An Open Response to the USPTO -- Physical Aspects of Mathematics

Sunday, September 26 2010 @ 10:32 PM EDT

Contributed by: PJ

An Open Response to the USPTO — Physical Aspects of Mathematics
by PolR, author of <u>An Explanation of Computation Theory for Lawyers</u>.

The USPTO has issued a <u>request for comments</u> on their new interim guidance, <u>Interim Guidance for Determining Subject Matter Eligibility for Process Claims in View of Bilski v. Kappos</u> [PDF]. They <u>invited comment</u> by Monday, September 27, from the public on three questions in particular:

> 1. What are examples of claims that do not meet the machine-or-transformation test but nevertheless remain patent-eligible because they do not recite an abstract idea?

> 2. What are examples of claims that meet the machine-or-transformation test but nevertheless are not patent-eligible because they recite an abstract idea?

> 3. The decision in *Bilski* suggested that it might be possible to "defin[e] a narrower category or class of patent applications that claim to instruct how business should be conducted," such that the category itself would be unpatentable as "an attempt to patent abstract ideas." Bilski slip op. at 12. Do any such "categories" exist? If so, how does the category itself represent an "attempt to patent abstract ideas?"

They ask in effect how to tell an abstract idea from an application of the idea. This article suggests answers to that question from the perspective of a computer professional.

The Supreme Court has <u>reaffirmed</u> the trilogy of cases <u>Gottschalk v. Benson</u>, <u>Parker v. Flook</u> and <u>Diamond v. Diehr</u> concerning the non-patentability of mathematical algorithms:

> In *Benson*, the Court considered whether a patent application for an algorithm to convert binary-coded decimal numerals into pure binary code was a "process" under §101. 409 U. S., at 6467. The Court first explained that "'[a] principle, in the abstract, is a fundamental truth;

an original cause; a motive; these cannot be patented, as no one can claim in either of them an exclusive right.'" *Id.*, at 67 (quoting *Le Roy*, 14 How., at 175). The Court then held the application at issue was not a "process," but an unpatentable abstract idea. "It is conceded that one may not patent an idea. But in practical effect that would be the result if the formula for converting . . . numerals to pure binary numerals were patented in this case." 409 U. S., at 71. A contrary holding "would wholly pre-empt the mathematical formula and in practical effect would be a patent on the algorithm itself." *Id.*, at 72.

In *Flook*, the Court considered the next logical step after *Benson*. The applicant there attempted to patent a procedure for monitoring the conditions during the catalytic conversion process in the petrochemical and oil-refining industries. The application's only innovation was reliance on a mathematical algorithm. 437 U. S., at 585586. *Flook* held the invention was not a patentable "process." The Court conceded the invention at issue, unlike the algorithm in *Benson*, had been limited so that it could still be freely used outside the petrochemical and oil-refining industries. 437 U. S., at 589590. Nevertheless, *Flook* rejected "[t]he notion that post-solution activity, no matter how conventional or obvious in itself, can transform an unpatentable principle into a patentable process." *Id.*, at 590. The Court concluded that the process at issue there was "unpatentable under §101, not because it contain[ed] a mathematical algorithm as one component, but because once that algorithm [wa]s assumed to be within the prior art, the application, considered as a whole, contain[ed] no patentable invention." *Id.*, at 594. As the Court later explained, *Flook* stands for the proposition that the prohibition against patenting abstract ideas "cannot be circumvented by attempting to limit the use of the formula to a particular technological environment" or adding "insignificant postsolution activity." *Diehr*, 450 U. S., at 191192.

Finally, in *Diehr*, the Court established a limitation on the principles articulated in *Benson* and *Flook*. The application in *Diehr* claimed a previously unknown method for "molding raw, uncured synthetic rubber into cured precision products," using a mathematical formula to complete some of its several steps by way of a computer. 450 U. S., at 177. *Diehr* explained that while an abstract idea, law of nature, or mathematical formula could not be patented, "an application of a law of nature or mathematical formula to a known structure or process may well be deserving of patent protection." *Id.*, at 187. *Diehr* emphasized the need to consider the invention as a whole, rather than "dissect[ing] the claims into old and new elements and then . . .

ignor[ing] the presence of the old elements in the analysis." *Id.*, at 188. Finally, the Court concluded that because the claim was not "an attempt to patent a mathematical formula, but rather [was] an industrial process for the molding of rubber products," it fell within §101's patentable subject matter. *Id.*, at 192-193.

In *Benson* and *Flook* the unpatentable algorithm is software. But it is recognized that an algorithm may take the form of hardware or software. In another case, In Re Alappat, the Federal Circuit had acknowledged the possibility that a patent on a circuit could read on a computer programmed with software and that this could make the algorithm unpatentable. They eventually ruled the circuit was patentable on the basis that they thought (among other things) the algorithm was not a mathematical algorithm in the sense of *Benson*, *Flook* and *Diehr* but this doesn't change the fact that they first considered the possibility of going the other way. (Also note the dissenting opinions in *Alappat*, which raised precisely the danger of patenting mathematical discoveries.)

The court in *Diehr* held an industrial process to cure rubber that uses an algorithm is patentable. This court made a distinction between the algorithm and a non-algorithmic process that comprises the algorithm as one of its elements. This article discusses a specific aspect of this issue. Could a patent on a circuit, as in *Alappat*, or a patent on physical computational activity, such as in *Benson* and *Flook*, be considered patents on an abstract idea? This is a notion that has a sound basis in mathematics. The purpose of this article is to explain this basis and show some possible problems that occur when one does not take it into consideration.

Physical Tools Are Necessary to the Practice of Mathematics

The relationship between mathematics and the physical world is often described, outside of the Supreme Court trilogy, with two notions:

- Abstract ideas are disembodied thoughts.

- Mathematics is often used as an abstract language to describe physical phenomena and laws of nature.

Taken together these ideas suggest that whenever we have a relationship between mathematics and the physical world, mathematics is always used to provide an abstract disembodied model describing the reality. This view, however, is not always accurate.

