Mathematics, Shaken but not Stirred

Kevin H Knuth

The last century saw the breakdown of the dream of the mechanical universe where it was imagined that given sufficient intellect and effort all truths could be discovered and known. Gödel's Incompleteness Theorems, which form the basis for the concept of Undecidability, revealed that there exist true statements that cannot be proven to be true. The analogue in computer science is the concept of Uncomputability, which is based on the proofs by Alonzo Church and Alan Turing that the *Entscheidungsproblem*, the Decision Problem, cannot be generally solved. This means that there exists no general algorithm that can logically decide whether a given statement is provable from a set of axioms.

While these revelations shook the mathematical community, it is reasonable to ask, in retrospect, whether this is at all surprising. These theoretical results were augmented by experimental results in quantum mechanics which demonstrated that the universe at the smallest of scales was inherently unpredictable in the sense that only the probabilities of outcomes could be predicted rather than the outcomes themselves. This inherent unpredictability of quantum mechanics was later supplemented with the unpredictability of a number of nonlinear dynamical systems which exhibited sensitive dependence on initial conditions. Together, these experimental demonstrations of inherent unpredictability in conjunction with the theoretical results of undecidability and uncomputability served to decimate our confidence in anything like a mechanical clockwork picture of the universe.

Mathematics has a long history of not dealing well with uncertainty. The drowning of Hippasus following his discovery of irrational numbers[1] may have been one of the first examples of Mathematicians desperately defending their dogmatic Worldview. Certainly, the tortuous history of Probability Theory, which includes Frequentism tying Bayesian Probability Theory up in the basement and holding it hostage for nearly a century, is another example of the Mathematician's strained relationship with uncertainty. But Kurt Gödel, a Mathematician, played their game well and delivered a theorem that proved undecidability with certainty. This undeniably shook the Mathematical Worldview, but was this Worldview shaken enough? Perhaps Mathematics should have been sufficiently stirred to follow the lead of the Physical Sciences and learn to embrace uncertainty when necessary.

The Unreasonable Effectiveness of Mathematics

When I was in graduate school, a question came to me that I couldn't answer. I wondered about addition and it bothered me that it wasn't quite clear to me how we knew that we should apply the operation of addition to numbers of things when they are combined.

> Was this a fact that had been proved at some point, and if so, how?
> Or was this discovered over time through experience, or rather, experiments?
> Or is the operation of addition simply defined to work?
> Or is the operation of addition somehow inherent to the concept of number?

I asked my fellow students, and it quickly became clear that none of them really knew. In fact, many of them had never really thought about it, which wasn't surprising since it wasn't something I had ever

---

[1] It is possible that this story is not entirely accurate.

noticed until it struck me. I then began to ask my professors, which is where this situation became very disappointing. None of them could answer the question, and only a handful admitted it. Instead, several of them declared that it was obvious and proceeded to make fun of me for asking the question.

For a short while, I was known as the graduate student who didn't understand why we add numbers. Fortunately, I was smart enough to recognize this as a sign of their ignorance. And it worried me. It worried me that such a fundamental question, which should really have a straightforward answer, had completely eluded all of these otherwise intelligent people. I could not help but feel that something was seriously wrong somewhere.

It was only a few years later that I stumbled on the following quote from the mathematician Richard Hamming in his 1980 paper that paid homage to Eugene Wigner's famous paper (Wigner 1990) with almost the same title, "The Unreasonable Effectiveness of Mathematics in the Natural Sciences", published twenty years earlier:

> "I have tried, with little success, to get some of my friends to understand my amazement that the abstraction of integers for counting is both possible and useful. To me, this is one of the strongest examples of the unreasonable effectiveness of mathematics. Indeed, I find it both strange and inexplicable."
>
> - Richard Hamming (Hamming, 1980)

I felt some relief and vindication that I was not alone in my wonderment. More importantly, I came to appreciate that the solution was not obvious and must be very deep. The question as to why we add the numbers that represent sets of objects when we combine those sets of objects lies at the very heart of Wigner's amazement of the unreasonable effectiveness of mathematics.

