Mathematics Challenges Originalism

Karl Manheim[1]

The 20th Century saw the development of revolutionary theories in physics and mathematics. These upended the classical view of reality that existed at the time the Constitution was written and adopted. Because the world as we now understand it is vastly different than that of the framers', we should be cautious in our interpretations and borrowings of 18th Century concepts, especially where they purport to bind current and future generations. Backward-looking ontologies for constitutional interpretation cannot stand in light of these modern developments.

This is especially true for originalism, a theory that purports to discover the original public meaning or communicative content of the 1787 Constitution. Mathematical analysis of constitutional text casts doubt on the claim that there ever was a received public meaning or intelligible communication. Rather, originalism simply tells a story of the founding, one that does not hold up to mathematical scrutiny.

This article tests the validity of originalism against recent developments in mathematics: 1) information and communication theory, 2) complexity theory, 3) statistical modeling of law, and 4) Kurt Gödel's incompleteness theorems. These theories provide us with insights that challenge the foundations of originalism.

The main takeaways from a mathematical analysis of the Constitution are: 1) the document inefficiently communicates meaning from drafter to reader; 2) some provisions are less comprehensible than others; 3) complexity analysis provides insights into meaning that rebut semantic interpretations; and 4) as a matter of formal logic, the Constitution is incapable by itself of resolving constitutional questions, thus necessitating resort to external factors and values.[2]

---

[1] Professor of Law (emeritus), Loyola Law School, Los Angeles. I am grateful for the many valuable comments I received during a Constitution Day 2020 workshop and from colleagues Jeffery Atik, Cornelia Dean, [others] and my research assistants, Alexis Ashjian and Skye Oyama.

[2] A forthcoming companion work similarly establishes that originalism cannot withstand scrutiny under modern understandings of reality as revealed by quantum mechanics and Einstein's theories of relativity.

I. Introduction

Mathematics is the language of the universe. That is clearly true for physical objects, space and time.[1] Is it also true for conceptual structures such as law or other intangible systems that are the product of human design? That might be a hard or metaphysical question if those structures were not subject to physical laws. But they are, and not simply because the forces that create those structures are themselves physical objects. Rather, ideas, models, philosophies, ontologies, and other "intangible" constructs of the mind are embodiments of information. As such, they do not exist in the absence of communication. Information can only be contained within physical objects, such as text on a page. Communication is a physical process subject to all the laws of physics and mathematics. Without information and communication, we cannot form concepts or create law. No part of reality, whether of natural or human design, is immune to the laws of physics and mathematics. Even pure logic

[1] Swedish physicist Max Tegmark takes the concept further, claiming "our physical world not only is described by mathematics, but that *it is* mathematics: a mathematical structure, to be precise." Max Tegmark, OUR MATHEMATICAL UNIVERSE, 6 (Knopf, 2014) (emphasis added).

– the original philosophy of science – is, at its core, an exercise in mathematics.[2]

This article offers a mathematical analysis of the Constitution. While perhaps a novel way to explore legal domains, everything, including textual analysis, is affected by mathematics.[3] One could hardly argue that the Constitution, written during the heyday of "natural law" theories, is exempt from such analysis.[4] Indeed, every legal discipline is subject to mathematical analysis, as when lawyers apply balancing tests and multifactor rules. Even fundamental rights jurisprudence requires Boolean operators to resolve equivalences, disparities, identicalities, and comparisons. One need only consider burdens of proof, degrees of scrutiny, standing, voting and representation, due process, and equal protection doctrines to see that avoiding mathematical analysis of legal rules and theories is impossible.[5] This article applies that proposition to constitutional originalism and finds it cannot stand up against rigorous mathematical scrutiny.

The emergence of computers, data analytics, and Natural Language Processing in recent years has given us new tools with which to "read" law. Abstract concepts and cultural artifacts can be represented in digital form, thereby facilitating forms of analysis that were previously unavailable. Quantitative analysis and powerful statistical tools provide new insights into content, meaning, and processes of law. The use of these quantitative tools has been called "distant reading" of text, and while it does not replace traditional "close reading" or analytical parsing of legal sources, it provides additional tools by which to extract meaning from text. It also gains importance as law becomes more interdisciplinary. As Holmes put it in 1897: "For the rational study of the law the blackletter man may

---

[2] Terminal degrees in mathematics and the sciences still get the designation *Doctor Philosophiae*, Ph.D or DPhil. The highest law degree is a JSD or SJD, *Scientiae Juridicae Doctor*. At least in common law systems, law is considered a science.

[3] *See* Michael I. Meyerson, THE LAW AND LARGE NUMBERS: POLITICAL NUMERACY: MATHEMATICAL PERSPECTIVES ON OUR CHAOTIC CONSTITUTION (W. W. Norton. 2002).

[4] *Id*. at 28 (the founders were influenced by mathematics' structure of formal deductive proof). *See also* Laurence H. Tribe, *The Curvature Of Constitutional Space: What Lawyers Can Learn From Modern Physics*, 103 Harv. L. Rev. 1, 20 (1989) (recounting early descriptions of the Constitution as "Newtonian in design").

[5] POLITICAL NUMERACY, *supra*, n.3, ch. 2-4.

be of the present, but the man of the future is the man of statistics and the master of economics."[6]

Distant reading can help with interpretation of legal texts as it allows for more robust use of the vast databases of sources and usages. This is the emerging field of "corpus linguistics," which draws from statistical analysis of text. Since the same word or phrase is often used in multiple ways (consider how dictionaries provide multiple definitions and historical usages of a term, as well as how certain words can be used as various parts of speech), quantitative analysis may be an indispensable tool in illuminating semantic meaning. Subtle changes or differences in usage by courts and society, in general, may go undetected with visual inspection but become obvious with computer analysis of large data sets. Livermore and Rockmore thus recommend the use of quantitative analysis in the assemblage of Restatements to complement existing qualitative analyses.[7] Empirical legal analysis already does this, but doctrinal analysis could benefit as well.

I posit that mathematics provides a useful analytical tool, complementing other theories of analysis, for the interpretation of the Constitution and the meaning of constitutional law. The text of the Constitution contains several mathematical concepts, as seen in apportionment, age requirements, equality principles, and structural ratios such as majority and supermajority rules. However, that sort of mathematical analysis is relatively trivial. Instead, I focus here on laws of information and communication, as well as the ability to derive meaning from constitutional text through computational, statistical and formal logical analysis.

The main takeaways from a mathematical view of the Constitution are 1) the document inefficiently and inadequately communicates meaning from drafter to reader, especially due to inevitable information loss; 2) some provisions are less comprehensible than others, prone to arbitrary interpretation; 3) complexity analysis provides insights into meaning that may rebut semantic interpretations made by scholars and courts; and 4) as a matter of formal logic, the Constitution is incapable by itself of resolving constitutional questions, thus necessitating resort to external factors. Each of these observations challenges the validity of backward-looking ontologies such as originalism that purport to be robust and credible theories of constitutional interpretation.

This article continues in Part II with a short review of originalist theory and its workings. Parts III through VI subject the Constitution to several

---

[6] Quoted in Michael A. Livermore and Daniel N. Rockmore, *Distant Reading the Law*, in LAW AS DATA: COMPUTATION, TEXT & THE FUTURE OF LEGAL ANALYTICS (Santa Fe Institute, 2019), at 4.

[7] *Id*. at 15.

major mathematical and logical theorems of the 20th Century, each demonstrating the impossibility of objective interpretation of constitutional text, or the inability to communicate its meaning over distant and noisy channels. If originalism cannot withstand mathematical rigor, it should not be relied upon for constitutional outcomes.

I concede an uphill battle; originalism is in vogue and an unorthodox approach, such as this mathematical accounting, is unlikely to displace it. Nonetheless, now that the tools and theories are available, such a critique should be mounted. By referring every constitutional question to a derived 1787 meaning of the text, originalism has become weaponized by its proponents to resist political, economic, cultural, and technological change.

The supposed exclusive route for maintaining constitutional relevancy is the amendment process of Article V.[8] However, that formalism instead reinforces anachronistic servings of constitutional law. Consider that we've had only 16 amendments (some simply overruling errant Supreme Court decisions) since the first Congress, or roughly one every 15 years of the Republic, and none since 1972, when the last proposed constitutional amendment was adopted.[9] The wheels of progress have accelerated in every venue except when it comes to the Constitution keeping pace with revolutionary upheavals in economics, transportation, communications, science, technology, and nearly everything else that defines modern society.

One such upheaval is in mathematics, with developments in the field casting doubt on the viability of originalism as a school of constitutional theory. While the tools described here were not available to the framers, they are to modern constitutional theorists. Their application should free us and allow us to focus on different theories of interpretation, rather than arguing with the past.

---

[8] U.S. Constitution, Article V reads: "The Congress, whenever two thirds of both houses shall deem it necessary, shall propose amendments to this Constitution, or, on the application of the legislatures of two thirds of the several states, shall call a convention for proposing amendments, which, in either case, shall be valid to all intents and purposes, as part of this Constitution, when ratified by the legislatures of three fourths of the several states, or by conventions in three fourths thereof, as the one or the other mode of ratification may be proposed by the Congress; provided that no amendment which may be made prior to the year one thousand eight hundred and eight shall in any manner affect the first and fourth clauses in the ninth section of the first article; and that no state, without its consent, shall be deprived of its equal suffrage in the Senate."

[9] Discounting the 27th Amendment; although ratified in 1992, it was proposed in 1791 as part of the original Bill of Rights.

II. The Pillars of Originalism

As a theory of constitutional interpretation, originalism has been wildly influential in recent years. It has been promoted by many conservative academics and adopted by like-minded Supreme Court justices. Given the recent and current composition of the Court, originalist interpretation has proven dispositive in many cases. This is most evident in opinions by Justices Scalia and Thomas, two of the most ardent supporters of originalism. But originalism is so ingrained in our constitutional discourse that politicians and lawmakers of all stripes invoke it when convenient to support their arguments. During congressional hearings on the impeachment of President Donald Trump in Fall, 2019, both his detractors and defenders, as well as their supporting witnesses, resurrected our long-dead framers to explain the meaning of controlling constitutional terms – "treason, bribery and other high crimes and misdemeanors."[10] As with most other constitutional terms, originalism enables one to derive competing and often incompatible definitions.

In this section I provide a general description of originalist theory, sufficient for application of the mathematical analyses that follow. More thorough treatments can be found in the works of my friends Lawrence Solum and Randy Barnett, among others, many of which are cited here. My analysis is not of their arguments *per se*, but of the constitutional text itself in terms of its ability to convey meaning, a prerequisite for originalism. Absent that, judges must consult external sources, including their own values, in determining the content of constitutional law.

