The study of a notion, viz. that of residue classes

The working mathematician introduces new concepts as he goes along, and he sometimes introduces new notations in connection with them. I think that each time he does so he should ask himself questions such as "Is this concept worth being introduced?" or "Do I really need this new notation?". Even if it cannot give the complete conclusive answers, a course on Mathematical Methodology should raise and discuss such questions. In this note we conduct a case study - knowing full well the limitations of such an approach - and have chosen the topic of residue classes.

Remark This choice is largely accidental. Some time ago - EWD770 - I seemed to be able to make good use of the concept, but being only faintly familiar with it, I needed to build up what theory I needed about it. While doing so I made a mental note that this might be a good candidate for such a "conceptual case study": it starts from a very modest mathematical base, it raises notational questions, and quickly leads to not entirely trivial results. To avoid any misunderstanding, the following development is not a historical reconstruction: in writing it I did not consult the literature. Also I would like to make it clear that I have very little interest in number theory per se. (End of Remark.)

For integers q and x we can raise for $q \neq 0$ the question whether q divides x. (If $q=0$, we can raise the question too, but then the answer is less clear; we skirt that issue.) Since divisibility by $-q$ is the same as divisibility by q, it is no loss to confine ourselves to positive q if that comes in handy, and we shall do so. (In a moment we shall see that a constraint to natural x does not come in handy.)

Number theory being what it is, we may expect the divisibility by q to be most interesting for some q that is prime. Should we immediately restrict our attention to prime dividers?

In our current explorative state the answer is "No" because divisibility by q is perfectly defined for non-prime q, and we can always decide later: at this stage we should keep our options open. (Note that I am not advocating "generality" for its own sake: if all interesting results are confined to q's that are prime and considering q's that are not prime only leads to ugly complications, we should restrict ourselves in the final presentation right from the start to prime dividers.)

For the boolean expression "q divides x" exists -since when? I have no idea- a more

or less accepted formal notation, viz.

$$q \mid x \quad .$$

(I say "more or less" because, though I have encountered it regularly in the literature, I do not remember a single author - myself included - being very explicit about its syntactic status. Do we admit $q \mid x+y$ for $q \mid (x+y)$ and $q \mid x \vee q \mid y$ for $(q \mid x) \vee (q \mid y)$? In this text I propose we do.)

What do we know about our relation $\mid$? Well, a lot. Because

$$x \mid y \equiv x = x \underline{\text{gcd}}\ y$$

we know - see EWD954 - :
because gcd is idempotent, $\mid$ is reflexive:

- $x \mid x$;

because gcd is associative, $\mid$ is transitive:

- $x \mid y \wedge y \mid z \Rightarrow x \mid z$;

because gcd is symmetric, $\mid$ is anti-reflexive:

- $x \mid y \wedge y \mid x \Rightarrow x = y$;

because 1 is a left-hand zero-element of gcd, it is a left-hand extreme of $\mid$:

- $1 \mid y$;

because 0 is a right-hand unit-element of gcd, it is a right-hand extreme of $\mid$:

- $x \mid 0$.

Well, that is very encouraging; it makes obvious

that our relation "divides" might be something worth pondering about. Let us look what it does with arithmetic operations and let us take the additive ones first. We have

- $q|x \wedge q|y \Rightarrow q|x-y$

and then as consequences of all the above

- $q|x \equiv q|-x$ and
- $q|x \wedge q|y \Rightarrow q|x+y$, and in particular
- $q|x-x$
- $q|x-y \equiv q|y-x$
- $q|x-y \wedge q|y-z \Rightarrow q|x-z$,

i.e. now viewed as relation between x and y – "for fixed q" so to speak – $q|x-y$ is reflexive, symmetric, and transitive, i.e. $q|x-y$ is what is known as an equivalence relation.

What about multiplication? Well,

- $q|x \Rightarrow q|x\cdot y$

is about the only thing I can think of. But it gives us

$$\begin{aligned} & q|x-x' \wedge q|y-y' \\ \Rightarrow\ & q|(x-x')\cdot y \wedge q|x'\cdot(y-y') \\ \Rightarrow\ & q|(x-x')\cdot y + x'\cdot(y-y') \\ =\ & q|x\cdot y - x'\cdot y' \end{aligned}$$, i.e.

