FORMAL ALGEBRA AS BOOKEEPING

The idea that the argument x is a variable that traverses continuously its values is foreign to algebra; it is just an indeterminate, an empty symbol that binds the coefficients of the polynomial into a uniform expression that makes it easier to remember the rules for addition and multiplication. 0 is the polynomial all of whose coefficients are 0 (not the polynomial that takes on the value 0 for all values of the variable x).

"Topology and Abstract Algebra as Two Roads
of Mathematical Comprehension"
— Hermann Weyl (1931)

Weyl describes a dialectic that runs thoughout algebra, a dialectic that we call "form versus function." Polynomials like $3x^3 - 7x^2 - 2x + 8$ have two faces in algebra. One face is that they encapsulate computation: $f(x) = 3x^3 - 7x^2 - 2x + 8$ describes a sequence of arithmetic calculations that can be performed to any number x. This is the function face, and it is at play when you think about x being a placeholder for a number. On the other hand, $3x^3 - 7x^2 - 2x + 8$ stands for a formal expression, an element of a mathematical system—the ring of polynomials in x with real coefficients, say—that has it's own rules for calculation and its own internal logic. When you are trying to factor the polynomial, writing, for example,

$$3x^3 - 7x^2 - 2x + 8 = (x - 2)(x + 1)(3x - 4)$$

you are thinking of x as "an empty symbol that binds the coefficients of the polynomial into a uniform expression."

Of course, like any dialectic, these two perspectives are not in contradiction and often complement each other. In the above factorization, for example, seeing that -1 is a zero of the function defined by the polynomial helps you find the factor $x + 1$ of the formal expression, and the fact that the two sides are equal as polynomials tells you that they'll produce the same value when x is replaced by any number (or polynomial, even). Along these lines, the remainder theorem in algebra 2:

The remainder upon division of polynomial $f(x)$ by $x - a$ is $f(a)$, the value of $f(x)$ at $x = a$.

shows both faces at once—"remainder upon division" is about form and "the value of $f(x)$ at $x = a$" is about function.

In a way, this movement between form and function is an example of the contextualization and decontextualization theme mention in MP.2. A colleague suggested a very helpful analogy for this program: Imagine a three-storey building. When students enter Algebra 1, they are on the first floor: algebraic expressions are shorthand for numerical calculations and the variables are placeholders for numbers—they think of expressions as defining functions, and the fact that $x^2 - 1 = (x - 1)(x + 1)$ is justified because it holds for every number x, a fact that can be established via the basic rules of arithmetic. The complete statement of this identity involves quantifiers: "for all real numbers x, …". Later, they climb to the second storey, and they begin to see the expressions and their algebraic structure as a system in its own right, with its own internal logic and as its own context. Polynomials are equal if they are equal coefficient by coefficient, and two expressions form an identity if one can get from one to the other via the basic rules for expressions. They then climb to the third storey and are able to re-contextualize the formal calculations, using them to generate numerical specializations, as in the Common Core example of generating Pythagorean triples from the identity

$$\left(x^2+y^2\right)^2=\left(x^2-y^2\right)^2+(2xy)^2$$

Common Core acknowledges this form-function dialectic through its separation of the conceptual categories of algebra and functions. One of the key uses of formal algebra is to use polynomial algebra as a bookkeeping mechanism: the use of calculations with formal expressions to keep track of numerical data. A good example of this: The coefficient of x^n in

$$(x+x^2+x^3+x^4+x^5+x^6)^k$$

is the number of ways you can roll a value of n if k dice are thrown (equivalently, it's the number of k-tuples of integers, each between 1 and 6, that sum to n). On the surface, this seems a little amazing.

