Mathematics as modelling toolkits: philosophical and practical aspects

1. Readers and objectives

This publication is for anyone interested in what mathematics is and how it can be better. This includes people who use mathematics, are learning mathematics, teach mathematics, and people who are making mathematical innovations, often working as professional mathematicians of some kind.

The idea of mathematical models is extended by focusing on practical matters of efficiency and flexibility. Along the way some important points are made about number systems and applications.

Philosophical old-chestnuts are answered from the perspective described. There are also suggestions for improved practice of mathematics.

2. Modelling

This section explains the perspective of mathematics as a collection of modelling toolkits, and illustrates it with some examples.

2.1 Models

Models are familiar in various forms. A toy car is a model of a real car. It is like the real thing but smaller and more convenient. This is an example of a physical model.

Various physical models were used by Barnes Wallis and his colleagues as they developed the famous bouncing bomb used in the Dambusters raid in 1943. These included a set up at home firing small balls at a tub of water, a larger version in a ship tank at the National Physical Laboratory, and model dams built outside at the Building Research Laboratory.

Wallis also used mathematical models to help with the design. These only represented certain important characteristics of the bomb's trajectory, for example, linking the way the bomb was dropped with its trajectory over the water. Johnson (1998) provides fascinating detail of this project and mathematical models of objects ricocheting off water.

If Wallis and his colleagues had been working today they would undoubtedly have used computers to simulate aspects of the bomb's behaviour, turning mathematical models into predictions that unfold over time.

A mathematical model is abstract, non-physical, and defined rather than built. Words and symbols are used to define the objects used in the model and their properties. There are also inference rules that enable people (or artificial computers) to make deductions from the model.

Mathematical models give us the ability to deduce what we should believe, but in a conditional way. For example, if we believe statements X, Y, and Z then what other statements should we believe because they can be deduced from what is already believed?

Just like physical models, mathematical models are usually convenient simplifications of reality and allow us to make inferences about the behaviour of the real system from the behaviour of the model.

Inference rules include some that just define our terms (e.g. definitions of AND, OR, NOT, and IMPLIES) and some that capture the behaviour of a system we wish to model.

Accurately representing particular aspects of the behaviour of a system we want to model is almost always a goal of modelling. The properties of the model are not an arbitrary matter. We select them in an effort to make accurate models.

Some models are more accurate than others. Some are so accurate that we hardly notice we are using a model. For example, if I count 10 sheep into a previously empty pen and then, later, count 11 sheep out of the pen then I will be surprised. Objects like sheep that can be counted and normally obey the rules of arithmetic, do so with such reliability that discrepancies are never attributed to a fault in the model. I would assume that I made a mistake in one of my counts, or that someone put an extra sheep into the pen while I was away, or that a new lamb was born.

In contrast, some systems are so hard to model accurately that differences between the model and reality are expected. Models of human behaviour, such as a model to predict if a person will commit a crime in the next year, are unreliable. They can still provide useful information, but we do not expect perfect predictions as we would with counting sheep.

2.2 Natural and artificial applications

Mathematics can be used to model natural and artificial systems. I suspect that modelling artificial systems has probably been the more important contribution to human well-being.

Mathematics is often called the queen of the sciences because it partners with science and is used extensively in modelling the natural world. It is used throughout astronomy, physics, chemistry, and biology, for example, where it is often possible to build models that very precisely match physical reality.

However, mathematical models can also match reality very well when we make things specifically to match a mathematical model. This is often their role in technology. For example, consider how often buildings feature straight lines, triangles, rectangles, squares, flat planes, circles, ellipses, catenary curves, and even parabolic curves. We usually strive to understand and model what is useful. We then usually build what we understand and can model.

Similarly, mechanical systems (with cogs, gears, pistons, and so on) also involve geometry we can model mathematically and use to fix the precise dimensions of components.

Boolean algebra can be useful for modelling electrical circuits. It makes it possible to restate logic in different ways, perhaps simpler, and to translate logic into circuitry.

More recently, electronic and computer systems have been designed to have the precise, logical behaviour that is easily captured in mathematics and controlled by languages that, while not quite mathematics, often have similar features.

Today's software also reflects mathematical ideas. Relational databases are called 'relational' because a relation, in mathematics, is a set of tuples, usually visualised as a table of data. Query languages almost always support 'Boolean' queries, which are queries built using the operators of Boolean algebra.

Number theory has been given a new importance in modelling algorithms that encrypt and decrypt data.

2.3 Deductions from models

Mathematical models can be used in many ways.

2.3.1 Deducing properties

Models can be used to deduce properties of something that exists or is being designed or planned.

For example, if a bridge is being designed then mathematical models and deductions can be used to work out such things as:

- how much concrete is needed
- how many bolts are needed
- how long the work will take
- how much the materials and labour will cost
- the forces acting on various parts of the bridge
- the strain within the materials of the bridge
- what loads the bridge can safely support
- how much wind force the bridge can cope with
- the maintenance costs of the bridge.

2.3.2 Sense making

There are some important situations where we have data but they are not entirely reliable and are difficult to interpret.

For example, tracking an aeroplane using radar can be done better if modelling techniques combine information from the past with current data. Predictions about where the aeroplane will be are combined with the latest indications of where it seems to be now. A famous example of this technique is the Kalman Filter.

Sophisticated mathematics is used for three dimensional medical scanning.

More generally, statistical techniques fit models to data to work out what the underlying truth is likely to be.

2.3.3 Predicting behaviour

There are surprisingly many different types of prediction that can be made.

Prediction might focus on performance in some kind of test e.g. maximum speed, fuel economy, time to completion.

Or it might predict the situation at a particular point in time (usually in the future but it is also possible to go back in time), or the complete trajectory of variables over time, or the situation in the long run.