Although profound, the solution really is simple. The effectiveness of mathematics is reasonable because mathematics is designed to work. But the fact that it is designed is subtle, in part because Mathematics masquerades as something that is discovered through absolute reason. In that Worldview, there is no room for design and there is certainly no room for uncertainty!

The mathematics of addition is designed to be consistent with the basic symmetries of combination, and it is unique up to isomorphism (invertible transform), which gives the illusion of it having been dictated by God. This aspect of the unreasonable effectiveness of mathematics was the topic of one of my earlier FQXi essays (Knuth, 2015), and it pays to look at it more closely here.

Measure, Probability, and Quantum

In science, we quantify things so that we can rank them: quantity, mass, volume, voltage, probability. To maintain such rankings, quantities must be assigned consistently, especially in situations in which things are combined or partitioned to form other numbers of things.

Consider an operator $\oplus$ that combines disjoint objects A and B into a composite object $A \oplus B$. In general, for objects that have N comparable qualities, we can quantify object A with an N-tuple of scalar quantities (numbers) $a = \{a_1, a_2, \cdots, a_N\}$, and quantify object B with the N-tuple $b = \{b_1, b_2, \cdots, b_N\}$. Since the composite object $A \oplus B$ is related to both objects A and B, we can expect that the N-tuple used to quantify $A \oplus B$ must a function of both N-tuples $\{a_1, a_2, \cdots, a_N\}$ and $\{b_1, b_2, \cdots, b_N\}$.

We have shown (Knuth, 2009; 2015; Skilling and Knuth, 2018) that the basic symmetries of the operator $\oplus$ suffice to constrain the quantification of $A \oplus B$. Specifically, if we have **Closure**, such that $A \oplus B$ is considered to be the same type of object as A and B, and the combination operator $\oplus$ satisfies:

Commutativity $$A \oplus B = B \oplus A \tag{1}$$

and

Associativity $$A \oplus (B \oplus C) = (A \oplus B) \oplus C \tag{2}$$

so that objects can be shuffled without changing the result, and we can conceivably continue to combine equivalent but disjoint objects, then one can show (Skilling and Knuth, 2018) that $A \oplus B$ is represented (up to isomorphism) by the component-wise sum $a + b = \{a_1 + b_1, a_2 + b_2, \cdots, a_N + b_N\}$.

In the case where the objects are fully commensurable so that there is only one relevant quantity then instead of an N-tuple quantification, we have a scalar quantification, resulting in *Measure Theory*. However, here we have additivity is a **theorem** rather than the usual axiom.
Hence the answer to my question!

We can consider logical statements and quantify the degree to which one logical statement implies another by a scalar quantity, which we call the probability. For example, we can consider the degree, $P(A \mid I)$, to which prior information I implies a logical statement A, and the degree, $P(B \mid I)$, to which prior information I implies a logical statement B. For disjoint statements A and B, the logical OR of statements A and B results in a logical statement, denoted $A \vee B$, (closure). Furthermore, the logical OR is both commutative and associative, which means that we can write the probability of the statement $A \vee B$ as the sum of probabilities:

$$P(A \vee B \mid I) = P(A \mid I) + P(B \mid I). \tag{3}$$

This is the *Sum Rule* of probability. If statements A and B are not disjoint, the Sum Rule takes the more general form (Knuth, 2009)

$$P(A \vee B|I) = P(A|I) + P(B|I) - P(A \wedge B|I), \tag{4}$$

where $A \wedge B$ denotes the logical AND of statements A and B.

We can also consider chaining logical statements, and again, *associativity* of chaining results in a quantification that is isomorphic to additivity. However, we also have *distributivity* over the combination, $\vee$, which results in a *Product Rule*

$$P(A|C) = P(A|B)\, P(B|C), \tag{5}$$

which can be expressed in the more familiar, more general form (Knuth, 2009)

$$P(A \wedge B|C) = P(A|B \wedge C)\, P(B|C). \tag{6}$$

As a result, we have a derivation of *Probability Theory*.

