can carry the investigation further and examine the structure of the set of numbers for which such a representation is possible. We can show that this set is closed under multiplication. For if $m = a^2 - b^2 = (a+b)(a-b)$ and $n = c^2 - d^2 = (c+d)(c-d)$, then

$$\begin{aligned} mn &= (a+b)(c+d)(a-b)(c-d) = (ac+bd+ad+bc)(ac+bd-ad-bc) \\ &= (ac+bd)^2 - (ad+bc)^2. \end{aligned}$$

We can write $1 = 1^2 - 0^2$, $4 = 2^2 - 0^2$, and any odd prime can be written as the difference of two squares, so that any number that is the product of a power of 4 and any number of odd primes can be written as the difference of two squares. This includes all numbers that are not twice an odd number and therefore corroborates our earlier finding.

Another possible avenue of investigation is to fix values of y and check out what happens when you multiply numbers of the form $x^2 - y^2$ where x takes different values. For example, the numbers $x^2 - 1$ are 0, 3, 8, 15, 24, 35, 48, $\cdots$, and it appears that the product of any two consecutive numbers in the sequences is in the same sequence. Can this be demonstrated? Empirically, students might arrive at the equation

$$[(x+1)^2 - 1][x^2 - 1] = [x(x+1) - 1]^2 - 1.$$

2. Sum of squares. We can now move to numbers of the form $x^2 + y^2$, where x and y are integers. The underlying theory is more complicated and formalizing this is beyond the range of most school students.

However, they can still be invited to make conjectures. One can address the question as to whether the set of such numbers is closed under multiplication. To discover this, one can follow a strategy similar to that in Section 1, if we have recourse to imaginary numbers. For

$$m^2 + n^2 = (m + ni)(m - ni),$$

so

$$\begin{aligned}(a^2 + b^2)(m^2 + n^2) &= (a + bi)(m + ni)(a - bi)(m - ni)\\ &= [(am - bn) + (an + bm)i][(am - bn) - (an + bm)i]\\ &= (am - bn)^2 + (an + bm)^2.\end{aligned}$$

Looking at the situation modulo 4, we find that no sum of squares can be congruent to 3 modulo 4. This includes in particular primes that are 1 less than a multiple of 4. However, it turns out that 2 and all primes that are 1 greater than a multiple of 4 are the sums of two squares. This is not an easy result to obtain, but there are a number of approaches. However, through investigation students may discover empirically that a number is the sum of two squares if and only if in its decomposition into prime factors, all primes congruent to 3 modulo 4 must occur to an even power.

There is an algorithm that will produce representations of primes congruent to 1 modulo 4 as the sum of two squares. Let (x, y, z) be a triple of integers. Define

$$U(x, y, z) = \begin{cases} (x - y - z, y, 2y + z), & \text{if } x > y + z\\ (y, x, -z), & \text{if } x < y + z.\end{cases}$$

(The possibility $x = y + z$ will not occur in the cases we are considering.) If $(X, Y, Z) = (x, y, z)$, then $4XY + Z^2 = 4xy + z^2$. Thus, $4xy + z^2$ is invariant under the transformation U. Suppose that $p = 4k + 1$ is prime. Start with $(k, 1, 1)$ and apply the operator U repeatedly. Eventually, you will get something of the form (u, u, v). The value of the invariant is $4k + 1 = 4u^2 + v^2$, so that $p = (2u)^2 + v^2$.

For example, if $p = 73$, then $k = 18$. Starting with $(18, 1, 1)$ leads us in turn to $(16, 1, 3)$, $(12, 1, 5)$, $(6, 1, 7)$, $(1, 6, -7)$, and in due course to $(4, 4, 3)$. Continuing past this stage will take us back to where we started, and looking at the result may give some idea as to why the method works.

As in Section 2, we can take $y = 1$ and investigate whether the product of $x^2 + 1$ for two consecutive integral values of x has the same form.

3. Pell's equation. A Pell's equation has the form

$$x^2 - dy^2 = k$$

where d is a positive nonsquare integer and k is an arbitrary integers. By expressing $x^2 - dy^2 = (x + y\sqrt{d})(x - y\sqrt{d})$, we can see that the set of k for which the equation is solvable is closed under multiplication. Furthermore, in this case, there are infinitely many ways to represent 1 in the form $x^2 - dy^2$. We can look at the solution with the small values of x and y, say $(x, y) = (u, v)$, and then get other solutions by taking the coefficients of 1 and d in the powers $(u + v\sqrt{d})^n$ where n is a positive integer. Thus, for example, we note that $(u + v\sqrt{d})^2 = (u^2 + v^2 d) + 2uv\sqrt{d}$, so that, from $u^2 - dv^2 = 1$, we can deduce

$$(u^2 + dv^2)^2 - d(2uv) = (u^2 - dv^2)^2 = 1.$$

The theory tells us that, for every positive nonsquare d, there is a solution of $x^2 - dy^2 = 1$ in positive integers x, y, and that every solution can be obtained from this using powers of surds. Students might investigate the truth of this for particular values of d. Some values of d, such as $d = 61$ present a tough situation, but East Indian mathematicians managed to find a solution over a millenium ago. However, finding solutions is a nice exercise for $d < 30$.