- $q|x-x' \wedge q|y-y' \Rightarrow q|x\cdot y - x'\cdot y'$

Similarly we have – proof omitted –

- $q|x-x' \wedge q|y-y' \Rightarrow q|(x+y)-(x'+y')$.

Now all this is notationally unattractive. Evidently, the assertion $q \mid x-y$ is an interesting assertion about x and y, but as written down it fails to express in a familiar manner its symmetry and a few other properties it shares with equality. One of the notations for $q \mid x-y$ that came in use is

$$x \equiv y \pmod q$$

- read as "x congruent y modulo q" -, in which "$\equiv$" figures as a somewhat weakened equality with the suffix "(mod q)" pinning that weakening down. I have two reasons to depart immediately from this notation: the personal one that I have decided to adopt "$\equiv$" for the associative equality between boolean arguments - in which rôle I shall need it in a moment - and the general one that there is little to be said in favour of a half-infix and half-postfix operator. For the purpose of this discussion I shall denote the symmetric relation $q \mid x-y$ with an infix $\underline{\text{con}}.q$:

$$x \mathrel{\underline{\text{con}}.q} y \quad .$$

In this notation we can summarize

$$\begin{aligned}
&x \mathrel{\underline{\text{con}}.q} x \\
&x \mathrel{\underline{\text{con}}.q} y \equiv y \mathrel{\underline{\text{con}}.q} x \\
&x \mathrel{\underline{\text{con}}.q} y \wedge y \mathrel{\underline{\text{con}}.q} z \Rightarrow x \mathrel{\underline{\text{con}}.q} z \\
&x \mathrel{\underline{\text{con}}.q} x' \wedge y \mathrel{\underline{\text{con}}.q} y' \Rightarrow x+y \mathrel{\underline{\text{con}}.q} x'+y' \\
&x \mathrel{\underline{\text{con}}.q} x' \wedge y \mathrel{\underline{\text{con}}.q} y' \Rightarrow x \cdot y \mathrel{\underline{\text{con}}.q} x' \cdot y' \quad ,
\end{aligned}$$

a bunch of properties of con.q that are shared by equality. Treating "being congruent" as "being equal, but not fully" is not very satisfactory, and we can do something about it in a few steps.

The first thing to do is to use a metaphor - "adopt a jargon" or "introduce a nomenclature" - that does justice to the fact that "being congruent" is - see the first three summarized properties - an equivalence relation and, hence, induces a partitioning of the domain - integers so far - that we have been talking about. We can do so by giving the partitions a generic name, and the one that has been introduced and generally accepted is "residue classes". It enables us to replace the expression "being congruent" by "belonging to the same residue class". It allows us to rephrase the last two summarized formulae as "the residue class to which a sum (product) belongs is determined by the residue classes to which the addenda (factors) belong".

Before continuing our linguistical and notational let us determine the number of residue classes. From

$$(\underline{A}\, x,y:\ 0<x<y<q:\ \neg\ q|\ x-y)$$

we deduce that there are at least q residue classes. From the fact that for each y the equation

$$x:\ (0\leq x<q\ \wedge\ q|\ x-y) \text{ is solvable,}$$

we conclude that there are at most q residue classes. Summarizing: there are precisely q residue classes, and the last equation puts them into one-to-one correspondence with the values x from the range $0 \leq x < q$.

More precisely, that equation having precisely one solution, it defines that solution as a function of y and q, the free parameters in the body of the equation. For that function a special notation has found its way in the literature, viz. an infix $\underline{\text{mod}}$:

$$x = y \underline{\text{mod}}\, q \;\equiv\; 0 \leq x < q \;\wedge\; q \mid x-y \quad .$$

Because

$$q \mid x-y \;\equiv\; x \underline{\text{mod}}\, q = y \underline{\text{mod}}\, q \quad ,$$

we can now express "being congruent modulo q" – or, equivalently, "belonging to the same residue class (modulo q)" – in terms of an equality, viz. yielding the same values under application of the (postfix) operator "$\underline{\text{mod}}\, q$".