As a final example, we consider the quantification of the interaction between a target (e.g. a beam of particles) and a probe. We postulate that this requires two numbers, which we call the Pair Postulate (Goyal, Knuth, Skilling 2010; Skilling and Knuth 2019). We can consider combining beams in parallel as well as series, and as before, the basic symmetries of shuffling hold, along with consistency with probability so that we can make inferences. This results in the *Feynman Sum and Product Rules* for quantum amplitudes:

$$u + v = \binom{u_1}{u_2} + \binom{v_1}{v_2} = \binom{u_1 + v_1}{u_2 + v_2} \tag{4}$$

$$u \circ v = \binom{u_1 v_1 - u_2 v_2}{u_1 v_2 + u_2 v_1} \tag{5}$$

along with the *Born Rule* $P(u) = |u|^2$ (Goyal, Knuth, Skilling 2010; Skilling and Knuth 2019).

We have outlined the derivation of three theories: Measure Theory, Probability Theory, and the Feynman Rules for manipulating Quantum Amplitudes. These derivations rely on consistent quantification, where one assigns numbers to objects in a way that is consistent with respect to underlying symmetries and the underlying order. The result is a set of constraint equations, typically recognized as laws, which enforce those symmetries. This is another way, distinct from, but equally profound as, Noether's theorem, in which symmetries give rise to laws.

Lattice Theory

Here we make a short detour into *order theory*, which is also known as *lattice theory*, so that we can think carefully about Hypothesis Spaces. A partially ordered set is a set of elements along with a binary ordering relation, generally denoted $\leq$, that obeys the following properties for elements x, y, and z:

For all x, $x \leq x$	Reflexivity
If $x \leq y$ and $y \leq x$, then $x = y$	Antisymmetry
If $x \leq y$ and $y \leq z$, then $x \leq z$	Transitivity

This structure is referred to as being *partially* ordered, since it is not assured that every element of the poset can be compared to every other element. That is, when there exist two elements x and w, such that $x \nleq w$ and $w \nleq x$, we say that x and w are incomparable, denoted $x \parallel w$. The ordering relation $x \leq z$ is read either as x is less than z, x is included by z, or z includes x. If $x \leq z$ and there does not exist an element y such that $x \leq y \leq z$, then it is said that z *covers* x.

A *lattice* is a poset such that for every pair of elements, u and v, there exists a least upper bound called the join, denoted $u \vee v$, as well as a greatest lower bound called the meet, denoted $u \wedge v$. While the join of two elements always exists in a lattice, there can be elements, called join-irreducible elements, that cannot be expressed as the join of two other distinct elements. The lattice meet and join are commutative and associative.

The consistency relation

$$x \leq y \ \ implies \ \begin{matrix} x \vee y = y \\ x \wedge y = x \end{matrix} \tag{6}$$

highlights the fact that there are two equivalent perspectives: focusing on the ordering relation and the partial ordering of the elements, one has a lattice; whereas by viewing the join and meet as algebraic operations, one has an algebra. Lattices are algebras.

Hypothesis Spaces

"Supposing we hit him by mistake?", said Piglet anxiously.
"Or supposing you missed him by mistake", said Eeyore.
"Think of all the possibilities Piglet before you settle down to enjoy yourselves."
— A. A. Milne

The foundation of a Hypothesis Space is an exhaustive set of N mutually exclusive logical statements called the atomic statements, which we denote as $\{x_1, x_2, \cdots, x_N\}$. The space is formed by taking the logical disjunction (OR) of all possible subsets of these atomic statements. Logical implication, denoted $x \rightarrow y$, is a binary ordering relation $x \leq y$, so that a set of logical statements ordered by implication is a poset.

