Comments on the Chinese remainder theorem

Minhyong Kim

According to D.Wells, the following problem was posed by Sun Tsu Suan-Ching (4th century AD):

There are certain things whose number is unknown. Repeatedly divided by 3, the remainder is 2; by 5 the remainder is 3; and by 7 the remainder is 2. What will be the number?

As you can guess, the solution is far from unique. However, one non-trivial part of the theorem is that a solution does exist, i.e., that

$$\mathbb{Z} \rightarrow \mathbb{Z}/3 \times \mathbb{Z}/5 \times \mathbb{Z}/7$$

is surjective. Of course, it also gives a precise measure of non-uniqueness, in that the kernel is precisely

$$(3) \cap (5) \cap (7) = (3)(5)(7) = (105).$$

That is to say, any two solutions differ by a multiple of 105. To get a definite answer, one could have asked for the smallest possible (positive) solution. On the other hand, the way it was posed does suggest that the existence by itself may have been regarded as surprising.

Exercise: Show (quickly, without solving) that there is a unique positive solution to Sun's question strictly less than 105.

Incidentally, from the perspective of the quotient map under discussion, the terminology 'remainder theorem' is appropriate when we concretely identify \mathbb{Z}/n with $\{0, 1, \ldots, n-1\}$, and the map $\mathbb{Z} \to \mathbb{Z}/n$ as associating to a number *m* the remainder after dividing by *n*. However, in modern times, after we have become accustomed to the concept of equivalence classes, we tend to think of the map as

$$m \mapsto [m]$$

sending *m* to its equivalence class with respect to the obvious(?) equivalence relation. One good reason for thinking in the latter way is that it allows us to generalize naturally to rings *R* (commutative with 1) and ideals *I*. That is, we can consider the map $R \rightarrow R/I$, where the 'remainder' interpretation can still be maintained (and is useful) sometimes, but at the expense of losing a certain flexibility of conceptualization and manipulation.

Now, if we consider a collection I_1, \ldots, I_n of ideals and the map

$$R \rightarrow \prod_i R/I_i,$$

then what is always true is that the kernel is $\cap_i I_i$. That is, we have an injection $R/(\cap_i I_i) \hookrightarrow \prod_i R/I_i$. If we make the 'co-maximal' assumption, i.e., $I_i + I_j = R$ for each i, j, then the surprising fact is that the map becomes surjective. Of course it's also true that $\cap_i I_i = \prod_i I_i$, giving us an isomorphism

$$R/\prod_i I_i \simeq \prod_i R/I_i.$$

To understand that this is indeed surprising, one needs to work out several special cases of the assertion, starting from the classical ones. So, for example, if $n = p_1^{m_1} p_2^{m_2} \cdots p_k^{m_k}$ is the prime factorization of n, then

$$\mathbb{Z}/n \simeq \prod_i \mathbb{Z}/p^{m_i}.$$

Recall that this isomorphism respects the *ring structure* and not just the group structure. This gives a quick structure theorem for the unit group of \mathbb{Z}/n ,

$$(\mathbb{Z}/n)^{\times} \simeq \prod (\mathbb{Z}/P^{m_i})^{\times}$$

For example,

$$(\mathbb{Z}/35)^{\times} \simeq (\mathbb{Z}/7)^{\times} \times (\mathbb{Z}/5)^{\times} \simeq \mathbb{Z}/6 \times \mathbb{Z}/4$$

In his regard, it is useful to recall that the structure of the group of units can be a rather tricky problem, starting from the data of the ring. For example,

$$(\mathbb{Z}[\sqrt{6}])^{\times} \simeq \mathbb{Z} \times \mathbb{Z}/2$$

but this fact includes a rather tricky assertion about the solutions of the Diophantine equation

$$x^2 - 6y^2 = \pm 1.$$

(What is this assertion?) Sometimes the unit group can be related to the ring in a natural way, but with a rather different flavor: For example,

$$M_n(\mathbb{C})^{\times} = GL_n(\mathbb{C})$$

or

$$(\mathbb{Z})^{\times} \simeq \mathbb{Z}/2.$$

In any case, the remainder theorem resolves the issue rather simply for these finite rings.

I like to consider the case of polynomial rings. In one variable, the situation is very similar to \mathbb{Z} , as emphasized many times in the lectures. In $\mathbb{C}[x]$, for example, (f(x)) and (g(z)) are co-maximal exactly when they have no common primes factors. The prime factors in this ring are exactly (x - a) for $a \in \mathbb{C}$. In this case, however, there is also a 'geometric interpretation' that they are co-maximal exactly when f and g have no common zeros on the complex plane.