4. A general quadratic result. Students can look at numbers representable by other quadratic forms, such as x^2+dy^2 where d is a positive integer and x^2+xy+y^2. One interesting fact that might emerge is that, when the coefficient of x^2 is 1 and y is set equal to 1, we get a quadratic of the form $f(x) = x^2+bx+c$ (with integers b and c), and for each integer x, $f(x)f(x+1) = f(z)$ for some integer z.

Let us look at the example $f(x) = x^2 + x + 1$. We can plough through a lot of heavy algebra to verify this is this case, but if we look at the situation more conceptually, we can readily verify our assertion. Note that $f(x-1) = x^2 - x + 1$. The left side of the equation asks us to consider the value of f at two consecutive integers, and generically we might as well take them to be $x-1$ and x. We find that

$$f(x-1)f(x) = [(x^2+1)-x][(x^2+1)+x] = (x^2+1)^2 - x^2 = (x^2)^2 + x^2 + 1 = f(x^2).$$

We can make another reduction to simplify things. Suppose we think of the result as having to show that, for each integer m, $f(m)f(m+1) = f(r)$ for some integer r. Let us define the "shifted polynomial" $g(x) = f(x+m)$ so that $g(0) = f(m)$ and $g(1) = f(m+1)$. Then we have to show that $g(0)g(1) = g(s)$ for some integer s. What does $g(x)$ look like? It has the same leading coefficient, 1, as f, so that

$$g(x) = x^2 + ux + v$$

.

Now let us compute:

$$g(0)g(1) = v(1+u+v) = v + uv + v^2 = v^2 + uv + v = g(v) = f(m+v).$$

Notice the dual role played by v here, as the argument in the polynomial and as the constant coefficient of the polynomial. This is a little subtle, but the student who becomes proficient at mathematics will have to become accustomed to such switches in standpoint.

A more general version of this result was given to some students in a competition, and here are a few of the solutions. Notice how the various solutions play off the different ways in which a quadratic can be represented.

Let $f(x)$ be a quadratic polynomial. Prove that there exist quadratic polynomials $g(x)$ and $h(x)$ for which

$$f(x)f(x+1) = g(h(x)) \ ,$$

Solution 1. [A. Remorov] Let $f(x) = a(x-r)(x-s)$. Then

$$\begin{aligned} f(x)f(x+1) &= a^2(x-r)(x-s+1)(x-r+1)(x-s) \\ &= a^2(x^2+x-rx-sx+rs-r)(x^2+x-rx-sx+rs-s) \\ &= a^2[(x^2-(r+s-1)x+rs)-r][(x^2-(r+s-1)x+rs)-s] \\ &= g(h(x)) \ , \end{aligned}$$

where $g(x) = a^2(x-r)(x-s) = af(x)$ and $h(x) = x^2 - (r+s-1)x + rs$.

Solution 2. Let $f(x) = ax^2+bx+c$, $g(x) = px^2+qx+r$ and $h(x) = ux^2+vx+w$. Then

$$f(x)f(x+1) = a^2x^4 + 2a(a+b)x^3 + (a^2+b^2+3ab+2ac)x^2 + (b+2c)(a+b)x + c(a+b-c)$$

$$\begin{aligned} g(h(x)) &= p(ux^2+vx+w)^2 + q(ux+vx+w) + r \\ &= pu^2x^4 + 2puvx^3 + (2puw+pv^2+qu)x^2 + (2pvw+qv)x + (pw^2+qw+r) \ . \end{aligned}$$

Equating coefficients, we find that $pu^2 = a^2$, $puv = a(a+b)$, $2puw + pv^2 + qu = a^2 + b^2 + 3ab + 2ac$, $(b+2c)(a+b) = (2pw+q)v$ and $c(a+b+c) = pw^2 + qw + r$. We need to find just one solution of this system. Let $p = 1$ and $u = a$. Then $v = a + b$ and $b + 2c = 2pw + q$ from the second and fourth equations. This yields the third equation automatically. Let $q = b$ and $w = c$. Then from the fifth equation, we find that $r = ac$.