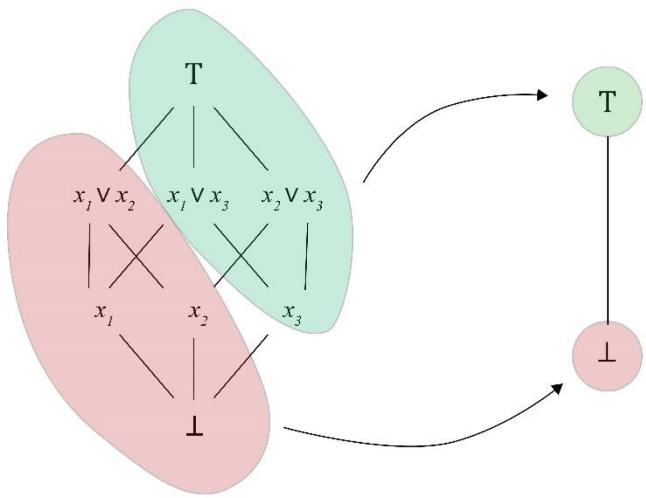

Figure 1. A Boolean lattice (left) formed from three atomic statements. If we are given the fact that x_3 is true this allows one to define an equivalence class of statements that are true (green), and a complementary equivalence class of statements that are false (red). Under this equivalence relation the lattice on the left collapses into the degenerate lattice on the right.

More importantly, this poset forms a Boolean lattice for which the lattice join, ∨, is the logical OR operation, and the lattice meet, ∧, is the logical AND, where the operations OR and AND comprise a Boolean algebra (Figure 1). The bottom element of the lattice, denoted ⊥, is called the *falsity*. It is a false statement, and as such it implies all others. The atomic statements that cover the bottom are join-irreducible. The top element, the truism, denoted T, which is formed from the join of all of the atoms is always true. Last, the Boolean lattice has a unary operation, *negation*, such that each statement x in the lattice has a logical complement $\bar{x}$, such that either the statement x OR the statement $\bar{x}$ is true, $x \vee \bar{x} = \mathrm{T}$, and the statement x AND $\bar{x}$ is false, $x \wedge \bar{x} = \perp$.

The resulting Boolean lattice is the Hypothesis Space. We are now prepared to discuss deductive logic and inductive logic in the context of a Hypothesis Space.

Deductive Logic versus Inductive Logic

Logical deduction works by following the ordering relation of implication. Proving that a logical statement x_T is true amounts to identifying an equivalence relation between the statement x_T and the truism T. Since any true statements must imply true statements, any statement that is implied by the true statement x_T is also a member of the equivalence class of true statements. Moreover, if x_T is true then its complement $\overline{x_T}$ is false, and the statement $\overline{x_T}$ is identified with the *falsity*. This true-false equivalence relation is illustrated in Figure 1 where $x_T \sim x_3$ and $\overline{x_T} \sim\ x_1 \vee x_2$.

An alternative technique to proving logical statements to be true, is proving that some logical statements are false. *Ruling out* possibilities, referred to as *eliminative induction*, is not only extremely useful in practice, as in high-quality machine learning algorithms, such as Nested Sampling (Skilling 2006), but it is the only way that learning progresses in the physical world. For example, The abbreviation, R/O, that medical doctors use when ordering a medical test means to 'Rule Out' a diagnosis. In such situations, one is reminded of the quote from Arthur Conan Doyle's Sherlock Holmes

> "When you have eliminated all which is impossible, then whatever remains, however improbable, must be the truth." —Sherlock Holmes

Of course, for this to apply, one needs to eliminate *all* of the other cases, regardless of their improbability. This is not surprising since the atomic statements in the hypothesis space must be exhaustive.

Deductive reasoning on the Boolean lattice of the Hypothesis space is carried out using Aristotle's two Strong Syllogisms:

Given:	If A is true, then B is true.	SS-1
Learn:	A is true.	
Deduce:	B must be true.	

and

Given:	If A is true, then B is true.	SS-2
Learn:	A is false.	
Deduce:	B must be false.	

