Math never seen

Johannes Küster

Abstract

Why have certain mathematical symbols and notations gained general acceptance while others fell into oblivion?

To answer this question I present quality criteria for mathematical symbols. I show many unknown, little-known or little-used notations, some of which deserve much wider use.

I also show some new symbols and some ideas for new notations, especially for some well-known concepts which lack a good notation (Stirling numbers, greatest common divisor and least common multiple).

1 Introduction

For TEX's "20th birthday it seems appropriate to present some fine points of mathematical typography and some ideas for new symbols and notations. Let's start with a quotation from *The METAFONTbook* [5, p. 8]:

> "Now that authors have for the first time the power to invent new symbols with great ease, and to have those characters printed in their manuscripts on a wide variety of typesetting devices, we must face the question of how much experimentation is desirable. Will font freaks abuse this toy by overdoing it? Is it wise to introduce new symbols by the thousands?"

We all know that METAFONT didn't become widely accepted. But even with other font editors, font freaks did not create new symbols by the thousands. So while maybe METAFONT was too complicated, and its way of thinking foreign to most designers, this can't be the real reason why only very few new symbols showed up. In fact, to design a new useful symbol is by no means an easy task, which I hope will become clear in the following. Just as we all do a lot more reading than writing, it is much easier to use existing symbols (e.g. with TEX) than to create good, useful new symbols (e.g. with METAFONT). So TEX with the character set offered by Computer Modern fonts (and the AMS fonts) shaped the typography of mathematics in the past 30 years.

This situation only changed with Unicode mathematics: Unicode now offers mathematical symbols literally by the thousands. But it gives little explanation and little usage information; many symbols are described only by shape, not by meaning. For many Unicode mathematical symbols it is not clear how to use them, and in many cases it is not clear whether there are any competing or superior notations.

2 Quality criteria

What makes a notation superior to another? What makes a symbol successful (in the sense that other mathematicians accept and adopt it)? The following list gives the most important quality criteria. A mathematical symbol or notation should be:

- readable, clear and simple
- needed
- international (or derived from Latin)
- mnemonic
- writable
- pronounceable
- similar and consistent
- distinct and unambiguous
- adaptable
- available

This list is certainly not exhaustive, but these are the most important points. Not all criteria are equally important, and some may conflict with others, so few symbols really fulfill all criteria. — Let me explain each point in turn.

Above all, a notation should be *readable* — but what constitutes readability? Certainly it comprises *clear* and *simple*. Also a notation should be short, at least it should make an expression shorter than writing out the same statement with words. Some of the other criteria contribute to readability as well.

When a good, widely accepted notation already exists, there is no need to invent a new one. So a new notation should be *needed* or *necessary*.

Most mathematical symbols are *international* (even if they are given different names in different languages and although there are different traditions in mathematical notation, e.g. the use of a dot or a comma as decimal separator). Of course a new notation should be international. In the case of an abbreviation (like "sin", "log", etc.), it should be derived from Latin, as most scientific terminology stems from Latin (and Greek), and so does the international vocabulary of mathematics.

A notation should be easy to learn, and its meaning should be easy to remember, at least after one has heard or read an explanation once; i.e. a notation should be *mnemonic*.

A lot of mathematics is still (and will be) written by hand (e.g. in a mathematician's research as the fastest way to denote his thoughts, on the blackboard, etc.). So a notation should be *writable*. In fact mathematical typography shows its close relation to handwriting in many places. But while written mathematics could always be explained by the writer (e.g. by the teacher at the blackboard), printed mathematics has to speak for itself. So in some cases it is desirable to go for greater differentiation in print than what is possible in handwriting.

A notation should also be *pronounceable*. Usually this is not a problem: for most notations there is a manner of speaking, although often language-specific and often not closely related to the notation (e.g. we call "$|a|$" the "absolute value of a", and we would do so whatever the notation would be). But we'll see an example below where a missing manner of speaking was a problem.

A new notation should be *consistent* with the general system of mathematical notation and *similar* to existing notations (e.g. for a symmetric relation one should choose a symmetric symbol, for a new kind of mapping one should choose some kind of arrow). In print, we can differentiate more than in handwriting, but still it is often preferable to stay close to existing notations.

As a special case of similarity, there are many concepts in mathematics which are *dual* or complementary to each other, and such dual concepts should be given dual notations (e.g. $<$ and $>$; $\wedge$ and $\vee$; $\cup$ and $\cap$; $\subset$ and $\supset$). Conversely, dual symbols should denote dual concepts.

In some cases dual symbols work against mnemonics. For many students it is difficult to remember which is which, so one has to use an additional memory aid (e.g. to remember which one of the the logic symbols $\wedge$ or $\vee$ denotes the "logical or", one might learn that $\vee$ reminds of Latin "vel", which means "or").

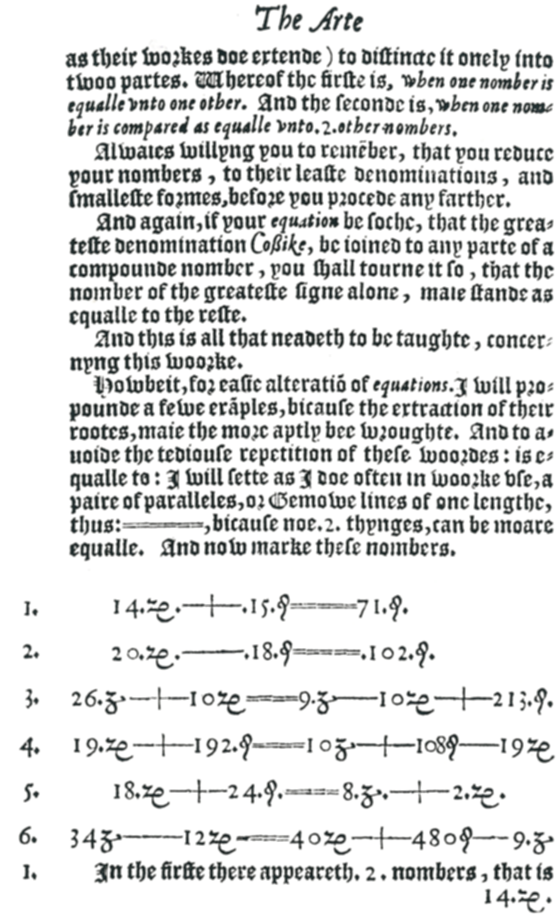
Of course a new notation should be *distinct* and *unambiguous*. Otherwise it will not be an improvement upon existing notations.

