“The Unreasonable Effectiveness of Mathematics”: Patterning of Patterns

Serge Dolgikh ^[0000-0001-5929-8954]

Dept. of Information Technology,
National Aviation University

Abstract. In this essay we continue the analysis of effectiveness of mathematics in description of natural phenomena. As was argued previously, the formation of foundations of mathematical thinking, including arithmetic and geometry can be determined to a large extent by the environment and the means of perception and processing of the sensory information about it. It is further argued that mathematics is an effective framework of concepts and logical instruments that allows to package patterns observed in the natural domains into compact logical structures: theories or “code” defined in the context of the logical model of mathematics. Predictive power of theories formulated within this framework is discussed and connections with epistemology principles established.

Keywords: Mathematics, Semantics, Epistemology, Intelligent Systems.

1 Introduction

In this essay the analysis of the underlying reasons of effectiveness of mathematics in description of natural phenomena is continued from the viewpoint of the “patterning of patterns”, interpreting mathematics as a framework of concepts and logical instruments allowing to “pack” patterns in the observed natural domains into compact logical structures, such as hypotheses, theories, conjectures and so on. This approach can be seen as related to the “model-code” paradigm [2] where logical structures expressed in the syntax and grammar rules of mathematics are considered as code whereas axioms, assumptions, principles, definitions and other structures of mathematical domains represent the model.

In the period since the question was raised, significant progress has been made in understanding the relations between models of description and observed patterns. We know for example, that shorter, more compact theories are more likely statistically to explain empirical experience [3]. However, statistical arguments may not provide all answers with respect to this particular theory (code) in this specific reality, for example, why a “longer” effective theory has been very successful in describing and predicting patterns, while a shorter one

“The Unreasonable Effectiveness of Mathematics in the Natural Sciences”, E. P. Wigner, [1]

was failed in experiment. Could there be more to the beauty statement and principle than probabilistic evaluation of possible outcomes? To approach these questions, we begin with an analysis of mathematics as an instrument and framework of organization of patterns. Immediately, one can observe that the framework of mathematics offers a number of significant advantages in description of natural phenomena including the following:

1. Non-ambiguity: a theory (code) expressed correctly in the syntax cannot produce logical contradictions.
2. Descriptiveness: a theory (code) describes all events in the area of applicability, with a possibility to determine an expected pattern, or observation (prediction) for any phenomena in the domain described by the theory.
3. Determinism: repetition of correct code produces compatible, non-contradictory result.
4. Completeness: all logical consequences of the code can be deduced by following the rules of syntax and operations.
5. Selectivity: essentially different patterns can be described by different code.
6. Expression and reduction of redundancy: capacity to produce compact code with the same scope of description of the natural domain.

From the perspective of mathematics as a logical framework to organize observable patterns, several questions arise naturally: is the effectiveness of mathematics subjective or objective? Is it natural (i.e. a property of the natural world) or perceptual (a characteristic, behavior or a method of the observer)? among others that will be considered in this work.

2 Standard Arguments: Beauty in the Eye and Occam's Razor

One possible explanation of the success of mathematical description of natural phenomena discussed in [1] relates to examples of theories with compact and clean formulation and effective predictive power (Newton's dynamics, Maxwell's electrodynamics; Einstein's General Relativity and other). It suggests that we may notice and select such examples as evidence of a connection between the "beauty" of the theory described as compact formulation in the model of mathematics and its predictive power, perhaps not taking into account other cases where "beautiful" formulations did not have such power or failed experimental verification.

It is known that selection bias or "look elsewhere effect" can produce incorrect conclusions about relationships between statistical variables, so unless some quantitative framework can be developed for the analysis, selected instances by themselves could not prove the existence of such correlation. However, it is clear that while selection bias is possible, examples of theories with both clear and precise mathematical formulation and astonishingly effective

and broad power of prediction do exist. And so, whereas these instances could not prove that beauty equals or correlates with predictive power, they still call for an explanation of causes of observed effectiveness of mathematics in description of the natural world.