This is the basic reasoning of the Mathematician. However, as Ed Jaynes (2003) points out that these syllogisms are too strong for practical use. One must rely on weaker syllogisms:

Given:	If A is true, then B is true.	WS-1
Learn:	B is true.	
Infer:	A becomes more plausible.	

and

Given:	If A is true, then B is true.	WS-2
Learn:	B is false.	
Infer:	A becomes less plausible.	

Clearly, in adopting weaker syllogisms we have now deviated from the deductive logic of the Boolean lattice. If it is any comfort, Jaynes reminds us that the mathematician George Polya wrote several books on the topic of plausible reasoning which obeys definite rules (Polya 1954; 1990), and Polya demonstrated that even pure mathematicians use the weaker syllogisms WS-1 and WS-2 most of the time only to buttress them later with deductive arguments of the stronger form, SS-1 and SS-2.

However, physical scientists usually do not have the luxury of knowing anything with certainty. This requires the application of an even weaker syllogism

Given:	If A is true, then B becomes more plausible.	WS-3
Learn:	B is true.	
Infer:	A becomes more plausible.	

Despite the weakness of the above syllogism, repeated application often allows one to achieve a precision approaching that of deductive reasoning. This is the essence of inductive reasoning.

Inductive reasoning can be quantified with real-valued functions, which we call probability. Probability is often thought of as quantifying the degree of truth. However, this conception has some difficulties since many of the models that scientists use are known to be approximations, which are strictly not true.

Instead, I think of probability as a generalization of implication to degrees of implication. As we learned above, implication flows up the lattice, with lower statements implying higher statements, as in $x_1 \rightarrow x_1 \vee x_2$ in Figure 1. But what we really want to quantify is the degree to which one of the higher statements, such as the truism, $\mathrm{T} = x_1 \vee x_2 \vee x_3$, implies a particular atom, such as x_3: `degree`$(x_1 \vee x_2 \vee x_3 \rightarrow x_1)$. Now, of course, it is not true that $x_1 \vee x_2 \vee x_3 \rightarrow x_1$. That is not how implication works. So, we introduce a function $P(y \mid x)$, called a bi-valuation, that generalizes binary implication by taking two logical statements to a real number, such that

$$P(y\,|x) = \begin{cases} 0 & \text{if} \quad x \wedge y = \bot \\ 1 & \text{if} \quad x \rightarrow y \\ 0 < p < 1 & \quad \text{otherwise} \end{cases} \tag{7}$$

This bi-valuation must be consistent with the symmetries of associativity and commutativity of ∨, and associativity of chaining inferences as well as the distributivity of chaining over ∨. This results in the familiar sum and product rules for probability, which act as constraint equations enforcing consistency with respect to the aforementioned algebraic symmetries resulting in the very reasonable effectiveness of mathematics! (Knuth 2005; 2009; Knuth & Skilling 2012; Skilling & Knuth 2019). This is Bayesian probability theory. The application of inductive inference typically proceeds by taking the Cartesian product of the Hypothesis Space (x) with a Data Space (d) of logical statements of observations with joint probabilities $P(x, d)$, with one particular data statement known to be true. Bayes' theorem is then derived by applying the product rule to this joint probability, and inference ensues.

Since this essay is focused on Gödel's Theorems, one ought to ask whether we could prove that probability theory is consistent. The sum and product rules are derived by enforcing consistency with the underlying symmetries of Boolean logic. However, the derivations we perform (Skilling & Knuth 2019) rely on eliminative induction where it is assumed that if there is a general theory, then it must apply to special cases. We then use special cases to constrain the theory to uniqueness, up to invertible transform, so that we then know that *if* there is a general theory of probability, *then* it must be the one we derived. The fact remains that we have not verified that a general theory exists. Such a proof may not be possible within the framework.