Originalism has itself undergone change.[11] The theory was first propounded as one of *original intent* by Robert Bork in 1971,[12] who some have dubbed the "father of originalism."[13] *Intentionalism* theory en-

---

[10] U.S. Constitution, Article II, Section 4. *See also Trump v. Mazars USA, LLP*, 140 S. Ct. 2019, 2045-2046 (2020) (Thomas, J., dissenting) (framers intended impeachment to be the sole means for Congress to subpoena presidential records).

[11] *See* Lawrence B. Solum, *What is Originalism? The Evolution of Contemporary Originalist Theory*, in Grant Huscroft and Bradley W. Miller, eds, THE CHALLENGE OF ORIGINALISM: THEORIES OF CONSTITUTIONAL INTERPRETATION (Cambridge 2011). For a review of the derivation of originalism and its varieties, see *Fixation Thesis, supra. n. 19*; Gregory E. Maggs, *A Concise Guide To The Records Of The State Ratifying Conventions As A Source Of The Original Meaning Of The U.S. Constitution*, 2009 U. Ill. L. Rev. 457, 461-63 (2009).

[12] *See* Robert H. Bork, *Neutral Principles and Some First Amendment Problems*, 47 Ind L J 1 (1971).

[13] Steven G. Calabresi and Lauren Pope, *Judge Robert H. Bork and Constitutional Change: An Essay on Ollman v Evans*, 80 U. Chi. L. Rev. Dialogue 155,

deavors to discover what the framers intended by the words and phrases they used in creating the Constitution. But the difficulty in "summing" individual subjective intents in the committee-driven document led many originalists to abandon that theory. In its place, "new originalism"[14] or "contemporary originalism,"[15] has developed, focusing on the *original public meaning* or *communicative content* of the Constitution.[16] This approach is concerned more with the meaning that the text conveyed to the politically engaged population of the late 1780s rather than the specific intentions of the founders themselves. After all, although the framers *proposed*, it was the people (through the agency of state convention delegates) that *adopted* the Constitution.[17] It was their understanding that should count. Still, most contemporary Supreme Court opinions still use the framers' intent, "expectations" and drafting history as a basis for their originalist interpretations.[18]

The major pillars of originalism are "fixation" and "constraint."[19] The former posits that meaning is fixed at the time of formation by the document's *communicative content*.[20] The latter works to constrain judges and other constitutional actors to abide by that fixed meaning in their official acts. For originalists, the Constitution does not evolve over time, except through the formal amendment process of Article V. For example,

---

155 (2013). Modern originalists may claim that Bork was a "proto-originalist" with incompletely formed ideas for the theory. *Fixation Thesis, supra*, n. 32, at 3. The actual term "originalism" seems to have been adopted somewhat later, by Paul Brest in 1980. See Paul Brest, *The Misconceived Quest for Original Understanding*, 60 B.U. L. Rev. 204 (1980).

[14] Randy E. Barnett and Evan D. Bernick, *The Letter and the Spirit: A Unified Theory of Originalism*, 107 Georgetown L.J. 1, 4 (2018).

[15] *Contemporary Originalism, supra*, n.11; Lawrence B. Solum, *Originalism and Constitutional Construction*, 82 Fordham L. Rev. 453, 459 (2013).

[16] As Solum explains, "there is a family of originalist constitutional theories." *Evolution of Originalism*, supra, n.1111 at 6.

[17] Solum describes a mid-way point between original intent and original understanding; namely the original understanding of the state conventions that ratified the constitution. *What is Originalism* at 10. This version suffers from similar defects as original intent; i.e., the inability to find coherent intentions of group proceedings. *Id*. at 11.

[18] See, e.g., *Trump v. Mazars USA, LLP*, 140 S. Ct. 2019, 2045-2046 (2020) (Thomas, J., dissenting); *Ramos v. Louisiana*, 140 S. Ct. 1390, 1400-1401 (2020) (Gorsuch, J.); Rogers v. Grewal, 140 S. Ct. 1865, 1874 (2020) (Thomas, J., dissenting) (equating framers' and the public's understanding of terms).

[19] Lawrence B. Solum, *The Fixation Thesis: The Role of Historical Fact in Original Meaning*, 91 Notre Dame L. Rev, 1, 3-6 (2015).

[20] *Id*. at 15.

whether a particular form of government surveillance (say, tracking one's web searches) constitutes a "search and seizure" depends on whether that or analogous action was understood as requiring a warrant back in 1791 when the Fourth Amendment was adopted.[21] Originalism applies to government powers as well as to individual rights. Thus, whether Congress had power to enact the Patient Protection and Affordable Care Act of 2010,[22] popularly known as "Obamacare," relied on the 18th century meaning of the commerce power.[23]

Originalism is based on the ontology that the semantic meaning of constitutional terms was known and fixed at the founding, and that those meanings can be discovered and enforced today.[24] But how does one ascertain the *communicative conten*t of the Constitution as of 1787? A simple lexical reading of text is inadequate and misleading because, among other things, the meaning of words have evolved over the centuries. Semantic "drift" is to be avoided. Instead, this exercise in interpretation necessarily entails subsidiary questions:

Who is doing the reading? As defective as intentionalism was, at least it focused on a discrete and identifiable population, namely the drafters of the document.[25] But original public meaning indicates that it was the "public" whose reading of the document counts. Which public? The entire population of the country? Eligible voters (*i.e.,* white, Protestant, freeholding men)? Delegates to the state ratifying conventions?[26] Other opinion leaders? In *District of Columbia v. Heller*,[27] Justice Scalia said

---

[21] *See, e.g., United States v. Jones*, 565 U.S. 400, 406 (2012) (we must "assure preservation of that degree of privacy against government that existed when the Fourth Amendment was adopted").

[22] Pub. L. No. 111-148, 124 Stat. 119, as amended by the Health Care and Education Reconciliation Act of 2010, Pub. L. 111-152, 124 Stat. 1029.

[23] See *Nat'l Fed'n of Ind. Bus. v. Sebelius*, 567 U.S. 519, 554 (1912) (Congress can not use its commerce power to require individuals to buy health insurance because "that is not the country the Framers of our Constitution envisioned"). *See also* Randy E. Barnett, *The Original Meaning of the Commerce Clause*, 68 U. Chi. L. Rev 101 (2001) (original meaning of commerce was narrow).

[24] *See* André LeDuc, *The Ontological Foundations of the Debate over Originalism*, 7 Wash. U. Jur. Rev. 263 (2015); *Fixation Thesis, supra*, n.19, at 23.

[25] There are selection problems even here. Do we focus on the 39 delegates that signed the document, or the 55 who attended the federal convention, or the 74 who were selected, some of whom chose not to attend out of concern over the enterprise? *See* Meet the Framers of the Constitution, https://www.archives.gov/founding-docs/founding-fathers.

[26] *See, e.g., Contemporary Originalism, supra*, n.11, at 10.

[27] 554 U.S. 570 (2008)

we should focus on the voters,[28] but the document was addressed to "the People," using that term multiple times, and it was their rights that were at stake. This would be a distinction without a difference if each of these subpopulations received the same communicative content and had the same understanding of the document's meaning. Given the controversies over ratification and its effects, such an assumption is unlikely.[29] To avoid this dilemma, originalists first must construct a "public," even if only a small slice of the population,[30] whose linguistic understanding is both unified and controlling.[31]

Once we know whose reading of the text counts, we next have to figure out how the reading is done. Discovering a text's *communicative content* is a fine theory, but can be a difficult and inconclusive exercise. For it to work, originalists must engage in historical research to find "linguistic facts about patterns of usage at the time;" *i.e.,* how a word or phrase was used, with focus on then prevailing "patterns of syntax and grammar."[32] These, in turn, are informed by context, structure, precedent, prudence, and even the "`ethos' of the American social order" at the time.[33]

This is an intensive fact-based inquiry, using lexiconic and historical sources. Those typically include notes of the convention, writings, letters, dictionaries, other surviving documents, as well as the structure of the text itself. One needs to see originalism in action to understand how the communicative content is actually identified.

---

[28] *Id*. at 576.

[29] Not only did several states reject the constitution on the first try, *see* http://teachingamericanhistory.org/ratification/overview, there was violence and claims of voter suppression and fraud in others. *See* THE DOCUMENTARY HISTORY OF THE RATIFICATION OF THE CONSTITUTION, Vol. III at 92.

[30] *See* Charles Beard, AN ECONOMIC INTERPRETATION OF THE CONSTITUTION, 239-252 (1939) (less than one-fifth of the 1787 population possessed the right of suffrage).

[31] McGinnis and Rappaport contend that any linguistic differences among 18th Century Americans are mostly irrelevant. To them, the constitution was not written in ordinary language, but in "legal language." JOHN O. MCGINNIS AND MICHAEL B. RAPPAPORT, ORIGINALISM AND THE GOOD CONSTITUTION 4-6 (Harvard Univ. Press 2013), at 117. While many of the framers were lawyers, most were not. Among them were: Roger Sherman, Elbridge Gerry, Benjamin Franklin, James Madison, George Mason, and George Washington. And some lawyers, such as Patrick Henry, were appointed as delegates but declined to attend the convention.

[32] *Id*. at 24-25.

[33] *Fixation Thesis*, *supra*, n.19 at 8, 18; *see also* Phillip Bobbitt, CONSTITUTIONAL INTERPRETATION 12–13 (1991).

A prime example of originalist mechanics is found in *Heller*,[34] the case that found a right of individual gun ownership in the Second Amendment. In his majority opinion, Justice Scalia relied on certain dictionary definitions and state constitutions of the mid- to late-Eighteenth Century where such a right was stated or implied. The dissents took their clues from other dictionaries and the lack of such a right in other state constitutions.[35] How could smart justices *qua* historians reach diametrically opposite conclusions? For one, in doing historical research there is "an unavoidable temptation to look over a crowd and pick out your friends."[36] It is curious that Justice Scalia cited the 4th edition of Samuel Johnson's Dictionary, published in 1773, for a definition of "arms," rather than the 6th (and last) edition published in 1785, just before the Constitution was drafted.[37] The definitions differed only slightly, but perhaps in a meaningful way.[38] As Dworkin noted, the selection of sources is "openly political."[39] Moreover, historical sources are grossly indeterminate, which is why there are as many histories as there are historians.[40] Interpretive regression is also necessary, as even Justice Scalia had to provide second-order histories of the documents he used.[41] Justice Thomas' re-

---

[34] *Supra*, n.30.