To see what's going on requires a different way to think about algebraic expressions. For example, the polynomial

$$x+x^2+x^3+x^4+x^5+x^6$$

can be interpreted as the "generating function" for the roll of one die: when you roll a fair die, each of the integers between 1 and 6 can show up once, and no other integer can show up. If you had an 8-sided die in which the faces were labeled $\{3,3,4,5,5,5,9,9\}$, you'd represent this by the polynomial

$$2x^3+x^4+3x^5+2x^9$$

Back to our fair die. When you square $x+x^2+x^3+x^4+x^5+x^6$, the coefficient of x^n is the number of ways n can be written as the sum of *two* integers between 1 and 6. And when you cube it, the coefficient of x^n is the number of ways n can be written as the sum of *three* integers between 1 and 6. The reason that this works is because polynomial multiplication is tailor made for this kind of bookkeeping—it's the generalized distributive law. And counting, say, pairs of numbers between 1 and 6 that add to n involves exactly this kind of "each with each" calculation. So, there's a general principle here: if the coefficient of each x^n in a polynomial is the number of ways that n can be represented by some function, the coefficient if x^n in the kth power of that polynomial is number of ways that the n can be written as a sum of k values of that function.

In our professional development work, participants use a CAS to experiment with the distribution of sums when several dice are thrown. The structural theme (MP.7) and MP.8 are the dominant habits of mind here. Two examples from the course are typical of how this generating function idea is used:

Example 1: The form-function dialectic can be used to get information from the algebra via substitution[1]. For example, the generating function for the distribution of sums on two dice is $\left(x+x^2+x^3+x^4+x^5+x^6\right)^2$. When expanded, this is

$$\operatorname{expand}\left(\left(x+x^2+x^3+x^4+x^5+x^6\right)^2\right)$$

$$x^{12}+2\cdot x^{11}+3\cdot x^{10}+4\cdot x^9+5\cdot x^8+6\cdot x^7+5\cdot x^6+4\cdot x^5+3\cdot x^4+2\cdot x^3+x^2$$

This tells you that there are, for example 4 ways to roll a 9. It also tells you that the most likely sum is 7 (there are 6 ways to roll 7). If you wanted to know how many possible rolls there are, you'd want to add up the coefficients in the expansion. But (here's where the function face emerges) you can add up the coefficients in a polynomial by replacing x by 1. And replacing x by 1 in the expanded version is the same as replacing x by 1 in the original factored version, and this is just

$$(1+1+1+1+1+1)^2 = 6^2 = 36$$

This reasoning shows that there are 6^k possible rolls when k dice are thrown. Replacing x by -1 gives the alternating sum of coefficients, and this is the excess of the number of even degree terms over the number of odd-degree term. In terms of the context, it gives the excess of number of even rolls over the number of rolls with an odd sum. But, looking at the factored form, this value is just 0, so there are as many even sums as there are odd ones, and this is true no matter how many dice are rolled.

Example 2: A more textured application shows how formal algebra can be used to find another way to label the faces of two dice with positive integers so that the distribution of sums is the same as that of two fair dice. This latter distribution comes from the coefficients of $\left(x+x^2+x^3+x^4+x^5+x^6\right)^2$. Looking at the prime factors of this polynomial[2], we can think of the two dice as having the same generating function:

die 1: $x(x+1)(x^2+x+1)(x^2-x+1)$ and die 2: $x(x+1)(x^2+x+1)(x^2-x+1)$

To get another pair of dice with the same sum distribution, we want to separate the prime factors of into two products that are legitimate "dice polynomials." The problem thus comes down to arranging the prime factors into two products subject to constraints:

(1) Each product has positive coefficients (so it needs to contain a factor of x).
(2) Each product's coefficients add to 6 (so the value of each product when $x=1$ is 6).

A little experimentation shows that the only possibility is

die 1: $x(x+1)(x^2+x+1)(x^2-x+1)^2$ and die 2: $x(x+1)(x^2+x+1)$

Expanding these:

[1]In the context of throwing dice, this information can also be obtained without use of the formal algebra, but we want to give a simple example of a very general method here.

