A Mindful Beauty

What poetry and applied mathematics have in common

JOEL E. COHEN

My grade-school education in mathematics included a strict prohibition against mixing apples and oranges. As an adult buying fruit, I often find it convenient to mix the two. If the price of each is the same, the arithmetic works out well. The added thrill of doing something forbidden, like eating dessert first, comes free. In any case, the prohibition against combining apples and oranges falls away as soon as we care about what two subjects, different in some respects, have in common.

I want to mix apples and oranges by insisting on the important features shared by poetry and applied mathematics. Poetry and applied mathematics both mix apples and oranges by aspiring to combine multiple meanings and beauty using symbols. These symbols point to things outside themselves, and create internal structures that aim for beauty. In addition to meanings conveyed by patterned symbols, poetry and applied mathematics have in common both economy and mystery. A few symbols convey a great deal. The symbols' full meanings and their effectiveness in creating meanings and beauty remain inexhaustible.

Consider the following examples, which involve a beautiful poem of A. E. Housman (1859–1936) and some applied mathematics from my own recent research.

In August 1893, Housman wrote:

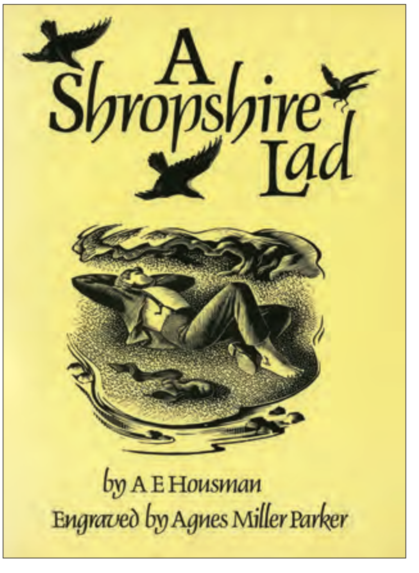
With rue my heart is laden
For golden friends I had,
For many a rose-lipt maiden
And many a lightfoot lad.

Joel E. Cohen is the Abby Rockefeller Mauzé Professor of Populations at The Rockefeller University and Columbia University, New York. He wishes to thank Helen Vendler and Rosanna Warren for their helpful comments on prior drafts of this article.

By brooks too broad for leaping
  The lightfoot boys are laid;
The rose-lipt girls are sleeping
  In fields where roses fade.

The surface meaning is simple: I regret that my friends, once young, have died. At that level of sophistication, the surface meaning of *The Odyssey* is equally simple: Odysseus has trouble getting home. Below the surface of Housman's poem, though, multiple meanings (social, personal, and allusive) interact.

The poem's social meanings arise from its time and place. The 63 poems in the collection *A Shropshire Lad* (of which this is number 54) describe the nostalgia of a country boy who moved to the big city. The poems, published in 1896, resonated widely in English society, where the population was rapidly urbanizing. By 1900, England would become the first country in the world to have most of its people living in cities.



1984 EDITION OF A SHROPSHIRE LAD, HARRAP LIMITED, COURTESY OF THE HOUSMAN SOCIETY

The poem also had personal meanings for Housman. The scholar Archie Burnett's 2003 essay "Silence and Allusion in Housman" showed that many of his poems were "for Housman a means of finding a voice for the love that dare not speak its name, a way of breaking silence, a veil for disclosure, at once catering to reticence and facilitating expression." In May 1895, Oscar Wilde was sentenced for the crime of "gross indecency" (homosexuality but not buggery) to two years' imprisonment with hard labor. Housman's *Shropshire* 54 seems benignly neutral about boys and girls, maidens and lads, and Housman went to great lengths from his youth onward to conceal his homosexuality. But his passionate objection to society's treatment of homosexuals, including Wilde, is clear in several poems in *A Shropshire Lad* and in his later writings, as the critic and scholar Christopher Ricks demonstrated in his essay "A. E. Housman and 'the colour of his hair'" in 1997. Among the personal meanings of "With rue my heart is laden" is what Housman dared not say.

This poem also has allusive meanings for those who read it with the literary background that Housman brought to writing it. In *Cymbeline* (act 4, scene 2), Shakespeare wrote a beautiful song of mourning for a boy, Fidele, thought to have died (but only drugged in a deep sleep):

Golden lads and girls all must,
As chimney-sweepers, come to dust.