[35] *See* David E. Young, *The American Revolutionary Era Origin of the Second Amendment's Clauses*, 23 Journal On Firearms & Public Policy (2011), Appendix I (only Pennsylvania and Vermont included a clear personal right to keep arms, while six others described the right as intended for the common defense). Many of those states used similar language when demanding a Bill of Rights as condition for ratifying the constitution. *Id*, Appendix II.

[36] Daniel A. Farber, *The Constitution's Forgotten Cover Letter: An Essay on the New Federalism and the Original Understanding*, 94 Mich. L. Rev. 615, 633 (1995) (quoting Justice Scalia in *Conroy v. Aniskoff*, 113 S. Ct. 1562, 1567 (1993).

[37] *Heller* at 581.

[38] The 1785 version contained an additional reference for arms, citing a poem by Alexander Pope:

"And seas, and rocks, and skies rebound

To arms, to arms, to arms!"

The quoted passage is from Pope's Ode on St. Cecilia's Day, written in 1730, in a verse that begins "But when our Country's cause provokes to Arms," thus providing further evidence of the military nature of "arms."

[39] Ronald Dworkin, *Law as Interpretation* in CRITICAL INQUIRY, 198 (U Chicago, 1982).

[40] *See* Suzanna Sherry, *The Indeterminacy of Historical Evidence*, 19 Harv. J. L. & Pub. Pol'y 437 (1995).

[41] 554 U.S. at 581 (citing Samuel Johnson, A DICTIONARY OF THE ENGLISH LANGUAGE (4th Ed. 1773) ("arms" defined as "[w]eapons of offence, or armour of defence")).

gression analysis is even more complex, going back to the early 14th Century.[42] Yet, "the historical' past ... is a complicated world … and support[s] no practical conclusions."[43]

Once the fine-grained historical and linguistic analysis is complete, an originalist will arrive at the "original communicative content" of constitutional text— what the words and phrases used conveyed to informed citizens (or subset) when they were called upon to ratify or reject the Constitution. The endeavor assumes that constitutional text in fact conveyed meaning. There can be no *original public meaning* without there first being a common understanding of meaning. If, as the following mathematical analyses conclude, many clauses were not coherent or comprehensible to the 1787 polity, then there never was a commonly held public meaning. In that case, originalism fails by its own premise.

III. Information Science and Originalism

One of the most significant mathematical advances of the 20th Century was the new field of information science. It provides insights into how human communication occurs and whether a signal can reliably be reproduced upon reception. Here we ask whether interpretations of communicative content made now can be faithful to the signals of that content sent more than two centuries earlier.

We should first grasp the centrality of *information* to the physical world and to human experience. Some contend that information is a more essential ingredient of reality, and instrumentality of it, than matter, energy or actions. In this view, the universe is mostly just a repository of information.[44] All other qualities and experiences are emergent, being derived from a vast sea of information that pervades all of space and time.

Information is any piece of knowledge concerning facts, data, rules, principles, theories or a description of something real or imaginary. It can be contained in physical particles and abstract ideas. It is also found in the oscillations of force fields that comprise the elements of the Standard Model of physics.[45] Oscillations are the carriers of information, as

---

[42] *See Rogers v. Grewal, supra*, n.18 at 1869 (Statute of Northampton (1328)).

[43] Emil A. Kleinhaus, *History as Precedent: The Post-Originalist Problem in Constitutional Law*, 110 Yale L.J. 121 (2000).

[44] James Gleick, THE INFORMATION: A HISTORY, A THEORY, A FLOOD, 10 (Vintage, 2011) (information "may be primary: more fundamental than matter itself").

[45] The Standard Model of particle physics describes all known particles and forces with the exception of gravity. Although gravity is a well-known force, we do not yet understand it well enough to include in the model. As a result, a "grand unified theory" remains yet to be found.

modulations of the electromagnetic field (e.g., radio waves) will attest. Since all particles can be described as oscillations in a force field (e.g., electrons are excitations of an electric field), those oscillations provide the substrate for the data that produces information. Accordingly, information is a physical quality that must obey all the laws of physics such as relativity and the speed of light *c*. Einstein gave us the equations for *c*, but later science describes *c* not as the speed of light, but as the speed of information transfer.[46] Information thus provides the substrate for space-time and everything in it.

Human thought is based on information.[47] When we sense, measure or interpret something, we are acquiring information about it. The Constitution, and law generally, is an information source, communicated across time and space to those who read and interpret it. Interpretation is simply a form of information processing. It consists of inputs, cognitive manipulations and outputs. As such, interpretation is a physical act, and like all physical acts, it can be described mathematically.

Information science began with mathematician Claude Shannon in the 1940s while he was working for Bell Laboratories, trying to improve signal propagation over telephone lines. Voice traffic was increasing at a faster rate than transmission lines were being built or upgraded. Could that traffic be condensed or encoded in a way that would increase the effective carrying capacity of the communication channel?

As Shannon discovered, information has no intrinsic meaning; it acquires meaning only upon reception and interpretation. Thus, *communication* is essential to the concept of information. While Shannon investigated the limits of communicating information across physical channels (such as wires and radio waves), his work applies to any form of knowledge management and transmission, including ancient mechanisms of writing and reading. Shannon's "classical" information theory did not involve the effects of quantum mechanics, but today *quantum information science* is a booming field with applications and devices such as quantum computing, already a multibillion-dollar industry.[48]

---

[46] *C* is also the speed of causality. If event A causes effect B, then the effect cannot be felt sooner than time *t*, which equals the distance between A and B divided by *c*. Any causative event propagates by way of information transfer.

[47] Emmanuel Desurvire, CLASSICAL AND QUANTUM INFORMATION THEORY: AN INTRODUCTION FOR THE TELECOM SCIENTIST ("Useful information is what intelligence looks after, with various degrees of expectation and priorities") location 1446.

[48] See George S. Greenstein and David Kaiser, QUANTUM STRANGENESS: WRESTLING WITH BELL'S THEOREM AND THE ULTIMATE NATURE OF REALITY (MIT, 2019), location 210-214

Information theory allows us to quantify the fidelity and reliability of communications, a question of engineering independent of particular semantic meaning. We focus not on applications per se, but on how well information can be communicated from point A to point B, or from source to receptor. As we'll see, information theory provides a way to examine the communicative content of the Constitution, or any text, and the meaning we are able to ascribe to it. More particularly, since originalism looks to the communicated meaning of constitutional text, it is vital that we develop a better understanding of that communication and whether it can be used as a basis for constitutional law. Accordingly, we must look to information theory, which "establishes a framework for any kind of communication and information processing."[49] We will see that information theory does not support the notion of communicating a fixed semantic meaning across the centuries without significant degradation and information loss.

The classical view of information treats objects and actions as real physical things and conceptually distinct from information about them. But even under classical interpretations, information is king. Take a gene, for instance, it is both a physical object (a nucleic molecule) and a repository of heritable information (a genetic sequence with its own 4-letter alphabet). While the gene may be destroyed or transformed, its information content still lives on.

Before the advent of formal logic and language, information about an object was usually contained only in the object itself. Thus, we would convey information about a hot rock by conveying the actual rock. But we eventually learned how to separate the two, abstracting away the information from the object. With the invention of language and other communication channels, information has become far more important than the object or event from which it was generated. As John Wheeler put it, the world is "constituted of information, so that information is more fundamental than what it describes."[50] One could go further and say that objects don't even exist in our experience; only information about them does.

A. Communication Channels

The Constitution is valued for its information content; less so for its substrate of vellum parchment (although the artifact itself may have historical or collectable value). If the artifact disappeared, as it did during the

[49] Dagmar Bruss and Gerd Leuchs, QUANTUM INFORMATION: FROM FOUNDATIONS TO QUANTUM TECHNOLOGY APPLICATIONS (Wiley-VCH, 2019).

[50] Lee Smolin, EINSTEIN'S UNFINISHED REVOLUTION: THE SEARCH FOR WHAT LIES BEYOND THE QUANTUM, location 264 (Penguin Press, 2019).

War of 1812,[51] its information content would be unaffected. Lawyers never produce the physical Constitution itself (which lies in the National Archives), or a statute offprint signed by the President, as the vehicle for argument and analysis. It's the information content that matters, not the physical object. An object can never be in two places at once, whereas information can be in an infinite number of simultaneous places (so long as we respect the speed limit for information transfer, *c*). Unlike physical objects, information is non-rivalrous and inexhaustible.[52] And, as described below, information never dies.

The Constitution is our legal system's primary information source. We no longer need to consult the physical object, but if the information is inaccessible, it serves no purpose. It must be disclosed and shared with others to have any meaning. Information loses all value if hidden.[53] Sharing requires transmission, or some form of communication. That, in turn, requires a communication channel. Thus, the science of information theory depends heavily on the types of communication channels that are available for any transmission of information.[54]

The channels that we use, and instinctively rely on in practicing or studying law, are writing and reading. Laying ink down on paper, or painting pixels on a screen, is communication that uses the channel of writing. Reading that paper or screen requires a different communication channel – modulated light waves. The physical channels in this enterprise – ink on paper for writing and electromagnetic radiation for reading – are physically quite distinct, but are nonetheless both treated as communication channels, subject to all the distortion, transmission loss, degradation and manipulation of channels generally. Such disruptions on a channel are called "noise."

It is impossible to avoid noise on a communication channel, or information processing generally.[55] That was the problem that Shannon set

---

[51] See Jessie Kraz, *PS: You Had Better Remove the Records*, https://www.archives.gov/files/publications/prologue/2014/summer/1812.pdf.

[52] That is why we have to create legal rules of excludability; *i.e.,* intellectual property laws.

[53] A corollary of Arrow's Information Paradox is that information has value only when shared with others.

[54] Communication theory is a subfield of information theory. The former focuses on the technical process of active information sharing, such as through channels, language and cryptography. The latter includes the creation, storage, processing and transformation of information, as well as distinguishing between data and information. I use the broader term – information theory – to apply to both fields.

[55] Michael A. Nielsen & Isaac L. Chuang, QUANTUM COMPUTATION AND QUANTUM INFORMATION (Cambridge, 2010) at 546.

out to solve - propagation of information (signal) across a noisy channel. The goal is to take the same information out of a channel as was put in; no more and no less. As stated by Shannon: "The fundamental problem of communication is that of reproducing at one point, either exactly or approximately, a message selected at another point."[56]

Shannon published two theories in 1948, one for communicating over hypothetical noiseless channels, and one for real-world noisy channels. All communication requires the use of code of some sort. To return to our hot rock for a moment: how does the information about the heat of the rock get conveyed apart from the rock? The heat itself, as a physical entity, could be communicated, as occurs during conduction, convection or radiation. But that involves the movement of tangible things (physical contact, air molecules or infrared radiation). It is far more useful to invent a symbolic language (coding system) to communicate information *about* the heat, rather than the heat itself, whether in English ("hot rock"), graphics (), alphanumeric (160°C) or any other symbolic system agreed upon by sender and recipient.