[2]The factorization is:

$$\left(x+x^2+x^3+x^4+x^5+x^6\right)^2 = x^2(x+1)^2\left(x^2+x+1\right)^2\left(x^2-x+1\right)^2$$

$\text{expand}\left(x\cdot(x+1)\cdot\left(x^2+x+1\right)\cdot\left(x^2-x+1\right)^2\right)$	$x^8+x^6+x^5+x^4+x^3+x$
$\text{expand}\left(x\cdot(x+1)\cdot\left(x^2+x+1\right)\right)$	$x^4+2\cdot x^3+2\cdot x^2+x$

we see that dice labeled $\{1,3,4,5,6,8\}$ and $\{1,2,2,3,3,4\}$ will have the same sum distribution as two fair dice.

More about generating functions. The field is vast and finds applications all over mathematics and science. For more detail, see [1, 2]. In this section, we close with an example that illustrates one use of formal algebra to investigate a context that has come up in many middle and high school curricula.

In the dice example, there are only so many sums that are possible. So, our generating function had only so many terms. For two dice, for example, there are only 11 possible sums (2–12), so the generating function has only 11 terms. But there's another way to think about it. We could think of any non-negative integer as a possible sum. For all but 11 of them, the number of ways an integer can show up on two dice is 0. In other words, we could think of a function, a, defined on the non-negative integers $\{0,1,2,3,4,5,\ldots\}$ by the rule:

$$a(n) = \text{the number of ways } n \text{ could show up as a sum on two dice}$$

The table for a gets kind of boring after awhile:

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20. . .
$a(n)$	0	0	1	2	3	4	5	6	5	4	3	2	1	0	0	0	0	0	0	0	0. . .

But now we have a function defined on all of the non-negative integers. And the *generating function* for a is the infinite formal power series

$$a(0) + a(1)x + a(2)x^2 + \cdots = \sum_{n=0}^{\infty} a(n)x^n$$

You may have run across infinite series before—perhaps in a calculus class. If you're unfamiliar with infinite series, think of this infinite sum as just a formal expression, an infinite polynomial, an algebraic device to "hang out a clothesline" (this image of an expression as clothesline is due to Herbert Wilf and is taken from his beautiful book, *generatingfunctionology* [1]).

One way to use generating functions is as a modeling tool—a tool to generate data. Consider, for example, the celebrated "post office problem:" the post office has only 5 and 8 cent stamps. What denominations can it make? Well, let

$$a(n) = \text{the number of ways you can make } n \text{ cents from 5 and 8 cent stamps}$$

You can check that, for $n < 28$, $a(n) = 0$ unless $n \in \{0,5,8,10,13,15,16,18,20,21,23,24,25,26\}$. After that, $a(n) \neq 0$. In fact, the first n for which $a(n) = 2$ is 40 (so, 40 cents is the first denomination you can make in two ways—what are the two ways?).

The generating function for a is, by definition,

$$f(x) = \sum_{n=0}^{\infty} a(n)x^n$$

but it has another, more familiar form.

To see this, we need a formalization of the high school methods for summing geometric series. These methods lead to several very useful formal identities that are, in fact, the generating functions that single out multiples of a specific integer, like:

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + \dots$$

$$\frac{1}{1-x^2} = 1 + x^2 + x^4 + x^6 + \dots$$

$$\frac{1}{1-x^3} = 1 + x^3 + x^6 + x^9 \dots$$

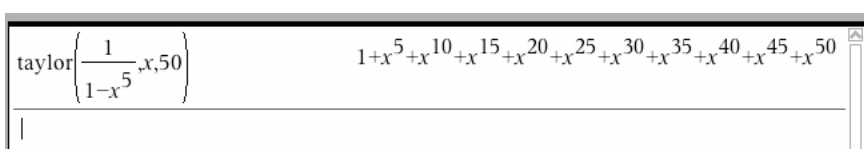
$$\vdots \quad \vdots$$

$$\frac{1}{1-x^k} = 1 + x^k + x^{2k} + x^{3k} \dots$$

$$\vdots \quad \vdots$$

Remember, these are formal identities. So, to prove these statements, multiply the denominator on the left side of each equation by the corresponding right side and watch everything fall apart.