It is sometimes possible to estimate the results that can be achieved with a given level of resources (e.g. fuel, money, people, time).

Simulation, in particular, can be useful for understanding the varieties of behaviour that are possible for a system. This might identify specific possibilities that are dangerous or that are particularly valuable.

For example, models of epidemics have revealed the crucial factors that determine if an infection will grow rapidly or dwindle to nothing.

2.3.4 Improving designs and plans

It is often possible to improve a design or plan using mathematical modelling.

Sometimes it is possible to calculate, directly, the input values needed to achieve the best output, or to achieve a particular output that is required.

More generally, if we can predict the performance of different designs then we can try alternatives and pick the one with the best predicted performance. Sometimes it is possible to try many variations, perhaps systematically, in a hunt for the best design parameters. This search can be automated and there are several well-known algorithms for doing this.

In simulations, control systems can be tested to see if they are able to defend against bad behaviour, or able to maintain rare but helpful behaviour. The simulation can include sudden, externally driven events to see how the control system responds. For example, this could be done for electronic currencies. Such simulations can also be used to train people to act as controllers.

2.3.5 Quantifying uncertainty

Many of the results referred to above cannot be determined exactly with certainty. Mathematics can be used to analyse the level of uncertainty in results.

For example, the numerical method of Monte Carlo simulation can take a model and information about uncertainty around inputs to that model and work out the uncertainty we should have around our predictions. Other techniques make it possible to identify how much uncertainty results from each part of a model, which is useful in directing research to improve the model.

2.4 Modelling toolkits

Mathematics is much more than individual models for particular situations. To tackle modelling problems as if they were all dissimilar tasks would be inefficient.

As mathematical innovators develop mathematics they tend to focus on particular *types* of model or particular methods for use with a variety of types of model. When they focus on types of model they learn to use them in different ways and create variations on that model. Models are often generalized and specialized, creating an array of somewhat different models for slightly different situations.

Typically, the model types developed and studied initially are simple but, as more is learned, mathematical innovators move on to more complicated models. In recent decades this has often involved developing software tools to apply numerical methods or for symbolic inferences. The numerical methods in particular have often made it possible to develop models that are too complicated or awkward to be used in any other way.

Although mathematical innovators and educators have rarely thought of themselves as developing modelling toolkits, *they have done so*. They have provided:

- Defined objects with often-useful properties (e.g. numbers, shapes, topological objects, axioms)
- Often-useful inference rules (e.g. identities that allow us to rewrite mathematical statements)
- Often-useful model structures (e.g. formulae with letters representing parameters of the model)
- Ready-made deductions (theorems, lemmas, solutions, etc.)
- Methods for making logical deductions from models to derive useful formulae
- Procedures (e.g. numerical methods) for making deductions in situations where a neat formula is not possible or just not available yet.

These tend to be grouped up in ways that point towards particular application areas. For example, geometry is good for modelling physical objects, arithmetic is useful for money, and probabilities are useful for decisions under uncertainty.

This can be viewed as a collection of toolkits, often containing more specialised toolkits. Each toolkit is potentially applicable to a range of modelling tasks.

As mentioned, in modern mathematical practice it is common to develop software that automates use of the mathematics. Examples include modelling packages and languages that contain commands that carry out mathematical operations.

> E.g. Animated movies are made on computers that contain models that are used to simulate movement and visual effects, and to calculate the appearance of scenes.

> E.g. Models used in engineering and architecture represent not only the appearance of objects but also the forces acting within them.

> E.g. Models used in weather forecasting represent the atmosphere and simulate how it will change in future, automating the mathematics.

2.5 Logical explanations

The way that mathematics is usually written reflects, to some extent, practices that help to build a reliable modelling toolkit.

One style, used for the most elementary foundations of mathematics, is to start with a set of axioms. These are statements assumed to be true, in the sense that no proofs are attempted or required. The idea is to start from statements so self-evidently true that nobody will want to argue over them.

The explanation then uses appropriate inference rules to prove more statements from just these axioms. Some of these new statements are considered important enough to highlight. They may be called theorems (the highest status), corollaries (add-ons to theorems), lemmas (subsidiary results established on the way to a theorem), or identities (rewriting rules).

These are then used to deduce more theorems, corollaries, lemmas, and identities. As this process continues the toolkit acquires more inference rules that shortcut new proofs and give users greater reasoning abilities.

Another way to start is to begin with an elementary model, usually a very general one with flexibility built in. The writer can then deduce some things from that generic model, but does not stop there. The next stage is to define more specialized versions of the generic situation and deduce things about each version.

This is almost explicitly a modelling toolkit with a range of similar models, ready-made inferences, and solution methods developed.

2.6 Idealization

Mathematical objects and their properties are typically idealized, even though they are often inspired by real-world phenomena that are familiar to everyone through daily experience. This is characteristic of models.

2.6.1 Idealized and practical shapes

The properties of geometrical objects are hard or impossible to achieve in practice. They are idealized versions of shapes we experience in the real world.

For example, a mathematical line has no thickness, unlike a line drawn on a piece of paper or a computer screen. A mathematical line is perfectly straight, whereas a drawn line will be seen under a microscope to have surprisingly ragged edges (as well as being wide). A mathematical line can continue without ending (and this is usually assumed by default).

Similarly, circles are perfectly round, squares are perfectly square, and so on. All this is true regardless of size, so if a shape is simply changed in size then it will retain its other properties even as it shrinks to a tiny speck or expands out across the universe.

Idealized objects can be used to make deductions and assess the accuracy of practical shapes created by humans or found in nature.