The theorem says that

$$\mathbb{C}[x] \to \prod_{i=1}^{d} \mathbb{C}[x]/(x-a_i)$$

is surjective, in so far as the a_i are distinct. This is true regardless of the cardinality d. Now we have an isomorphism $\mathbb{C}[x]/(x-a_i) \simeq \mathbb{C}$ induced by the evaluation map $f \mapsto f(a_i)$. So we interpret the theorem as the statement that we can specify arbitrary values c_1, \ldots, c_d at the d points a_1, \ldots, a_d , and there will be a single polynomial f satisfying $f(a_i) = c_i$ for all i. Since the kernel of the map is $((x-a_1)\cdots(x-a_d))$ and hence generated by a polynomial of degree d, we see that any solution can replaced by its remainder after dividing by this generator. Hence, there is always a solution of degree $\leq d-1$. It is important to learn the algorithm for actually producing solutions, in \mathbb{Z} or in $\mathbb{C}[x]$. But let me outline one approach that will bring up an interesting connection. By the information we have gained already, we can take a polynomial $f(x) = b_0 + b_1 x + \cdots + b_{d-1} x^{d-1}$ with undetermined coefficients, and then solve for the b_i so that the conditions are satisfied. But these conditions are

$$b_0 + b_1 a_1 + \dots + b_{d-1} a_1^{d-1} = c_1$$

$$b_0 + b_1 a_2 + \dots + b_{d-1} a_2^{d-1} = c_2$$

$$\vdots$$

$$b_0 + b_1 a_d + \dots + b_{d-1} a_d^{d-1} = c_d$$

or

$$4\underline{b} = \underline{c}$$

where \underline{b} and \underline{c} are column vectors formed by the fixed c_i and unknown b_i respectively, while A is the matrix

$$\left(\begin{array}{ccccc} 1 & a_1 & \cdots & a_1^{d-1} \\ 1 & a_2 & \cdots & a_2^{d-1} \\ & & \vdots \\ 1 & a_d & \cdots & a_d^{d-1} \end{array}\right)$$

That is, what we are required to solve is in fact a linear system. Since the theorem says we can always solve for <u>b</u> regardless of what <u>c</u> is, this implies that the $d \times d$ matrix A has full rank, and hence, is invertible. You should know the classical formula from linear algebra, that says

$$\det A = \prod_{i>j} (a_i - a_j),$$

a so-called Vandermonde determinant. This also gives the invertibility, via a precise expression. But it is interesting that just abstract algebra suffices for just the invertibility part. (In the construction of codes, being able to write down such invertible matrices with flexibility in the a_i 's is important.) The remainder theorem might thus be expected to comprise many other apparently different results in linear or non-linear algebra as special cases, but we will not dwell on working out more of these. However, you should consider the case of the ideals $(x - a_1)^{n_1}, \dots, (x - a_d)^{n_d}$ which are still co-prime when the a_i are distinct. In this case, we are saying that not just values, but a finite collection of derivatives can be specified as each point. There may be an invertible matrix underlying this assertion as well. Can you work it out?

The several variable case is also interesting. Inside $\mathbb{C}[x, y]$, take I = (x - 1, y) and J = (x). Then I + J = (x - 1, x, y) contains 1 and is therefore the unit ideal. We know then that

$$\mathbb{C}[x,y] \rightarrow [\mathbb{C}[x,y]/(x-1,y)] \times [\mathbb{C}[x,y]/(x)]$$

is surjective. But $\mathbb{C}[x,y]/(x-1,y) \simeq \mathbb{C}$ via evaluation at (1,0) while $\mathbb{C}[x,y]/(x) \simeq \mathbb{C}[y]$ via $f(x,y) \mapsto f(0,y)$. Thus, we are allowed to specify any value c at (1,0) and any polynomial g(y). Then there exists f(x,y) such that f(1,0) = c and f(0,y) = g(y). It is harder to visualize the co-maximal condition in two variables. For the maximal ideals (x - a, y - b) and (x - c, y - d), co-maximality is easily seen to be equivalent to $(a,b) \neq (c,d)$ as pairs. This is an instance of the general fact that two maximal ideals are co-maximal iff they are distinct (Why?). You are invited also to try a finite collection of ideals of the form

$$(x-a_1, y-b_1)^{n_1}, \cdots, (x-a_d, y-b_d)^{n_d}$$

corresponding to distinct points (a_i, b_i) and work out the interpretation of the Chinese remainder theorem in terms of pre-specified *partial* derivatives at those points. It seems you should get some interesting non-vanishing determinants.

