A Pretty Identity

An old but beautiful identity

Joseph Lagrange was both a mathematician and an astronomer, who made significant contributions to just about everything. Yet like all of us he could make mistakes: he once thought he had proved Euclid's parallel postulate. He wrote a paper, took it to the Institute, and as they did in those days, he began to read it. But almost immediately he saw a problem. He said quietly:

> Il faut que j'y songe encore.

"I need to think on it some more." He put his paper away and stopped talking.

Today Ken and I thought we would talk about a beautiful identity of Lagrange, not about the parallel axiom.

Many of you may know the identity, almost all of you probably know one of its consequences, the famous Cauchy-Schwarz inequality:

$$(a, b)^2 \leq ||a||^2 ||b||^2.$$

Here (a, b) is the usual inner product and $||a||^2$ is the square of the L_2 norm,

$$\sum_k a_k^2.$$

The finite-dimensional case of this inequality for real vectors was proved by Cauchy in 1821. Then his student Viktor Bunyakovsky obtained the integral version, by taking limits. The general result for an inner product space was proved by Hermann Schwarz in 1888.

Bunyakovsky worked in theoretical mechanics and number theory. He conjectured, in 1857, a very natural result, one that is surely true, but even after its sesquicentennial anniversary, it remains completely untouched. The conjecture is:

> For each integer polynomial $f(x)$ there are an infinite number of primes in the sequence
>
> $$f(0), f(1), f(2), \dots$$
>
> provided the polynomial satisfies certain trivial constraints.

These are: (i) it must be a polynomial that tends to positive infinity—primes are positive, (ii) it must be irreducible, and (iii) it must not always be divisible by a fixed prime. The first two constraints are trivial and the last avoids polynomials like $x(x+1)$. It is still open for all polynomials of degree at least two.

Cauchy's student made a beautiful conjecture, but a very hard one. We wonder how it was to work with Cauchy as an advisor?

The Identity

Lagrange's identity is quite pretty. For any real numbers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$:

$$\left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right) - \left(\sum_{k=1}^{n} a_k b_k\right)^2 = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} (a_i b_j - a_j b_i)^2$$

Using notation for the norms and inner product we can shorten it to:

$$||a||^2||b||^2 - (a \cdot b)^2 = \sum_{1 \le i < \le j \le n} (a_i b_j - a_j b_i)^2$$

The importance of the identity is that it immediately implies the famous Cauchy-Schwarz inequality over the real numbers.

For a historical note, the case $n = 2$ was known going back to antiquity:

$$(a_1^2 + a_2^2)(b_1^2 + b_2^2) - (a_1 b_1 + a_2 b_2)^2 = (a_1 b_2 - a_2 b_1)^2.$$

Wikipedia names this for Brahmagupta, whom we recently mentioned, and for Fibonacci, but notes that it goes back (at least) to Diophantus. Brahmagupta actually stated and proved this identity for any n as well:

$$(a_1^2 + na_2^2)(b_1^2 + nb_2^2) = (a_1 b_1 + na_2 b_2)^2 + n(a_1 b_2 - a_2 b_1)^2.$$

Incidentally, this shows that not only are numbers that are sums of two squares closed under multiplication, but also numbers of the form $x^2 + ny^2$, for any fixed n. Also we recently mentioned Lagrange's use of a lemma that numbers that are sums of *four* squares are also closed under multiplication, on the way to proving that they encompass all whole numbers.