Take a pocket calculator. If one calculates the energy contained in matter with the formula $E=mc^2$, does the formula describe the calculator? Of course not. The relationship goes the other way round. The calculator is a tool to do

mathematics, not an application of the formula. I like to say the calculator is a physical model of the abstract mathematical calculation. Instead of doing the calculation in our head, we enter the numbers in the machine and watch what it does. The calculator shows us how the mathematics work.

The idea that mathematics can model the physical world is sometimes used to discount the use of mathematics in a patent claim. The argument is that it is not because an invention can be described mathematically that it is abstract and non-patentable, because otherwise everything is mathematics and nothing would be patentable. This is a fair point, but what if the modeling relationship goes the other way round? What if, like a pocket calculator, the physical entity is used to represent the mathematical abstraction? Isn't this the very issue raised by the Supreme Court trilogy? It is hard to see how the questions raised by *Benson*, *Flook* and *Diehr* can be answered without first getting an understanding of which direction the modeling relationship goes and what are the consequences.

So to summarize, when there is a relationship between mathematics and a physical entity this relationship may occur in two different ways:

> 1. Mathematics is used to provide a model of the physical reality. This is what happens, for example, when we use mathematics to describe the laws of physics.

> 2. A physical device of process is used to carry out an abstract mathematical operation. This is what happens, for example, when we use a calculator to carry out a calculation.

The task of sorting out whether or not a patent claim covers an algorithm requires that we understand which of these two alternatives is the correct one. These two situations are what I call the directions of the modeling relationship. In the first situation the direction is from the abstract to the concrete. In the second situation the direction is from the concrete to the abstract.

Any notion that the mathematics of computing are related to the computer in the same way as the laws of physics are related to physical phenomenon is erroneous[1]. In computation theory, the modeling relationship goes the other way round. This follows from common sense by observing a pocket calculator and the use of pencil and paper. This is also supported by established knowledge in abstract mathematics, computer science, and philosophy.

Another way to make the same point is to establish that the use of a physical symbolic representation is an essential requirement of the practice of mathematics with references to mathematical literature. Without some physical representation no mathematical activity can occur at all. If one is not careful, it

is possible that a patent on a physical device happens to be a patent on mathematics. This is the outcome that *Benson, Flook* and *Diehr* are meant to avoid.

The simplest explanation of this point is that mathematics is a written discipline. One cannot use mathematical concepts without writing them down except in the most simple situations. The written text is necessarily physical. But mathematicians don't care whether the text is written on paper or is electronic signals such as bits. As long as there is a usable physical representation mathematical activity can occur.

But this simple explanation doesn't do justice to the actual mathematical truth. This is not just about the ability of text to express meaning. It is also about the ability to conceive. One cannot think of or use a mathematical idea in the first place if it cannot be written. There is a considerable body of mathematical research on the topic of the connection between what can be written and which mathematical truths can be understood by humans. There are some areas of mathematics that are inaccessible because the written language is insufficient even though its use is unavoidable.

I provide later in this article a selection of the main points with reference to mathematical literature in case lawyers would like authoritative evidence. But it is not the details of the mathematical research that matter, although some of them are legally relevant for purposes other than the main topic of this article. It is that the research has actually been made. We have factual evidence that one cannot think of or use a mathematical idea in the first place if it cannot be represented physically, and this is what gives *Benson, Flook* and *Diehr* a sound mathematical basis.

Application to Patent Law: Analysis of *In Re Alappat*

Although software is a recent technology it doesn't bring a new type of relationship between the abstract and the concrete. On the contrary, the same sort of relationship has existed since antiquity. The Antikythera mechanism is an ancient mechanical computer from the first century BC. The Romans used abacuses. More modern devices are the slide rulers and the differential analyzers. All these devices are physical models for abstract mathematical computational procedures. Although the range of applications and the economic consequences of software are arguably new[2], the relationship of the physical world and the abstract mathematics is not.

To my knowledge, this is a fact that hasn't been acknowledged by the courts. This has led to erroneous analysis of the subject matter claimed in patents. A good example is the *In re Alappat* case. Its examination in the light of a correct understanding is instructive. The litigated patent is a simple one, and the

mathematical aspects are clear cut. It is also a historically important case because it introduced the useful, concrete and tangible result test. This test was found to be unworkable for empirical reasons. After years of usage, it was found that this test allows patents that shouldn't be. When we are equipped with a correct understanding of the relationship between mathematical abstraction and the physical circuit, it is possible to see *why* the useful, concrete and tangible result test doesn't lead to a correct analysis of a patent.

The *Alappat* claim covers a rasterizer for an <u>oscilloscope</u>. The oscilloscope is a device for measuring electromagnetic signals and showing the waveforms visually on a screen. With modern digital electronics, the picture is drawn pixel by pixel on a rectangular grid. Without the rasterizer, the curved lines of the waveform have an awkward visual appearance because the pixels are drawn either entirely lit or entirely black and must be located at the intersections of rows and columns on the grid. These restrictions give the curve a discontinuous or jagged appearance. The rasterizer makes the curve appear smooth by averaging the luminosity of neighboring pixels. This makes the waveform more visually pleasing to the eye.

The patented claim is directed to a digital electronic circuit. The *Alappat* court quotes the claim as follow:

> When independent claim 15 is construed in accordance with Section 112 Para. 6, claim 15 reads as follows, the subject matter in brackets representing the structure which Alappat discloses in his specification as corresponding to the respective means language recited in the claims:

> A rasterizer [a "machine"] for converting vector list data representing sample magnitudes of an input waveform into anti-aliased pixel illumination intensity data to be displayed on a display means comprising:

>> (a) [an arithmetic logic *circuit* configured to perform an absolute value function, or an equivalent thereof] for determining the vertical distance between the endpoints of each of the vectors in the data list;

>> (b) [an arithmetic logic *circuit* configured to perform an absolute value function, or an equivalent thereof] for determining the elevation of a row of pixels that is spanned by the vector;

>> (c) [a pair of *barrel shifters,* or equivalents thereof] for normalizing the vertical distance and elevation; and

(d) [a *read only memory (ROM)* containing illumination intensity data, or an equivalent thereof] for outputting illumination intensity data as a predetermined function of the normalized vertical distance and elevation.