A notation should be *adaptable,* it should allow for manipulation. Also mathematical concepts are often generalized, and thus notation is often stretched to more general cases. A good notation allows for that.

A historical example is given by the competing notations $\dot{x}$ of Newton and $\mathrm{d}x$ of Leibniz. While Newton's notation was similar to existing notations and better fitted into the general system, the novel notation of Leibniz was superior, as it was more versatile and allowed for manipulation and generalization.

To give another example, the greatest common divisor of two integers a and b could be denoted as $\gcd(a, b)$; alternatively one might think of an infix notation, e.g. $a \top b$. When applied to three arguments both notations still work: $\gcd(a, b, c)$ and $a \top b \top c$. But one could also take the gcd of all elements of a set S. With the first notation, we can write this as $\gcd(S)$. Yet the alternative notation fails, it is not adaptable enough.

And last on our list, a symbol should be *available*. This is not really a criterion for quality, but rather for acceptance. The best notation does not help much if other people are not able to use it. In former times, this mainly meant availability at the printer's office — nowadays it means availability in a font, then a clear and simple shape which can be added to other fonts with ease, and of course inclusion in Unicode mathematics.


The Arte

as their workes doe extende) to distincte it onely into twoo partes. Whereof the firste is, *when one nomber is equalle vnto one other.* And the seconde is, *when one nomber is compared as equalle vnto .2. other nombers.*

Alwaies willyng you to remēber, that you reduce your nombers, to their leaste denominations, and smalleste formes, before you procede any farther.

And again, if your *equation* be soche, that the greateste denomination *Cosike*, be ioined to any parte of a compounde nomber, you shall tourne it so, that the nomber of the greateste signe alone, maie stande as equalle to the reste.

And this is all that neadeth to be taughte, concernyng this woorke.

Howbeit, for easie alteratiō of *equations.* I will propounde a fewe exāples, bicause the extraction of their rootes, maie the more aptly bee wroughte. And to auoide the tediouse repetition of these woordes : is equalle to : I will sette as I doe often in woorke vse, a paire of paralleles, or Gemowe lines of one lengthe, thus: ═══, bicause noe .2. thynges, can be moare equalle. And now marke these nombers.

1. In the firste there appeareth. 2. nombers, that is

Figure 1: Robert Recorde, *The Whetstone of Witte* (London, 1557). Recorde's explanation for his symbol "=" is given in the lines just above the display formulae.

3 Historical examples

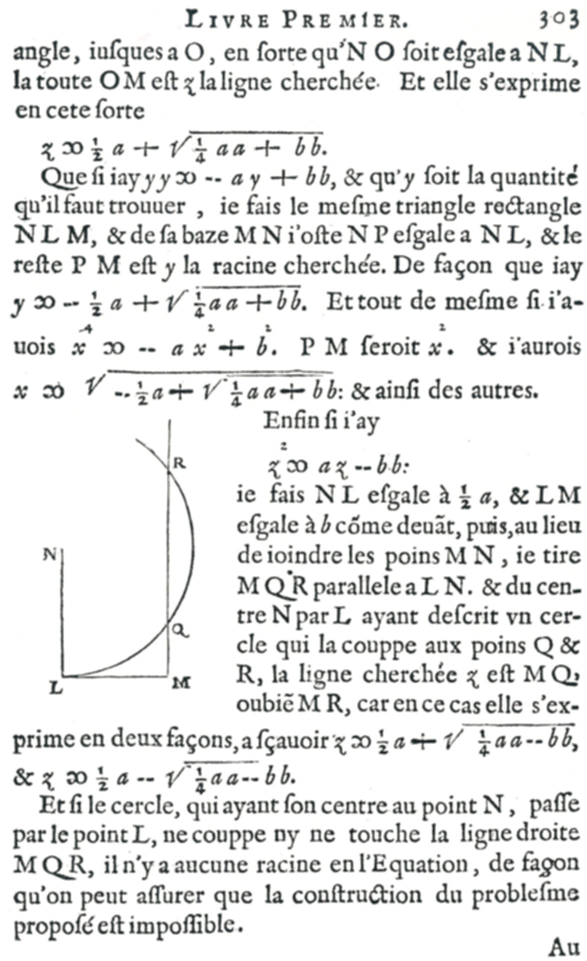
To illustrate these quality criteria, I will give a few historical examples, some unsuccessful, some successful. The historical information is mainly taken from [1].

3.1 Symbols for equality

Our modern symbol for equality "=" was introduced by Robert Recorde in 1557 in his book "*The Whetstone of Witte*" (see figure 1). Recorde explained his choice thus:

> "And to avoide the tediouse repetition of these woordes : is equalle to : I will sette as I doe often in woorke use, a paire of paralleles, or Gemowe lines of one lengthe, thus: ═══, bicause noe .2. thynges, can be moare equalle."

("Gemowe" means "twin"). This is quite a famous example, as it is one of the very few cases where an author not only introduced a new symbol, but explained why he chose its particular form.


LIVRE PREMIER. 303

angle, iuſques a O, en ſorte qu'N O ſoit eſgale a N L, la toute O M eſt z la ligne cherchée. Et elle s'exprime en cete ſorte

$z \propto \frac{1}{2} a + \sqrt{\frac{1}{4} aa + bb}$.