Remark The very attentive reader will have noticed that, rather implicitly, we did something else: we derived from the binary operator "$\underline{\text{mod}}$" the unary operator "$\underline{\text{mod}}\, q$". More than a century after the discovery of residue classes, this device would become known as "Currying". (End of

Remark.)

We have done more: the postfix operator "mod q" assigns to each integer a value that is characteristic for the residue class to which the integer belongs, and we should observe that this is a general way of characterizing partionings of any domain: choose a function that assigns the same value to elements in the same partition and different values to elements in different partitions.

What does it do to our five propositions we listed in terms of con.q ? The first three yield:

$$x \bmod q = x \bmod q \quad ,$$

which is not very interesting;

$$x \bmod q = y \bmod q \;\equiv\; y \bmod q = x \bmod q \quad ,$$

which is no more than the symmetry of the equality;

$$x \bmod q = y \bmod q \;\wedge\; y \bmod q = z \bmod q \;\Rightarrow\; x \bmod q = z \bmod q \quad ,$$

which boils down to the transitivity of equality.

The last two propositions become more interesting if we choose $x' = x \bmod q$ and $y' = y' \bmod q$ and realize that $(x \bmod q) \bmod q = x \bmod q$. The antecedents are then true, and we get the almost nice consequents

$$(x+y) \underline{\text{mod}}\ q = ((x \underline{\text{mod}}\ q) + (y \underline{\text{mod}}\ q)) \underline{\text{mod}}\ q$$

$$(x \cdot y) \underline{\text{mod}}\ q = ((x \underline{\text{mod}}\ q) \cdot (y \underline{\text{mod}}\ q)) \underline{\text{mod}}\ q .$$

Almost nice. Had it not been for the final "$\underline{\text{mod}}\ q$" applied to the right-hand sides, $\underline{\text{mod}}\ q$ would have distributed over + and · !

The problem with the integer → integer operator $\underline{\text{mod}}\ q$ is that it characterizes each residue class by one of its members, and – somewhat arbitrarily – we have chosen its one and only member x satisfying $0 \leq x < q$. (Hence the need for that additional "$\underline{\text{mod}}\ q$" at the far right: its argument could be "out of range".)

The way out of this dilemma is to remove the observed arbitrariness by replacing "$\underline{\text{mod}}\ q$" as characteristic function of the partitioning by one of type integer → residue class. No longer bothering to mention q explicitly – consider it as a global parameter – we denote it by a pair of braces. Being a characteristic function for the partitioning induced by the equivalence relation $q \mid x-y$ it satisfies

$$q \mid x-y \equiv \{x\} = \{y\} .$$

In the preceding paragraph we made really two steps – jumps, if you like –. Besides the new function type integer → residue class we introduced "residue class" as a (q-valued)

type in its own right! So far, our only benefit is that the "normal" equality is defined on it, and that is already something: it allows us, for instance, to formulate now the theorem

$$\{x+q\} = \{x\} \quad .$$

In view of the above it stands to reason to define for our new type addition, subtraction, and multiplication by

$$\begin{aligned} \{x\} + \{y\} &= \{x+y\} \\ \{x\} - \{y\} &= \{x-y\} \\ \{x\} \cdot \{y\} &= \{x \cdot y\} \end{aligned} \quad ,$$

thus seeing to it that $\{\}$ nicely distributes over the arithmetic operators considered so far.

Remark We have "overloaded" $+$, $-$, and $\cdot$, since at the right-hand side their operands are integers, at the left-hand side their operands are residue classes. We should not be frightened by this: firstly, the overloading is necessary if we wish to see the above three formula as distribution laws, secondly we should be quite used to it, when we interpret

$$a \cdot (b+c) = a \cdot b + a \cdot c$$

as the distribution of the scalar product over the vector sum: the $+$ at the right-hand side adds reals. (End of Remark).