Two Complex Versions

The identity shows that an inner product can be reduced to the computation of only sums of positive quantities. This works also

in the complex case, but there are two interesting versions that are not immediately interchangeable. First, let

$$S = \sum_{k \in [N]} a_k b_k$$

where we will have in mind that a, b are unit vectors. Then by the identity we get that S^2 is equal to

$$\left(\sum_{k=1}^{N} a_k^2\right)\left(\sum_{k=1}^{N} b_k^2\right) - \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} (a_i b_j - a_j b_i)^2$$

This consists of only two sums, each of which is a sum of squares. Now we may interchange each a_k by its conjugate $\bar{a}_k$, so that we get a proper complex inner product

$$S' = \sum_{k \in I} \bar{a}_k b_k,$$

which is thereby equal to

$$\left(\sum_{k=1}^{N} \bar{a}_k^2\right)\left(\sum_{k=1}^{N} b_k^2\right) - \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} (\bar{a}_i b_j - \bar{a}_j b_i)^2.$$

Note that the conjugates do not go away even though they are squared, so we do not obtain sums of *positive* quantities. To get this, we need a different version that was also found by Lagrange:

$$\left(\sum_{k=1}^{N} |a_k|^2\right)\left(\sum_{k=1}^{N} |b_k|^2\right) = \left|\sum_{k=1}^{n} \bar{a}_k b_k\right|^2 + D,$$

where

$$D = \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} |a_i b_j - a_j b_i|^2 .$$

This is the same as Wikipedia's formula with a_k replaced by $\bar{a}_k$, which in turn fixes an apparent typo in its source. We note that the brute-force proof of this for $n = 2$ leads to Lagrange's four-square lemma. Namely, let

$$\begin{aligned} a_1 &= s + it \\ a_2 &= u + iv \\ b_1 &= w + ix \\ b_2 &= y + iz. \end{aligned}$$

Then the left-hand side becomes

$$(s^2 + t^2 + u^2 + v^2)(w^2 + x^2 + y^2 + z^2),$$

which represents an arbitrary product of sums of four squares that we want to write as a sum of four squares. The right-hand side D becomes

$$|sw + xt + uy + vz + i(sx + uz - tw - vy)|^2 + |sy + vx - tz - uw + i(sz + ty - ux - vw)|^2 .$$

This is a sum of four squares, and to verify that it equals the left-hand side, one need only see that all twenty-four cross-terms in D *cancel*.

Toward New Applications?

Written more compactly using inner product and norm notation, what we have is

$$P = |S'|^2 = |(a, b)|^2 = ||a||^2 ||b||^2 - D.$$

Thus we have written the squared inner product as a difference of two positive real terms, where each term is a sum of squared real subterms or a product of the same. Ken and I have been trying to use this to improve some known simulations of quantum algorithms. So far we have no new results, but the above manipulations make a tighter connection seem plausible. The question for the applications we seek is:

> Under what conditions can good approximations to the squared subterms yield a good approximation to P?

In general, the sticking could be the minus sign, together with the presence of situations where each of the two terms on the right-hand side has magnitude much higher than P. Thus approximations of these terms to within $(1 + \epsilon)$ will not help unless ϵ is truly tiny.

However, in quantum algorithms, we can expect a and b to be unit vectors. Depending on the algorithm, we may be able to arrange for the squared inner product $|(a, b)|^2$ to yield the acceptance probability, so that P is a number between 0 and 1. Then $||a||^2 ||b||^2$ simply equals 1. Thus we have:

$$P = 1 - D = 1 - \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} |a_i b_j - a_j b_i|^2 .$$

Thus we can hope that good approximations to the squared terms in the sum can yield a good approximation to P. Moreover we are helped by the promise for a BQP algorithm that either P is close

to zero (so the sum is close to 1) or P is close to one (so the sum is close to 0). What can go wrong?

The sticking point again could be the minus sign, together with what could be exponentially many complex "cross terms" of the form $a_i b_j$. It may be hard to get enough of a handle on those terms to approximate all their (squared) differences. Two further complications are that sometimes our acceptance probability involves not just one inner product (a, b) but many, though in some cases we can arrange polynomially many, and that the indices i, j, k may be limited to some subset I of $[N]$. What could help us most would be a further manipulation of D, taking into account the BQP "promise" condition that D be near 0 or 1.

Open Problems

Did you know this identity? Besides Cauchy-Schwarz, what useful inequalities and estimates can be derived—in the presence of certain promise conditions?