For example, imagine taking a pair of compasses and opening them to about 5 cm. Draw a circle then put the point on the circumference somewhere and mark off a point further along the circumference with the compasses (still open to the same size). Then move the point to that mark and repeat the move, over and over. The sixth mark will be over the point you started at, or very nearly.

The reasons for it being near but not perfect will be your small mistakes in positioning the point and perhaps also your compasses having eased open a little more.

How do we know these are the reasons for the slight difference? Using geometrical reasoning with an imaginary ideal circle and some equilateral triangles we can prove that the drawing process, if executed perfectly, would have brought you back to your starting point exactly.

2.6.2 Idealized numbers and practical numerals

Similarly, the fundamental number systems are defined as idealized systems, but with properties familiar to us through daily experiences. We can understand this by distinguishing clearly between:

- **Number** systems, which are idealized; and
- **Numeral** systems, which are systems for permanently labelling (some of) the numbers and doing computations, and which have practical limitations.

There are two fundamental types of number system:

- **Integers**, which are the counting numbers; and
- **Reals**, which are the measuring numbers.

Here is an explanation of the properties of these two idealized number systems. This is followed by an explanation of practical numeral systems.

The Integers are the idealized numbers we use to say 'how many'. Their meaning is intuitively obvious to people and reflected in the English language in the distinction between 'how many' and 'how much'. We determine 'how many' by counting, which matches numbers to the items to be counted, one-to-one.

If you were counting people walking past you on a path you might count up for people coming from one direction and down for people coming from the other direction. If you start at zero you could end the day with a negative or a positive total, or end back at zero.

(Imagine the idealized numbers as being each unique by some invisible means. They are not named. Names are applied later by a practical numeral system.)

The Integers allow counting, including counting backwards past their starting point. They are assumed to have one number that is the starting point (usually to be labelled 0) and to extend up and down without gaps and without end.

Having 'no gaps' means that, however many items you need to count, there is an Integer that corresponds to that many items (negative or positive). Going on 'without end' just means that we never worry about running out of Integers, no matter how far we go in either direction.

Other properties sometimes associated with Integers are really linked to operations using them. The most fundamental are comparison (which is bigger?) and addition. Others can be derived from these.

Subsets of the Integers include Natural numbers, prime numbers, and square numbers.

The Real numbers are used for measuring. They answer the question 'how much'. They are qualitatively different from the numbers used to say 'how many'.

The Real numbers have a starting point (usually to be labelled 0) and extend up and down without gaps and without end. In this context 'without gaps' means that any extent you wish to measure has a corresponding Real number, with no exceptions. Again, 'without end' means that we never worry about running out of Real numbers, no matter how far in the positive or negative direction we go.

Other properties sometimes linked to Real numbers should really be linked to operations using them. The most fundamental are comparison (which is bigger?) and addition. Others can be derived from these.

Subsets of the Reals include Rational numbers, irrational numbers, transcendental numbers, and computable numbers.

The two fundamental idealized number systems[1] (for counting and measuring) are logically different but related to the practical numeral systems that have been

[1] Other mathematical objects using numbers include:

- Bundled numbers e.g. vectors, matrices, determinants.
- Mistaken inventions that can be retired e.g. hyperreals, surreals, complex numbers (which should be replaced with vectors and suitable vector operations, avoiding the square root of -1, which does not exist or have any real-world meaning, see Leitch (2017a)).

developed to label the numbers permanently and facilitate computations.

Numerals[2] include labels such as 1, 3, 3.1, 3 x 10^{17}, 10010110, VIII, and 2A3. On computers we label numbers using a variety of systems. Different numbers of binary bits may be used. Negative numbers may or may not be represented and there are alternative techniques for doing so. Floating point is another technique to label numbers that can be larger or smaller than otherwise would be the case.

On computers there is almost always a limit to how many labels can be created by a numeral system. Usually, Integers are labelled without gaps but there are numbers above and below that are out of range and not labelled. With Real numbers there are, in effect, frequent gaps between labelled Reals as well as regions above and below with no labels at all. When the floating point technique is used the gaps between labelled Reals are not of fixed size but tend to grow as the numbers themselves grow.

Numerical software has to check frequently for calculations that go out of range and may also check for and try to reduce rounding errors.

When we write numerals by hand (usually using decimal notation) the limit on number size is not specific but still we cannot, in practice, work with very large numbers or very small numbers. Too much paper and concentration are required.

While the designs of the idealized number systems are very simple, the designs of practical numeral systems are complicated and have improved greatly over the centuries. For example:

- Roman numerals are awkward to compute with compared to the Arabic system with place value.
- Binary numbers have become important because they can more easily be represented in electronic circuits.

There is still scope for improvement in the design and use of numeral systems.

Idealized number systems can be used to deduce conclusions that are exact and reliable. One use of these is to assess the accuracy of computations using practical numeral systems.

For example, using Real numbers it is always true that if

$$a + b = c$$

then

$$c - a = b.$$

However, computer arithmetic can fail to replicate this when a is very large and b is very small. On Excel 2010, if

$$a = 100{,}000{,}000$$

and

$$b = 0.000{,}000{,}1$$

then $c - a$ is returned as zero, not $0.000{,}000{,}1$. This is because of rounding.

(Python 3.7 on my laptop does better. Although it does not give $0.000{,}000{,}1$ as the final result, it gets close at $0.000{,}000{,}104{,}308{,}128{,}356{,}933{,}6$.)

There are two other ways that numbers get labelled, but both are different in principle from the permanent labels provided by numeral systems.

When writing mathematics it is common to introduce letters to represent numbers or other objects that are either variable or currently unknown. For example, x_1 and

[2] Wikipedia offers an interesting page listing numeral systems through history. https://en.wikipedia.org/wiki/List_of_numeral_systems

x_2 might be introduced as the names of two Real numbers. This is one type of label.