Que ſi iay $yy \propto -ay + bb$, & qu'y ſoit la quantité qu'il faut trouuer, ie fais le meſme triangle rectangle N L M, & de ſa baze M N i'oſte N P eſgale a N L, & le reſte P M eſt y la racine cherchée. De façon que iay $y \propto -\frac{1}{2} a + \sqrt{\frac{1}{4} aa + bb}$. Et tout de meſme ſi i'auois $x^4 \propto -ax^2 + b^2$. P M ſeroit x^2. & i'aurois $x \propto \sqrt{-\frac{1}{2} a + \sqrt{\frac{1}{4} aa + bb}}$: & ainſi des autres.

Enfin ſi i'ay

$z^2 \propto az - bb$:

ie fais N L eſgale à $\frac{1}{2} a$, & L M eſgale à b cõme deuãt, puis, au lieu de ioindre les poins M N, ie tire M Q R parallele a L N. & du centre N par L ayant deſcrit vn cercle qui la couppe aux poins Q & R, la ligne cherchée z eſt M Q, oubiẽ M R, car en ce cas elle s'exprime en deux façons, a ſçauoir $z \propto \frac{1}{2} a + \sqrt{\frac{1}{4} aa - bb}$, & $z \propto \frac{1}{2} a - \sqrt{\frac{1}{4} aa - bb}$.

Et ſi le cercle, qui ayant ſon centre au point N, paſſe par le point L, ne couppe ny ne touche la ligne droite M Q R, il n'y a aucune racine en l'Equation, de façon qu'on peut aſſurer que la conſtruction du probleſme propoſé eſt impoſſible.

Au

Figure 2: René Descartes, *La géométrie* (Leiden, 1637). The symbol "$\propto$" for equality appears throughout this page, e.g. as the second symbol in the first displayed formula.

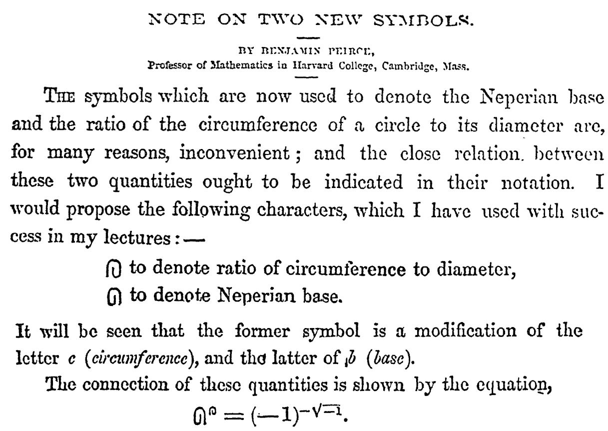
But 80 years later, René Descartes introduced a different symbol for equality, namely "$\propto$", in his book *La géométrie* (see figure 2). Descartes didn't give an explanation, so it is not clear why he invented a new symbol nor why he chose this particular form. Most likely, he was in need for a new symbol as he already used "=" for "plus or minus" (i.e. "±" in modern notation) elsewhere in his writings. The symbol of Descartes might stem from the ligature "æ", a common abbreviation for the Latin word "aequalis", but rotated 180 degrees. Typographically, it rather resembles a rotated "œ", or maybe it's even the astrological symbol for Taurus, turned sideways.

When we compare the two symbols (with our quality criteria in mind), we see that both symbols are mnemonic. Yet Recorde's symbol is simpler, and it is simpler to write. Equality is of course a symmetric relation, but the symbol of Descartes is not symmetric, and this is its main disadvantage. So it seems clear that "=" is the superior symbol.

But in fact these two symbols (and a few competing symbols as well) struggled for supremacy throughout the 17th century. Descartes was the more eminent mathematician, and with his important works his notation also spread. General adoption of "=" as the symbol for equality came only in the early 18th century, mainly because Leibniz and Newton both used it.

3.2 Symbols of Benjamin Peirce

In 1859, Benjamin Peirce introduced the symbols "ᑎ" and "ᑎ" to denote the numbers 3.14159… and 2.71828… (see figure 3). To my knowledge, these were the first significant symbols of American origin.


NOTE ON TWO NEW SYMBOLS.

BY BENJAMIN PEIRCE,
Professor of Mathematics in Harvard College, Cambridge, Mass.

The symbols which are now used to denote the Neperian base and the ratio of the circumference of a circle to its diameter are, for many reasons, inconvenient; and the close relation between these two quantities ought to be indicated in their notation. I would propose the following characters, which I have used with success in my lectures:—

ᑎ to denote ratio of circumference to diameter,

ᑎ to denote Neperian base.

It will be seen that the former symbol is a modification of the letter *c* (*circumference*), and the latter of *b* (*base*).

The connection of these quantities is shown by the equation,

$$ᑎ^{ᑎ} = (-1)^{-\sqrt{-1}}.$$

Figure 3: Benjamin Peirce's symbols for the numbers 3.14159… and 2.71828… (from J. D. Runkle's *Mathematical Monthly*, Vol. I, No. 5 (February, 1859), p. 167–168).

Peirce's symbols were used by some of his pupils (among them his sons Charles Sanders Peirce and James Mills Peirce), but they weren't generally accepted, and they were never used in Europe. By checking our quality criteria, we can see a number of possible reasons.

First of all, the symbols were not really necessary: π and e were already widely used to denote these two numbers, and this was good enough for most mathematicians. Also they are not consistent with the general system of mathematical notation: constants and special numbers are usually denoted with letters, not with special symbols. Then these symbols were not readily available at the printer's office (of course this difficulty was often overcome with other symbols when demand was high enough). More importantly, the symbols ᑎ and ᑎ are not really mnemonic:

> "It will be seen that the former symbol is a modification of the letter *c* (*circumference*), and the latter of *b* (*base*)."