Have we achieved something? Well, we have introduced for residue classes a (symmetric) multiplication that distributes over the (symmetric) addition:

$$
\begin{aligned}
& \{x\} \cdot (\{y\} + \{z\}) \\
= & \quad \{\text{definition of } + \text{ for residue classes}\} \\
& \{x\} \cdot \{y+z\} \\
= & \quad \{\text{definition of } \cdot \text{ for residue classes}\} \\
& \{x \cdot (y+z)\} \\
= & \quad \{\text{distribution of arithmetic } \cdot \text{ over arithmetic } +\} \\
& \{x \cdot y + x \cdot z\} \\
= & \quad \{\text{definition of } + \text{ for residue classes}\} \\
& \{x \cdot y\} + \{x \cdot z\} \\
= & \quad \{\text{definition of } \cdot \text{ for residue classes, twice}\} \\
& \{x\} \cdot \{y\} + \{x\} \cdot \{y\} \quad .
\end{aligned}
$$

Encouraging, though from another point not impressive: one could conclude that the arithmetic with residue classes is so similar to that with integers that you don't really need very much to distinguish between them and that you don't really need to bother about the braces at all. But let's keep them.

Whereas, without the addition of residue classes, the equation in the integer x

$$x : (\{x+y\} = \{z\})$$

has no unique solution, the one in the residue class $\{x\}$

$$\{x\} : (\{x+y\} = \{z\})$$

has a unique solution, viz. $\{z-y\}$.

Is this important? Yes, I think so, for it invites us to look at equations like

$$x: (\{x \cdot y\} = \{z\}) \quad , \qquad \text{and}$$

$$\{x\} : (\{x \cdot y\} = \{z\}) \quad .$$

Let us look at the first one, which we may rewrite as

$x: (q \mid x \cdot y - z)$, from which we conclude that it may have no solutions at all - e.g. for $q, y, z = 6, 2, 3$ - or may have solutions less than q apart - e.g. for $q, y, z = 6, 2, 4$ - .

When has the second equation a unique solution (if it has one)? From

$$\{x \cdot y\} = \{z\} \ \wedge \ \{x' \cdot y\} = \{z\}$$

we would like to conclude $\{x\} = \{x'\}$.

Note

$$\begin{aligned} & \{x\} = \{y\} \\ = & \quad \{\text{arithmetic}\} \\ & \{x\} = \{y + 0\} \\ = & \quad \{+ \text{ for residue classes}\} \\ & \{x\} = \{y\} + \{0\} \\ = & \quad \{\text{arithmetic}\} \\ & \{x\} - \{y\} = \{0\} \\ = & \quad \{- \text{ for residue classes}\} \\ & \{x - y\} = \{0\} \end{aligned}$$

. (End of Note)

We can conclude

$$\{(x - x') \cdot y\} = \{0\}$$

but that does not suffice, e.g. $q, x, x', y = 6, 2, 4, 3$. The question of uniqueness raises the question under which circumstances we have - the usual! -

$$\{x \cdot y\} = \{0\} \Rightarrow \{x\} = \{0\} \vee \{y\} = \{0\} \quad .$$

This implication does not hold if q is composite! With $q = 6$, for instance, for $x, y = 2, 3$, the antecedent is true, whereas the consequent is not. So, for composite q we get a theory in which a product can be zero without any of its factors being zero. Surely something worth investigating, but we won't do so here and will from now on restrict ourselves to values of q equal to some prime p, under which constraint the above implication holds.

Consequently, for $\{y\} \neq \{0\}$ and $\{x\}$ ranging over the p residue classes, the values $\{x \cdot y\}$ are all distinct, i.e. also range over the p residue classes. Consequently, for $\{y\} \neq \{0\}$

$$\{x\} : (\{x \cdot y\} = \{z\})$$

has a unique solution, which we may denote by $\{z\}/\{y\}$, thus introducing a quotient of residue classes, provided the "denominator" differs from $\{0\}$.

Remember that we had introduced for residue classes an addition and a multiplication satisfying

$$\{x\} + \{y\} = \{x + y\}$$

$$\{x\} \cdot \{y\} = \{x \cdot y\} \quad .$$

Is there a sense in which our newly introduced quotient for residue classes can satisfy the similar

$$\{x\}/\{y\} = \{x/y\} \quad \text{for } \{y\} \neq \{0\} \text{ ?}$$

In other words, may the braces distribute over division as well?

