Noah Monk

1 April 2024

# On Selecting Good Notation

Selecting good notation is a crucial and underrated aspect of solving problems in mathematics. Good notation mitigates errors and is convenient, compact, unique, and descriptive. The benefits of good notation can influence one's ability to solve a problem. Famous mathematician Terence Tao has described in his thought process for solving mathematical problems that one should select good notation to represent data and objectives efficiently (3). One must first understand what problems poor notation brings.

Starting with a common cause of frustration and tears: $Sin^{-1}[x]$ and $(Sin[x])^{-1} = \frac{1}{Sin[x]}$. For those unfamiliar with these two strange-looking collections of letters, these are known as trigonometric functions (Nelson 442). Trigonometry is just a fancy word for talking about triangles. It is about the relationships between angles in a triangle and the sides of a triangle (Nelson 442). Functions are just decorative rules for how something changes or is to be changed (Nelson 182). In the case of the two trigonometric functions above, we see the letter '$x$'. In this case, the letter '$x$' is a variable. A variable describes things like the list of even numbers, $x = 2,4,6,8,10,...$ or some number like 343 (Nelson 455-456). This variable '$x$' can denote any number one chooses.  'Sin' is the trigonometric function acting on $'x'$. We denote this as $Sin[x]$. Various operations can apply to variable $'x'$. We can multiply itself to itself, $x * x$. The nice thing about this notation is that when we raise '$x$' to the power of some positive number like two, it tells us how many times we have to multiply '$x$' by itself. For example, $x^2 = x * x$ (Nelson 160). When we raise '$x$' to a negative number, it tells us how many times someone would have to divide one by '$x$'. For example, $x^{-1} = \frac{1}{x}$ or $x^{-3} = \frac{1}{x * x * x} = \frac{1}{x^3}$ (Nelson 160). Naturally, by this convention, if one had viewed $Sin^{-1}[x]$, they would likely interpret this as the same thing as saying $\frac{1}{Sin[x]}$. In other words, that $Sin^{-1}[x] = \frac{1}{Sin[x]}$. This identity, however, is not true at all. $Sin^{-1}[x]$ is not the same as $\frac{1}{Sin[x]}$. $\frac{1}{Sin[x]}$ is the definition of a different trigonometric function known as Cosecant, in which one denotes $\frac{1}{Sin[x]} = Csc[x]$ (Nelson 443). $Sin^{-1}[x]$ refers to an inverted trigonometric function known as the arcsine or $ArcSin[x]$ (Nelson 443). $Sin^{-1}[x]$ will not apply appropriately to solving for a side of a triangle, leading one away from the correct answer due to this inconsistency. Arcsine and Cosecant are classic examples of the inconsistency of how bad notation can lead to an incorrect path in mathematics.

Another frequent example of how poor notation can lead one astray is when letters or variables look too similar. For example, in electricity and magnetism physics, one often sees ‘V’ denoting a volume (Griffith 41). Nevertheless,’$V$’ is also for voltage, and ‘$v$’ is the magnitude of velocity and speed (Griffith 79). Having all of these variables present can lead to confusion about which one is denoting what. This issue occurs when solving Laplace equations in 3-D for electrostatic potentials (Griffiths 119). In that situation, both voltage, ‘V’, and volume, ‘$V$’, are in the same equation. Accidentally equating the two, V=$V$, would be a critical mistake and undoubtedly lead to an incorrect answer. A more straightforward example is ‘e’ and ‘e’ in physics. ‘e’ often denotes an electron's electrical charge, $e = 1.602176634 * 10^{-19}$ Coulombs (Knight 607). However, e often denotes Euler's number, $e = 2.718281828459045...$ (Nelson 161). The difference between these two numbers is substantial. Confusing the two is a miscalculation. One may face temptations to denote Euler's number as capital ‘E’ for Euler. The variable ‘E’ already denotes energy in physics (Knight 226). This lack of uniqueness is problematic for solving mathematics problems. One should avoid duplicates whenever possible.