Here are Housman's "golden" "lads" and "girls." John Sparrow in 1934 noted echoes of Shakespeare's dirge in this and two other poems of Housman's. Beyond the specific words, Housman echoes Shakespeare's point that mortality masters all. But there is more to this allusion, as the poet and critic Rosanna Warren has pointed out. Fidele is in fact a young woman, Imogen, in boys' clothing. In Shakespeare's time, the female role of Imogen would have been played by a boy or young man, giving the audience a male actor playing a female (Imogen) pretending to be a male (Fidele). Given what is now clear about Housman's sexual orientation, it seems plausible that Housman, consciously or not, identified with the doubly cross-dressing Imogen/Fidele, and nurtured a hope that his poems, if not he, would live.

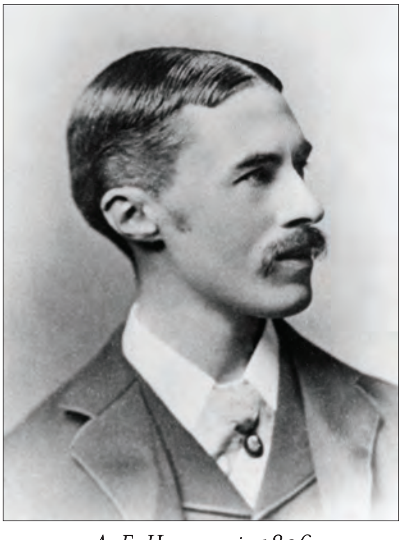
The economy of the poem is evident in the many images compressed into eight lines and in the questions left unanswered. As the poem opens, the narrator speaks of his rue-laden heart, raising the question: Why is he sad? The second line explains why: his friends are gone. Where have they gone? Wait till the second stanza. Who were the rose-lipt maidens (echo of *Othello* act 4, scene 2, as noted in Archie Burnett's 1997 edition of Housman's poems) and lightfoot lads in lines 3 and 4, and what were his relations with them? What happened in those friendships? The narrator never says. Instead he speaks of brooks too broad for leaping, evoking not slender streams easily leaped but a broader, slower descent to the sea in the fullness of time. Only in the last three lines, where the boys and girls are laid in death, in fields where roses fade, do we finally learn that his friends are gone not only as a result of migration (possibly) but also as a result of mortality. Additional questions remain. Why is the narrator's mourning plural and anonymous? (*The Oxford English Dictionary* does not support any sexual interpretation of "laid" at the end of line 6, as the earliest sexual use of "lay" or "laid" dates from 1932.)

Turning from meanings to patterns, we face another mystery. How do the combined patterns of the symbols, on the page or spoken, evoke so much beauty? The patterns in these eight lines interweave meter, rhyme, ending accent, internal repetition, play on the letters *r* and *l,* alliteration, du-bi-du consonants, and two layers of chiasmus, within two symmetrical stanzas. Each line is written in iambic trimeter, and each set of four lines constitutes a quatrain.

With rue | my heart | is laden
For gold | en friends | I had,
For man | y a rose | -lipt maiden
And man | y a light | foot lad.

By brooks | too broad | for leaping
The light | foot boys | are laid;
The rose | -lipt girls | are sleeping
In fields | where ros | es fade.

The irregularity of the anapests in lines 3 and 4 relieves the repetitious symmetry of the other lines. The rhyme scheme is equally simple: *abab.* The ending accent alternates feminine ("laden," "leaping") and masculine ("had," "fade"). There is an extraordinary amount of internal repetition. The first syllable of "laden" reappears in "laid" and, with a slight change in vowel, in "lad." The second syllable of "laden" reappears in "golden" and "maiden." Lines 3 and 4 repeat "many a" exactly. The phoneme "rōz" in "rose-lipt" and "roses" appears in lines 3, 7, and 8. "Lightfoot" appears in lines 4 and 6. Every even-numbered line ends with "d" (the initial consonant of "death," a word that does not appear in the poem), preceded by one or another variant of the vowel-sounds that the letter "a" can exhibit. The poem uses two liquid consonants: *r* 10 times and *l* 12 times. Alliteration crosses lines: "many," "maiden," "many"; "lipt," "lightfoot," "lad," "leaping," "lightfoot," "laid," "lipt," "(s)leeping"; "friends," "fields," "fade"; "brooks," "broad," "boys." Another pattern, for which I do not know a technical name, I have called du-bi-du consonants. It is the pattern illustrated by the consonants in its name, du-bi-du: "lightfoot lad" (*l-f-l*), "brooks too broad" (*b-t-b*), "lightfoot boys are laid" (*l-b-l*), "fields where roses fade" (*f-r-f*).