Another type of label is exemplified by π and e. These are particular Real numbers with great significance in mathematics that have been dignified with names of their own outside the usual framework of number labels. In these two cases, that is necessary because decimal notation cannot represent these values exactly no matter how many digits follow the decimal point.

Other, less prestigious, numbers that have no exact representation in decimal notation are often referred to using a formula for calculating their value, e.g.:

- $\sqrt{2}$, $\sqrt{3}$, $\sqrt{5}$, $\sqrt{6}$, $\sqrt{13}$
- $\frac{1}{3}$, $\frac{1}{6}$, $\frac{1}{7}$, $\frac{1}{9}$, $\frac{1}{11}$, $\frac{1}{33}$
- $\sin[20]$, $\cos[125]$, $\tan[12]$

There is also the dotted notation for recurring decimal digits. This refers to the mathematical limit that would be approached if more and more decimal digits were added according to the repeating pattern.

2.7 Example: Counting and arithmetic

This is the first of five examples of mathematical modelling toolkits.

The ability to count and use basic arithmetic is useful in many ways. It allows some useful predictions.

> E.g. If you are preparing party bags for a child's party then you can predict if there will be the right number at the end of the party if you know how many children are coming and count the bags.
>
> E.g. If you have 54 sheep and buy 23 more then you will expect to have 77 sheep in total and when you count them all you should find your prediction is correct.
>
> E.g. You can identify when an item has been lost or stolen by repeatedly counting, because your model predicts no change.
>
> E.g. If you have 25 fish to divide between 5 people then you can give 5 to each person as they come forward. This predicts, correctly, the number you would eventually give each person if you went round in a circle giving fish one at a time until they were all handed out.
>
> E.g. To equip 200 archers with bows and a stock of 10 arrows each will require 200 bows and 2,000 arrows. If you know more you can predict how long it will take to do this and how much it will cost.

The modelling is very simple but extremely accurate in the situations where we usually do it. We often feel as if we are dealing directly with reality, not a model.

2.8 Example: Differential equations

Models based on differential equations are collections of one or more equations where at least one of the terms of the equation is a derivative (i.e. a rate of change of one variable with respect to another). These equations imply the values of their variables at particular points, usually in time, but do not allow them to be calculated directly.

'Solving' a system of differential equations usually means finding formulae that can be used to calculate directly what values variables will have at any particular time.

Although many of these models were developed to represent specific phenomena (e.g. movement of an object subject to forces, flow of heat) they are also classified by their mathematical properties.

For example, the model:

$$\frac{dy}{dx} = 3y + x^2$$

is classified as an ordinary differential equation, first order, linear, and heterogeneous. It is an example of a general type that looks like this:

$$\frac{dy}{dx} + f[x]y = g[x]$$

The next model is not of that type. The model:

$$\frac{dy}{dx} = yx^2$$

is an ordinary differential equation, first order, non-linear, but separable. 'Separable' means it is also an example of the more specialized form:

$$\frac{dy}{dx} = f[x]g[y]$$

and a particular solution method is applicable, which is to solve:

$$\int \frac{1}{g[y]}\,dy = \int f[x]\,dx.$$

The mathematical properties of the models help to identify which equations and systems of equations can be solved, and how. Sometimes an equation or system of equations cannot be solved but properties of its solution can be deduced.

It is typical of mathematical models that, as they get more realistic, they also get more complicated and, quite soon, it is impossible to solve them symbolically. In other words, there is no simple formula that will deliver the predictions and other answers desired.

So, instead, we now use computers to apply numerical methods that provide the required answers, approximately but very accurately.

One of the first and simplest such methods is Euler's method. It is a form of simulation. The algorithm starts with the variables of the system having their known starting values. It then moves ahead in small steps of time, calculating the various derivatives approximately and applying them to work out the values of the variables at the next time step.

Although errors tend to accumulate, they can be kept smaller by using very small time steps. There are also more sophisticated alternatives to Euler's method that use techniques to reduce those errors.

Such methods are very flexible and able to simulate the behaviour of very complicated systems of differential equations. Mathematical analysis has been used to understand how the errors grow as the simulation goes further into the future.

Development of the mathematical modelling toolkit(s) built around differential equations began over 300 years ago. At that stage the toolkit contained the ideas of continuous variables, derivatives, and a tiny number of differential equation forms for which solutions were known.

Over subsequent years and centuries more and more types of differential equation and system were studied and solutions to some types were worked out[3]. The toolkit expanded and one could also see parts of it as sub-toolkits.

The modern situation is that there are toolkits for specific types of equation but beyond that is a more general purpose toolkit that features:

[3] A sense of how far this has progressed can be gained by scrolling to the bottom of the Wolfram MathWorld page on ordinary differential equations: https://mathworld.wolfram.com/OrdinaryDifferentialEquation.html

- Variables
- Differentials
- Function types
- The ability to put these together as systems of differential equations of almost any form
- Software to simulate systems numerically using the model, even when it is complicated and there is no known symbolic solution.

The toolkits also extend to equations that feature other elements beyond derivatives.

Courses and textbooks teaching differential equations typically do not reflect this idea of toolkits and instead move through the various types of equation in only a partly systematic manner.

2.9 Example: The Kelly Criterion

The Kelly Criterion is an objective for repeated bets. The idea is to find the fraction of your current wealth to bet on each occasion that leads to the highest expected value of the logarithm of your final wealth. This is the Kelly fraction. Risking only a fraction of your wealth each time ensures that you do not run out of money completely, which would prevent further betting.