Another problem that arises from poor notation is information loss. Velocity and speed denote the same symbol, ‘v’ (Knight 9). However, this convention leads to information loss as velocity and speed differ. The main difference is that, unlike speed, velocity has a speed and a direction in which the object moves (Knight 9). For example, 80 km/h is a speed. It explains that the object in question is moving at a distance of 50 kilometres every hour. 50 km/h N is a velocity and says something entirely different. It is saying that the object in question is moving a distance of 50 kilometres every hour in the direction of the North. This distinction is essential when talking about moving objects. If one were doing calculations involving speed and velocity, one could mistake velocity for speed and vice versa, resulting in some answer that makes no sense.

A critical aspect of good notation is that good notation is convenient. Sometimes, in mathematics, one sees equations that are far too long to write down entirely in unrestricted form. For example, the exact solution of the hydrogen atom wavefunction from the famous Shrödinger Equation in quantum mechanics in complete unrestricted form looks like this:

$$\psi_{nlm} = \sqrt{\left(\frac{2}{n\frac{4\pi\epsilon_0\hbar^2}{m_e e^2}}\right)^3 \frac{(n-l-1)!}{2n(n+l)!}}\, e^{\frac{-r}{n(\frac{4\pi\epsilon_0\hbar^2}{m_e e^2})}} \left(\frac{2r}{\frac{4\pi\epsilon_0\hbar^2}{m_e e^2}}\right)^l \frac{\left(\frac{2r}{n(\frac{4\pi\epsilon_0\hbar^2}{m_e e^2})}\right)^{-2l+1} e^{\left(\frac{2r}{(n(\frac{4\pi\epsilon_0\hbar^2}{m_e e^2}))}\right)}}{(n-l-1)!} \left(\frac{d}{d(\frac{4\pi\epsilon_0\hbar^2}{m_e e^2})}\right)^{n-l-1} \left(e^{-\frac{2r}{(\frac{4\pi\epsilon_0\hbar^2}{m_e e^2})}} \frac{4\pi\epsilon_0\hbar^2}{m_e e^2}^{2l+1+n-l-1}\right) \sqrt{\frac{(2l+1)}{4\pi}\frac{(l-m)!}{(l+m)!}}\, e^{im\phi} \frac{1}{2^l l!} \left(\frac{d}{dCos[\theta]}\right)^l (Cos[\theta]^2 - 1)^l$$

Other forms are more thorough than this (Griffith and Schroether 151). Putting in all of the exact numbers would make this solution unreadable. Furthermore, in this solution, some parts have e's denoting electron charge and others denoting Euler's number. Separating which one is which is almost impossible unless one already knows every part of this solution and has it memorised. Keeping something as long and tedious as this more compact by denoting different parts can significantly simplify the expression and reduce confusion. Typing that solution took about half an hour. A good notation will reduce writing time because it is convenient, compact, readable, and necessary for equations as unsightly as hydrogen wavefunction.

The efficiency of using a good notation can speed up communication between readers, speakers, and writers. For example, in Einstein's Theory of General Relativity, one writes repeated indices or contractions frequently; one example could look like this $\eta_0 j^0 + \eta_1 j^1 + \eta_2 j^2 + \eta_3 j^3$. Writing such a thing repeatedly a couple of thousand times will reduce efficiency significantly (Schutz 34). As a result, Einstein created what is now known as the Einstein Summation Convention notation, $\eta_\alpha j^\alpha = \sum_{i=0}^{3} \eta_i j^i = \eta_0 j^0 + \eta_1 j^1 + \eta_2 j^2 + \eta_3 j^3$, so that one could write something much more efficiently (Schutz 34). Now, all one needs to write under this convention is $\eta_\alpha j^\alpha$, which denotes the same as this $\eta_0 j^0 + \eta_1 j^1 + \eta_2 j^2 + \eta_3 j^3$ from earlier.