A. E. Housman in 1896

The pattern of chiasmus is central to this example of the connection I want to make between poetry and applied mathematics. In poetry, chiasmus refers to the statement of two words or ideas and then their restatement in reverse order. A subtle example of chiasmus is the appearance of maiden and lad (female, male in elevated language) in lines 3 and 4 followed by boys and girls (male, female in demotic language) in lines 6 and 7. The change in language from elevated to demotic suggests that even the highborn are brought to earth, and the reversal of order suggests that any precedence in life ("ladies first," maidens before lads) may be undone in death. The second stanza gives another example of chiasmus. Line 5 tells where (by brooks too broad for leaping), and line 6 tells who (the lightfoot boys). Line 7 tells who (the rose-lipt girls) followed by line 8, which tells where (in fields where roses fade). Again the pattern conveys a meaning: the brooks (line 5) and fields (line 8) enclose the boys and girls (lines 6 and 7) as a coffin contains its cadaver. The patterns of the symbols and the messages of the poem are inextricable.

In mathematics (pure or applied), a strict definition would limit "chiasmus" to the exact repetition of two symbols (*a, b*) in reverse order (*b, a*). A more inclusive definition includes any repetition of a sequence of symbols in a permuted order.

Like poetry, applied mathematics combines multiple meanings, economy, pattern, and mystery. In its scientific or practical applications, applied mathematics points to something external. It also alludes to prior mathematics. Its few symbols convey a lot. Its use of symbols often involves internal repetition, symmetry, and chiasmus. It is replete with unexpected truths, unexpected applications, and diverse proofs that illuminate different aspects of a single truth.

My example from applied mathematics comes from work I did with two outstanding colleagues, Johannes H. B. Kemperman, retired from Rutgers University, and Gheorghe Zbăganu, University of Bucharest. In 2000, Zbăganu published a fact new to mathematics:

If *n* is a positive integer (a counting number like 1, 2, 3, ...), and a_1, a_2, ..., a_n and b_1, b_2, ..., b_n are any nonnegative real numbers (any fractional or whole number larger than or equal to zero, such as 17 or 0.333333 ... or 3.14159 ...), then

$$\sum_{i,j} \min\left((a_i \times a_j),(b_i \times b_j)\right) \le \sum_{i,j} \min\left((a_i \times b_j),(b_i \times a_j)\right)$$

No understanding of the meanings of this beautiful formula is necessary to appreciate that it is an intricately patterned array of symbols. Whatever it means, the formula has a left lobe $\sum_{i,j} \min\left((a_i \times a_j),(b_i \times b_j)\right)$ and a right lobe $\sum_{i,j} \min\left((a_i \times b_j),(b_i \times a_j)\right)$ mediated by $\le$. The symbols in the left lobe are exactly the same as the symbols in the right lobe but the letters *a* and *b* appear in different order; this is chiasmus in the broad sense. In the right lobe, (*a, b, b, a*) is an example of chiasmus in the strict sense as the sequence (*a, b*) is repeated in reverse order (*b, a*). That is about as far as one can go without having any idea of the meanings of the symbols.

Understanding the formula's meanings only enhances one's sense of its beauty, economy, and mystery. The connective $\le$ between the two lobes means that the quantity on the left is less than or equal to the quantity on the right. The expression $\sum_{i,j}$ means sum (add) for all pairs *i, j*, where *i* and *j* are positive whole numbers from 1 to *n*. Finally, $\min\left((a_i \times b_j),(b_i \times a_j)\right)$ means the minimum (smaller) of $a_i \times b_j$ and $b_i \times a_j$, and similarly for $\min\left((a_i \times a_j),(b_i \times b_j)\right)$.

The words "If *n* is a positive integer, and a_1, a_2, ..., a_n and b_1, b_2, ..., b_n are any nonnegative real numbers, then," which precede the formula, declare that the inequality holds for any such numbers. Therein lies the immense and surprising power of Zbăganu's inequality. A numerical example illustrates its economy of expression. If n=2 and if a_1=2, a_2=5, b_1=4, b_2=3, then the expression on the left side of the inequality equals min(2×2, 4×4) + min(2×5, 4×3) + min(5×2, 3×4) + min(5×5, 3×3) = 4 + 10 + 10 + 9 = 33, while the expression on the right side equals min(2×4, 4×2) + min(2×3, 4×5) + min(5×4, 3×2) + min(5×3, 3×5) = 8 + 6 + 6 +15 = 35. As the formula predicts, 33 < 35. Zbăganu's inequality asserts that, no matter what natural number n you may pick, and no matter what nonnegative real numbers $a_1, a_2, ..., a_n$ and $b_1, b_2, ..., b_n$ you may pick, the value of the left side will be less than or equal to the value of the right side.