Thus, when $f(x) = ax^2 + bx + c$, we can take $g(x) = x^2 + bx + ac$ and $h(x) = ax^2 + (a+b)x + c$.

Solution 3. [S. Wang] Suppose that

$$f(x) = a(x+h)^2 + k = a(t - (1/2))^2 + k \ ,$$

where $t = x + h + \frac{1}{2}$. Then $f(x+1) = a(x+1+h)^2 + k = a(t + (1/2))^2 + k$, so that

$$\begin{aligned} f(x)f(x+1) &= a^2(t^2 - (1/4))^2 + 2ak(t^2 + (1/4)) + k^2 \\ &= a^2t^4 + \left(-\frac{a^2}{2} + 2ak \right) t^2 + \left(\frac{a^2}{16} + \frac{ak}{2} + k^2 \right) \ . \end{aligned}$$

Thus, we can achieve the desired representation with $h(x) = t^2 = x^2 + (2h+1)x + \frac{1}{4}$ and $g(x) = a^2x^2 + (\frac{-a^2}{2} + 2ak)x + (\frac{a^2}{16} + \frac{ak}{2} + k^2)$.

Solution 4. [V. Krakovna] Let $f(x) = ax^2 + bx + c = au(x)$ where $u(x) = x^2 + dx + e$, where $b = ad$ and $c = ae$. If we can find functions $v(x)$ and $w(x)$ for which $u(x)u(x+1) = v(w(x))$, then $f(x)f(x+1) = a^2v(w(x))$, and we can take $h(x) = w(x)$ and $g(x) = a^2v(x)$.

Define $p(t) = u(x+t)$, so that $p(t)$ is a monic quadratic in t. Then, noting that $p''(t) = u''(x+t) = 2$, we have that

$$p(t) = u(x+t) = u(x) + u'(x)t + \frac{u''(x)}{2}t^2 = t^2 + u'(x)t + u(x) \ ,$$

from which we find that

$$\begin{aligned} u(x)u(x+1) &= p(0)p(1) = u(x)[u(x) + u'(x) + 1] \\ &= u(x)^2 + u'(x)u(x) + u(x) = p(u(x)) = u(x + u(x)) \ . \end{aligned}$$

Thus, $u(x)u(x+1) = v(w(x))$ where $w(x) = x + u(x)$ and $v(x) = u(x)$. Therefore, we get the desired representation with

$$h(x) = x + u(x) = x^2 + \left(1 + \frac{b}{a}\right)x + \frac{c}{a}$$

and

$$g(x) = a^2v(x) = a^2u(x) = af(x) = a^2x^2 + abx + ac \ .$$

Comment. The second solution can also be obtained by looking at special cases, such as when $a = 1$ or $b = 0$, getting the answer and then making a conjecture.

5. When is a polynomial a composition of other polynomials? One Ottawa, ON high school student, James Rickards, carried the result much further and actually got a note published in the *American Mathematical Monthly*, the flagship jounral of the *Mathematical Association of America.* When $f(x) = x^2 + bx + c$ is a monic, quadratic polynomial, it is straightforward to check that $f(0)f(1)) = f(c) = f(f(0))$. By a translation of the variable, it follows that $f(x)f(x+1) = f(x + f(x))$ identically in x. This can be generalized to general quadratic polynomials to obtain $f(x)f(x+1) = g(h(x))$ where $f(x) = ax^2 + bx + c$, $g(x) = x^2 + bx + ac$ and $h(x) = ax^2 + (a+b)x + c$.

There is more significance to this result than a chance computation. If the roots of the polynomial $f(x)$ are r and s, then the roots of $f(x+1)$ are $r - 1$ and $s - 1$, and we note that $f(x)f(x+1)$ is a quartic

polynomial for which the sum of two of its roots is equal to the sum of the other two. This is the critical observation. For, it turns out that a quartic polynomial can be expressed as the composite of two quadratics *if and only if* two of its roots have the same sum as the other two. We generalize this to polynomials of higher degree.

For a set $X = \{x_1, x_2, \cdots, x_n\}$, we denote by $\sigma_k(X)$ the kth symmetric function of the variables x_i, the sum of $\binom{n}{k}$ products of k of the x_i:

$$\sigma_k(X) = \sum\{x_{\alpha_1}x_{\alpha_2}\cdots x_{\alpha_k} : 1 \le \alpha_1 < \alpha_2 < \cdots < \alpha_k \le n\}$$

for $1 \le k \le n$. We define $\sigma_0(X) = 1$. We recall that if $f(x) = g(h(x))$ for polynomials f, g, h, then the degree of f is the product of the degrees of g and h.

The criterion for composition. Suppose that $f(x)$ is a polynomial with complex coefficients of degree mn where m and n are integers exceeding 1. We begin by assuming that the leading coefficient of $f(x)$ is 1 and generalize later.