A good notation makes things more convenient and efficient and reduces the chances of transcription or copying errors. For example, the number 100000000000000000000000000000000000000000000000000 is a gigantic number with many zeros. One can imagine writing this number out a few times. Any human, even just reading such a number, will struggle to correctly know the exact number of zeros present in this number. When this number is rewritten or read, one or two zeros will likely be lost or gained. This kind of error can change everything.

The number 100000000000000000000000000000000000000000000000000 and the following number 100000000000000000000000000000000000000000000000 look identical. The difference between these two numbers is three orders of magnitude. If one accidentally removes three zeros from the first number, this would be the equivalent of one house suddenly going from 10 metres long to 10 millimetres long. It may be comical, but it is a significant misrepresentation. The difference between these numbers is substantial, considering the house length shrinking to 10 millimetres. A better way to represent these large numbers is $10^{50}$ and $10^{47}$. This notation reduces the chances of losing information and makes writing more compact and efficient.

Another important thing about good notation is that good notation is unique, such as duplicated variables from earlier. A good notation rarely has duplicates and keeps notations unique. In trigonometry, labelling the triangle sides ‘A’, ‘B’, and ‘C’ gives each side of the triangle a unique letter. This notation means the letters chosen are distinguishable. However, denoting the sides of the triangle by ‘a’, ‘$a$’, and ‘$\alpha$’ is likely to lead to some errors. It would be a semantic nightmare trying to verbally describe to someone the trying with sides labelled with ‘a’, ’$a$’, and ‘$\alpha$’. It is much easier to describe the ‘ABC’ triangle than the ‘a$a\alpha$’ triangle despite both being the same triangle.

Good notation being unique is one of many things to consider when selecting notation. Good notation is unique and descriptive. One could denote Euler's number by ‘g’, making it unique from ‘e’. Changing variable naming to be more descriptive needs explanation and sound reasoning. ‘e’ for Euler makes sense as Euler's name starts with an ‘e’. An example of an exceptional descriptive naming convention is transistor naming conventions in physics. For example, JFET means junction field-effect transistor, MOSFET is metal oxide semiconductor field-effect transistor, and GAAFET is gate all around field-effect transistor (Jaeger and Blalock 146). The good thing about all these notations for each transistor is that they describe the type of transistor, field-effect transistors, and their constructions simultaneously (Jaeger and Blalock 146-147). For example, GAAFET, more specifically, "gate all around," states that the gate of the transistor extends all around the transistor. These notations are unique and descriptive. Naming groups 463r234 and 792658t does not describe what the group has in it. Non-descriptive notation is poor notation where information perishes about its subject.

It is essential to select good notation. Poor notation is prone to errors and, as a result, leads one away from the desired answer. Good notation is also convenient, compact, unique, and descriptive. Choosing good notation over bad notation can be the determining factor in making progress in a problem in mathematics or being able to communicate effectively with others.

Works Cited


Tao, Terence. Solving Mathematical Problems: A Personal Perspective. Oxford: Oxford University Press, 2006. Print.

Nelson, David. Penguin Dictionary of Mathematics. 4th ed. London: Penguin Books Ltd, 2008. Print

Griffith, David. Introduction to Electrodynamics. 4th ed. Noida: Pearson India Education Services Pvt, Ltd, 2015. Print.

Knight, Randall. Physics: For Scientists and Engineers; A Strategic Approach; With Modern Physics. 4th ed. Boston: Pearson Education INC, 2017. Print.

Griffith, David and Schroether, Darrell. Introduction to Quantum Mechanics. 3rd ed. Cambridge, Cambridge University Press. Print

Schutz, Bernard. A First Course in General Relativity. 2nd ed. Cambridge, Cambridge University Press, 2009. Print.

Jaeger, Richard and Blalock, Travis. Microelectronic Circuit Design. 4th ed. New York, The McGraw-Hill Companies INC, 2011. Digital.