As is evident, claim 15 unquestionably recites a machine, or apparatus, made up of a combination of known electronic circuitry elements.

This circuit carries out a computation. The symbols are digital 0s and 1s. They are used to represent numbers that have real-world meanings, specifically the sample magnitudes of an input waveform, the computed anti-aliased pixel intensity data, and other numbers that are required in the course of the computation. The circuitry performs the computation according to mathematically defined rules. The rules are not stated in the patent claim although they can presumably be found in the patent specification[3]. Examining the components reveals their function is to carry out ordinary arithmetic. The two arithmetic logic circuits perform absolute value functions to determine distance of numbers. Barrel shifters multiply (or divide) by powers of two and the ROM stores a lookup table containing values of illumination intensity data that are precomputed according to a mathematical function. There is nothing in the circuit that is not doing an arithmetical calculation. However this arithmetic is thinly veiled by the mention of the meanings of the numbers in the claimed functions.

The issue in front of the court was whether this claim was patentable subject matter because it is a circuit or whether it is non-patentable because it would pre-empt some mathematics as per *Benson*, *Flook* and *Diehr*.

My view as a computer professional is that we can answer the issue in front of the court just by looking at the circuit. It does arithmetic calculations and produces numeric answers. Arithmetic and numbers are clearly mathematics.

The court didn't look at the circuit the way I would. They looked at the text of the claim. Upon analyzing the claim the court concluded the rasterizer is not doing a mathematical calculation with this reasoning. (emphasis in the original):

Given the foregoing, the proper inquiry in dealing with the so called mathematical subject matter exception to Section 101 alleged herein is to see whether the claimed subject matter *as a whole* is a disembodied mathematical concept, whether categorized as a mathematical formula, mathematical equation, mathematical algorithm, or the like, which in essence represents nothing more than a "law of nature," "natural phenomenon," or "abstract idea." If so, *Diehr* precludes the patenting of that subject matter. That is not the

case here.

> Although many, or arguably even all, of the means elements recited in
> claim 15 represent circuitry elements that perform mathematical
> calculations, which is essentially true of all digital electrical circuits,
> the claimed invention as a whole is directed to a combination of
> interrelated elements which combine to form a machine for
> converting discrete waveform data samples into anti-aliased pixel
> illumination intensity data to be displayed on a display means. This is
> not a disembodied mathematical concept which may be characterized
> as an "abstract idea," but rather a specific machine to produce a
> useful, concrete, and tangible result.

Please notice the use of the expression "disembodied mathematical concept".
This is a phrase that is inconsistent with the notion that an abstract
mathematical calculation must be carried out physically in order to be used at
all. Then the court concludes the circuit is not an abstract idea because it yields
a useful, concrete and tangible result.

What is the useful, concrete and tangible result of the circuit? Again my answer
as a computer professional is to look at the circuit. It carries out an arithmetic
calculation and produces numeric answers in the form of bits. This is the result.
They are the electronic equivalent of the symbols a mathematician writes on
paper. The bits are electronic signals that are useful, concrete and tangible. And
still they are numbers produced by an arithmetic calculation. When I look at the
circuit I don't see how the useful, concrete and tangible result test separates
mathematics from something that is not mathematics, *because the symbols
produced by any computation will always be useful, concrete and tangible.*

The court didn't follow my approach. They didn't look at the circuit. They looked
at the text of the patent claim. The court found that the display of the bits on a
display is relevant. They also find that the meaning of the numbers as quoted in
the patent text are relevant. The above analysis contains this text:

> the claimed invention as a whole is directed to a combination of
> interrelated elements which combine to form a machine for
> converting discrete waveform data samples into anti-aliased pixel
> illumination intensity data to be displayed on a display means.

This is a very different from my look-at-the-circuit approach. Does the display
really matter? When I look at the circuit I don't find any display. According to
the claim, the rasterizer is comprised of four elements listed from (a) to (d), and
the display is not among them. By itself the rasterizer has no capability of
displaying anti-aliased pixel illumination intensity data on a display. For this you
need an apparatus such as the oscilloscope where the rasterizer is connected to

a display. But this claim, as I read it, is claiming just the rasterizer and not the entire oscilloscope. Do I misread something?[4]

Of course one may connect a rasterizer to a display as the claim preamble suggests. This is only an option because nothing in the laws and principles of electronics requires it. The rasterizer may as well be connected to a hard disk to record the result of the calculations if someone wishes. Or some other creative use could be found. The useful, concrete and tangible result contemplated by the court is one that may occur but will not necessarily occur. Its occurrence depends on how exactly the rasterizer is used.

Perhaps I am overly focused on the display. Perhaps they just point out that the bits being produced are not any random bits but they are bits representing pixel illumination intensity data. Perhaps the meaning of the bits somehow makes the bits a useful, concrete and tangible result in the eyes of the court. But then the correct understanding of the modeling relationships doesn't support the court conclusion. There are two modeling relationships as follows:

> 1. The rasterizer circuit is a physical model for the abstract calculation. This is a relationship from the concrete to the abstract.

> 2. The calculation is an abstract mathematical model (formula) for the illumination of pixels on a screen. This is a relationship from the abstract to the concrete.

These are two distinct modeling relationships going in opposite directions and their concrete elements are different. This court is using one modeling relationship to reach a conclusion about the other. But the two relationships are independent because the circuit doesn't have to be connected to a display and the mathematical calculation doesn't have to be about pixel illumination data. The court's logic doesn't follow from how the modeling works.