The answer is "no" unless we are willing to extend the braces from a function formerly only defined on integers to a function defined on rationals. The answer is "yes" if we do extend the function to rationals whose denominator y satisfies $\{y\} \neq \{0\}$. The reason is that $\{x\}/\{y\}$ - defined as unique solution of a linear equation - indeed only depends on the (normal) value of the fraction x/y, more precisely:

$$\{w\} \neq \{0\} \wedge \{y\} \neq \{0\} \Rightarrow \{x\}/\{y\} = \{w \cdot x\}/\{w \cdot y\}$$

To demonstrate this we observe that for any w, x, y, and z

$$\begin{aligned} & \{z \cdot w \cdot y\} = \{w \cdot x\} \\ = & \quad \{\text{residue arithmetic}\} \\ & \{z \cdot w \cdot y - w \cdot x\} = \{0\} \\ = & \quad \{\text{arithmetic}\} \\ & \{w \cdot (z \cdot y - x)\} = \{0\} \end{aligned}$$

$$
\begin{array}{ll}
= & \{ p \text{ is prime} \} \\
 & \{w\} = \{0\} \ \vee \ \{z \cdot y - x\} = \{0\} \\
= & \{ \ \{w\} \neq \{0\} \ \} \\
 & \{z \cdot y - x\} = \{0\} \\
= & \{ \text{residue arithmetic} \} \\
 & \{z \cdot y\} = \{x\} \quad ,
\end{array}
$$

i.e. the same z solves the equations for $\{x\}/\{y\}$ and for $\{w \cdot x\}/\{w \cdot y\}$ respectively.

Since all properties of rational arithmetic (such as $a/(b/c) = (a \cdot c)/b$, $a/b + c/d = (a \cdot d + b \cdot c)/(b \cdot d)$, etc.) can be derived by defining a/b as the unique solution of

$$x : (a = b \cdot x)$$

all these properties carry over to the rational arithmetic of residue classes, the only difference being that the null-element differs: in the normal case the numerator should differ from 0 , in the residue arithmetic it should differ from $\{0\}$. In both cases the requirement guarantees the existence of a unique solution.

And now we can start building up a theory, coming up with theorems such as

- $\{x\}^2 = \{y\}^2 \quad \equiv \quad \{x\} = \{y\} \ \vee \ \{x\} = -\{y\}$
- for $p = 2 \cdot n + 1$ the p residue classes consist of n non-squares and $n+1$ squares, being $\{i^2\}$ with $0 \leq i \leq n$
- for $p = 2 \cdot n + 1$
  $\{(\underline{\Sigma} i : 1 \leq i \leq n : i^2)\} = \{(\underline{\Sigma} i : 1 \leq i \leq n : i^{-2})\}$, etc.

which, for instance, provides a way of demonstrating that for $p \geq 5$

$$(\underline{S}i: 1 \leq i < p: (p-1)!/i) \ \underline{\text{mod}}\ p^2 = 0 \quad .$$

(This is not the place to show that demonstration, which uses the last bullet. It is mentioned as example of how extending the "brace function" to rational arguments gives our vocabulary an enrichment without which some arguments would be extremely hard to conduct.)

* * *

What did I learn? One trivial thing: square brackets are more convenient to write than braces. (Earlier I used square brackets; in this note I replaced them by braces and, at least in my hand, that was not an improvement.)

We encountered the very general technique of characterizing a partitioning of a domain by a function on that domain: belonging to the same partition equivales yielding the same function value.

More important was the discovery that for the range of that function we could introduce a type - residue class in this case - for which we could define + and · in such a way that the function distributed over these arithmetic operators. I think that the possibility of that distribution was a very encouraging indication that we were heading for something worthwhile. After all, we would like

to manipulate, and in order to be able to do so one needs laws.

Encouraging as that may be, the possibility of manipulation is not enough: the manipulations could be nothing more than insipid rewritings back and forth between different notations. The possibility of extending the domain to rationals was, I think, the crucial test. (It immediately suggest an extension to the complex plane; this is left as an exercise to the reader.)

Finally I think it worth noticing that we only got off the ground after abolishing the notation $x \equiv y \pmod{q}$, which is ugly anyhow.

Closing remark This EWD took a very long time to complete and I am not fully pleased with its completion. I have the uneasy feeling that the exercise has taught me more than I can formulate now. (End of Closing Remark.)

Austin, 6 January 1987

prof. dr. Edsger W. Dykstra
Department of Computer Sciences
The University of Texas at Austin
Austin, TX 78712-1188
United States of America