The rock is a physical thing; so too are representations about the rock. "Information is physical."[57] Because it is a physical thing, it requires channels for transmission, which in turn are affected by noise (which disturbs physical things). With noisy channels, error correction becomes necessary. Without error correction, messages become lost or distorted. So, if a waiter tries to warn a customer of a hot plate, but the restaurant is noisy, the customer may not receive the message and might get burned anyway. One way to correct for this, is to repeat the warning several times, or require an acknowledgment (a "handshake" in communication parlance) before setting down the dish. Error correction slows the flow of information, but without it, the message might not get through at all. Information is very vulnerable in that way. The noisier the channel, the greater the distortion of the message, or the slower the communication if error-corrected. This was Shannon's second theory—how to calculate the information-carrying capacity of a noisy channel using error-correction coding.

We need to encode a data stream both to limit loss across a communication channel and to turn the data into useful information. A random stream of 0s and 1s across a communication channel carries little information unless it is encoded at the transmission end and decoded at the

---

[56] Claude Shannon, *A Mathematical Theory of Communication* at 1, The Bell System Technical Journal, Vol. 27, pp. 379–423, 623–656, July/Oct, 1948.

[57] QUANTUM COMPUTATION, at 1. "More concretely, possessing information allows us to extract useful work from a system in ways that would have otherwise been impossible." Sean Carroll, FROM ETERNITY TO HERE: THE QUEST FOR THE ULTIMATE THEORY OF TIME, 189 (Dutton, 2010)

reception end. Replace the 0s and 1s with any "language" you choose, and you find the same problem: the sender and recipient need to use the same "codec" (**co**ding/**dec**oding system) for effective communication to occur.[58]

It is assumed by many originalists and historians that writers and contemporary readers of the Constitution used the same codec – a dialect of 18th Century written American English – to communicate across 2 different physical channels (quill/paper and light waves). Or to state it simply, they were speaking the same language. And, more fundamentally, their communications systems were not infected with noise. If either premise fails – different codec or noisy communication channel – then we can't be sure of what information was conveyed from drafter to voter.[59]

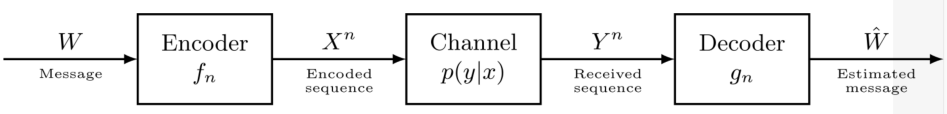
B. The Codec of the Constitution

Do you speak the same language as engineers, politicians, scientists, geeks or bureaucrats? How about teenagers? At what level of comprehension do you understand them, and they you? Many of the words used may be the same, but it is a stretch to claim a single set of meanings. Do you Twitter, newspeak, rap, use emojis, slang or body language? The English language and its symbols are so rich as to enable multiple means of communication within the vast linguistic umbrella of what we call English. The breadth of language variants may have been narrower in the 18th century, but there were still widely different usages and meanings. The assumption that the Renaissance men who wrote the Constitution spoke the same language as the New England merchants and southern plantation owners who ratified it is highly questionable. Even in the 18th century, stylistic conventions distinguished legal writing from ordinary English, much less the everyday vernacular of average citizens.

American English traces its roots to the regional dialects and flexible standards of the London area around 1600, sometimes referred to as

---

[58] The basic mathematical model for a communication system is graphically represented in Wikipedia as follows:



where the probability of error

$$P(\text{error}) = P(\text{error}|W=1) \leq P(E_1^c) + \sum_{i=2}^{2^{nR}} P(E_i)$$

[59] Shannon's Theorem does postulate fully error-correcting methods of communicating across noisy channels. That requires precision in code and limiting the rate of transmission to below that of channel capacity. The theoretical Shannon limit is not realistically achievable.

"London Standard English." But the diglossia of American English also included regional dialects that occasionally deviated from the orthographic norm. "Older Southern American English" was especially noteworthy in this regard as it incorporated dialects from poor English communities in Southern and Northern England and Ireland, as well as the Creole speech of African slaves. Social class was often determined by one's dialect, vocabulary and pronunciation.

There likely was no "standard" English, or uniform codec, at the time the Constitution was written. By the early 18th Century, many scholars considered the language "degenerate," "chaotic and in desperate need of some firm rules." In 1789 Noah Webster declared that "Great Britain, whose children we are, and whose language we speak, should no longer be our standard; for the taste of her writers is already corrupted, and her language on the decline." Accordingly, many British scholars of that era wrote instead in Latin or French. Affairs began to stabilize somewhat with the publication of Samuel Johnson's Dictionary of the English Language in 1755. But disagreement over grammar, spelling and rhetoric gave rise to over 200 other books on the English language in the latter half of the century. Most of these were prescriptive, advocating for "correct" English usage, rather than descriptive – how the language was actually being used. This strongly suggests that voters of different backgrounds and in different states may have understood the text (if they did) differently from one another. The "communicative content" of the constitution was hardly uniform, making it difficult to extrapolate a universally common meaning. Despite this, originalists posit a political community of early Americans that was fully informed, literate and subscribed to modern theories of philosophy and language.

Quantitative analysis of legal writing suggests otherwise. A recent work by Marion Dumas applied computational linguistics to legislative and judicial documents.[60] She found enough divergence in vocabularies and linguistic patterns to be able to identify ideological and partisan affiliations. "The two [political] parties now speak in different languages" including "robust splits in political language in elite discourse."[61] These language differences extend to judicial interpretations of text and even disagreement over "what the law is at a given point in time."[62] The current political polarization in America amplifies these trends. Thus, while Dumas' study was of modern legal discourse, we might imagine similar polarities arising during the 1787-88 constitutional debates. Linguistic differences among partisan or even unaffiliated voters would likely lead

[60] Marion Dumas, *Detecting Ideology in Judicial Language*, in LAW AS DATA, *supra* n.6, at 38.

[61] *Id*. at 384.

[62] *Id*. at 388.

them to have vastly different understandings of the Constitution presented to them, particularly in the case of idiomatic language that was not in wide use at the time, such as "keep and bear arms."[63] At the very least, "Americans were just as deeply divided over questions of constitutional methodology [in 1787] as they are now."[64] So when Justice Scalia says in *Heller* that the Constitution's "words and phrases were used in their normal and ordinary as distinguished from technical meaning," that does little to resolve the actual meaning received by the public.[65]

C. Noise

In communications theory, noise is any unwanted influence on the integrity of a communication, negatively affecting the reliability of a perceived signal. Take a musical note for instance. The human ear may hear a C although the composer intended a B flat due to any number of causes. The music sheet may have been incorrectly translated; the musical instrument may be out of tune; the recording device may be faulty; the transmission medium may be distorted; the air may be turbulent; etc.[66] The difference between an input sequence and the corresponding output sequence is known as "the Hamming distance," named after the mathematician Richard Hamming. It measures the number of operations needed to recover the original signal from the received one. As such, it is a useful metric for the design of error-correction mechanisms.

Noise takes many forms. It can be perturbations in the physical communication channel itself, such as fading paper or ink, or static on radio waves. Or it can be in how the information is coded, as with imperfect translation of pre-linguistic semantic concepts into words. In modern language theory, we may call that form of noise "ambiguity." You probably see where this is going. Ambiguities in constitutional text might foreclose any effort at reliable communication and make interpretation more difficult.

Nonetheless, some claim that the language of our founding documents was both deliberate and clear. Referring to the Declaration of Independence, one Supreme Court Justice wrote:

---

[63] *Cf.* Neal Goldfarb, *Corpora and the Second Amendment: "arms,"* https://languagelog.ldc.upenn.edu/nll/?p=41900

[64] Saul Cornell, *The People's Constitution vs. The Lawyer's Constitution: Popular Constitutionalism and the Original Debate over Originalism*, 23 Yale J.L. & Human. 295, 296 (2011).

[65] *Heller, supra*, n.30, at 576.

[66] This assumes a typical listener can differentiate neighboring notes. If not, as often the case, we have an analogous case to the inability of most voters and ratifiers to distinguish among fine legal terms in the Constitution.

> [T]he men who framed this declaration were great men,—high in literary acquirements, high in their sense of honor, and incapable of asserting principles inconsistent with those on which they were acting. They perfectly understood the meaning of the language they used, and how it would be understood by others….

That passage appears in the opinion by Chief Justice Roger Taney in *Dred Scott v. Sandford*,[67] as he endeavored to explain away the promise of the Declaration that "all men are created equal." Taney continued:

> [those great men] knew that it would not, in any part of the civilized world, be supposed to embrace the negro race, which, by common consent, had been excluded from civilized governments and the family of nations, and doomed to slavery. They spoke and acted according to the then established doctrines and principles, and in the ordinary language of the day, and no one misunderstood them.

If this is the argument for coincidence of purpose and understanding by the founding generation, it is a slender reed indeed. As with the holding of *Dred Scott* itself, we can probably dispense with Taney's assertion that public meaning of foundational texts was clear. No matter the context, language is just too noisy to allow for that.

There are other forms of noise on the reader's communication channel affecting the reproducibility of the original message. We know that the actual text of the constitution changed between the last session of the convention on September 15 and its final engrossment. At the signing on September 17, the scrivener had to include *errata* at the end of the document to reflect late changes.[68] The September 17 print has disappeared from history, but the September 18 print contained its own errors. Prints later sent to the states contained "small variations in punctuation, spelling, capitalization, and other stylistic details."[69] There was no official version of the Constitution for decades after its drafting. "In the absence of a recognizable official print, the nation's printers were obliged to copy whatever print might be at hand."[70] The errors and irregularities multiplied. A supposedly corrected and certified official version was not published until 1847. How are we to know whether slight variations pre-

---

[67] 60 U.S. 393, 410 (1857) (holding that the Constitution did not intend to treat either free or enslaved black Americans as citizens).

[68] See https://www.archives.gov/founding-docs/constitution-transcript.

[69] Henry Bain, *Errors in the Constitution—Typographical and Congressional*, PROLOGUE, Fall 2012, Vol. 44, No. 2.

[70] *Id*.

sented in different states during the ratification debates of 1787 and 1788 generated different understandings of meaning?

So, even if we are concerned solely with conveying the literal words on the page, there inevitably will be some distortion.[71] But if we also want to transmit the original meaning of those words across a two and one-third century long communication channel, then noise becomes a serious problem, as many perturbations are introduced along the way, not the least of which are cultural and environmental changes and intervening observations.