The court further explains (bold added):

> The fact that the four claimed means elements function to transform one set of data to another through what may be viewed as a series of mathematical calculations does not alone justify a holding that the claim as a whole is directed to nonstatutory subject matter. *See In re Iwahashi,* 888 F.2d at 1375, 12 USPQ2d at 1911. **Indeed, claim 15 as written is not "so abstract and sweeping" that it would "wholly pre-empt" the use of any apparatus employing the combination of mathematical calculations recited therein.** *See Benson,* 409 U.S. at 68-72 (1972). Rather, claim 15 is limited to the use of a particularly claimed combination of elements performing the particularly claimed combination of calculations to transform, i.e.,

rasterize, digitized waveforms (data) into anti-aliased, pixel illumination data to produce a smooth waveform.

Furthermore, the claim preamble's recitation that the subject matter for which Alappat seeks patent protection is a rasterizer for creating a smooth waveform is not a mere field-of-use label having no significance. Indeed, the preamble specifically recites that the claimed rasterizer converts waveform data into output illumination data for a display, and the means elements recited in the body of the claim make reference not only to the inputted waveform data recited in the preamble but also to the output illumination data also recited in the preamble. Claim 15 thus defines a combination of elements constituting a machine for producing an anti-aliased waveform.

This court worked from the idea that the real-world meaning of the bits defines the actual computation. We can immediately see how this causes a problem. Please refer to the sentence in bold. How does this work? If an electronic engineer wants to build an apparatus employing the combination of mathematical calculations recited therein, he has to build the same circuit as the one described in the *Alappat* claim. Or in the alternative he needs some other circuit that has the same functionality and will be subject to the treatment patent law applies to functional equivalents. This is the very situation *Benson* is meant to avoid.

Here is an example of how this problem would happen. Suppose our hypothetical engineer designs a new oscilloscope whose display draws the waveforms in yellow over a neutral gray background[5]. There is no change in pixel illumination intensity value between the background and the waveform. It is a change of color. The anti-aliasing is done by altering the color saturation value to make the yellow color more or less grayish. A fully saturated color is pure yellow and a totally unsaturated color is pure gray. If we use the same formula for the anti-aliasing, the rasterizer no longer computes pixel illumination intensity data. It computes pixel color saturation data. But the same circuit is being used.

This is a change from the overall function of the rasterizer, because the preamble says "converting vector list data representing sample magnitudes of an input waveform into anti-aliased pixel illumination intensity data". This doesn't happen any more. This is also a change from element (d) which will turn out not to be present. This element says "means for outputting illumination intensity data". This will not occur. The new rasterizer will output color saturation data.[6]

The arithmetic formula used to perform the anti-aliasing is the same. But the

meaning of the bits is changed. In order to carry out this calculation, the same circuit is used. The question is, does it infringe? If the answer is yes, then how is it that "claim 15 as written is not 'so abstract and sweeping' that it would 'wholly pre-empt' the use of any apparatus employing the combination of mathematical calculations recited therein"? Conversely, if the circuit doesn't infringe, what exactly is being covered by the patent? Isn't it supposed to claim a circuit?

All the language about waveform data samples, anti-aliased pixel illumination intensity data and the likes imposes no limitations to the circuit beyond constraining the structure to represent the correct arithmetic calculation according to the intended interpretation of the numbers[7]. The fact that the numbers mean pixel illumination intensity data is not recorded with the bits. If the same calculations are required for another purpose the bits will be the same, the arithmetic operations will be the same, and the circuit carrying out the task will be the same. Therefore anyone who needs this calculation has no choice but to use the same circuit (or an equivalent).

The reason is this: the calculation is done with mathematical symbols. When I write 24, nobody can tell just by looking at the written digits whether this is the number 24 in the abstract, a count of 24 apples, or 24 monkeys. This is contextual information that is not recorded with the written digits. When someone adds numbers, the pocket calculator doesn't care if the numbers mean dollars in a bank account or something else. It just adds the numbers. The real life meaning of the numbers is not stored with the symbols, and it is not used by the computing process.

The same thing happens with bits. They are the electronic version of mathematical symbols. Whether they represent pixel illumination intensity data or or the quantity of jet fuel required by a plane at take off is not recorded with the bits. The bits just represent numbers, and the digital electronics components work just the same whatever the meaning is. If an engineer needs to carry out the same computation for another purpose, the circuit that will fulfill this purpose is the same circuit[8].

Of course when an engineer uses the rasterizer circuit for another purpose, the circuit is no longer a combination of elements constituting a machine for producing an anti-aliased waveform as the court puts it. In this sense, it is no longer the machine claimed by *Alappat* claim 15. This doesn't change the fact that the same circuit is used.

I think that in addition to claiming an abstract mathematical computation, this claim is somehow faulty. I am not a lawyer. This is a hunch and not knowledge. My hunch is that Section 112 paragraph 2 might be applicable. (emphasis

added):

> The specification shall conclude with one or more claims particularly pointing out and **distinctly claiming** the subject matter which the applicant regards as his invention.

My hunch is that reliance on the real-world meaning does not distinctly claim the subject matter unless the real-world meaning is actually used by an element of the claim and not merely referred to. If I understand the law correctly, patent claims that do not conform to section 112 paragraph 2 are invalid and cannot be enforced.

A Note on the Machine or Transformation Test

Among the questions asked, the USPTO requests examples of patents that pass the machine or transformation test but are patents on abstract ideas. We have just seen such an example.

The *Alappat* rasterizer passes the machine or transformation test, because it is a patent on a specific circuit. But it is a patent on a mathematical computation as surely as the *Benson* patent is. Such a remark is applicable every time a specific circuit is built for the purpose of making a computation pursuant to a mathematical algorithm and there is no invention outside of the computation.

The machine or transformation test should work fine in most if not all circumstances where the abstractions involved are models for a real-world invention, like the application of the laws of physics to engineering. Then this test will provide a strong clue as to whether the patent is on the abstract model or on the real-world application. But this test doesn't work when the modeling relationship is reversed, because in such case the machine or process is the tool used to understand the abstract ideas. Then it is possible that a patent on the tool is effectively patenting the idea as was the case in *Benson* and *Flook.*

It is very difficult to make sense of the notion that a patent on a circuit is a patent on an abstract idea when we assume the modeling relationship always works from the abstract to the physical and never the other way round. The trilogy *Benson*, *Flook* and *Diehr* is best understood after the modeling relationship has been worked out properly.