The connection between ᑎ and *c*, and between ᑎ and *b* is hard to see, and it is difficult to remember which is which. To make matters worse, James Mills Peirce used variations of his father's symbols (and also a special symbol for the imaginary unit, see figure 4), but the supposedly mnemonic connection to c and b does not get any clearer.

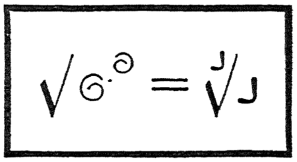


Figure 4: Variations of Benjamin Peirce's symbols (James Mills Peirce, *Three and Four Place Tables* (Boston, 1871)). In modern notation, this formula reads as $\sqrt{\mathrm{e}^{\pi}} = \sqrt[\mathrm{i}]{\mathrm{i}}$.

But on two of our criteria these symbols really fail: firstly, how should we pronounce these? The symbols do not provide a manner of speaking:

> "ᑎ to denote ratio of circumference to diameter,
> ᘎ to denote Neperian base."

Should we always say "ratio of circumference to diameter" and "Neperian base"? In comparison, to pronounce "π" and "e" is easy and fast.

Secondly, the symbols are dual, but the underlying concepts are not. Of course, 3.14159… and 2.71828… are connected in many interesting ways, but they are not dual to each other. So there are good reasons why these two symbols were not generally accepted.

3.3 Symbols for "floor" and "ceiling"

To denote the floor function (i.e. rounding a real number to the largest previous integer), Gauß introduced the bracket notation "$[x]$" (C. F. Gauß, *Theorematis arithmetici demonstratio nova* (1808)). This remained standard for a long time, and is sometimes even used today. But in 1962, Kenneth E. Iverson (in his book *A Programming Language*) introduced new notations

$\lfloor x \rfloor$ for the floor function, and

$\lceil x \rceil$ for the ceiling function.

These notations were readily accepted and are the standard notations today. Also they have been available in TeX and Computer Modern fonts right from the beginning, which certainly helped them to spread. Instead of the ambiguous $[x]$ (as brackets are used for many different concepts, not only for "floor"), we get a new, unambiguous notation $\lfloor x \rfloor$, and also a new, dual notation $\lceil x \rceil$ for the dual concept "ceiling" which didn't have a standard notation before.

These new notations are definitely very mnemonic, almost self-explanatory, and still they are not too far from the old notation, so they are consistent with the general system. Anyone used to the notation $[x]$ could learn and accept the new notations without difficulty.

These were very successful innovations indeed, and they meet all our quality criteria.

4 Unknown and little-known notations

Now I will discuss some important existing notations which deserve to be better known or to be used more often. All of these improve readability, but some only work in print, not in handwriting.

4.1 Usage of roman and italic letters

By careful usage of roman letters (or upright glyph shapes) one can greatly improve the readability of mathematical formulae. Instead of "roman" and "italic" I prefer to use the terms "upright" and "oblique" (or "slanted") here as these terms apply to all kind of glyphs, not only to letters. There's a little-known rule, best stated as

> Operators and constants with a fixed meaning should be set upright.

Important here is *"with a fixed meaning"*. Note that this rule only applies to operators and constants, not to functions or other concepts. Of course this only works in print, not in handwriting. This rule is seldom applied properly in TeX, probably because Computer Modern fonts did not supply upright lowercase Greek.

This rule applies at least to the following *constants with a fixed meaning:* Euler's number e, circle number π, imaginary unit i, Euler's constant γ (or C in European tradition), golden ratio ϕ; and at least to the following *operators with a fixed meaning:* differential operator d and partial differential operator ∂, difference Δ, Kronecker symbol δ_{ij}, and Christoffel symbols $\Gamma^{\kappa}_{\mu\nu}$. For consistency, all "ordinary" Greek uppercase letter must be italic then: $\mathit{\Gamma}, \mathit{\Delta}, \mathit{\Theta}, \ldots$

This list is not exhaustive, and the actual scope of this rule might depend on context. An author could extend the scope to some constants and operators which carry a fixed meaning throughout his text. In an encyclopedia of mathematics (with a wide range of topics and notations), applying this rule greatly improves readability, while e.g. in a monograph about all the fascinating properties of Euler's number, using an italic e might seem preferable, to separate it better from surrounding text — but even here I would apply this rule, with some careful spacing and kerning. Matters are more complicated when typesetting physics, as upright type is used here also for units, indices with a fixed meaning, particles, quanta, and quantum states; but even here this rule is useful.

My suggestion is to apply this rule to integral symbols as well: an upright integral symbol and an upright differential operator d serve as a kind of delimiters around the integrand:

$$\int_a^b f(x)\,\mathrm{d}x.$$

This is not the case when the integrand is a fraction: here the differential operator is often written in the numerator, but still, using an upright "d" increases readability.

When we look at a few examples, we see that this rule gives more structure and more clarity to formulae:

$$\bar{z} = a - ib = \rho\,(\cos\varphi - i\sin\varphi) = \rho e^{-i\varphi}$$

$$\bar{z} = a - \mathrm{i}b = \rho\,(\cos\varphi - \mathrm{i}\sin\varphi) = \rho \mathrm{e}^{-\mathrm{i}\varphi}$$

$$\left(\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr}\right)\ln\psi_0(r) = h(r)$$

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)\ln\psi_0(r) = h(r)$$

$$\int_0^\infty \frac{e^{-at^2}\,dt}{t+x} = e^{-ax^2}\left(\sqrt{\pi}\int_0^{\sqrt{ax}} e^{t^2}\,dt - \frac{1}{2}\,\mathrm{Ei}(ax^2)\right)$$

$$\int_0^\infty \frac{\mathrm{e}^{-at^2}\,\mathrm{d}t}{t+x} = \mathrm{e}^{-ax^2}\left(\sqrt{\pi}\int_0^{\sqrt{ax}} \mathrm{e}^{t^2}\,\mathrm{d}t - \frac{1}{2}\,\mathrm{Ei}(ax^2)\right)$$

For the most important constants and operators, I suggest to use the following TEX macros (somewhat analogous to the way to input these in some computer algebra systems):

`\E` for e, `\PI` for π, `\I` for i, `\df` for d.