Kelly (1956) was trying to derive information theory by a different method. Along the way he introduced the Kelly Criterion, which has been more influential than his alternative derivation of the formulae for information.

Over time, the variety of situations in which this criterion can be applied has increased. In other words, the toolkit has expanded and now covers more situations.

Kelly's first paper covered three situations involving discrete bets with sensible odds. Later work covered two-outcome bets with any odds and investments in securities whose values fluctuated lognormally – but only for a portfolio of one type of stock and only if we can continuously and without cost adjust our holding.

Other work considered the situation where we do not know the long run relative frequencies of the outcomes in the repeated bets, but can perhaps learn them over time. Others looked at discrete bets with more than two possible outcomes, and at share portfolios with multiple stocks.

Some approaches to multiple discrete outcomes require numerical methods to solve the resulting equations. Another approach is to simulate hundreds of thousands of bets using different fractions of wealth and select the best performing. This is often quite practical and quick, and makes it possible to consider awkward real-world issues like transaction costs.

The Kelly criterion is a much more recent development than differential equations and has been studied far less. Work on particular situations has emerged in individual papers and there is a need for a systematic taxonomy of applications and techniques. Software to support this would also be helpful.

2.10 Example: Z

In the 1980s there was a surge of interest in using mathematical models as specifications of computer systems. It was thought that mathematics offered a way to write more precisely and less ambiguously about systems and that this would be advantageous, which it is if you have the intellect and skill to do it.

One prominent example was the Z style of specification. Like similar approaches this used discrete mathematics, especially set theory, to build specifications of the state of a system and operations that changed this state.

The approach offers clever solutions to problems that only arise when mathematical models get big. These clearly draw on computer programming practices. In particular:

- Object names are words and abbreviated words chosen for being memorable, rather than single letters.
- The mathematical type of every object is specified in advance of its use.
- Sections of specifications are bundled into 'schemas', which have a distinctive style of three-sided box.
- The sections of specifications can be named so that the names can be used later to reference (in effect, to include) those blocks of mathematics in later schemas.
- Mathematical text is interleaved with plain English statements of the same points. The English is more accessible but the mathematical text removes doubts over interpretation.

The skills needed from users of this approach include wise decisions on how to decompose complicated models into smaller objects, with memorable names and memorable behaviour, that will be needed often during modelling.

A useful book documenting Z is the reference manual by Spivey (1988). Chapter 4 is about what he calls its 'mathematical toolkit', but this term could have been used for much more of the content.

2.11 Example: CSP

Communicating Sequential Processes (CSP) is a notation for specifying (modelling) the behaviour of systems that operate in parallel (Hoare, 1978, 1985). The idea is that such systems can be seen as separate processes that operate in parallel with others and communicate with them by sending signals along channels.

CSP is a mathematical notation that allows compact specifications of such systems.

In some ways it is similar to Z but CSP shows how mathematical techniques can be applied to yet another area where precise modelling and logical reasoning are useful.

3. Philosophical aspects

The perspective described above answers some of the best known philosophical questions about mathematics.

3.1 Invented or discovered?

In this perspective, mathematics is invented but, usually, to mimic properties discovered in the real world. Mathematical inventions are brought into being by writing definitions and building on them by deduction[4].

Because of the desire to mimic reality, this modelling is not unconstrained invention, although there are design alternatives and clever choices can make a big practical difference.

Key discoveries have already been made. These concern the permanence of reality and the way it maintains counts and measurements over time. The correspondence between model and reality in the case of counting and measuring things is so good that it is often hard to see that our calculations are using models at all.

There is also invention in the choice of notation, the details of definitions, and the selection of statements to prioritize as

[4] At least, this is how it is presented. In practice, surely there have been instances where the definitions were crafted to result in the deductions desired.

theorems. There are some choices between alternative conventions that could be made differently, though convenience usually argues for one convention in particular. For example, probabilities add up to 1 because of a convention, well chosen, hundreds of years ago.

Some statistical techniques are also just one approach of many that could have been chosen. For example, standard deviation is not the only possible measure of spread.

Measures of performance are also invented. For example, many alternative rules for scoring probabilistic forecasts have been proposed.

Methods for reasoning from models often involve substantial efforts of invention.

Almost certainly some mathematical models have been created without an effort to mimic reality. When this is done there is a higher risk that the model will be useless. There is also a higher risk that some people will form wrong beliefs about the world, reassured by the apparently logical derivation of the model's predictions.

The extent to which models have been created without a deliberate attempt to capture real world phenomena is sometimes obscured by the way mathematics is often presented. A model that was originally developed for a real world situation gets generalized and linked to other theories, then stripped of references to its original context, then taught to undergraduates as if the theory came first.

3.2 Construction of numbers

In the perspective described above, the counting numbers (Integers) and the measuring numbers (Reals) are based on intuitive notions of quantity so familiar and universal that justification is not needed. These notions agree with our experience with the real world and everyone knows it. A foundation for these concepts that is less familiar to everyone would be a weaker, less useful foundation.

The counting numbers are qualitatively different from the measuring numbers and so the counting numbers are not a subset of the measuring numbers. It only seems that way because we have re-used numeral systems developed to label counting numbers to label points on the Real number line. If a completely different form of labelling had been used for the Reals we would not notice any apparent similarity.

And yet, despite the simplicity and intuitive appeal of the counting and measuring numbers, some late 19th and early 20th century mathematicians thought it would be helpful to, somehow, deduce numbers from something logical and (they thought) even more fundamental.

One famous attempt to define Natural numbers involves a recursive definition:

$0 = \{\} = \emptyset$

$1 = \{0\} = \{\emptyset\}$

$2 = \{0,1\} = \{\emptyset,\{\emptyset\}\}$

$3 = \{0,1,2\} = \{\emptyset,\{\emptyset\},\{\emptyset,\{\emptyset\}\}\}$

and so on.