Appendix: Mathematical References

Here are the references to published literature that establish that mathematics require a physical representation of the concept in order to be able to understand, use, or even conceive mathematical ideas. Physical tools impose fundamental limitations on the ability of human beings to reach mathematical

truths. My hope is that lawyers can use these references to locate authorities to support this point in court.

The very first two paragraphs of the very first chapter[9] of *Introduction to Metamathematics*[10] by [Stephen Cole Kleene] are (emphasis in the original)

> Before turning to our main subject, it will be appropriate to notice briefly Cantor's theory of sets.
>
> A flock of four sheep and a grove of four trees are related to each other in a way in which neither is related to a pile of three stones or a grove of seven trees. Although the words for numbers have been used to state this truism on the printed page, the relationship to which we refer underlies the concept of cardinal number. Without counting the sheep or trees, one can pair them with each other, for example by tethering the sheep to the trees, so that each sheep and each tree belongs to exactly one of the pairs. Such a pairing between the members of two collections or 'sets' of objects is called a *one-to-one* (1-1) *correspondence.*

The first thing Kleene does is to use a concrete and tangible procedure to represent the abstract concept of cardinal number in Cantor's theory of sets. What is a cardinal number? On page 9 Kleene quotes Cantor:

> Cantor describes them thus: "The general concept which with the aid of our active intelligence results from a set M, when we abstract from the nature of its various elements and from the order of their being given, we call the 'power' or 'cardinal number' of M."

This is definitely an abstract concept. Cantor's theory of sets is a mathematical theory. Although this concept of cardinal number is described using English there is a more formal definition given mathematically by means of an operation of pairing which is called "a one-to-one correspondence" [11].

A numerically challenged shepherd may use pairing to keep track of his sheep. The shepherd may be unable to count past three, but he knows that if a tree ends up without a sheep, then a sheep is missing. It doesn't matter if the pairing is done "in the abstract" in the shepherd's mind or if it is done in the concrete by physically tethering the sheep to trees. Both ways of pairing count sheep.

In this little example the abstract theory of sets is not used as a model of the act of tethering sheep to trees. It goes the other way round. The act of tethering is a physical model for the mathematical theory of cardinal numbers. It allows the shepherd of the example to overcome his limited mind and perform an abstract

operation he would otherwise be unable to do.

Most people are able to count normally. All of us are subject to ordinary human limitations. There are abstractions that cannot be conceived or handled without physical assistance. This is why we use pocket calculators or pencil and paper for all but the simplest calculations.

The ancient Greeks practiced mathematics, especially geometry. They used figures drawn in the sand or on papyrus to help them track the geometrical abstractions. As [Plato puts it](#) in book 7 of The Republic (around 380 BC):

> You also know how they make use of visible figures and discourse about them, though what they really have in mind is the originals of which these figures are images: they are not reasoning, for instance, about this particular square and diagonal which they have drawn, but about *the* Square and *the* Diagonal; and so in all cases. The diagrams they draw and the models they make are actual things, which may have their shadows or images in water; but now they serve in their turn as images, while the student is seeking to behold those realities which only thought can apprehend.

Nowadays we use the written text to achieve the same result. Like the drawn figures, written text is a physical tool.

There is more at play than the ability of text to express meaning. The tools also enable procedures, the algorithms, for solving problems. The ancient Greeks carried out their mathematical algorithms mostly with compass and straightedge. Modern mathematicians use either pencil and paper or a computer. For example we all learned in school the algorithms to carry out additions, subtractions, multiplications and divisions by manipulations of decimal digits written with pencil and paper. Algebra, calculus and other mathematical disciplines provide methods for solving problems based on the written text. These methods can be automated by means of digital electronics devices like the pocket calculator and the computer.

The choice of the physical tool is important. This choice limits which abstractions we are able to use and determines how easy the computational process will be. If you have any doubt please try to multiply MDCCVIII by CCXXXIV with pencil and paper using roman numerals for all manipulations. Please don't cheat. No translation into decimal numbers is allowed.

Here is another example. Try solving $x^2 + 3x - 2 = 0$ like you were taught in school but using only English words instead of the algebraic notations. You need to use sentences like "the square of the value added three times the value from

which you subtract two yield nothing." According to Kleene[12], this is how mathematical problems were solved before the invention of the algebraic notation by Vieta and others. The discovery of new notations is one of the ways mathematics makes progress. Our ability to use algorithms that solve the problem at hand is contingent on the availability of the adequate physical tool.

This issue goes deeper than mere convenience and practicality. Some problems are fundamentally unsolvable with some categories of tools regardless of whether we may resolve the practical aspects. A famous example is the squaring of the circle problem in Euclidean geometry. As Howard Delong explains[13], when discussing the power of modern algebraic methods when applied to geometry: (emphasis in the original)

> Using this powerful method of analysis, nineteenth century scientists were finally able to solve the three famous outstanding geometric problems which the Greeks passed on to posterity: the duplication of the cube, the squaring of the circle, and the trisection of an arbitrary angle. The problem was to make an exact construction with the aid of a compass and straightedge alone. It was shown in each case that it is impossible to solve the problem under the condition stated. *Impossible* here means *logically impossible*. That is, given certain assumptions, it was shown that it is as impossible to solve these problems as it is to draw a square circle. The proofs are complex — especially the one for the squaring of the circle — but they nevertheless settle the problems once and for all.

Here is another example. There are real numbers that can't be defined by any mathematical formulas. The reason is that formulas are countably infinite while real numbers are *un*countably infinite. This is a limitation of the algebraic notation as physical symbols written on paper.