Another standard argument, Bayesian inference and its formal extensions [3] has been briefly discussed earlier. It provides strong statistical basis for the expectation of effectiveness of compact, smaller theories, and has been discussed at length in the literature.

3 Patterning of Patterns

“Patterning of patterns” means a framework of logical rules, procedures and conventions aimed at organizing patterns observed in the natural world, including but not limited to: enumeration; packaging; description of time and spatial environment; and symmetry. It allows to express patterns and their relations observed in the natural world or specific domain thereof as compact logical structures satisfying the conventions and rules defined by the model i.e. mathematics, and producing, based on its syntax, grammar and additional structures such as definitions, assumptions, postulates, etc. a theory, “code” or meta-pattern.

A priori, one has no strong reasons to expect that such framework of patterning would be 1) universal or 2) unique. Indeed, the progress of mathematics and natural science has seen on more a few occasions, fundamental changes to descriptive frameworks that were considered unique and universal, while simple logical experiments provide arguments [4] that foundations of descriptive frameworks may trace to fundamental relations of the observer with the environment that can be established long before creation of a formal framework is attempted.

These arguments suggest that mathematics as a descriptive patterning framework may not be unique, universal or “natural”, in the sense that it may not exist as an objective entity in the natural world but rather emerges from the ways and patterns in which we interact with it, perceive it and describe it.

This conclusion is in a way confirmed by the uneven, patchy character of progress in the discipline; some areas perhaps stimulated by practical problems experience near-explosive growth; whereas a number of long-standing problems in many areas, for example, the number theory to this day have no solutions, and even possible directions to solutions may not always be clear. This character of development is compatible with the general pattern of progress of life that develops in the directions and ways that are beneficial to its needs rather than equally and uniformly in all directions. If, on the other hand, mathematics was an objective entity that emerges and exists by its own rules, perhaps more uniform progress could be expected, determined by the effort invested in investigation, not unlike exploration of natural areas?

4 Unreasonable Connections

Observations have been made about unexpected and in some instances, little short of amazing connections between different areas of mathematics. The subject deserves a dedicated study but perhaps without many words it can be illustrated by the famous Euler's identity:

$$e^{i\pi} = -1 \quad (1)$$

Even recently such far-reaching connections were made between problems in number theory and algebraic geometry [5] and intrinsically related algebraic structures emerging in different areas of mathematics.

Attempting a rational explanation of these and similar examples short of attributing them to unnatural and unreasonable qualities, one could recall apparent parallels in biology, where unexpected similarities and connections have been found in organs and functions or entirely different and unrelated organisms. Commonly, an explanation of such instances is attempted via one of the two arguments:

1. Common ancestry.
2. Similarity in the important characteristics of the environment, with some commonality in the underlying structure, causing emergence of similar functions.

Within the patterning of patterns paradigm similar arguments can be made in explanation of unexpected connections in mathematics. Indeed, having roots in the basic patterns of enumeration, packaging and spatial modeling, evolving fields of mathematics, like organisms may not realize or forget, after long periods of isolated development of the common ancestry in the patterns that lie in the foundations and are shared by some or many or most fields.

This viewpoint can be illustrated on the example above (1) that provides as expressive as succinct connection between several areas of basic mathematics: exponential and logarithmical functions and series; trigonometry; complex numbers and whole / integer numbers.

One can begin with infinite series that provide enormously effective tool for storing and calculation of patterns of sums of terms of similar structure. Next, one can notice that two basic types of series, exponential and trigonometrical can be considered as same (i.e. combined into a single pattern) by extending the concept of number from scalar to vector form, a complex number. Finally, geometric interpretation of trigonometric functions provides a direct connection to whole numbers. Thus, combination of a small number of basic patterning concepts can produce very expressive, compact and powerful code.

As a more complex and recent example, let us consider the connections that were discovered around the turn of the 20th century between algebraic geometry and number theory.