Zbăganu's inequality has social or referential meanings because it answers a question Zbăganu was considering in the mathematical theory of information systems: If one of two messages must be sent over a channel with only two input symbols, A and B, and with n output symbols, 1, ..., n, is the chance of error in transmission smaller if the first message is sent as AA and the second message as BB, or if the first message is sent as AB and the second message as BA?

Johannes Kemperman and Gheorghe Zbăganu in 2001

COURTESY OF JOEL E. COHEN

The left lobe of Zbăganu's inequality represents the probability of an error in transmission if the first message sent is AA and the second is BB, while the right lobe represents the probability of error in the alternative. Hence the inequality says that coding the two messages by AA and BB gives a lower risk that the wrong message will be received than coding by AB and BA. (The tongue in cheek message for teachers might be: If you're trying to teach your students one of two messages, it's better to convey the message twice in the same way than to convey it once in each of two different ways. But don't take this interpretation too seriously; some students have memories, unlike the communication channels in this theory.)

Nothing prevents us from playing formally with Zbăganu's inequality as long as we remember that such formal play yields only questions, not answers. For example, if we exchange multiplication and addition, we get another formula:

$$\prod_{i,j}\left(\min\left[\left(a_i+a_j\right),\left(b_i+b_j\right)\right]\right) \le \prod_{i,j}\left(\min\left[\left(a_i+b_j\right),\left(b_i+a_j\right)\right]\right).$$

Is this formula always true if n is a positive integer, and $a_1, a_2, ..., a_n$ and $b_1, b_2, ..., b_n$ are any nonnegative real numbers?

Pushing the same idea further, we produced 64 possible generalizations of Zbăganu's inequality by replacing each occurrence of addition or summation, minimum, and multiplication by each of four operations: addition, multiplication, minimum, and maximum, then looking for a direction of the inequality ≤ or ≥ that would make the statement true whenever n is a positive integer and $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ are any nonnegative real numbers. Zbăganu's inequality is one of these 64. We produced another 32 formal generalizations of Zbăganu's inequality by swapping in other mathematical functions (for cognoscenti, the spectral radius of a square nonnegative matrix, and quadratic forms). The mathematical question for us became: are these other 95 inequalities besides Zbăganu's inequality always true or true under some interesting conditions?

Here are a few examples of the resulting generalizations of Zbăganu's inequality, which we proved to be true under appropriate conditions:

$$\sum_{i,j} \max\big((a_i + a_j),(b_i + b_j)\big) \geq \sum_{i,j} \max\big((a_i + b_j),(b_i + a_j)\big),$$

$$\prod_{i,j} \min\big(\max(a_i, a_j), \max(b_i, b_j)\big) \leq \prod_{i,j} \min\big(\max(a_i, b_j), \max(b_i, a_j)\big),$$

$$\sum_{i,j} \min\big((a_i a_j),(b_i b_j)\big) x_i x_j \leq \sum_{i,j} \min\big((a_i b_j),(b_i a_j)\big) x_i x_j,$$

$$\rho\big(\min((a_i a_j),(b_i b_j))\big) \leq \rho\big(\min((a_i b_j),(b_i a_j))\big),$$

$$\iint \log\big[(f(x) + f(y))(g(x) + g(y))\big] d\mu(x) d\mu(y)$$

$$\leq \iint \log\big[(f(x) + g(y))(g(x) + f(y))\big] d\mu(x) d\mu(y)$$

You are supposed to be saying, or at least thinking, ooooooh, aaaaaaaaah, how beautiful these formulas are.

Every mathematician knows, and every student of elementary mathematics has to learn, that swapping the operations of addition, multiplication, minimum, and maximum does not generally produce true formulas.

But after a few years of hard, exciting work, Zbăganu, Kemperman, and I were astonished to find that when $n=2$, all 96 formulas we had invented by purely formal manipulations are true. When $n>2$, 62 of our first 64 generalizations are true but two are false in general. Some of the inequalities involving the spectral radius and quadratic forms are false in general, some are true. We were able to distill several of our inequalities into a more abstract so-called matrix-norm inequality that is valid for any one of the four operations of addition, multiplication, maximum, and minimum—the kind of mathematical discovery that brings joy to the heart of applied mathematicians (or at least to the hearts of the three of us).

Our inequalities have meanings both referential and allusive. I described earlier Zbăganu's construction of his inequality in the context of information theory. Among our other inequalities, several have interpretations in operations research with potential applications for scheduling transportation and manufacturing. In addition, our inequalities allude by similarity of form to a beautiful inequality by Augustin Louis Cauchy (1789–1857):

If n is a positive integer, and $a_1, a_2, ..., a_n$ and $b_1, b_2, ..., b_n$ are any non-negative real numbers, then

$$[a_1^2+...+a_n^2][b_1^2+...+b_n^2] \geq [a_1b_1+...+a_nb_n]^2.$$

Equality holds if and only if for some real constants λ and μ with $\lambda^2 + \mu^2 > 0$ we have $\lambda a_j + \mu b_j = 0$ for $j = 1, ..., n$.