Here `\df` could be defined as `\mathop` with an argument, which takes care of proper spacing, e.g.

```
\def\df#1{\mathop{\mathrm{d}{#1}}}
```

(proper font-specific spacing and kerning could be added to these macros with `\mspace` or `\mskip` and `\mkern`; in "newmath" encodings, the upright "d" is contained in "Math Core" to allow for kerning with math italic letters).

4.2 *O*-notation and Vinogradov symbols

For the well-known O-notation (invented by Paul Bachmann in 1894 and made popular by Edmund Landau), there is a little-known alternative with the so-called Vinogradov symbols, named after the Russian number theorist Ivan Matveevich Vinogradov (1891–1983). Unfortunately, I could not find when and where he introduced this notation.

So instead of $f(x) = O(\log n)$, equivalently we can write $f(x) \preccurlyeq \log n$, or we could use the symmetric variant of "⩽" and reverse the order: $\log n \succcurlyeq f(x)$. This notation is used mainly in number theory. The two Vinogradov symbols are included in Unicode:

`uni2AA1`	⪡	"double nested less-than",
`uni2AA2`	⪢	"double nested greater-than".

In my opinion, the Unicode character names are misnomers. At least additional information is missing in Unicode that these two symbols are used as Vinogradov symbols.

The obvious TEX macro names for these symbols are `\subord` for "⪡" and `\supord` for "⪢", analogous to `\subset` and `\supset`.

The Vinogradov symbols must not be confused with

`uni226A`	≪	"much less-than",
`uni226B`	≫	"much greater-than".

Alas, very often "≪" is used instead of "⪡", either because authors are unaware of the difference, or because Computer Modern fonts do not provide the Vinogradov symbols.

When we compare Vinogradov's notation and the O-notation we see that both have their advantages; neither is superior to the other.

Vinogradov's notation does not require additional parentheses. With its symmetric variant, it works in two ways: $f ⪡ g$ and $g ⪢ f$. It better fits the general system of mathematical notation, and it better fits with other symbols, especially with Hardy's symbol "≍" for asymptotic equivalence:

$$(f ⪡ g) \wedge (g ⪡ f) \iff f \asymp g.$$

O-notation is similar to other Bachmann-Landau notations, namely o-, ω-, Ω-, and Θ-notation. Also it can be used in terms in arithmetic expressions:

$$f(x) = \frac{x}{\log x}\left(1 + O\left(\frac{1}{\log x}\right)\right),$$

with the downside that the O might be overlooked in a longer expression.

But O-notation makes strange use of "=", it is somewhat foreign to the general system. In fact, here "=" does not stand for "is equal to", but rather for "is of the order of" or "is a member of the class". So it would be more correct to use "∊". Of course this is well-known and has often been discussed. Still it is annoying, and so this might be a case where we should use greater differentiation in print: i.e. to keep "=" in handwriting as a short and fast notation, but to use an unambiguous special variant of "=" in print, maybe by creating a new special symbol.

4.3 Intervals

In exercise 18.14 in *The TEXbook* [4, p. 171], Knuth says *"Some perverse mathematicians use brackets backwards, to denote 'open intervals'"*, and the following formula is given as an example:

$$]-\infty, T[\,\times\,]-\infty, T[\,.$$

This notation for open intervals is taught in school at least in some countries (e.g. in Germany), and it is also recommended by a German DIN standard and an international ISO standard. So I prefer to be a perverse mathematician — but only in handwriting.

The answer to this exercise [4, p. 322] states *"Open intervals are more clearly expressed in print by using parentheses instead of reversed brackets"*, and the given formula is then written as

$$(-\infty, T) \times (-\infty, T).$$

But the notation "(a, b)" is overloaded with meanings: it is used to denote an ordered pair, coordinates, the greatest common divisor, etc. So this cannot be the best way to denote open intervals, neither in handwriting nor in print.

One simple way to improve the ambiguous notation (a, b) is to use a semicolon instead of a comma to separate the endpoints: $(a; b)$. This is especially useful when the decimal separator is a comma (which is the standard notation in some countries, e.g. in Germany): $(1{,}9; 3{,}8)$ is much more readable than $(1{,}9, 3{,}8)$.

This improvement works in handwriting as well, and it adds a lot of clarity for the reader, with minimal effort on the writer's side.

Still we can do better in print, namely by using special delimiters, already available in Unicode:

`uni2997`	⦗	"left black tortoise shell bracket",
`uni2998`	⦘	"right black tortoise shell bracket".

If there is such a thing as an "unknown standard", this certainly is one: at least one German manual of style [12] recommends these special delimiters for intervals, and one important German book [9, 10] uses these to very good effect. Of course, I also recommend these delimiters in my own writings about typography of mathematics [7, 8].

Just as brackets, these delimiters are reversed to denote open intervals:

⦗$a; b$⦘, ⦗$a; b$⦗, ⦘$a; b$⦘, ⦘$a; b$⦗;

in an example formula, this looks like this:

⦘$0; 1$⦘ $= \{\, x \in \mathbf{R} \mid 0 < x \leq 1 \,\}$

(note that we keep the semicolon as separator, as suggested above). The formula from above is now written as

⦘$-\infty; T$⦗ $\times$ ⦘$-\infty; T$⦗

— admittedly, this is still not very readable, but it is not a nice example anyway (it might be preferable here to introduce an abbreviation for the given interval, say U, and to denote the formula as $U \times U$ or even as U^2).