Countably infinite means an enumeration like the natural numbers 0, 1, 2, 3 … If an infinite set can be enumerated in such a way each member can be associated with a natural number, then it is countably infinite[14]. Formulas can be put in lexicographical orders, this is something like a, b, … z, aa, ab, ac, … zz, aaa, aab etc. except that the order must be defined on the mathematical symbols rather than the Latin alphabet. You group formulas by their length first, and then you sort the formulas of the same length in the lexicographical order. The result is an enumeration that could be numbered[15]. Formulas are countably infinite.

It is proven that real numbers are not countably infinite. The proof is given by the Cantor diagonal method[16]. Assuming you have an ordering of real numbers

you prove there is a real number not of the list as follows.

1. 4.**3**5447646 ...
2. 2.4**5**790209 ...
3. 0.75**8**90000 ...
4. 5.982**3**7515 ...
5. 2.8342**5**897 ...

You write the infinite decimal expansion of each real number in the enumeration as in the above example. If some real number doesn't have an infinite expansion then you pad it with infinitely many zeros (like the third number in the example). Then you take the first digit of the first number, the second digit of the second number and you continue like this over the diagonal. You make a new number with the rule: every time the diagonal digit is 5, you make the new number a 6, otherwise you make it a 5. With the example this gives:

- 0.56556 ...

The result is a real number guaranteed not to be in the list. In other words, there can be no enumeration of all the real numbers, because given any enumeration, we can always find at least one real number that is not included.

The consequence of the Cantor diagonal method is that there are real numbers we cannot define and comprehend because there is no formula able to define them. This is a limitation of relying on a mathematical language made of written symbols.

Physical tools do more than representing meaning and computing solutions to problems. They are also the embodiment of the rules of logic. This is a topic that has been studied by philosophers and mathematicians since the nineteenth century. The name of the discipline is symbolic logic. Its most prominent features are propositional calculus and predicate calculus. They are synthetic languages made of written symbols whose purpose is to provide a logical foundation to mathematics. The formulas and expressions in these languages must follow a rigidly defined syntax, like a programming language. This formalism allows to subject the language itself to mathematical study. This is called metamathematics, or mathematics about mathematics[17].

Some of the greatest mathematical discoveries of the twentieth century arose from metamathematics. For instance see this series of theorems[18].

- *Gödel's first incompleteness theorem*: in every formal system of mathematics that is powerful enough to represent the arithmetic of natural numbers there is a formula that is true but cannot be proved.

- *Church theorem*: consider a formal system that represents pure ordinary mathematical logic (this is known as *predicate calculus*), there is no algorithm that will decide whether or not a formula in this system can be proven as a theorem.

- *Skolem theorem*: if a formal system is powerful enough to represent the arithmetic of natural numbers then it simultaneously supports an alternative interpretation where there are "unnatural numbers", that is number other than those in the familiar sequence 0, 1, 2, 3... to infinity.

I have translated into plain English the mathematical terms of arts and notations in order to make their meaning accessible to laymen. Please go by the mathematical text for any use that requires mathematical accuracy.

These theorems and others like them are called limitative theorems. They establish fundamental epistemological limits to what can be known by man. This is a fascinating topic of deep philosophical consequences. This is an area of knowledge where mathematics, philosophy and computer science overlap. Delong, philosopher and author of *A Profile of Mathematical Logic,* describes these epistemological limits as follow[19]: (emphasis in the original)

> What the limitative theorems represent then is the discovery of an abstract structure which is of such a sort that it is impossible for any human to make systematically complete and correct assumptions about it. No matter how hard we try, we cannot talk about just the natural numbers, but must always talk about the unnatural numbers, too. In other words, just as our sensory perceptions have limits which can be extended but not eliminated by certain technique of science (for example, the microscope), so our abstract conceptions are limited and the methods of mathematics which are intended to extend them (for example, mathematical induction) are at best partial in their effect. Our powers of conceptual discrimination are no less limited than our powers of perceptual discrimination.

> There doesn't seem to be any way around these conclusions. But it should be noted that the above argument depends on what is known as *realism* (or *platonism*) in mathematics. That is, it is assumed that the abstract structure of arithmetic exists independently of human conceptions about it. The assumption is embedded in classical mathematics, but we might question it.

Then Delong proceeds to examine what happens when we question this assumption. The topic is quite fascinating, but I don't want to stray too far away from the legal issues. This quote will be enough for our purposes.

It is not so much the theorems that are relevant but the way they have been derived. Howard Delong explains[20]: (emphasis in the original)

> [It can be argued that: "I object,] the impossibility of squaring a circle or trisecting an angle is not philosophically interesting. It is perfectly possible to do it, but not with just a straightedge and a compass. Similarly, when Gödel proves that there are undecidable formulas, or Church that there is no decision procedure, or Skolem that a calculus doesn't categorically represent the natural numbers, all this comes to is that *using the means they have selected* (just as Euclid selected a straightedge and compass), their conclusions follow. There is no philosophical import to this: it just suggests that mathematicians and logicians must look for other means."

> This objection would have a great deal of sting were it not for one circumstance: *There do not appear to be any other means*. To see this, let's consider briefly the requirements we make in order to claim deductive knowledge. We first require that all the "symbols" which are used be publicly and effectively recognizable. The word "symbols" is in double quotes because it is not required that they be symbols in the usual sense of the word, that is, marks on paper or on a blackboard, etc. They might be sounds or colors or electric charges. The requirement includes not only being able to tell the difference between a symbol and something else (for example we should be able to tell the difference between a parenthesis and a broken fishhook), but being able to effectively recognize repeated occurrences of the same symbol type [for example, '(' is of the same symbol type as '(']. The 'all' in our requirement means that the number of symbols be finite (for example, a zillion symbols) or, if infinite, that there be an ordering (called the alphabetic order above) such that only a finite number of symbols precede any given symbol in that ordering.