In 2004 the mathematician Michael Steele called Cauchy's inequality "one of the most widely used and most important inequalities in all of mathematics." The opening words and symbols are identical to those of Zbăganu's original inequality except that Zbăganu's restriction to non-negative numbers is absent from Cauchy's inequality. Here the letters a and b which appear in the left lobe in the order (a, a, b, b) appear in the right lobe in the order (a, b, a, b), whereas in Zbăganu's inequality they appear in the right lobe in the order (a, b, b, a). For readers with the requisite mathematical background, Zbăganu's inequality alludes to Cauchy's inequality no less clearly than Housman's verse alludes to Shakespeare's.

A deeper and more substantial connection lies beneath the similarity in the pattern of the symbols. After Zbăganu's inequality appeared in 2000, Ravi Boppana of New York University proved it in a new way. In 2006, Titu Andreescu of the University of Texas at Dallas and Gabriel Dospinescu of Lycée Louis-le-Grand, in Paris, found new proofs that strengthen Boppana's intermediate result on the path to Zbăganu's inequality. Some of these proofs depend on the Cauchy-Schwarz inequality, a generalization of Cauchy's inequality. So among the lines of logical genealogy that lead to Zbăganu's inequality are lines that pass through Cauchy's inequality.

In applied mathematics, as in poetry, at the end of the analysis, things not yet understood remain. Why are so many of our formulas true? They were generated by a formal process that any mathematician, pure or applied, regards with total disbelief. In the first group of 64 putative inequalities, why are the two particular failed inequalities false when $n>2$? Why are the proofs of the true inequalities so extraordinarily diverse? Why can't we find a unifying approach that separates the sheep from the goats, the true from the false? Perhaps most mysteriously, why do several of our formulas have meanings for potential practical applications?

The critic and scholar Helen Vendler has shown me a precedent for the mixing of poetry and applied mathematics in *Seven Types of Ambiguity,* written by William Empson (1906–1984) at the age of 22 and published when he was 24. At Cambridge, Empson won firsts in mathematics and English. His book repeatedly cites the commonalities of poetry and math. For example, Empson quotes from George Herbert's book *The Temple* (1633) the eight-line poem "Hope," which alternates lines of iambic pentameter and iambic trimeter. Empson wrote:

> One can accept the poem without plunging deeply into its meaning, because of the bump with which the short lines, giving the flat, poor, surprising answer of reality, break the momentum of the long hopeful lines in which a new effort has been made; the movement is so impeccable as to be almost independent of the meaning of the symbols.
>
> And, indeed, the symbols themselves seem almost to be used in a way familiar to the mathematician; as when a set of letters may stand for any numbers of a certain sort, and you are not curious to know which numbers are meant because you are only interested in the relations between them.

As apples do differ from oranges, poetry does differ from applied mathematics, despite their commonalities. For example, in poetry, but not mathematics, sound and "mouthfeel" (Galway Kinnell) matter. In applied mathematics, unlike poetry, calculation and shared scientific concepts and data, rather than intimate experience, lend conviction. Examples of differences could be multiplied. They do not undermine the significance of the similarities.

Poetry and applied mathematics, with mysterious success, both use symbols for beautiful, economical pointing and patterning. Pointing establishes a relation between symbols and a world beyond the domain of symbols. Patterning establishes a relation between symbols and other symbols in the same domain. Poetry and applied mathematics fall along a continuum between pointing and patterning.

The pointing of symbols to something else is most important.

Formal prose	Numerical data collection
Poetry	Engineering
Songs with words, program music	Applied mathematics
Abstract music	Pure mathematics

The patterning of symbols themselves is most important.

The same continuum runs in the visual arts from journalistic photography at the extreme of pointing to purely abstract art at the extreme of patterning. Between those extremes lies most of the world of art, mixing apples and oranges, mixing meanings and patterns, along with poetry and applied mathematics.

The differences between poetry and applied mathematics coexist with shared strategies for symbolizing experiences. Understanding those commonalities makes poetry a point of entry into understanding the heart of applied mathematics, and makes applied mathematics a point of entry into understanding the heart of poetry. With this understanding, both poetry and applied mathematics become points of entry into understanding others and ourselves as animals who make and use symbols. ❖