It begins with algebraic methods in geometry developed by Decart that produced immense improvement in descriptiveness of the mathematical code in geometry. Where long and precarious proofs by construction were needed to demonstrate that two curves e.g. parabolas belonged to the same class, algebraic notation allowed to express it in a single line of code. Algebraization also led to direct connections between geometry and certain problems in algebra and number theory for example, Diophantine equations.

Next, in the broad large field of algebraic equations we found a subset that 1) had relatively simple formulation, such as polynomial Diophantine equations and 2) possessed symmetry structure, such as elliptic curves. At the same time, in another field (or submodel) of complex number analysis, functional structures with symmetry properties were introduced and studied, namely, modular forms.

With two types of symmetry patterns developed on the basis of more general patterns only a small number of logical options is possible:

1. The patterns are related: one implies the other (implication), or each one implies the other (equivalence). Essentially it means that patterns relate to the same logical entity described with different submodels, elliptic curves; modular forms. This is a case of stratification of the model, allocation of submodels to related and equivalent classes.
2. The patterns are independent. It means that detalization of the existing submodels can produce further distinct submodels, such as, in the previous example, elliptic curves associated with modular forms; and perhaps, an entirely different type of curves that cannot have associated modular forms. This case is of extension of the model.
3. The last case can be seen as a combination of two previous cases, patterns can be partially related where certain specific subtypes of each type can be related while others, not. This is a case of detalization with extension of the model.

Note that given that all patterns were developed on the same foundation set of patterns, none of the possibilities just described would appear to be astounding. As it stands, the first option was proven in the case of elliptic curves and modular forms as the modular theorem [5]; as a consequence, any pattern p or its derivative under the model (i.e. the implication tree of the pattern $I(p)$) that contradicted that result could not be valid under the model. That was shown to be the case with the counterexample to the Last Fermat Theorem, finally proving it.

To summarize, a formally proven relation between submodels based on enumeration and symmetry eliminated a set of incompatible patterns, including in one case, a counter-example to a well-known long-standing conjecture. While it still inspires awe of the delicacy and complexity of patterns and their relationships, the need for out-of-this-world, unnatural quality needed to understand and explain such examples diminishes in the perspective of patterning of patterns. As living organisms share basic foundations such as DNA producing adaptations of unlimited complexity at times with amazing patterns of similarity between ostensibly unrelated species, fields and areas of mathematics are rooted in the same general patterns of enumeration, packing, spatial sense and symmetry, producing amazing similarities and connections in the logical structures it creates.

5 Predictive Power and Epistemological Principle

Over the decades and centuries, some theories have demonstrated astounding power to predict natural phenomena that were not observed previously. The well-known examples are Newton's dynamics and discoveries of new planets; Maxwell's theory and discovery of electro-magnetic waves followed by technological revolution in communications; quantum mechanics to name only a few of well-known examples.

5.1 Beauty and Minimality

Clearly, these examples go beyond just packaging of already established patterns and relations into compact logical structures; what makes a planet appear in the point where its trajectory was predicted by an equation of dynamics; why electron invariably splits between two possible trajectories always satisfying the pattern of probabilities? All such examples appear to suggest that the Universe itself is governed by and must obey laws expressed in the language of mathematics.

For illustration, let us consider as example evolution of Solar system astronomy from prehistoric to modern times. It can be divided into the following periods:

1. Symbolic (marks, tables): before 10,000 BC
2. Geocentric model: 200 A.D.
3. Heliocentric model: c. 1600
4. Gravitational dynamics: 1687
5. General relativity: 1915

First, one can notice the difference between descriptive models, that summarize and organize previous observations, or patterns, from inductive ones that

describe any pattern within the domain of the theory, regardless of whether it was observed or not. Descriptive theories deal only with a specific set of patterns and cannot be extended to include a new pattern without modification of the theory (code) whereas inductive ones describe any pattern that belongs to the domain of the theory. For example, to describe the trajectory of a new object such as planet, asteroid or comet, descriptive models in p.p. 1-3 above would need to create a new chapter of code, a table, symbolic map, or an elliptic model in pre-Newton heliocentric astronomy.