In arbitrary dimensions, I would like to relate the remainder theorem to our discussion about ideals and spaces. That is recall that to an ideal I in $\mathbb{C}[\underline{x}] = \mathbb{C}[x_1, \ldots, x_n]$, we associated its zero set Z(I). Now it is a fact that $I + J = \mathbb{C}[\underline{x}]$ if and only if $Z(I) \cap Z(J) = \phi$. However, $Z(I) \cap Z(J) = Z(I+J)$. (By the way, what is $Z(I \cap J)$? What about Z(IJ)?) So this is merely the statement that an ideal Iis the unit ideal iff $Z(I) = \phi$. Note that one direction is easy. But the converse is a rather difficult theorem called *Hilbert's Nullstellensatz*. In any case, we get the interpretation of the co-maximal property as the *disjointness* of the associated spaces. Meanwhile, recall our interpretation of $\mathbb{C}[\underline{x}]/I$ as the ring of *algebraic functions* on Z(I). Of course we think of $\mathbb{C}[\underline{x}]$ itself as the algebraic functions on the ambient space \mathbb{C}^n . So now we restate the remainder theorem in a particularly natural geometric form: pre-specify any algebraic functions on f_I on Z(I) and f_J on Z(J). Then as long as these spaces are disjoint, there exists a function f on the ambient space that restricts to f_I and f_J on Z(I) and Z(J), respectively. There is, of course, the stronger result involving any finite collection of disjoint spaces $Z(I_i)$. In this generality, solving for the function f *effectively* in the input data in a way analogous to the single variable case (or the Euclidean algorithm in \mathbb{Z}) involves rather recent research in computational algebra, and goes by the name of 'effective nullstellensatz.'

To see how well the geometric intuition works, consider what happens if the zero sets are not disjoint. Of course in that case, we have no right to specify the functions on Z(I) and Z(J) arbitrarily. After all, any restriction from the ambient space would yield two functions f_I and f_J that agree on the intersection $Z(I) \cap Z(J)$. So let us try to build this condition into a theorem. That is we would like

Given f_I on Z(I) and f_J on Z(J) such that $f_I|Z(I) \cap Z(J) = f_J|Z(I) \cap Z(J)$, there exists an f on the ambient space restricting to f_I and f_J .

Is this true with our interpretation of algebraic functions? Well recall that $Z(I) \cap Z(J) = Z(I+J)$. So the functions on the intersection are exactly $\mathbb{C}[\underline{x}]/(I+J)$. Then we are requiring that f_I and f_J map to the same element in $\mathbb{C}[\underline{x}]/(I+J)$ under the natural maps $\mathbb{C}[\underline{x}]/I \to \mathbb{C}[\underline{x}]/(I+J)$ and $\mathbb{C}[\underline{x}]/J \to \mathbb{C}[\underline{x}]/(I+J)$. (By the way, have you noticed that such natural projection maps to quotient rings correspond to restriction of functions?)

Put differently, we would like to have an exact sequence:

$$\mathbb{C}[\underline{x}] \longrightarrow \mathbb{C}[\underline{x}]/I \times \mathbb{C}[\underline{x}]/J \longrightarrow \mathbb{C}[\underline{x}]/(I+J)$$

where the last map takes (f_I, f_J) to the class $[f_I] - [f_J]$. Recall that exactness here just means that the kernel of the second map exactly equals the image of the first map. In the usual remainder theorem, the last ring is zero, so the exactness amounts to a surjectivity. This general statement is in fact true! Prove it yourself. It is actually very easy. And then formulate the correct statement with several ideals. Finally, reflect on how geometric intuition has led us to a very natural generalization of the remainder theorem.

What happens if you now apply this generalization to \mathbb{Z} ? Anything interesting? Consider the ideals (6) and (10), for example. Is there a number that gives you remainder 5 when divided by 6 and 4 when divided by 10? What about 2 and 5? 3 and 5?

Another fun ring to test your skills on is $\mathbb{Z}[x]$ which should roughly behave like $\mathbb{C}[x, y]$. Look for some co-maximal ideals and formulate concrete instances of the Chinese remainder theorem in this ring.