For use in TEX, I recommend the following macros (with two arguments):

`\ivc{a}{b}`	for ⦗$a; b$⦘	("interval, closed"),
`\ivo{a}{b}`	for ⦘$a; b$⦗	("interval, open"),
`\ivco{a}{b}`	for ⦗$a; b$⦗,	
`\ivoc{a}{b}`	for ⦘$a; b$⦘.	

These macros can take care of proper kerning and spacing and of the semicolon as separator. We can alter these as necessary, e.g. whenever the special delimiters are not available in the used font. For larger versions, we can define macros as `\bigivc` etc. For automatic extension of delimiters (i.e. using `\left` and `\right`), we can define macros starting with an uppercase letter: `\Ivc` etc. In an similar way we can define macros for other delimiters with special meaning, e.g. `\abs` for absolute value $|a|$ or `\norm` for norm $\|a\|$.

5 New symbols and new notations

In this last section I will show some of my ideas for new symbols (even though some of these are not successful). The first few examples are rather minor points, but the last one seems quite important, at least in my opinion.

5.1 Vega

In mathematical finance (with the pricing of stock options) the so-called Greeks occur: Gamma, Delta — and Vega (these are quantities representing the sensitivities of derivatives):

$$\Delta_c = \frac{\partial C}{\partial S} = \phi(d_1) \quad \text{and} \quad \Delta_p = \frac{\partial P}{\partial S} = -\phi(-d_1)$$

$$\Gamma = \frac{\partial^2 C}{\partial S^2} = \frac{\phi(d_1)}{S\sigma\sqrt{T-t}}$$

$$\text{Vega} = S\sqrt{T-t}\,\phi(d_1)$$

While Gamma and Delta are denoted by Greek letters, Vega is either written out, or a script letter $\mathcal{V}$ is used, or sometimes even a lowercase Greek "nu" (ν) — a horrible misuse of notation. So it seems appropriate to design a new pseudo-Greek letter for Vega:

$$Ɣ = S\sqrt{T-t}\,\phi(d_1).$$

This works well in uppercase (roman Ɣ and italic *Ɣ*), but it would be hard to find a new distinct lowercase shape: this is overcrowded territory, with v, ν and upsilon υ.

Yet people in mathematical finance are quite inventive: there's not only Vega, but also "Vanna" and "Volga" (also called "Vomma") — all derived from or related to volatility, so all start with "V" — and then also "speed", "color", "charm", and "zomma". So while the idea for a special letter for Vega might be nice, it seems quite hopeless to design proper letter-like symbols for all these quantities.

5.2 Field extension

In algebra, a common way to denote a field extension "L over K" is by $L{:}K$, alternatively L/K or $L|K$ is used. All three notations are over-used: "$:$" for index of a subgroup; "$/$" for quotient ring, quotient group, division; "$|$" for "divides".

To get an unambiguous notation, my idea is a special "field extension colon", formed by two small triangles, thus: LːK (the international phonetic alphabet IPA contains a similar idea: a colon of two triangles pointing towards each other is used to denote length of a vowel).

By its asymmetric form it shows that L is the extended field. This is close to the current notation. It does not disturb the reader, but it is there to help when he is in doubt. Of course this can't be used in handwriting, and admittedly it is not very visible in print (and it needs high-quality printing). But it might work well in online documents, where the reader could magnify the text — yet a properly tagged pdf file might be more helpful.

5.3 Algebraic substructures

One can often read sentences like

> Let $H \subset G$ be a subgroup of G.

This is logically wrong, as it tries to express two statements in one sentence (*"let H be a subset of G"* and *"let H be a subgroup of G"*), and the "$\subset G$" part is redundant, as any subgroup is a subset *ipso facto*. So it suffices to say

> Let H be a subgroup of G.

But obviously people like to use symbols (it seems that to many people the logically correct version without symbols feels somewhat weaker), so I thought of a way to express "is a subgroup of" with a symbol. I suggest to use "$\subset$" with a small "G" set atop, or alternatively below on the right (preferably, the "G" should be upright and sans-serif, as this better separates it from other letters and makes clear it is part of the symbol):

$$H \stackrel{\mathsf{G}}{\subset} G \quad \text{or} \quad H \subset_{\mathsf{G}} G.$$

Thus the sentence from above is shortened to

> Let $H \stackrel{\mathsf{G}}{\subset} G$.

This works for other algebraic structures as well:

$$S \stackrel{\mathsf{R}}{\subset} R, \qquad E \stackrel{\mathsf{F}}{\subset} F, \qquad U \stackrel{\mathsf{V}}{\subset} V, \qquad B \stackrel{\mathsf{A}}{\subset} A, \quad \text{etc.}$$

(ring, field, vector space, algebra). The obvious TeX macro names for these would be `\subgroup`, `\subfield`, etc., and one can easily construct these symbols in TeX with `\stackrel` and appropriate font switches.

But some names are not international, e.g. "field" (Latin "campus", French "corps", German "Körper") and "ring" (Latin "anellus", French "anneau") — according to our quality criteria, abbreviations should come from Latin, but "anellus" and "algebra" would both require an A. So to make this notation really useful, we need a list of standard abbreviations for algebraic structures.

5.4 Stirling numbers

Stirling numbers are used to convert from factorial powers to ordinary powers, and vice versa. For definition and properties see [6, pp. 66–69] or [3, pp. 257-267].

Stirling numbers of the first kind are notated with brackets:

$$\begin{bmatrix} n \\ k \end{bmatrix} = (n-1)\begin{bmatrix} n-1 \\ k \end{bmatrix} + \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} \quad (\text{for } k > 0),$$

with the initial conditions

$$\begin{bmatrix} n \\ 0 \end{bmatrix} = \delta_{n0} \quad \text{and} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0.$$

Stirling numbers of the second kind are notated with braces:

$$\begin{Bmatrix} n \\ k \end{Bmatrix} = \begin{Bmatrix} n-1 \\ k-1 \end{Bmatrix} + k\begin{Bmatrix} n-1 \\ k \end{Bmatrix}$$

with

$$\begin{Bmatrix} n \\ 1 \end{Bmatrix} = 1 \quad \text{and} \quad \begin{Bmatrix} n \\ n \end{Bmatrix} = 1.$$

These notations are similar to those for binomial coefficients and Eulerian numbers (denoted by $\binom{n}{k}$ and $\left\langle {n \atop k} \right\rangle$, respectively). According to Knuth:

> "These notations [...] have compelling advantages over the many other symbolisms that have been tried." [6, p. 66].