The distinction between descriptive and inductive models is closely related to predictive power of theory. Whereas descriptive models literally describe only already seen and established patterns, inductive ones describe a domain that can and commonly includes both known patterns and those that have not yet been observed. At this point one can notice that the distinction between observed and predicted patterns can be being accidental, relative to the state of the observer, whereas the complete set of implications, implication domain of the code can be a more objective factor.

One can also define the factors of the size, and minimality of the code. Between codes describing the same domain, the minimal one can be defined as such that no smaller code exists that describes the entire domain, i.e. every pattern that belongs to it. The size of code requires more rigorous definition, perhaps related to algorithmic complexity [6] and will be considered elsewhere, but for now we will use intuitive interpretation as “compactness” of the theory, i.e. its capacity to limit key assumptions, postulates and independent parameters and derive implications via the syntax and grammar of the model.

One can conjecture that for any finite domain, once defined, the minimal code must exist, that can be proven by the standard arguments of finiteness of the code and reduction. Does minimal code have to be unique, down to isomorphism of patterns? That can be an interesting topic for a further discussion in epistemology.

Based on the characteristics of inductive power and minimality in classification of possible theories, the cases of “beautiful” theories belong to the class of minimal code with very high inductive power as illustrated below:

Criteria	Minimality: No	Minimal code
Inductive power: Low	Descriptive theory	Tiny domain (less interesting)
High	Effective theory	Beautiful theory (minimal code with high inductive power)

Let us recall that up until very recently in the development of science very large amounts of empirical data was available in many scientific domains for verification of theories, not only as formal scientific evidence, but also everyday experience and observations. If one considers a set of minimal theories with high inductive power, this observation entails a strong selection on the basis of known experience, the empirical sieve through which the set of minimal theories is filtered. We are not interested in beautiful theories that obviously contradict empirical observations or experience, they fall into the domain of fiction. Consequently, minimal theories and logical and empirical foundations they are based on already passed strong empirical selection.

Consider for an example a thought experiment: suppose what following formulation of Maxwell's electrodynamics theory, experiments failed to confirm the existence of electromagnetic waves. Such an experiment would have been an upset, not an extension of a beautiful (minimal code with very high inductive power) theory, something that does not happen very often. To that one could counter that a world without electromagnetic waves would have been very different to begin with, and intelligent observers in it perhaps created very different mathematical model (for example, many of basic geometrical patterns are directly related to empirical experience of direct vision) in which the equivalent of Maxwell's theory would have been formulated very differently.

The logical conclusion from this experiment is that in empirical science with a large domain of verified empirical implications, an upset of the minimal code could result in an essential change of the model; or, alternatively, formulation of a theory in the syntax of specific model (mathematics) implicitly carries an assumption of validity of a number of basic patterns and rules that define the model that has been verified over a large body of empirical evidence by passing through the empirical sieve. The latter then performs two essential functions: selection or filtering of code that is relevant to the observed reality; and secondly, self-consistency of the model in which the code describing observations is expressed.

5.2 Epistemological Principle

The observations in the previous sections elicit a question: is the epistemological principle of empirical science discussed in [1] equivalent to, and possibly, almost tautological to the drive for minimal code? Let us recall that in addition to perceived beauty, minimality has essential practical value to an intelligent observer as an effective instrument of organization and communication of information about observed patterns by reducing volume or bandwidth of information that needs to be transmitted in the context of the shared model in which the code is expressed. Clearly, there would be a natural benefit in such a quest.

To attempt to answer it, let us consider the earlier experiment in Section 5.1 under further assumption of self-consistency of the model i.e. that an upset of the minimal code C_i describing certain implication domain X by an observation δ does not cause an essential change of the model in which the code is expressed. Let C_j be the minimal code for the domain $X' = X + \delta$, that is presumed to exist. For a “beautiful” theory, the implication domain of the code C_i , X can be presumed very large, $X \gg \delta$.