To generate them on the TI-nspire you ask for the "taylor expansion about 0" and type, for example



Now the generating function for the post office problem has coefficients $a(n)$, where $a(n)$ is the number of ways you can write n as a sum of a non-negative multiple of 5 and a nonnegative multiple of 8. But, if you work it out, this is the coefficient of x^n when you multiply the series that single out multiples of 5 and 8:

$$\left(1 + x^5 + x^{10} + x^{15} + \dots\right)\left(1 + x^8 + x^{16} + x^{24} + \dots\right)$$

This is best seen by computing a few terms by hand and focusing on the regularity of the calculations (MP.8). But replacing each factor above with it's "closed form", we have

$$\left(\frac{1}{1-x^5}\right)\left(\frac{1}{1-x^8}\right) =$$

$$\begin{aligned} &1+x^5+x^8+x^{10}+x^{13}+x^{15}+x^{16}+x^{18}+x^{20}+x^{21}+x^{23}\\ &+\ x^{24}+x^{25}+x^{26}+x^{28}+x^{29}+x^{30}+x^{31}+x^{32}+x^{33}\\ &+\ x^{34}+x^{35}+x^{36}+x^{37}+x^{38}+x^{39}+2\,x^{40}+x^{41}+x^{42}+x^{43}\\ &+\ x^{44}+2\,x^{45}+x^{46}+x^{47}+2\,x^{48}+x^{49}+2\,x^{50}+\dots\\ &=\sum_{n=0}^{\infty} a(n)x^n \end{aligned}$$

And a CAS can generate as many terms as one needs:

$$\text{taylor}\left(\frac{1}{1-x^5}\cdot\frac{1}{1-x^8},x,50\right)$$

$$1+x^5+x^8+x^{10}+x^{13}+x^{15}+x^{16}+x^{18}+x^{20}+x^{21}+x^{23}+x^{24}+x^{25}+x^{26}+x^{28}+x^{29}+x^{30}+x^{31}+x^{32}+x^{33}+x^{34}+x^{35}+x^{36}+x^{37}+x^{38}+x^{39}+2\cdot x^{40}+x$$

This provides a genuine mathematical model for the dice situation, one that can be used as an experimental tool to generate data, make conjectures, and even inspire proofs.

This circle of ideas—using formal calculations as modeling tools—highlights how many of the standards for mathematical practice are brought together to bear on the same mathematical territory. Using appropriate tools, building mathematical models, abstracting regularity from repeated actions, moving between form and function, and (especially) seeking and using algebraic structure are all used in inseparable ways to create a set of general-purpose tools that have many different applications from using formal matrix calculations to analyze Markov chains to using the binomial theorem to investigate coin flips.

This "structural approach" to school algebra was often neglected in the last century, with more emphasis on the functional approach to algebra. Common Core calls for a more balanced approach, equally highlighting algebra, functions, and the interaction among their many moving parts.

Postscript. Formal algebra is used for much more than generating data, and the references at the end are full of examples. Just one here: A theme that runs through many articles in the *Mathematics Teacher* is the resolution of recurrence equations. That is, one is given a recurrence, like the celebrated Fibonacci function:

$$F(n)=\begin{cases}0 & n=0\\ 1 & n=1\\ F(n-1)+F(n-2) & n>1\end{cases}$$

and the goal is to come up with a "closed form" definition that generates the function values without recursion (for Fibonacci, one gets the "Binet formula" involving the golden mean). One of the most flexible methods for doing this is to use generating functions:

$$f(x) = \sum_{n=0}^{\infty} F(n)x^n$$

Briefly, the reason this is so useful is that multiplying the series above by a power of x just shifts the terms up, so you can replicate the recurrence with an algebraic calculation. This is developed, slowly and clearly, in [1, 2]

References

[1] Wilf, H. *generatingfunctionology.* Academic Press. New York, 1994.
[2] Graham, R., Knuth, D., and Patashnik, O.
Concrete mathematics. Addison Wesley. Reading, 1989.