> Second, we require that the length of any given sentence be finite and that a sentence be effectively recognizable as such. We make the same requirement of rules of inference. As for the axioms, we require that they be either finite or, if infinite, that there be an ordering such that only a finite number of axioms precede any given axiom in that order. For the theorems, we do not require that there be an effective method to recognize them, but only that there be a constructive program which, if carried out, will produce any given theorem. Finally, we require that for any given sentence we be able to effectively recognize a proof of it if one should be presented. The above are all syntactical requirements. There is only one semantical requirement: that the theorems all be true under every interpretation which makes the

axioms true.

The point is that fundamental limits on the extent of mathematical truths that can be known to mankind follow from the physical tools that are necessary to the practice of mathematics.

Metamathematics gives a mathematically exact definition to each element identified as a requirement of mathematics by Delong. Then mathematical methods are applied to these definitions to prove the limitative theorems and other results. This is an inquiry on the very nature of the tools required to gain deductive knowledge, essentially the symbols and the various syntactical elements that are made with them. This inquiry goes well beyond the basic observations that the written text has a meaning, like is done in the printed matter legal doctrine. It extends the analysis to computational and deductive processes which exist in the real world as manipulation of symbols and syntactic elements.

This is the sort of thing the law should look at in the context of *Benson*, *Flook*, and *Diehr*. It is not possible to fully understand the relationship of mathematical abstractions and physical reality without knowing what is found in this part of mathematics. It not only relates to why a patent on a circuit may be a patent on an abstract idea, it also relates to why software is speech because it involves the study of the relationship between text, computation, and meaning. Its very topic is the ability of mankind to reach knowledge by means of written symbols.

Among the features of metamathematics there are objective definitions of what is a mathematical proof[21]. These definitions define 'proof' by syntactic means. It is the organizing of the formulas as syntactic elements in the proper sequence that makes a mathematical proof valid. Semantics, in the sense of looking at the meaning of the formulas and verifying their truths, plays no part in these definitions. A proof is correct when it satisfies syntactic rules. The consequence is that mathematical reasoning is amenable to machine processing. Kleene observes[22]:

> [C]omputers may be applied in metamathematics to seeking proofs of theorems, or to checking proposed proofs etc.

This is known as automatic theorem proving. This is the automation of operations of abstract logic. Quoting from the linked page: (emphasis and bold in the original)

> **Automated Theorem Proving** (ATP) deals with the development of computer programs that show that some statement (the *conjecture*) is a *logical consequence* of a set of statements (the *axioms* and

hypotheses). ATP systems are used in a wide variety of domains. For examples, a mathematician might prove the conjecture that groups of order two are commutative, from the axioms of group theory; a management consultant might formulate axioms that describe how organizations grow and interact, and from those axioms prove that organizational death rates decrease with age; a hardware developer might validate the design of a circuit by proving a conjecture that describes a circuit's performance, given axioms that describe the circuit itself; or a frustrated teenager might formulate the jumbled faces of a Rubik's cube as a conjecture and prove, from axioms that describe legal changes to the cube's configuration, that the cube can be rearranged to the solution state. All of these are tasks that can be performed by an ATP system, given an appropriate formulation of the problem as axioms, hypotheses, and a conjecture.

This is an example of using the real-world device to model abstract logical processes. The linked page discusses more, pointing to actual programs and applications.

Another feature of metamathematics is the precise definition of what is computable. This is what *Benson*, *Flook*, and *Diehr* refer to as a "mathematical algorithm". This is knowledge which should be used by the law. This definition follows from the same type of epistemological inquiry as the limitative theorems. Delong explains[23] (emphasis in the original, the occurrence of patent statute section numbers is pure coincidence):

> We begin by imagining some human who is faced with a specific computational problem. It might be some such problem as computing the sum of 101 and 102 and 103 and . . . and 198 and 199, where the three dots indicate one occurrence of all the natural numbers between 103 and 198. We assume that he is working according to a finite set of rule which have been fixed before the problem was given and that he is using pencil and paper. We also assume that after a finite amount of time he stops with the correct answer.
>
> If we examine this paradigm case of computing with a view toward eliminating inessential features, a number of such features comes to mind, such as the use of pencil and paper or the particular computational problem chosen. However the most striking one appears to be the *human*: the computer might as well be a machine.

Here too we see the prominent position of the use of symbols in mathematical processes. The analysis of how a human computer can be replaced by an automatic device has been carried out by Alan Turing. The outcome is the

Turing machine[24].

John von Neumann was familiar with Turing's work and used it to derive the stored-program architecture which forms the basis of nearly all modern programmable computers[25]. Jack Copeland reports on [The Turing Archives for the History of Computing]:

> In 1944, John von Neumann joined the ENIAC group. He had become 'intrigued' (Goldstine's word) with Turing's universal machine while Turing was at Princeton University during 1936-1938. At the Moore School, von Neumann emphasised the importance of the stored-program concept for electronic computing, including the possibility of allowing the machine to modify its own program in useful ways while running (for example, in order to control loops and branching). Turing's paper of 1936 ('On Computable Numbers, with an Application to the Entscheidungsproblem') was required reading for members of von Neumann's post-war computer project at the Institute for Advanced Study, Princeton University (Julian Bigelow in personal communication with William Aspray, reported in the latter's *John von Neumann and the Origins of Modern Computing* Cambridge, Mass.: MIT Press (1990), pp. 178, 313).

Digital computers relate to written text in a way analog technologies implementing computations don't. Digital information is made of symbols, the bits, the 0s and the 1s which represent the boolean values False and True respectively. By design, the same class of computational processes that could, in principle, be carried out by means of pencil and paper are automated by the computer. This is in accordance to the analysis of computation done by Alan Turing and the engineering decision made by John von Neumann and his colleagues.

This is more than enough evidence that physical devices and computations have the ability to represent abstract ideas. Their use is essential for human beings to even be able to conceive, let alone use, these ideas. In particular the computations described by algorithms must be carried out physically in order to be used at all. If you patent the physical means to carry out the computation then you may end up patenting the computation itself if you are not careful. This is the mathematical basis for the non-patentability of algorithms as per *Benson*, *Flook* and *Diehr*.