Further, we will require that the pattern δ is a regular pattern in the code C_i as opposed to the case of implicit extension of the domain to a new area where certain assumptions may or may not hold true. The examples of the latter case are the “ultraviolet catastrophe” in classical thermodynamics caused by implicit assumption of applicability of classical description to microscopic processes or the assumption of possibility of infinite speed in classical dynamics. In these cases of domain extension, the code remains valid within certain implication domain, whereas examples of the former, such as the electromagnetic waves experiment, or a hypothetical discovery of an asteroid of special mass violating the law of gravity would invalidate the entire theory (code) as failing empirical test.

Under these assumptions, two logical possibilities exist: 1) $C_j \gg C_i$ or 2) $C_j \sim C_i$. In the first case, the minimal code of X' is significantly larger than that of X , whereas in the second case they are similar (see the earlier note on quantification of code). Because δ is a regular pattern, the first case would result in a scenario where the code, itself a pattern of patterns would change significantly as a result of a single regular observation i.e. instability of observed patterns. As can be seen, the second case would imply the existence of multiple logically independent near-minimal descriptions of the same domain.

Thus, a counter-example to the principle “Nature avoids upsetting beautiful theories (that is, minimal code with great inductive power)” would require violation of one of the assumptions:

1. Stability of observed patterns
2. Uniqueness of minimal code.

The former can be seen as generalization of the *principle of reproducibility*, fundamental for the empirical method in science. Indeed, the minimal code can be considered a pattern itself, a meta-pattern, with regular patterns within the implication domain of the code constituting observations. Failing minimal code theory on a regular observation would not be unlike obtaining wrong result after multiple successful runs of an experiment.

Consequently, it can be argued that in the framework of patterning of patterns, the epistemological principle of [1] would be equivalent to the postulates of stability of observations; and uniqueness of the minimal code.

In conclusion of this section a brief comment on the appearance of a resemblance of the principle of self-consistency of the model mentioned earlier in this section, and the *anthropic principle* [7]. It does not extend beyond the appearance though because self-consistency relates to the model as a framework of organizing patterns and observations by an intelligent observer, and has no implications or claims to explanation, rationalization or selection of the observed reality itself.

5.3 Post-Empirical Science

As discussed earlier, the conclusions of the previous sections relate to theories, or code, with very large verified area of implication domain, both scientifically and in regular experience. While minimal code theories were very successful in a number of classical domains, in some areas of contemporary science this assumption or expectation may no longer be the case. Well-known examples of such post-empirical theories can be found in the domain of quantum gravity [8], a long-standing effort to unite the successful theories of matter and gravity, essentially, a quest for the minimal code of the implication domain that contains both domains of the theory of matter (QFT) and the theory of gravity (General relativity). If D_m, D_g are the respective implication domains of each theory, then quantum matter-gravity theory C_{mg} has to be the minimal code of a larger implication domain D_{mg} that contains both D_m, D_g :

$$C_{mg} = \text{Min}(D_{mg}); \quad D_m, D_g \subset D_{mg}$$

The problem here is that while there is sufficient implication area for empirical verification of partial codes D_m, D_g , with the effective code of the current status quo (though perhaps not the minimal one) $C_e = C_m \cup C_g$, the relationship between the partial domains and the implication domain of the resulting minimal code theory D_{mg} is not known:

$$D_{mg} - (D_m \cup D_g) = ?$$

Then, the problem of finding the minimal code describing D_{mg} can become uncertain as illustrated below as there is not a well-defined sufficient domain for the minimization problem to be defined on.

Fig. 1 Minimal code challenge in post-empirical science.

In this scenario, assuming additional factors such as “beauty” with any of possible interpretations in selection of candidate theories can be, in our view, unwarranted and risky. It may divert attention to some lines and directions of research in favor of the others, perhaps novel and “orthogonal” ones and delay progress in the field. In our view, the attitude of the Renaissance age researchers can be more productive: observe, investigate, measure and describe with an open mind in the hope of finding and establishing patterns that with time would lead to formation of new minimal and beautiful theories.