This is certainly true for existing symbols. But brackets and braces are used for many concepts, so how about new, distinct delimiters?

My idea was to keep close to the notation with braces and brackets, but to make it mnemonic by including an "s" form. But this proved to be unsuccessful. It seemed nice as an idea, it still seemed possible in handwriting, but when tried in print, it becomes clear that this is not working (at least it would need some reworking to make it useful).

First I tried new special brace-like delimiters, with an "s" in the top for Stirling numbers of the first kind, and in the bottom for those of the second kind. This looks just too obtrusive, too distracting:

$$\left\{ {n \atop k} \right\} \quad \text{and} \quad \left\{ {n \atop k} \right\},$$

especially when tried in a formula:

$$\left\{ {n \atop k} \right\} = \left\{ {n-1 \atop k-1} \right\} + k \left\{ {n-1 \atop k} \right\}.$$

Instead of helping the reader, it hampers readability.

Then I tried to keep brackets and braces, but now including a small "s" form (again in the top for the first kind, in the bottom for the second kind):

$$\begin{bmatrix} n \\ k \end{bmatrix} \quad \text{and} \quad \begin{Bmatrix} n \\ k \end{Bmatrix};$$

used in a formula:

$$\begin{bmatrix} n \\ k \end{bmatrix} = (n-1)\begin{bmatrix} n-1 \\ k \end{bmatrix} + \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}.$$

I consider this worse than the first version. Both just do not work. So it seems best to keep brackets and braces.

5.5 Greatest common divisor, least common multiple

My last example is one where I think a new notation is really needed. For "greatest common divisor" and "least common multiple", a standard notation is missing. To me, this is the most severe shortcoming in mathematical notation in general. So far mathematicians have failed to come up with a good notation, which is completely incomprehensible, as the concept of greatest common divisor is important and in wide use. It seems each language just uses some abbreviations:

English:	$\mathrm{gcd}(a, b)$	and	$\mathrm{lcm}(a, b)$
French:	$\mathrm{pgcd}(a, b)$	and	$\mathrm{ppcm}(a, b)$
German:	$\mathrm{ggT}(a, b)$	and	$\mathrm{kgV}(a, b)$
Dutch:	$\mathrm{ggd}(a, b)$	and	$\mathrm{kgv}(a, b)$
Polish:	$\mathrm{NWD}(a, b)$	and	$\mathrm{NWW}(a, b)$
Spanish:	$\mathrm{mcd}(a, b)$	and	$\mathrm{mcm}(a, b)$

The French version needs four letters (while almost all abbreviations in mathematics use three letters at most), the German version gives us the additional ugliness of mixed-case abbreviations. The Spanish version makes clear why Latin abbreviations would not work here: the "m" stands for "máximo" in one case and for "mínimo" in the other.

Also a very common notation for greatest common divisor, especially in number theory, is just (a, b). As I said above when discussing interval notation, this is ambiguous and over-used, and a reader always has to check the context to make sure what is meant. And there isn't any matching fixed notation for the least common multiple, so either $[a, b]$ or $\{a, b\}$ is used, but both notations always require an explanation in the text.

Formulae written in any of these notations are either lengthy, or not very readable, or not self-explanatory:

$$\gcd(F_m, F_n) = F_{\gcd(m,n)} \qquad (F_m, F_n) = F_{(m,n)}$$

$$\varphi\bigl([d, k]\bigr) = \varphi(d)\varphi(k) / \varphi\bigl((d, k)\bigr)$$

and often they need additional explanation (this example with mixed notation is taken from [11, p. 63]):

Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be sequences such that $\bigl((\operatorname{lcm} \mathcal{D}_1), (\operatorname{lcm} \mathcal{D}_2)\bigr) = 1$, where lcm *denotes the least common multiple of the members of a sequence, and the outer parentheses denote greatest common divisor. Then [...]*

Therefore I suggest the following new notation with new, special delimiters:

$\overline{\backslash} a, b \overline{/}$ for the greatest common divisor,

$\underline{/} a, b \underline{\backslash}$ for the least common multiple.

Many of our quality criteria are easily checked: obviously, these notations are *readable, needed, international, distinct and unambiguous* (at least in print).

They are also *writable,* although in sloppy handwriting they might be confused with 1 or 7 (depending on writing style) and possibly also with "floor" $\lfloor x \rfloor$ and "ceiling" $\lceil x \rceil$, but these two functions always have just one argument, while $\overline{\backslash} a, b \overline{/}$ usually has two or more arguments.

Of course these notations are *pronounceable:* in English, $\overline{\backslash} a, b \overline{/}$ is pronounced as the "greatest common divisor of a and b", and so analogously in any other language (just as it is the case with many notations which have their language-specific names).

They are *adaptable:* taking our example from above, we see that we could stretch the notation to more than two arguments: $\overline{\backslash} a, b, c \overline{/}$, but also to a single argument: $\overline{\backslash} S \overline{/}$ (e.g. when taking the greatest common divisor over all members of a set S).

They are *similar and consistent:* they fit into the general system, $\overline{\backslash} a, b \overline{/}$ is close to the common notation (a, b), and now we have *dual* symbols for dual concepts.