As quantum mechanics establishes, observation of an object or event inevitably changes the thing observed.[72] The more powerful the observation, the greater the resulting change. That leads us to decisions of the Supreme Court, the ultimate arbiter of constitutional meaning. The Court has said in no uncertain terms that Americans are not entitled to interpret the Constitution for themselves. Rather, the Court is the final and only authoritative interpreter of constitutional meaning. That result emerges from several cases, most notably *Cooper v. Aaron*,[73] the only opinion of the Court ever signed by all nine justices.

From the standpoint of information theory, an "authoritative" judicial interpretation of the Constitution constitutes an external influence on the text-to-observer information channel. It adds extraneous information (noise) to the primary signal that would otherwise be received by the observer. An authoritative interpretation inevitably distorts that original signal, no matter how agreeable the Court's opinion may be. We can recover the original information at the receiving end in one of two ways: 1) wall off the noise in the first instance, or 2) use Shannon's error-correction theorem.

Walling off the noise would require, *inter alia*, depriving Supreme Court opinions of authoritative effect. Interestingly, the Constitution does not

---

[71] In addition to changes in the text during the ratification period, some symbols used in the 18th century to construct English language words are slightly different than those in use today. Thus, we must interpret those symbols. Consider the double "s" (e.g., Congrefs), or the modern substitution of "s" for "c" (e.g., defence). The use of punctuation was almost certainly different, creating interpretive problems in the Second Amendment and elsewhere.

[72] This is the "observer effect." It is the observer who sets an object's parameters simply by looking. There is no "objective reality" that exists independent of observation. *See* generally Werner Heisenberg, PHYSICS AND PHILOSOPHY: THE REVOLUTION IN MODERN SCIENCE (1958).

[73] 358 U.S. 1 (1958) (state officials were bound to the Court's decision in *Brown v. Board of Ed*., 347 U.S. 483 (1954), and could not implement their own interpretation of the Equal Protection clause).

explicitly give the Court that power in the first place. Its preeminence is simply inferred, and by who else – the Supreme Court.[74] Or we could adopt the theory of "nullification," by which each state or state actor is empowered to interpret the Constitution for itself.[75] At least it would reduce the noise from judicial opinions. Or we can move from a common law to a civil law system, with its *ad hoc* interpretations of primary text in each case.

I don't propose that we do either of those things. Perhaps "noise" isn't all that bad for our legal system after all. That has not been the claim in any event. It is simply that noise, in the guise of authoritative constitutional decisions, further distorts the signal such that it is impossible to reconstruct the exact communicative content of the 1787 Constitution. We would have to ignore every Constitutional case ever decided and denude the Supreme Court of its interpretive power, except for the instant originalist opinion that is now about to issue. Then later, we'd have to ignore that one too and start over again.

That is what Justice Clarence Thomas advocated in *Saenz v. Roe*,[76] in explaining his originalist approach to interpretation. He rejected the majority's reading of the Fourteenth Amendment's Privileges or Immunities Clause – that it protected the right of interstate migration. He also rejected the earliest (nearly contemporaneous) interpretation of the Clause in *The Slaughter-House Cases.* "Legal scholars agree on little beyond the conclusion that the Clause does not mean what the Court said it meant in 1873."[77] In other words, the Court's first interpretation of the 1868 enactment was faulty, even though made only five years later. We cannot know which interpretation, Justice Miller's in 1873 or Justice Thomas' in 1999 was truer to original public meaning; we can only note that they differed. Of course, that is inevitable when trying to reconstruct a message sent years earlier over a noisy channel. Compare the similar phenomenon with biblical texts and authoritative canonical interpretations. Defy the Vatican, Westminster, or the Ayatollah at your peril.

A second way to recover the signal as sent (in originalists' terms, the original public meaning communicated by the Constitution) is to use some error-correction mechanism per Shannon. Mathematically this is done either through repetition or interaction ("handshake") between

---

[74] *Marbury v. Madison*, 5 U.S. (1 Cranch) 137 (1803).

[75] After the Affordable Care Act was enacted, a constitutional amendment was proposed that would allow states to "repeal" federal law, http://archive.nytimes.com/www.nytimes.com/2010/12/20/us/politics/20states.html, essentially codifying the Confederate Constitution.

[76] 526 U.S. 489 (1999)

[77] *Id*. at 522.

sender and receiver. For repetition, the same signal would be sent multiple times over a noisy channel. Where the signals agree (constructive interference), they get amplified; where they disagree (destructive interference), they cancel each other out.[78] Thus, if the same communicative content were contained in multiple sources, or transmitted multiple times, error correction might be possible. The receiver (that would be us, receiving the intended meaning) could then be confident of the message's integrity.

We do have multiple sources for original meaning. These include: the Constitution's actual text, Madison's notes of the convention, the ratification debates, contemporaneous writings, the election for delegates to the state conventions, and the votes and actions of those conventions. But each one of these is a distinct signal from a different source, rather than repetition of the same signal. Sending different signals is not the type of repetition that serves as error correction. And since many of those distinct signals disagree, they destructively interfere with the message. Thus, noise on the communication channel is amplified, rather than abated, by that effort at error correction.

So instead, originalists engage a second type of error-correction— interactive dialog with the source. In essence, a contemporary reader at the end of the communication channel interrogates the sender to confirm the integrity of the signal and ensure its original meaning. The Internet and other communications systems utilize measures of this sort to verify data integrity, such as "checksum," hash functions, fingerprints and parity checks. However, if the information source is no longer available to us, the needed "check" interaction cannot occur. In the case of originalism, we cannot seek clarification or verification of meaning from the authors, voters or ratifying delegates. So, originalists purport to talk to the Constitution itself; a self-referential process that Douglas Hofstadter calls "a strange loop."[79] Originalists thus attain agreement, not by uncovering what was actually meant, but by creating that meaning in the first place. Validation becomes tautological in that case. Error-correction becomes error-creation.

But conversations with the Constitution, or at least its authors, did occur during the ratification period. In some instances, these were more than polite dialogs; they were full blown, loud disagreements over the mean-

---

[78] Bruss, *supra*, n.49, at location 1544-49 ("The simplest code that can be used to detect or correct errors is the repetition code. A repetition code with rate transmits every symbol twice. At the receiver, the two symbols are compared, and if they differ, an error is detected.") For error correction, the signal has to be sent three or more times. *Id*. at 1549.

[79] Douglas R. Hofstadter, I AM A STRANGE LOOP (Basic Books, 2007) at Location 426.

ing of the text. Because of the ability to interrogate the Constitution's sources between the time it emerged from the old Pennsylvania State House on September 17, 1787 and the time it received its ninth state ratification by New Hampshire on June 21, 1788, some back-propagation and error checking occurred. Perhaps uncertainty in meaning as received by some of the public or state convention delegations was clarified by interactive dialog with Philadelphia delegates and their supporters during the nine-month interregnum.[80] The *Federalist Papers*, essays appearing in *The New York Packet* and *The Independent Journal* between October 1787 and May 1788, are a well-known example of that. Of course, one would also need to consult the *Anti-Federalist Papers*, written by other pseudonymous founders, to get a fuller picture of the error-correction and debate that occurred at the time in constitutional meaning.[81]

As a consequence of these dialogs, particular Constitutional provisions probably absorbed revised meaning during the ratification period. Do we know the results of those error-corrections? The *Federalist Papers* and a few other contemporaneous documents are often cited to support originalist meaning. But while some Constitution ratifiers did have a conversation (communication "handshake") with Alexander Hamilton, John Jay, and James Madison, authors of *The Federalist*, subsequent generations cannot, nor with the ratifiers to check their received meaning.

Thus, while some noise in the Constitution's original meaning may have been error-corrected prior to ratification, other clarification opportunities have been assiduously avoided. Originalists tend to be choosy in the historical documents they consult, as we see in Justice Scalia's selective citation of dictionaries in *Heller*.[82] Other examples are more poignant. The *Treaty of Paris* of 1783 – the document transferring sovereignty from the Crown to the United States – is especially notable for its omission from originalist investigation. The Treaty is what gives formal legal and diplomatic recognition to the United States of America. Yet, it is seldom mentioned in Supreme Court opinions, and never for the purpose of resolving questions of federalism and sovereignty, for which it is especially suited.[83] Other "forgotten" documents include the Constitu-

[80] Solum disagrees. "As a practical matter, these differences are likely to be minor: framing and ratification are likely to be proximate in time, separated by a few years at most." *Fixation Thesis, supra*, n. 32, at 7-8.

[81] Herbert J. Storing, Murray Dry, eds., THE COMPLETE ANTI-FEDERALIST, (University of Chicago Press, 1981).

[82] *Supra*, n.38.

[83] It would seem relevant for instance that the treaty was negotiated by representatives of Congress and not the states, and ratified by the former even though it resolved land claims in the latter. *See* LIBRARY OF CONGRESS, *Treaty of Paris*, https://www.loc.gov/rr/program/bib/ourdocs/paris.html (last visited Nov. 7, 2016).

tion's cover letter[84] and the Preamble to the Bill of Rights.[85] Although only 39 of the convention's 55 delegates signed the Constitution, they unanimously joined the convention's president, George Washington, in signing the letter transmitting the document to Congress.[86] Because the "letter represented an effort by some of the Convention's most distinguished members to explain the nature of the final product,"[87] it "has a virtually unique claim to our attention."[88] It would be hard to find a more authoritative source for the intent of the framers. Since the letter was also transmitted to the state ratifying conventions, it is also relevant to the other form of originalism— original understanding. Yet, it is never cited by originalists, perhaps because it challenges the "original public meaning" they prefer.

These contemporaneous documents raise interesting historical questions, but they may not help us with the interpretive problems we encounter today. Error-correction from the foundational period to the present time can no longer occur. If a modern receptor of a communication interrogates only her own interpretation of the sender's signal, that is not error-correction, it is error-confirmation.

In sum, to accurately receive the information content of Constitutional meaning 233 years after it was sent along the noisy channel that is the real world of constitutional law, we have to filter out the noise or correct for its inevitable distortions. Neither is feasible. We cannot subtract the noise that is the Court's many "authoritative" interpretations, or influential legal commentary and opinion. And we cannot "handshake" with the 18th century polity that erected the Constitution. The "Hamming distance" is just too great. Accordingly, the signal that ultimately emerges today from the communication channel of reading ancient text may bear little or no resemblance to the information content of the original.

There are two kinds of information channels: classical and quantum. Measurement of information content, such as occurs with legal interpretation or any text analysis, is a classical information channel. While quantum channels perfectly preserve information flowing across them,[89]

---

[84] Farber, *supra*, n.36 at 618-26.