5.4 The Catalogue Experiment

Empirical verification of the connection between “beauty” i.e. minimality of a theory with its predictive power can be attempted with a project in cataloguing known scientific theories by the factors of minimality and inductive power. Quantitative methods would need to be developed to quantify the size of the code, i.e. formal quantification of the description of the theory that can be an interesting and fruitful undertaking in its own right.

For the resulting minimal and near theories with high inductive power the postulates of the empirical scientific method can be verified. Qualitative examples of entries in the catalogue, without attempt at completeness are shown below.

The data of the catalogue project can be used to test and verify the postulates of the epistemological principle associated with minimality of the code, as discussed earlier in Section 5.2.

Table 1. Scientific code catalogue.

Theory	Minimal?	Inductive power	Implication domain
Aristotle-Ptolemy astronomy	No	Low	Solar system astronomy
Heliocentric astronomy	No	Low	Solar system astronomy
Gravitational dynamics	Yes	High	Classical dynamics, astronomy
General relativity	Yes	High	Classical and relativistic dynamics, astronomy, cosmology
Classical electromagnetism	Yes	High	Electric and magnetic fields
Quantum mechanics	Yes	High	Quantum matter and dynamics
Classical thermodynamics	Yes	High	Macroscopic heat and energy exchange
Standard model	Effective theory?	High	Relativistic quantum theory of matter and interactions (with QFT)
Supersymmetry	Post-empirical theory	Low ¹	Quantum theory of matter and interactions
String / brane theory	Post-empirical theory	Low ¹	Quantum theory of matter, interactions and gravity up to Planck scale
Physical chemistry	No	Medium	Macroscopic chemical processes
Organic chemistry	No	Medium	Organic chemical processes
Evolutionary biology	No	Low	Living systems and organisms

¹ See the discussion of post-empirical science in Section 5.3.

6 Conclusion

In this work we investigated connections between perceived beauty of theories of the natural world expressed in the language and model of mathematics with the concepts of patterning of patterns, minimality and predictive power of natural theories.

The effectiveness of mathematics in description of the natural world can be related to its ability to create compact theories or “code” with very large domain of applicability; this ability to effectively “package” everyday experience and observations into compact conceptual structures appears intrinsically attractive to us not in the least due to advantages it provides in sharing and communicating information about the observed environment.

From the perspective of patterning, the noted apparent preference of the Nature for beautiful theories can be associated with epistemological postulates of stability of observed patterns and uniqueness of minimal code. The next

question that can be asked is, do they apply to the Nature itself or to our perception of it, the ways and instruments we use, sometimes subconsciously and instinctively to observe and organize the observed patterns? While such questions are commonly relegated to the domain of philosophy, it would be both exciting and challenging to explore the possibilities of their empirical application and verification.

References

1. Wigner, E. P.: "The unreasonable effectiveness of mathematics in the natural sciences", *Communications on Pure and Applied Mathematics*, vol. 13, pp. 1–14, 1960.
2. Rissanen, J.: "Modeling by shortest data description". *Automatica*, vol. 14 (5), pp. 465–471, 1978.
3. Solomonoff, R.J.: "A formal theory of inductive inference part I", *Information and Control*, vol. 7 (1), pp. 1–22, 1964; "A formal theory of inductive inference part II", *Information and Control*, vol. 7 (2), pp. 224–254, 1964.
4. Dolgikh, S.: "The unreasonable effectiveness of mathematics": a universal language of contact? *Cambridge Open Engage*. doi:10.33774/coe-2021-t64xx, 2021.
5. Zagier, D.: "Modular points, modular curves, modular surfaces and modular forms", *Lecture Notes in Mathematics*, 1111, Springer, 1985.
6. Kolmogorov, A.N.: "Three approaches to the quantitative definition of information", *Problems in Information Transmission*, vol. 1 (1), pp. 1–7, 1965.
7. Barrow, J. D., Tipler, F. J.: *The Anthropic Cosmological Principle* (1st ed.). Oxford University Press, 1986.
8. Kiefer, C.: "Quantum Gravity: general introduction and recent developments", *Annalen der Physik*, vol. 15 (1), pp. 129–148, 2005.