The good news is that probability theory will detect contradictions. So that if contradictions are ever found, they would be detected. That is, the theory will tell us if contradictions exist, which in a sense, is a way out.

Probability Theory and Tarski's Theorem

Related to Gödel's Incompleteness Theorems is Tarski's Undefinability Theorem, which states that within a formal system the concept of truth cannot be defined within that system. At first blush, such a revelation is disconcerting. However, we find that this fact is essential to the success of probability theory.

Probability theory itself consists of the sum and product rules, from which Bayes theorem is readily derived. In the past, there have been complaints that probability theory does not provide any clues as how to assign the probabilities of the atomic (join-irreducible) statements in the hypothesis space. Of course, this is a good thing because it makes it possible for one to apply the theory to a wide variety of problems. The probability assignments are problem-specific—not dictated by the theory of inference. In fact, if probability theory dictated the probability assignments, then the theory would be too constrained to be useful.

From the perspective of Tarski's theorem, probability theory cannot be used to define the concept of truth within probability theory. Probabilities must be assigned based on other application-specific constraints. In addition, despite the fact that the atomic statements are exhaustive, so that their logical disjunction, the truism, is true, it is the application, which also dictates which atomic statements belong in the Hypothesis Space.

Tarski's theorem, which says that probability theory cannot dictate truth values, does not tie our hands. Quite the opposite, Tarski's theorem ensures that probability theory can be used to solve inference problems in general where the concept of truth depends on the specific application!
Undefinability makes Probability Theory useful!

What Kinds of Truths are Undecidable?

Consider a hypothesis space generated by an exhaustive set of N atomic independent, mutually-exclusive logical statements denoted $\{x_1, x_2, \dots, x_N\}$. Consider a logical statement u in the hypothesis space that is undecidable so that it is impossible to prove that u is true or false.

It is clear that if there exists a true logical statement v that implies u, so that $v \rightarrow u$, then u must also be true, and hence decidable. However, since u is undecidable, then there can be no true statement, v, that implies u. This means that either u is an atomic statement such that there are no statements, other than the falsity, that imply u, or there exists a set of atomic statements that imply u, for which at least one of those atomic statements is undecidable and the rest are false.

Moreover, since it is not possible to prove u to be false, this implies that the logical complement of u, denoted $\bar{u}$, is also undecidable. However, since either u or $\bar{u}$ is true, $u \vee \bar{u} = T$, and the truism T is the join of all of the atomic statements, then the atomic statements can be separated into two subsets, each of which has at least one undecidable statement with the rest being false. We find that the existence of one undecidable statement u in the hypothesis space implies that there are at least two undecidable atomic statements in the hypothesis space with the rest being false.

However, it really shouldn't be too surprising that a hypothesis space cannot determine the truth values of the logical statements comprising it. This is especially obvious when one considers that assigning truth values amounts to defining equivalence classes that reduce the Boolean lattice with 2^N elements to the 2^1 lattice consisting of a truism and a falsity, as in Figure 1. One would be surprised if a logical structure with 2^N elements would collapse into a degenerate structure by itself.

Instead, when perceived from the context of information, to assign truth values to logical statements in a hypothesis space would require information additional to the hypothesis space itself. It is then not at all surprising that the logical relationships among statements do not reveal their truth values. From an information perspective, Gödel's Theorem appears to be a truism itself.

Embracing Uncertainty

Gödel's revelation exposed a fundamental flaw in mathematical reasoning and shook mathematics to its core resulting in the end of Russell and Whitehead's program of *Principia Mathematica*. But that isn't all. Gödel showed that in mathematics there are fundamental uncertainties, information that we have no access to. In some ways, this is reminiscent of the problem of *information isolation* seen in quantum mechanics (Schumacher & Westmoreland 2010) where the phase of a target is inaccessible. When interacting, the target and the detector belong to the same system. Both the state of the target and the state of the detector are unknown (two unknowns), yet only one piece of information can be learned from the interaction. That final piece of information, the absolute phase of the target, necessarily eludes us. Quantum Mechanics accepts and accommodates this fact. To make progress, one must accept, embrace, and work with uncertainty.