[85] *See* Joint Resolution of Congress Proposing 12 Amendments to the U.S. Constitution (1789).

[86] As explained by Farber, the letter was drafted by the Committee on Style, the most influential of the convention's committees. *Supra*, n.36 at 627.

[87] *Id*.

[88] *Id*. at 635.

[89] This is how quantum computers process information. See QUANTUM COMPUTATION, supra n. 55 at 554.

classical channels cannot. Any measurement results in information loss. The information content at the receiving end is far less robust than the information content at the sending end. It is true that information never dies, due to the law of conservation of information (which is comparable to other conservation laws; *e.g.,* for matter and energy). However, much of the information about the original state (constitutional meaning) dissipates into the environment and can never be recaptured or recompiled. Time reversal occurs in quantum information channels, but not in classical ones. Thus, in a classical channel, "information about the past is not preserved into the future."[90] It is hard to do historical analysis when that happens.

D. Entropy and Information Loss

The Second Law of Thermodynamics describes the orderliness of a system. It is the only physical principle that treats time as unidirectional. Time is invariant in all other physical laws; *i.e.,* operations can flow backward or forward in time. Under the Second Law, systems always tend to go from an ordered state to a disordered state. This is known as *entropy*. It is what gives us the sense of time.[91] Were it not for the irreversible tendency toward disorder, time would proceed in both directions, just as each of the spatial dimensions does (up/down, forward/backward, right/left).[92] The past might lie ahead of us, not behind us. We could remember the future but be uncertain about the past.

Entropy is mainly a measure of the accessibility of information and the ability of a system to do work.[93] As entropy increases, information about the system becomes less accessible, and its energy becomes less useful. Due to the law of conservation of energy, the amount of energy remains fixed (subject to Einstein's mass-energy equivalence). However, not all energy is the same; some energy has the ability to do work, while other energy is functionally useless.[94] As entropy increases, the ability of the energy in a system to do work decreases. In other words, as systems be-

---

[90] Sean Carroll, *supra*, n.57 at 141.

[91] Entropy is the measure of the number of possible states for a system (e.g., arrangement of molecules). "All else being equal, an orderly arrangement will naturally tend toward increasing disorder," and not the other way. Since "there are more ways to be disorderly than to be orderly," entropy is seen as an irreversible process, giving us the sense of time. Carroll, *id*, at 2.

[92] "The technical way to say this is that there is a symmetry in the laws of nature—every direction in space is as good as every other." Carroll, *id*, at 30.

[93] Entropy is expressed as the accumulated heat lost in a system divided by the system's temperature.

[94] The Greek root "energie" means "to work."

come more disordered (increasing entropy) their ability to be productive decreases.

Shannon originally called his theory "information entropy."[95] Systems with low entropy have greater information content than systems with high entropy. Although information never dies, as time progresses and systems become more disordered, information about them becomes harder to retrieve. The information is still there, thus obeying the conservation law; it just cannot be accessed. "Entropy is a measure of how much information is hidden in the details—details that for one reason or another are too hard to observe. Thus, entropy is hidden information."[96]

As entropy increases (inevitably so with the passage of time), our ability to extract information from a system decreases. Examining a system now will yield less useful information than had we examined it yesterday, last week, or last century. Thus, what we know about a system today is less than its information content at its origin. Indeed, if we wait long enough as entropy increases, we will be unable to know anything about the initial state of the system, other than that it had low entropy. Just think of how much detailed and verifiable information we have about Genghis Kahn, or any historical figure; even less the further back we go, say to Jesus of Nazareth. The stories we tell are as much imagination as they are historical discovery.

Let's now apply Shannon information entropy to originalist and historical analysis of the Constitution. The document starts off in 1787 in a tightly organized low entropy state. Its concepts were actually in a higher entropy state before they were compressed into elegant text by the framers, because the words and phrases used emerged from complex ideas that had been evolving over years. Precise language was debated at the convention, merging the many, and often competing, concepts into discrete words and phrases. However, despite such careful deliberation, words cannot capture the full meaning of the concepts they express. Such is the limit of language.

How is it that entropy *decreased* with the coalescing of the Constitution? Doesn't the Second Law require that entropy always increase? Yes, it does, for any *isolated* or closed system. But the process of Constitution-making is hardly an isolated one. It is subject to strong external forces, which bakes a single low-entropy product. For example, a new deck of cards will typically come packaged with the cards segregated by suit, all numbers in order. It is in a state of low entropy. Once the deck is shuffled, however, it will be in a higher entropy state with most of its infor-

---

[95] Shannon's Theorem can also be described as Shannon Entropy. QUANTUM COMPUTATION, *supra*, note 55 at 572.

[96] Leonard Susskind, THE BLACK HOLE WAR, at 132 (Little Brown, 2008).

mation hidden.[97] Looking at the shuffled deck, or at any one of the cards, tells you almost nothing about where other cards are. Of course, the deck can be reassembled into its initial state of high order/low entropy. You'd have to wait a long time for that to happen randomly by shuffle, but only a minute or so if the external force of human hands reassembled the deck in suit and number order. You have to add work from outside the system in order to decrease entropy. The same is true regarding the high entropy state from which the Constitution was crafted. A lot of work went into distilling pre-formation concepts into agreed language. The problem is, once the system of the Constitution goes through these phase transitions (high entropy to low, to high again), it loses all sense of its history, which can never be recaptured. Upon looking at a newly shuffled deck of cards, we cannot tell what order they were in before they were shuffled. That history is forever lost to us.

Words diffuse, just as molecules and playing cards do, spreading out into a larger container that is our ever-expanding technological-social-economic-political environment. When words diffuse, they are able to do less work on their own and must be augmented, often by judicial opinion. With decreasing amounts of useful information derivable from the text, information from outside the Constitution has to be consulted in order to create a body of constitutional law. This also adds complexity because we can no longer rely simply on the highly organized low entropy text for guidance. The complexity of Constitutional law probably needs no argument. We've gone from 4,543 words in the original Constitution to thousands of opinions, each typically longer than the Constitution itself. Complexity in constitutional law is a gross understatement.

The Supreme Court overtly operates in this way. The shortest route to *certiorari* is through conflicting opinions from lower courts. The Court takes up those cases in an effort to restore order to the constitutional system. In doing so, it adds work to decrease the entropy of constitutional law and overcome system complexity (difficulty in describing the content of law). The added work is external to the system, it does not derive from the energy or meaning of original text. In short, the Supreme Court *adds* meaning to constitutional law with each originalist opinion; it does not derive meaning from it. Adding meaning, just like adding elements of any sort, converts an original closed system into a new larger one. It

[97] There are 52 factorial (80,658,175,170,943,878,571,660,636,856,403,766,975,289,505,440,883,277,824,000,000,000,000) different ways to order a deck of cards; a number far larger than the number of stars in the Universe. Each shuffle will yield one of those ways. We can extract more information from an ordered deck of cards than from one that is randomly shuffled. In a new deck, we know what each card is without looking, or seeing just one; whereas in a shuffled deck, we know nothing about each card, other than it has a 1/52nd chance of being the Queen of Spades (or any other card).

is now the larger system whose entropy will increase over time. In other words, each decision of the Supreme Court changes the Constitution and, in doing so, resumes its inexorable path toward disorder.

In sum, the entropy of the Constitution increases over time, both because that is the inevitable consequence of the evolution of its broader environment, and because of repeated interpretations by the Supreme Court and others. Increased entropy means decreased accessible information and a reduction in the document's useful energy. We cannot recover the lost information or its ability to do work by looking backward along a historical communication channel.

IV. Complexity

A concept closely related to entropy and information theory is complexity.[98] Complexity theory originates with theoretical computer science as a way of measuring the amount of computational resources needed to solve a particular problem. But the utility of complexity science transcends computational issues. It can help explain why both natural and social systems behave the way they do. Many have applied its tools to better understand legal systems.[99]

In general terms, a system is complex if many independent agents interact with one another in many ways.[100] It is hard to model *a priori* the interaction of multiple factors. If those factors can be reduced to data points, then there are statistical tools, such as multiple regression analysis, to approximate their interaction. But statistics yields only possible explanations and probabilities, not definite answers. It is also important to distinguish correlation from causation. Statistical analyses can offer up the former, but not the latter. That may be one reason, for instance, why the Supreme Court insists on proof of intentional discrimination in equal protection cases, rather than simply disparate effects, even when highly correlated with protected classifications.[101]

---

[98] This is more formally known as Kolmogorov complexity, the length of a shortest computer program that produces the object as output. Kolmogorov, A. N. *On tables of random numbers*. *Sankhyā Ser. A* 25 (1963), 369–376.

[99] *See, e.g.*, J.B. Ruhl, *Law's Complexity: A Primer*, 24 Ga. St. U.L. Rev. 885, 887 (2008) and authorities collected at n.11.

[100] M. Mitchell Waldrop, COMPLEXITY: THE EMERGING SCIENCE AT THE EDGE OF ORDER AND CHAOS (loc. 62) (Simon & Schuster, 1992)

[101] See *McCleskey v. Kemp*, 481 U.S. 279 (1987) (statistical study showing a risk that racial considerations entered into capital sentencing determinations did not establish equal protection violation).

Nonetheless, complex systems often exhibit a distinctive property called *emergence*, known colloquially as "the whole is greater than the sum of its parts." Quantum mechanics provides a theoretical explanation for the phenomenon.[102] Because of uncertainty in measuring individual components, it is often easier to gain knowledge of an entire system than any of its constituent parts standing alone. Thus, we can identify overall themes of the Constitution, such as representative democracy, fairness and justice, with greater confidence than we can than by interpreting individual clauses.

Another way to look at complexity is that a simple situation is easy to describe, while a complex situation is hard to describe. Low entropy systems tend to have low complexity because the individual elements of the system are highly organized and thus can be described more simply (or more elegantly) than those in a disorganized system. If the system under consideration is a set of rules, then simpler (more generalized or less variable) ones are easier to express than complex, detailed or idiosyncratic rule sets. This observation holds true even when run in reverse. Simple expressions of a rule tend to indicate low complexity of the underlying concept, whereas complex expressions (such as those with dependent or prefatory clauses; *e.g.,* the Second Amendment) correspond to more complex and divergent pre-textual concepts. Such expressions have high entropy and thus can do little work.

Complexity theory complements the Second Law of Thermodynamics in describing the entropy or disorder of a system. Consider the Constitution as a macrostate (complete and closed system) comprised of its several individual clauses or microstates. Those microstates diffuse over time (*i.e.,* with each "clarifying" interpretation and implementation), while the macrostate remains mostly stable. The latter's system values are much easier to describe than the formers' many distinct interpretations. Even if each microstate interpretation is supposed to look back to an earlier time of fixed meaning, one can never recover that meaning.