This defines numbers using just empty sets, which is so bizarre it could be some kind of joke. Where do these ideas come from? What is the point? The lack of any straightforward connection to our universal experience of the world is a fundamental flaw in this approach.

Various attempts have been made to 'construct' Real numbers, often from Natural numbers or Integers. They tend to be complicated and to include elements of operations with the numbers as well as attaching the labels 0 and 1. Some ideas are baffling and bizarre.

In the perspective described earlier, the Reals are just as fundamental as the Integers, being equally strongly based in our intuitive thinking. There is nothing to be gained from trying to build one from the other.

3.3 Infinity

In the perspective explained earlier, there is no infinity, either as a number or anything else. The idealized number systems simply continue without end and we never check for going out of range or for rounding errors.

The numbers can be as big as we like or, in the case of Reals, as small as we like.

For practical suggestions on eliminating references to infinity from mathematical writing, see Leitch (2016).

4. Practical aspects

The idea of mathematics as modelling toolkits suggests several exciting areas for improvement in practices[5].

This would involve a conscious attempt by anyone involved to create or contribute to modelling toolkits that others can understand, learn to use, and apply efficiently (usually with computer support). This means going beyond just writing proofs, defining objects, and developing methods. It means making them available to others in a form that promotes their more rapid take up and efficient use. It means thinking of users and useful applications from the very start.

This might have a number of effects.

4.1 Toolkit quality

More attention could be paid to the design and presentation of toolkits. This would cover improved understanding of the desirable attributes of toolkits and how to achieve them.

The most fundamental design choices are ontological i.e. they concern what types of thing are considered to exist in the model. In Z specifications these include the fundamental sets/types used in the model.

These things should be clearly linked to the common experiences of users. If this is not done it is hard to apply the models reliably.

Names used in a toolkit (e.g. for objects, operators, theorems, solution methods) should be memorable and not misleading. For historical reasons, mathematics has many examples of names that are misleading today. An example is the term 'random variable', which refers to objects that are neither random nor variables.

Notation should be compact yet also unambiguous, consistent, understandable, memorable, and scalable to large models. These requirements often conflict. It is usually necessary to accept less compact notation when models get larger. It is occasionally necessary to accept slightly less compact notation to get rid of logical inconsistencies.

There will also be opportunities to make wise choices about what model types to focus on, how to craft definitions, which operators to define and use, and which deductions to highlight (as theorems, identities, etc.). This is a point that is easier to understand if you have written large models or computer programs. Often, the early stages involve some decisions that get revised, perhaps more than once, as you search for a structure that will make the work easier.

[5] For a broader look at ways that mathematics could make a better contribution to society, see Leitch (2017b).

These choices should help to give models that are compact and understandable, and help users become adept at efficient reasoning and problem solving.

An example of a poor initial choice is the invention of quaternions. When Hamilton thought of them in 1843 he thought he had made a breakthrough in modelling mechanical systems in three dimensions. However, his methods were rapidly displaced in the 1880s by simpler, clearer, more logical vector methods.

There are large opportunities to improve the way model toolkits are structured when presented and taught. It will help to have structured breakdowns of model types and the situations where they can be applied. (This is discussed in more detail in the next sub-section.)

It would also be useful if all the elements needed to use each tool were identified and presented in a standardised way for each tool. Leitch (2019) presents ideas on specifying regression methods that show the elements needed in this case and suggests how standardised presentation might look.

4.2 Organization and presentation

Many areas of mathematics would benefit if someone wrote organizers for the model types and other elements of toolkits. Such organizers might reflect the various uses of the models, the model structures, and the methods for their use.

Typical organizers might include:

- Breakdowns of model types (in the form of structured lists or trees), starting from the most general and working down to more specialized versions. (Tables might be useful as another way to structure the breakdown.)
- Lists or tables of alternative ways to represent models.
- Breakdowns of model application situations.
- Tables or trees mapping modelling situations to suitable models.
- Breakdowns of tasks using models (e.g. make prediction, work out optimal parameter values).
- Breakdowns of methods for doing tasks.
- Tables mapping models and tasks to methods for doing those tasks with those models.

Two small examples showing organizers for school-level topics in mathematics are provided in Appendix A and Appendix B.

Such organizers could be used in various ways:

- Placed on a website and used as an index into relevant books, research papers, and other websites. Anyone trying to use the modelling toolkit who needs more information about something can use the organizers to find it.
- Mathematical innovators might be encouraged to do work in areas identified by organizers where there are gaps. Well-designed organizers might help to generalize problems and identify variations in modelling that still need attention.
- Educators could use them to structure courses and textbooks. Learners would feel they are being taken, systematically, through a collection of useful modelling tools.

Mathematical innovators writing about their innovations might be motivated to give more practical tips on using their developments.

4.3 Automation

Although there are already many amazing computer tools for mathematics, further opportunities to develop computer tools to support model toolkits may become more obvious.

The computer tools could include libraries/packages and applications.

4.4 Choice of development area

Some people looking to contribute to the development of mathematics might choose to go back to existing developments and repackage them as well-organized modelling toolkits, with organizers, tips for use, software support, and clearer explanations[6].

In other cases, the idea of promoting toolkits might inspire choices that focus on important application areas where there is a good prospect for progress.

4.5 Education

4.5.1 Teaching toolkit development

Teachers of mathematics at school and university might look to go beyond solving individual problems within a brief academic examination or assignment.

Professional use of mathematics often involves tackling a series of similar but different problems, looking to improve over time. This requires:

- use of existing modelling toolkits; and
- development of specialised modelling toolkits for particular types of problem.