References

Cases

[Bilski v. Kappos](#) [PDF] [[text](#)] The Supreme Court opinion

[Diamond v. Diehr](#), 450 U.S. 175, 182 (1981)

[Gottschalk v. Benson](#), 409 U.S. 63, 71-72 (1972)

[In re Alappat](#), U.S. Court of Appeals Federal Circuit, 33 F.3d 1526 July 29, 1994

[In re Bilski](#) the Court of Appeals for the Federal Circuit decision

[In re Gulak](#), 703 F.2d 1381

[Parker v. Flook](#) 437 U.S. 584 (1978)

In Print

[Davis 2000] Davis, Martin, *Engines of Logic, Mathematicians and the Origin of the Computer*, W.W. Norton and Company, 2000. This book was originally published under the title *The Universal Computer: The Road from Leibnitz to Turing.*

[Delong 1970] Delong, Howard. *A Profile of Mathematical Logic*. Addison-Wesley Publishing Company. 1970. Reprints of this book are available from Dover Publications.

[Gödel 1986] Gödel, Kurt, 1986, *Collected Works*, vol. 1, Oxford: Oxford University Press.

[Kleene 1952] Kleene, Stephen Cole, *Introduction to Metamathematics*, D. Van Nostrand Company, 1952. I use the 2009 reprint by Ishi Press International 2009.

[Kleene 1967] Kleene, Stephen Cole, *Mathematical Logic*, John Wiley & Sons, Inc. New York, 1967. I use the 2002 reprint from Dover Publications.

[Turing 1936] Turing, Alan, *On Computable Number with and Application to the Entscheidungsproblem*, Proceeding of the London Mathematical Society, ser. 2, vol 42 (1936), pp. 230-67. Correction: vol 43 (1937) pp. 544-546.

This paper could be ordered from the publisher here [[Link](#)]

This paper is available on-line here [[link](#)]

Biographies

Biographies are from the [MacTutor History of Mathematics Archive](#)

Kurt Gödel

Stephen Kleene

Alan Turing

Footnotes

1 See Donald Knuth's 1994 letter to the US Commissioner of Patents and Trademark: "An algorithm is an abstract concept unrelated to physical laws of the universe."

2 One shouldn't underestimate the importance of the written text for the storage, communication and processing of information. Before the advent of computers written text was as important as computers are nowadays. It still is. The social, scientific and economic issues raised by software are not nearly as new as students of software patents may think. Computers represent an expansion of existing capabilities and not the creation of new capabilities that didn't exist before. This expansion is formidable, I will agree, but an expansion of existing capabilities no matter how formidable is not the same thing as the creation of new capabilities.

3 The text of the court ruling contains a description of the mathematical formula.

4 The phrase "for converting vector list data representing sample magnitudes of an input waveform into anti-aliased pixel illumination intensity data to be displayed on a display means" implies that the act of displaying occurs after the rasterizer has produced its output. It appears to me that the display is explicitly excluded.

5 Whether or not the engineer has an awful taste in matters of color is beside the point.

6 This design may have an amusing twist. In this particular setup the blue component of the video signal happens to have a well defined mathematical relationship with color saturation. When the blue component has a high value it indicates the pixel corresponds to the gray background because pure yellow doesn't have any blue component. And conversely a zero value of the blue component indicates pure yellow. What if the engineer modifies the display to monitor the blue component and adjust the pixel illumination intensity inversely proportional to the blue component? Low blue value raises the intensity to the brightest yellow while high blue values drop the intensity to pitch black. Then we display the waveform as a yellow line on a black background just as the *Alappat* rasterizer might do. The color saturation rasterizer would control pixel

illumination intensity without computing pixel illumination intensity data. I am bringing up this point just for its amusement value. This scenario is beside the point of this article. Lawyers may wish to discuss whether this modified display makes a functional equivalent of the *Alappat* circuit and whether this will lead to infringement.

[7](#) If the point needs further hammering, such calculations can be done by a standard off-the-shelf pocket calculator when only the human being operating the calculator knows the meaning of the numbers.

[8](#) Had the patent been on the process manipulating the bits then by the same logic assigning a different meaning doesn't make a physical process distinct from the base mathematical computation.

[9](#) See [Kleene 1952] p. 3.

[10](#) [Kleene 1952] in the reference section.

[11](#) See [Kleene 1952] p. 9 where the exact mathematical definition is given.

[12](#) [Kleene 1952] p.61.

[13](#) [Delong 1971] p. 69.

[14](#) Kleene [1952] pp. 3-6.

[15](#) [Kleene 1967] p. 178. This is assuming there is a finite number of allowable symbols in a formula. If the number of allowable symbols is infinite (not a practical possibility) then the formulas are still countable but the ordering requires the use of Gödel numbers. See [Gödel 1986]

[16](#) Kleene [1952] pp. 6-8, [Kleene 1967] pp. 180-183 or [Delong 1971] pp. 75-76.

[17](#) This the the topic of [Kleene 1952] which is titled *Introduction to Metamathematics*. Chapter IV to X are especially applicable. This is also covered in [Kleene 1967] chapters I to III and in [Delong 1971] chapters 3 and 4.

[18](#) [Delong 1971] p. 195.

[19](#) [Delong 1971] p. 204

[20](#) [Delong 1971] pp. 193-194.

[21](#) This is known as proof theory. See for example [Kleene 1967] pp. 33-58, 107-134. There are several proof theories depending on the system of logic

being used. Each of them is mathematically unambiguous.

22 [Kleene 1967] p. 201.

23 [Delong 1971] p. 197.

24 Turing's analysis of the use of symbols by humans is found in section 9 of [Turing 1936]. His thesis is that all computations done with pencil and paper can be replicated by Turing machines, therefore Turing machines are a mathematically accurate description of what is computable.

25 This history is told in [Davis 2000] chapters 7 and 8.

284 comments

http://www.groklaw.net/article.php?story=2010092621054289