In this sense, complexity theory supports the philosophical notion of *holism* over *reductionism*. Reductionism posits that the behavior of a system can best be explained by understanding its constituent parts. In contrast, holism is a top-down view. Phenomena may appear at the system level which cannot be explained simply in component-level terms.[103]

---

[102] Quantum emergence describes an ontology where a system's behavior is independent of its parts. Indeed, the system is just as likely to exert influence on individual components as the components are to influence the system. See Rudolfo Gambini and Jorge Pullin, *Event ontology in quantum mechanics and the problem of emergence,* http://philsci-archive.pitt.edu/12231/1/eventontology.pdf.

[103] *Law's Complexity, supra*, n. 99, at location 275.

Legal language has peculiar functionality. That functionality is usually parsed through semantic analysis of text. But computational analysis offers additional insights. By applying complexity theory to the Constitution, we will see that while the document may have conveyed to the ratifying public a rich meaning at the system level, its components (individual clauses) were not as successful. Accordingly, an ontology that endeavors to discover original public meaning in individual clauses is not likely to succeed.

A. Algorithmic Information Content

Computational complexity describes the amount of computing power needed to solve a problem. Not surprisingly, complex problems require more resources and longer algorithmic runtime than simple problems. Physical and logical systems can also be complex, depending on how much information they contain. As used in this context, information is more than just data; it is "surprising" data. In other words, a data stream containing expected values, say "1, 2, 3, 4, 5," etc., contains little information, and is accordingly simple to compute. Unexpected symbols, in a sequence or elsewhere, contain more information, and are correspondingly more difficult to compute and are more difficult to communicate. Humans differ from other animals in the complexity of their communications; intricate sounds forming "languages," as opposed to monotone chirps, barks, or hums. Complex communications require a lot of computing power at both ends.

The complexity of a system or communication can be quantified according to its algorithmic information content. A repeating string of numbers, say 33333333, even if an infinite sequence, can be simply described – "print out the number 3, repeat." We can digitize the description (everything can be digitized) in binary form and count the number of binary 0s and 1s needed to describe the system. The translated string of 0s and 1s is the string's algorithmic information content. A sequence of random numbers, on the other hand, is difficult to describe or explain. Consider 305810396842555067732276l0. Assume for this exercise that this is a random number and can be described only by repeating each digit verbatim. There is no way to reduce the description to a shorter construct or algorithm. Not all long strings are complex. Take the sequence 3.1415926535…. That irrational number has an infinite non-repeating sequence of digits but can be easily described – "the circumference of a circle divided by its diameter." Even simpler – "π." In short, the length of a string of alpha or numeric characters is a measure of the string's complexity, unless it can be compressed into a simpler expression.

The algorithmic information content of a communication is integral to Shannon's Law. If the information is complex, it cannot easily be compressed or shortened, and requires greater bandwidth to communicate. A

Van Gogh painting is very complex. One must describe each arbitrarily small pixel separately according to its location on the canvas, along with its color temperature, hue, density, and the chemical composition of the oil paint used, among other qualities. A description of Minimalist Art on the other hand, such as Kazmir Malevich's "White on White," can easily be reduced, at least where there aren't many unique pixels. Compression algorithms such as mp3, jpeg, and similar use this ability to provide rich content with a digital signature that is far smaller than the uncompressed source. However, these programs are unable to significantly reduce file size if the content is disordered (i.e., has high entropy) or unexpected. Thus, it takes more megabytes to transmit or accurately describe a copy of Van Gogh's *Starry Night* than it does a picture of an actual starry night. That is computational complexity theory in a nutshell.

It seems intuitive that broad systems are more complex than their constituent parts. As Locke understood, even without the mathematics, "Ideas thus made up of several simple ones put together, I call Complex." Complexity theory applies not only to ideas, but also to every phenomenon in the real world. It is how Shannon discovered communication channel capacity and how a neuroscientist figures out the firing of neurons in the brain. Similarly, we can measure the complexity of text statements and other expressions according to their algorithmic information content. For instance, some provisions of the Constitution are complex, while others are not. Note, I am not talking about complexity in its legal sense – difficulty of application. Rather, this complexity analysis measures the degree of difficulty in communicating the ideas embodied in the text.

Efficiency in communication is important, both under Shannon's Theory and with transmission of law. Efficiency increases the likelihood of reliable reception by listeners, viewers and those involved in interpretation. By that I mean comprehension of communicated meaning. However, if the signal source is complex, or encoded in an inefficient manner, any communicated meaning is less likely to be properly understood. That is because communication consumes more resources, including cognitive resources, in proportion to its level of complexity.

With this in mind, I propose to measure the Constitution's information content with the goal of inquiring why particular statements may have been expressed in algorithmically complex manners instead of simple and efficient ones. Complexity theory provides a useful tool with which to analyze the semantic meaning of legal rules. Mathematical analysis of the Constitution thus supplements other analytical methods in deriving meaning from text. For instance, researchers have applied computational text analysis techniques to the work of the Supreme Court, including

both stylistic and semantic analysis of opinions,[104] revealing insights not previously available in the development of constitutional law.

B. The Constitution's Complexity

Analyzing the algorithmic complexity of the Constitution is a quantitative exercise, as it does not require the interpretation or examination of semantic meaning. Yet, it may shed light on those matters. As argued above, a complex provision is harder to communicate and understand than a simple one. If an interpretive method, such as originalism, depends upon the received public meaning of text, then the complexity of that text is a relevant consideration in evaluating the method's validity. In the following analysis I argue that the Constitution's complexity varies across the document; some provisions are much harder to process than others, and therefore may not have acquired any settled "public meaning" upon adoption. To the extent that a provision is quantitatively complex, the notion of original meaning of that provision is questionable.

In this empirical observation, I divide the Constitution into conceptual clauses; where each clause conveys a meaning separate from its neighbor. By my count, there are 313 such clauses in the original un-amended Constitution. The Bill of Rights contains 36 additional clauses, and the remainder of the amendments another 125 clauses, for a total of 474 distinct logical clauses. Defining a clause for this exercise is admittedly subjective. Different protocols will generate different computational results. The rules I adopt for this exercise are: 1) clauses are presumptively delimited by punctuation; but 2) where punctuation does not divide a unified thought, it is disregarded; and 3) a condition placed on the application of a provision is treated as a separate clause, whether marked by punctuation or not. An example of these rules in practice is shown as follows with Article I, section 3, paragraph 2, regarding the selection of U.S. senators in the case of vacancies:

> and if Vacancies happen by Resignation, or otherwise, during the Recess of the Legislature of any State,
>
> the Executive thereof may make temporary Appointments until the next Meeting of the Legislature,

The first part sets forth a condition, while the second creates a power. Thus, I treat them as 2 separate logical clauses. The first two commas are ignored (as they contain dependent grammatical clauses), while the third divides the clauses between the stated condition and power. Also, I exclude the leading "and" in my algorithmic count since it is just a con-

[104] An especially thorough explanation and application of these techniques is found in Keith Carlson, et al, *Style and Substance on the US Supreme Court*, in LAW AS DATA, *supra* n.6.

necting word and not meaningfully part of the clause in this case. I claim no authority for this protocol other than that it provides one workable means for measuring the Constitution's complexity.

The measure of complexity that I adopt is simply the number of digits in a binary conversion of each clause's English language text.[105] Converting into binary provides a convenient metric for analysis, although there are surely other ways to measure complexity. The shortest clause is in Article I, Section 8, paragraph 5 – Congress' power to "coin money" (80 binary digits). The longest is in the Fourteenth Amendment, Section 2--reduced representation in Congress when voting rights are denied (3,496 digits). This is truly a complex clause, unlikely to be understood by many when adopted in 1868, or now.[106] It is also an outlier, as the median binary length of the Constitution's clauses is 592 digits. It is even shorter for the original un-amended Constitution – 560 digits per clause.

Using this protocol, we can evaluate the communicated content and comprehensibility of each of the Constitution's clauses. A few are described here, with the full data source posted online for verification and further exploration.[107]

The grant of federal power in Article II, Section 8 is mostly simple and efficiently stated rather than complex. It has low algorithmic information content and its concepts are already highly compressed into a short list. Consider the following grants of power, with their algorithmic binary lengths in parentheses.

[105] There are various coding and conversion methods. I use ASCII conversion since it is the most common and yields the most consistent results. ASCII-to-binary conversion calculators are easily found on the internet.

[106] The clause contains a condition for reducing the number of Representatives: "But when the right to vote at any election for the choice of electors for President and Vice President of the United States, Representatives in Congress, the Executive and Judicial officers of a State, or the members of the Legislature thereof, is denied to any of the male inhabitants of such State, being twenty-one years of age, and citizens of the United States, or in any way abridged, except for participation in rebellion, or other crime…."

[107] The data set, charts, formulae and assumptions are posted in an Excel spreadsheet at [to be added]

- Congress shall have power to: (224 digits)[108]
- Lay and collect taxes (168 digits)
- Pay the debts (104 digits)
- Regulate commerce with foreign nations (304 digits)
- and among the several states (192 digits)
- and with the Indian tribes (176 digits)
- coin money (80 digits)

The rights statements in the First Amendment also have low complexity.

- Congress shall make no law: (208 digits)
- respecting an establishment of religion (312 digits), or
- prohibiting the free exercise thereof (296 digits)
- abridging the freedom of speech (248 digits), or
- of the press (96 digits)

These powers and rights statements have low complexity as is, and thus do not need to be compressed into simpler expressions. They often are, however, both colloquially and in legal discourse, such as "the right of free speech," which reduces algorithmic complexity somewhat.[109] The Second Amendment, in contrast, is very complex in terms of its algorithmic information content.

- A well-regulated militia being necessary to the security of a free state, the right of the people to keep and bear arms shall not be infringed (1,152 digits)

Its complexity suggests that greater resources are required, whether computational or cognitive, in order to extract meaning from the clause. In my analysis, I have chosen binary lengths exceeding 2^{10}, or 1,024, as an indication of complexity. One could select a different boundary, but 2^{10} is mathematically convenient and produces clear results. By this

[108] The actual converted binary string of this clause contains 224 digits: 01000011 01101111 01101110 01100111 01110010 01100101 01110011 01110011 00100000 01110011 01101000 01100001 01101100 01101100 00100000 01101000 01100001 01110110 01100101 00100000 01110000 01101111 01110111 01100101 01110010 00100000 01110100 01101111. That is how a computer would store or process the text.