The skills of developing, extending, refining, packaging, and automating toolkits are important to this but are rarely if ever taught explicitly at school or at university.

That should change.

[6] See Leitch (2009) for ideas on writing mathematics more clearly.

4.5.2 Inspiring problems

Students, especially at school level, often question if the mathematical techniques they are learning will ever be useful to them outside academic examinations. They are right to raise this issue because a large proportion of questions children are asked, especially at secondary school level, are contrived puzzles of no direct use to anyone. This surely reduces motivation for at least some students, especially those who want to do something useful with their lives.

Instead, students could be taught, explicitly, that they are learning to use (and develop) tools from mathematical modelling toolkits and students could be tested using questions that *only* test useful skills.

Ideal questions have a real-world setting, a task that a person might really want to complete, and a story that puts the student into that situation as a modeller using models to calculate useful results.

It is not necessary that all questions have a realistic setting, task, and story. Many questions simply require methods for reasoning with models. However, the inclusion of many questions with fully realistic elements should help to focus learning and testing as well as motivate students.

Here are some groups of similar questions to illustrate these ideas. The first two examples illustrate the familiar difference between realistic and unrealistic settings.

> **Unrealistic:** A linear sequence starts:
>
> $$a + b, a + 3b, a + 5b, \ldots$$
>
> Its second term has the value 11. Its fifth term has the value 23. Work out a and b.
>
> **Realistic:** Records of faults in a factory show them to have risen linearly for 5 days. The number of

faults on day one was 7 and on day 5 it was 23. You want to create a model that will predict the number of faults in future days if this trend is not corrected and linear increases continue. Work out this model and predict the faults for the 6^{th} and 20^{th} days.

To succeed with the realistic problem above the student has to recognize that there is the opportunity to apply an arithmetic sequence model and then apply it. Predictions are asked for that might be wanted in a real situation. The desire for these predictions is described as coming from the learner.

Here is another group of example questions, this time adding a variant of the question that develops the skill of toolkit making.

Unrealistic: A sphere has radius $2x$ cm and a cone has radius $3x$ cm. If the sphere and cone have the same volume, what is the ratio of the cone's radius to its perpendicular height?

Realistic: A lump of clay rolled into a perfect sphere has a radius of 20 cm. If it is made into a perfect cone with radius 30 cm, what will be the perpendicular height of the cone?

Toolkit making: A large number of lumps of clay of different sizes are to be made into cones in a workshop. Each lump is to be rolled into a sphere and measured before being reshaped into a cone with perpendicular height twice its radius. You want to work out formulae that give the radius and height of each cone from the diameter of the ball that makes it. Choose suitable notation.

And here are questions for much younger students.

Unrealistic: What is 12 – 3?

Realistic: If dad has baked 12 muffins and says he has eaten only 3 of them, how many should be left?

In this last example the realistic problem has the elements of modelling, prediction, and motivation in a very simple, natural form.

Other questions might use pre-existing, scientifically derived models to calculate interesting and useful numbers such as calorie and nutrient requirements for different people, health risks, and the average time needed to learn useful skills.

4.5.3 Teaching about numerical issues

It would also be worth considering giving more prominence to the design of practical numeral systems and computation processes from an earlier age. Rounding errors first become important in school mathematics when children learn about different ways to round numbers and realise that some numbers cannot be represented exactly using decimal notation.

There is a particular type of two-stage question at GCSE level (a UK exam usually taken at age 16) where a student who rounds excessively at the first stage may take that error into the second stage, resulting in an error that is too large at the final stage. This is a major teaching point but currently given little attention.

Much more could be taught about accumulation of rounding errors in larger calculations. There are also some simple skills to learn that use the best precision an electronic calculator can provide.

Realistic modelling tends to require complicated models, which usually can only be solved by numerical methods, wherein rounding errors and going out of range are important problems. That is why these issues are important for practical use of mathematics.

Topics to consider for school and undergraduate level teaching include:

- Design criteria for numeral systems.
- The properties of commonly used numeral systems, including the range and accuracy limitations of common computer representations of numbers.
- Distribution of labels across the Reals and alternative designs (floating point, tapered floating point, logarithmic).
- Management of rounding errors in calculations. For example, is there a best order for adding up a large number of numbers? How do errors propagate through calculations? What algorithms can reduce these errors?
- Iterative solution and ways to know when you are close enough to the exact answer.

5. Conclusions

Mathematics is, already, a collection of modelling toolkits, but with some innovation and effort we can embrace that idea and do it better.

There are opportunities to develop, present, teach, and use mathematics in better ways by applying this perspective.

6. References

Hoare, A. (1978). Communicating sequential processes. *Communications of the ACM*. 21 (8): 666-677. Available at: https://www.cs.cmu.edu/~crary/819-f09/Hoare78.pdf

Hoare, A. (1985). *Communicating sequential processes*. Prentice Hall. Available at: http://lass.cs.umass.edu/~shenoy/courses/677/readings/Hoare.pdf

Johnson, W. (1998). The ricochet of spinning and non-spinning spherical projectiles, mainly from water. Part II: An outline of theory and warlike applications. *International journal of impact engineering*, *21*(1-2), 25-34. Available at: http://www2.eng.cam.ac.uk/~hemh1/dambusters/ricochet_johnson2.pdf

Kelly, J.L. (1956). A new interpretation of information rate. *The Bell System Technical Journal*, vol. 35, no. 4, pp. 917-926. Available at: https://cpb-us-w2.wpmucdn.com/u.osu.edu/dist/7/36891/files/2017/07/Kelly1956-1uwz47o.pdf

Leitch, M. (2009). *How to write mathematics clearly*. Available at: http://www.workinginuncertainty.co.uk/clearmaths.shtml