Going back to Undecidability, while the truth of some propositions cannot be proven through deductive reasoning, it is still possible to make *some* progress with inductive reasoning using probability theory. This is what is done in the Physical Sciences. Cannot Mathematics embrace uncertainty and employ inductive reasoning when deductive reasoning fails? Could that possibly work?

We have evidence that it could work. And that evidence is my not-so-silly question "Why addition?". It appears that no one knew the answer to that question, and more importantly, few people thought to ask it. It is precisely because the experimental evidence was so overwhelming that acceptance was automatic—proof seemed unnecessary. Mathematicians had such confidence that in measure theory, additivity was taken as a postulate. Of course, obtaining a proof is better, mainly because it leads to additional insights, such as Probability Theory and the Feynman Rules.

When deduction is not possible, inductive inference can serve in its stead and form a different kind of base. What progress could be made in Mathematics if uncertainty was embraced and inductive inference employed to its maximum potential? One could hypothesize mathematical relations and verify them to some degree through both constraints and data (results of experiments). One cannot help but envision a daring New Branch of Mathematics, which could then join the other Natural Sciences in their careful treatment of information in the quest for knowledge.
But could the Mathematicians ever accept it?

References

Bailey, D. H.; Borwein, P. B.; and Plouffe, S. "On the Rapid Computation of Various Polylogarithmic Constants." Math. Comput. 66, 903-913, 1997.

Hamming, R.W., 1980. The unreasonable effectiveness of mathematics. The Amer. Math. Monthly, 87(2), pp.81-90.

Jaynes, E.T., 1990. Probability theory as logic. In Maximum Entropy and Bayesian Methods (pp. 1-16). Springer, Dordrecht.

Jaynes, E.T., 2003. *Probability theory: The logic of science*. Cambridge Univ. Press.

Knuth K.H. 2005. Lattice duality: The origin of probability and entropy. *Neurocomputing*. 67C: 245-274. doi: 10.1016/j.neucom.2004.11.039

Knuth K.H. 2009. Measuring on lattices. P. Goggans, C.-Y. Chan (eds.), AIP Conference Proceedings 1193, American Institute of Physics, Melville NY, 132-144. arXiv:0909.3684 [math.GM]

Knuth K.H. 2015. The deeper roles of mathematics in physical laws, FQXi 2015 Essay Contest Entry (Third Place Prize): "Trick or Truth: The Mysterious Connection between Physics and Mathematics". http://fqxi.org/community/forum/topic/2382

Knuth K.H. 2018. Lattices and their consistent quantification. *Annalen der Physik* (Invited Submission): 1700370. doi: 10.1002/andp.201700370

Knuth, K.H., Skilling, J. 2012. Foundations of Inference. *Axioms* 1:38-73. doi:10.3390/axioms1010038, arXiv:1008.4831 [math.PR]

Pólya, G., 1954. Mathematics and plausible reasoning: Induction and analogy in mathematics (Vol. 1). Princeton Univ. Press.

Polya, G., 1990. Mathematics and Plausible Reasoning: Patterns of plausible inference (Vol. 2). Princeton Univ. Press.

Schumacher, B. and Westmoreland, M., 2010. Quantum processes systems, and information. Cambridge Univ. Press.

Skilling, J., 2006. Nested sampling for general Bayesian computation. Bayesian Analysis, 1(4), pp.833-859.

Skilling, J., Knuth, K.H. 2018. The symmetrical foundation of measure, probability and quantum theories. Annalen der Physik, 1800057. https://doi.org/10.1002/andp.201800057. https://arxiv.org/abs/1712.09725

Wigner, E.P., 1990. The unreasonable effectiveness of mathematics in the natural sciences. In Mathematics and Science (pp. 291-306).