Proposition 1. The monic polynomial $f(x)$ can be written as the composite $g(h(x))$ of a polynomial h of degree n and g of degree m if and only if R can be partitioned into m sets S_1, S_2, $\cdots$, S_m, each with n elements (not necessarily distinct) such that, for each integer j with $0 \le j \le n-1$,

$$\sigma_j(S_1) = \sigma_j(S_2) = \cdots = \sigma_j(S_m) \ .$$

Proof. Suppose that the set R of roots of f can be partitioned as indicated. Let $R = \{r_1, r_2, \cdots, r_{mn}\}$ be the set of roots of f, each listed as often as its multiplicity and indexed so that $S_1 = \{r_1, r_2, \cdots, r_n\}$, $S_2 = \{r_{n+1}, r_{n+2}, \cdots, r_{2n}\}$, $\cdots$, $S_m = \{r_{nm-n+1}, r_{mn-n+2}, \cdots, r_{mn}\}$. For $1 \le i \le m$, let

$$y_i(x) = (x - r_{(i-1)n+1})(x - r_{(i-1)n+2})\cdots(x - r_{in})$$

be the monic polynomial whose roots are the elements of S_i. Then, if i and j are distinct positive ingers not exceeding m, the condition that the corresponding symmetric functions of S_i and S_j are equal except for the nth implies that $y_i(x) - y_j(x)$ is a constant. Define $z_i = y_1(x) - y_i(x)$ for $1 \le i \le m$.

Let $h(x) = y_1(x)$ and

$$g(x) = (x - z_1)(x - z_2)\cdots(x - z_m) \ .$$

Then

$$g(h(x)) = (y_1(x) - z_1)(y_1(x) - z_2)\cdots(y_1(x) - z_m) = y_1(x)y_2(x)\cdots y_m(x) = \prod_{i=1}^{mn}(x - r_i) = f(x) \ .$$

Now we prove that the condition on the roots of f is necessary. Suppose that we are given polynomials g and h of respective degrees m and n for which $f(x) = g(h(x))$. Let

$$g(x) = (x - t_1)(x - t_2)\cdots(x - t_m) \ .$$

For each positive integer not exceeding m, let

$$u_i(x) = h(x) - t_i = (x - r_{(i-1)n+1})(x - r_{(i-1)n+2})\cdots(x - r_{in}) \ ,$$

say, where each linear factor is listed as often as the multiplicity of the corresponding root of u_i. Then

$$\begin{aligned} f(x) = g(h(x)) &= (h(x) - t_1)(h(x) - t_2)\cdots(h(x) - t_m) \\ &= (x - r_1)(x - r_2)\cdots(x - r_{mn}) \ , \end{aligned}$$

so that the r_j are the roots of f.

For each index i with $1 \leq i \leq m$, $u_i(x) = h(x) - t_i$ so that all the coefficients of u_i except the constant are independent of i. It follows that all the symmetric functions of the roots of the polynomials u_i agree except the nth. Thus, we obtain the desired partition where s_i consists of the roots of u_i. $\square$

Let us deal with polynomials in general. Suppose that $f(x)$ is a polynomial of degree mn and leading coefficient a, so that $f(x) = au(x)$ for some monic polynomial $u(x)$. Then we show that $f(x)$ is a composite of polynomials of degrees m and n if and only if $u(x)$ is so. Suppose that $f(x) = g(h(x))$ where $g(x)$ is of degree m with leading coefficient b and $h(x)$ is of degree n with leading coefficient c. Then, by comparison of leading coefficients, we have that $a = bc^m$. It can be checked that $u(x) = v(w(x))$ where $v(x) = (bc^m)^{-1}g(cx)$ and $w(x) = c^{-1}h(x)$.

On the other hand, suppose that $u(x) = v(w(x))$ for some monic polynomials $v(x)$ and $w(x)$ of respective degrees m and n. Then $f(x) = g(h(x))$ with $g(x) = au(x)$ and $h(x) = v(x)$.

We note that, even for monic polynomials, the decomposition of $f(x)$ as a composite $g(h(x))$ is not unique. For example, for arbitrary values of a and d, the pairs $(g(x), h(x)) = (x^2 + d, x^2 + ax + 1)$ and $(g(x), h(x)) = (x^2 + 2x + d + 1, x^2 + ax)$ both yield

$$f(x) = x^4 + 2ax^3 + (a^2 + 2)x^2 + 2ax + 1 + d = (x^2 + ax + 1)^2 + d \ .$$

Reference

James Rickards, *When is a polynomial a composition of other polynomials? American Mathematical Monthy* **118:4** (April, 2011), 358-363