Leitch, M. (2016). *Eliminating infinity*. Available at: http://www.workinginuncertainty.co.uk/infinite.pdf

Leitch, M. (2017a). *Eliminating imaginary numbers*. Available at: http://www.workinginuncertainty.co.uk/imaginary.pdf

Leitch, M. (2017b). *Efficient mathematics*. Available at: http://www.workinginuncertainty.co.uk/efficient_maths.pdf

Leitch, M. (2019). *Specifying regression*. Available at: http://www.workinginuncertainty.co.uk/regrspec2.pdf

Spivey, J.M. (1988). *The Z Notation: a reference manual*. Prentice Hall. Available at: http://spivey.oriel.ox.ac.uk/corner/Z_Reference_Manual

7. Appendix A

The following example of organizers for a very small toolkit is based on the topic of sequences at GCSE level in the UK.

First, the various models can be represented with an nth term formula, or using a term-to-term formula (and one or two initial values). The coverage at this level is patchy.

Models:

Type	**Nth term**	**Term-to-term**
Arithmetic / linear	$an + b$	$U_{n+1} = U_n + d$
Quadratic / triangular	$an^2 + bn + c$	-
Cubic	n^3	-
Geometric	-	$U_{n+1} = U_n \times r$
Fibonacci	-	$U_{n+2} = U_n + U_{n+1}$

Students are taught to do various tasks with these models.

Tasks:

- Inferring models
  - Terms → n^{th} term defn.
  - Terms → term-to-term defn.
  - Infer a sequence from constraints
  - Infer parameters to arrive at an n^{th} term with a particular value
- Calculating terms
  - n^{th} term defn. → terms
  - Term-to-term defn. → terms
- Converting between model forms
  - n^{th} term defn. → term-to-term defn.
  - Term-to-term defn. → n^{th} term defn.
- Miscellaneous
  - Is a number a term?

Applications of sequences are not explored very far but a couple are mentioned.

Applications:

- Geometric
  - Compound interest
  - Repeated % losses

8. Appendix B

The following examples of organizers for a small toolkit are based on the topic of differential equations in the Further Mathematics A level offered by Edexcel in the UK.

Model types mapped to solution methods:

Single equations

- Linear
  - General
    - $a_n[x]y^{(n)} + a_{n-1}[x]y^{(n-1)} + \cdots + a_1[x]y' + a_0[x]y = Q[x]$
  - First order
    - General
      - $a_1[x]y' + a_0[x]y = Q[x]$
    - No y term
      - Homogeneous
        - $\frac{dy}{dx} = 0$, trivial, y is a constant
      - Non-homogeneous
        - $\frac{dy}{dx} = f[x]$, solve by integrating f
    - Includes a y term
      - Homogeneous
        - $\frac{dy}{dx} + f[x]y = 0$, solve with separation method
      - Non-homogeneous
        - $\frac{dy}{dx} + f[x]y = g[x]$, solve with integrating factor method
  - Second order
    - General
      - $a_2[x]y'' + a_1[x]y' + a_0[x]y = Q[x]$
    - Only y'' term
      - Homogeneous
        - $\frac{d^2y}{dx^2} = 0$, trivial, solve by finding line
      - Non-homogeneous
        - $\frac{d^2y}{dx^2} = f[x]$, solve by integrating f twice
    - All terms or y' only is missing or y only is missing

Homogeneous

General form

Not covered – no general solutions

Constant coefficients (no functions of x)

$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$, solve with discriminant

Non-homogeneous

General form

Not covered – no general solutions

Constant coefficients (no functions of x)

$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f[x]$, solve with $y = CF + PI$

Non-linear

First order

No y term

$\frac{dy}{dx} = f[x] \times f[y]$, solve by separation method

$\frac{dy}{dx} = f[x] + f[y]$, no general solution (yet)

Simultaneous coupled pair of differential equations

First order equations

Homogeneous

$\frac{dx}{dt} = ax + by$

$\frac{dy}{dt} = cx + dy$, solve by eliminating to form 2nd order

Non-homogeneous

$\frac{dx}{dt} = ax + by + f[t]$

$\frac{dy}{dt} = cx + dy + g[t]$, solve by eliminating to form 2nd order

Tasks using models:

Find solutions

Find form of solution

Find particular solution

Verify given solutions

Verify form of solution

Verify particular solution

Modelling situations:

Fluids

- Filling and emptying containers with chemical mixing
- Transfer of pollutants/nutrients from one body to another (coupled differential equations)

Living things

- Growth of bacteria
- Predation of one species by another (coupled differential equations)

Moving particles

- In lines/curves
  - Velocity a function of time
  - Acceleration a function of time
  - Velocity a function of displacement
  - Acceleration a function of time and velocity
- Harmonic motion (particles on springs, swinging, rotating)
  - Simple harmonic motion (no resistance or forcing)

    $\frac{d^2x}{dt^2} = -\omega^2 x$, $x = R\ \sin[\omega t + \alpha]$, $\ddot{x} = v\frac{dv}{dt}$
  - Damped harmonic motion (with resistance proportional to speed)

    $\frac{d^2x}{dt^2} + k\frac{dx}{dt} + \omega^2 x = 0$, ($2^{nd}$ order, homogeneous)
    - Heavy damping ($k^2 > 4\omega^2$)
    - Critical damping ($k^2 = 4\omega^2$)
    - Light damping ($k^2 < 4\omega^2$)
  - Forced harmonic motion (with resistance and pushing a function of time)

    $\frac{d^2x}{dt^2} + k\frac{dx}{dt} + \omega^2 x = f[t]$, ($2^{nd}$ order non-homogeneous)