But are these new delimiters *available?* Well, yes — at least in my mathematical fonts ("Minion Math", as used here). But then it is such a simple, almost primitive design that it does not pose any problem to a font designer who wants to add these to his font. For the METAFONT sources of Computer Modern, it is a very simple addition: just take the code for "slash" or "backslash", add a third point whose coordinates are already known by the other two points, and connect the three points.

So the last remaining point is *"mnemonic",* and the mnemonic aspect is the reason for the particular form I chose. For two positive integers a and b, the following chain of inequalities holds:

$$\overline{\backslash} a, b \overline{/} \le \min(a, b) \le \max(a, b) \le \underline{/} a, b \underline{\backslash}.$$

Therefore $\overline{\backslash} a, b \overline{/}$ should remind of two vertices of a triangle pointing downwards to a lesser number (lesser than $\min(a, b)$, that is), and $\underline{/} a, b \underline{\backslash}$ is meant to point upwards to a greater number. Let's illustrate this with an example:

$$\begin{array}{l} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad 42 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\;\; / \quad \backslash \\ \overline{\backslash} 14, 21 \overline{/} \le \min(14, 21) \le \max(14, 21) \le \underline{/} 14, 21 \underline{\backslash}. \\ \;\backslash \quad / \\ \;\;\; 7 \end{array}$$

Of course the concept of greatest common divisor applies not only to positive integers. It could be generalized to negative integers, to polynomials, or to elements of a commutative ring. Then the above inequalities do not hold in general. Still the mnemonic is correct as it only serves to remember which is which of our notations.

When we apply these new notations to the above examples, we see that the formulae get shorter and do not need any additional explanation anymore:

$$\gcd(F_m, F_n) = F_{\gcd(m,n)} \qquad \overline{\backslash} F_m, F_n \overline{/} = F_{\overline{\backslash} m,n \overline{/}}$$

$$\varphi\bigl(\underline{/} d, k \underline{\backslash}\bigr) = \varphi(d)\varphi(k) / \varphi\bigl(\overline{\backslash} d, k \overline{/}\bigr)$$

We could even agree upon saving parentheses, so the last formula might be written as

$$\varphi \underline{/} d, k \underline{\backslash} = \varphi(d)\varphi(k) / \varphi \overline{\backslash} d, k \overline{/},$$

but this would be overdoing it maybe. Our last example above could now be shortened to

Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be sequences such that $\overline{\backslash} \underline{/}\mathcal{D}_1\underline{\backslash}, \underline{/}\mathcal{D}_2\underline{\backslash} \overline{/} = 1$, where $\underline{/} \ldots \underline{\backslash}$ denotes the least common multiple of the members of a sequence. Then [...]

or even, when it's clear what we mean by $\underline{/} \ldots \underline{\backslash}$ with a single argument, simply to

Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be sequences such that $\overline{\backslash} \underline{/}\mathcal{D}_1\underline{\backslash}, \underline{/}\mathcal{D}_2\underline{\backslash} \overline{/} = 1$. Then [...]

This is considerably shorter, and yet clearer than the original version.

One objection to the particular form of these new delimiters is that the notation seems counter-intuitive, especially compared with "floor" $\lfloor x \rfloor$ and "ceiling" $\lceil x \rceil$, where the "serifs" or little bars of the delimiters point to smaller and greater numbers just in the opposite way. But I do not see this as a contradiction: while the entire form $\backslash a, b /$ should remind one of a lesser number, the serifs remind one of the *greatest* number which divides both a and b.

I think that mathematicians can all agree that a good, distinct, international notation is really necessary here. The exact form of the symbols is up for discussion, but I can't think of a more suitable form. I consider the notations $\backslash a, b /$ and $/ a, b \backslash$ necessary and important innovations, and I hope that mathematicians will adopt these symbols.

References

[1] Florian Cajori. *A History of Mathematical Notations.* Dover, New York, 1993. Reprint. Originally published in two volumes in 1928 and 1929 by The Open Court Publishing Company, Chicago.

[2] Friedrich Forssman and Ralf de Jong. *Detailtypografie*. Hermann Schmidt Verlag, Mainz, fourth edition, 2008.

[3] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics*. Addison-Wesley, Reading (MA), second edition, 1994.

[4] Donald E. Knuth. *The TEXbook*, volume A of *Computers and Typesetting*. Addison-Wesley, Reading (MA), 1986.

[5] Donald E. Knuth. *The METAFONTbook*, volume C of *Computers and Typesetting*. Addison-Wesley, Reading (MA), 1986.

[6] Donald E. Knuth. *Fundamental Algorithms*, volume 1 of *The Art of Computer Programming*. Addison-Wesley, Reading (MA), third edition, 1997.

[7] Johannes Küster. *Mathematischer Formelsatz*, pages 203–233. In *Detailtypografie* [2], fourth edition, 2008.

[8] Johannes Küster. *Sonderzeichen: Mathematikzeichen*, pages 382–389. In *Detailtypografie* [2], fourth edition, 2008.

[9] Fritz Reinhardt and Heinrich Soeder. *dtv-Atlas Mathematik*, volume 1: Grundlagen, Algebra und Geometrie. Deutscher Taschenbuch Verlag, München, twelfth edition, 2001.

[10] Fritz Reinhardt and Heinrich Soeder. *dtv-Atlas Mathematik*, volume 2: Analysis und angewandte Mathematik. Deutscher Taschenbuch Verlag, München, eleventh edition, 2003.

[11] R. J. Simpson and Doron Zeilberger. Necessary conditions for distinct covering systems with square-free moduli. *Acta Arithmetica*, 59(1):59–70, 1991.

[12] Friedrich Wilhelm Weitershaus, editor. *Duden Satz- und Korrekturanweisungen (Duden Taschenbuch Band 5).* Bibliographisches Institut, Mannheim, fifth edition, 1986.

◇ Johannes Küster
typoma GmbH
Karl-Stieler-Str. 4
D–83607 Holzkirchen
Germany
`info (at) typoma dot com`
`http://www.typoma.com`