[109] The compressed expression "right of free speech" has a binary value of 152.

metric, the Second Amendment is algorithmically complex, exceeded in the Bill of Rights only by the Tenth Amendment at 1,256 binary digits. Can it be parsed or compressed so as to allow for more efficient communication across a noisy channel? Perhaps the Amendment can be subdivided as I have done with the rights in the First Amendment:

- A well-regulated militia [is] necessary to the security of a free state, (568 digits),
- The right of the people to keep and bear arms shall not be infringed (544 digits)

Indeed, if we were to treat the first part simply as a condition or qualifier for the right in the second part, we would have two clauses that are close to the Constitution's median. But Justice Scalia rejected such treatment.

> The Second Amendment is naturally divided into two parts: its prefatory clause and its operative clause. The former does not limit the latter grammatically, but rather announces a purpose. The Amendment could be rephrased, "Because a well regulated Militia is necessary to the security of a free State, the right of the people to keep and bear Arms shall not be infringed."[110]

Thus, Scalia treats the amendment as a singular clause despite its grammatical separation, which produced interpretive difficulties later in the opinion. Also, instead of compressing the clause, *i.e.,* maintaining its meaning with a shorter string, Scalia's version actually lengthens it to 1,184 digits. So, I will stick with the actual text length of 1,152 digits.

Scalia's rephrased the amendment in order to maintain a logical connection between the two parts. He allowed the first clause no operative effect because that would impact the outcome.[111] If instead, the first clause had effect independent of the second, one would not expect to find it among a series of rights statements. Rather it would be combined with the clauses conferring power over the militia; to Congress in Art. I, section 8, and to the President in Art. II, section 2. Plus, the trick of replacing *being necessary* (conditional) with *is necessary* (declarative) is an example of "lossy" compression, rather than "lossless," meaning whatever protocol was used to simplify or compress the full statement resulted in a loss of information. It is that lost information that is so vital to the interpretation of the Second Amendment.

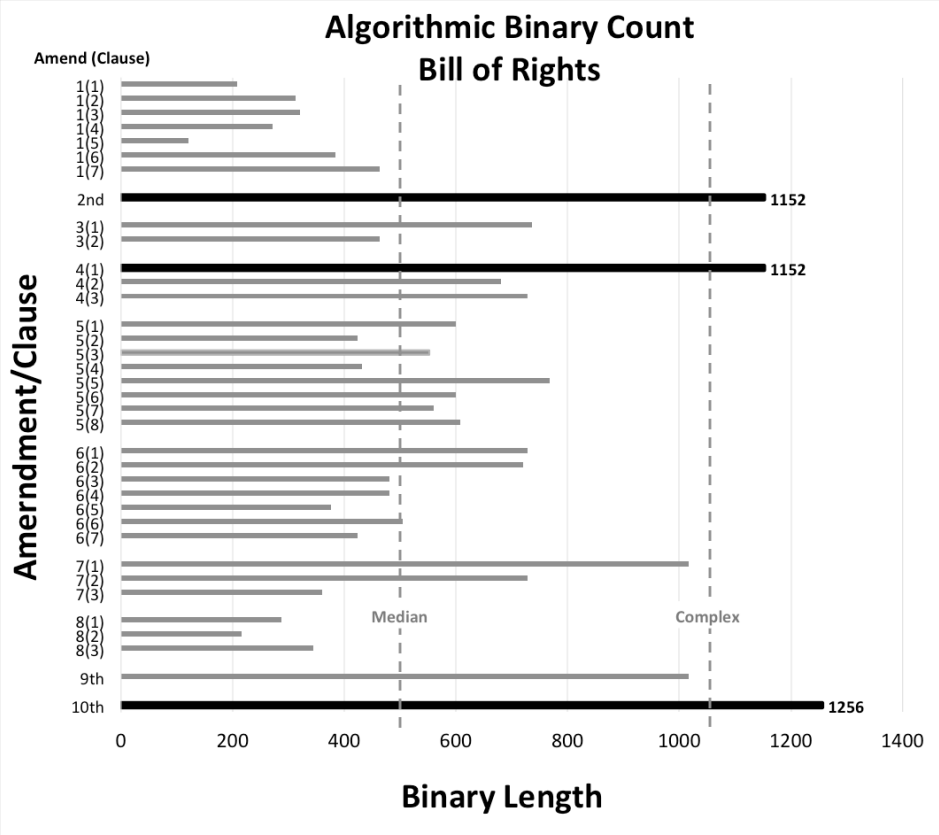
If the second clause were treated separately from the first, it would be about average complexity for the Bill of Rights (median length 492). But, if independent of the first clause, one might expect to find it among

[110] *Heller*, at 577.

[111] *Id*. at 579.

the "Congress shall nots" in the First amendment,[112] rather than having its own separate article of amendment. Since compression does not appear to be feasible with the "unique" structure of the Second Amendment,[113] we must take the algorithmic information content in its entirety as measured above (1,152 digits); which is significantly more complex than most other rights statements in the Constitution.

It may be helpful to visualize the complexity of the Constitution's clauses. The chart below does that for the Bill of Rights. It serves as a tool to identify unusually complex clauses and will permit us to see if the complexity I describe is supported by interpretive difficulties. In other words, if ordinary readers, and even Supreme Court justices, encounter difficulty in interpreting algorithmically long clauses, it supports the notion that complex clauses lack uniform public meaning.



[112] If it did, it might look as follows: "Congress shall make no law … infringing the right to keep and bear arms." As so construed, the right could be expressed in 384 binary digits, comparable in length to the Establishment Clause.

[113] *Heller*, at 577.

This chart shows the binary length of each clause in the first ten amendments.[114] Two full amendments, the 2nd and 10th, and the Search and Seizure clause of the 4th Amendment, exceed 1,024 digits in length, although the 7th and 9th are also long. The Fourth and Tenth Amendments are some of the most litigated provisions in the Constitution. The Second similarly suffers from the infirmity I claim here; namely, lack of clarity and comprehensibility.[115]

While the first ten amendments were even more compact than the original Constitution, with a median clause length of 492 binary digits, compared to the original's 560, the later amendments are quite prolix. One can infer that the Constitution became more complex as it aged, with the median clause length of the later Amendments (11 to 27) increasing to 728 digits, with a maximum of 3,496.[116]

At this point you're probably asking yourself several questions. Why do we care about the algorithmic complexity of legal text? What is the relevance of the fact that the construed right of individuals to bear arms is more complex than any of the rights contained in the First Amendment? And why am I belaboring the point?

The answer to the first question is simple. For law to work, it must be communicated. As Shannon showed, communication is a physical process that is fundamentally mathematical. Complexity analysis allows us to determine whether a signal can be sent over a noisy communication channel and, if so, how robust the error correction needs to be to ensure the message's integrity. It follows that the more algorithmically complex a constitutional provision is, the harder it is to communicate its meaning without information loss, especially in the absence of a reliable error-correction mechanism (such as the ability to interrogate the source). Moreover, information extraction is a computational process, whether done in our brains or on a computer. As readers and citizens, we derive meaning by computing the algorithm that is the Constitution.

Computational complexity is an important consideration in computer science. The longer an input string is, the more processing power that is required (independent of the actual content of the clause). Long inputs are more complex than short ones and require more resources to compute, regardless of the actual complexity of message the clause conveys. Is it

---

[114] The notation 1(1) means 1st Amendment, Clause 1. The 2nd, 9th and 10th Amendments each have a single clause.

[115] The Ninth Amendment appears to have little meaning to anyone and is seldom invoked. The first clause of the Seventh Amendment (civil jury trials) may stand as a counterexample since there are few disputes as to its meaning.

[116] Charts for other articles in the Constitution are found in the posted data source noted above.

any different if the computation is occurring *in vivo* rather than *in silico*; *i.e.*, in the brain versus in a computer? Unlikely, especially since our cognitive processes occur by electrochemical firing of neurons. It is no accident that artificial neural networks in artificial intelligence are modeled after human neural networks. It just simply takes more effort to read and understand long textual strings. As a consequence, the risk of misinterpretation is higher.

The response to the second question I posed above– why compare algorithmic complexities of the amendments – is that it leads us to inquire why the framers opted for simplicity in some rights statements and complexity in others. If, for instance, gun rights were as straightforward as speech rights, they could have been easily so stated in a similar linguistic format, and perhaps even included in the menu of the First Amendment. Complexity theory allows us to examine the information content of constitutional provisions in a way that semantic analysis (and certainly historical analysis) does not.

One explanation for why the algorithmic information content of the Second Amendment is far more complex than the rights statements in the First Amendment is that the amendment could not be compressed; *i.e.,* stated in a more parsimonious fashion. If that is the case, as I suspect it is, then the right it contains is fundamentally different than other rights in the Bill of Rights. No other amendment has a prefatory or explanatory clause. The framers were able to avoid complexity in the rest of the Bill of Rights (save the Tenth) because it was unnecessary to express those rights, whether free speech, just compensation, due process, or any other in a complex fashion. Instead, they observed Occam's Razor – start with the simplest solution to any problem.[117] Instead, the Second Amendment is complex, rather than a concise rights statement. Perhaps the prefatory clause is not that at all, but a category condition for the enjoyment of the right. In short, complexity theory supports the dissent's interpretation of the Second Amendment in *Heller*, but not the majority's.

Originalism suffers from complexity theory in another way. "The act of obtaining information through measurement … requires energy and must produce at least as much entropy as is decreased"[118] by resolving the information source. In the case of constitutional interpretation, the seeming order created by a judicial opinion is offset by an increase in entropy and algorithmic complexity of the underlying text. Each judicial opinion increases the disorder of the constitutional system, even while it purports to resolve a constitutional issue. While this notion seems intuitive to lawyers and law students; the more cases there are on a constitutional provision, the more complex are the contents of the law. Complexity

[117] *See, e.g*., http://math.ucr.edu/home/baez/physics/General/occam.html.

[118] *Complexity, supra,* n.100 at (loc. 62) (quoting Leo Szilard).

theory adds mathematical proof to what is known intuitively. With originalism, the mere act of interpreting (obtaining information about original meaning) muddies the water. We no longer have a short, elegant and simple Constitution, but instead have a massive, complex and often undecipherable body of constitutional law. Concomitant with increasing the Constitution's complexity, originalist interpretation results in information loss compared to the source.[119]

There are other forms of complexity that are arguably relevant to a proper understanding of constitutional terms.[120] For example, *communication complexity* examines how two or more parties, each with different inputs, can communicate mutual understanding with the least amount of raw data transmitted between them. If the framers approached constitutional drafting with different ideas of salient points, they could have written complex clauses, capturing each of those points, or at least those that survived debate. Instead, we see the economy of language nearly everywhere in the Constitution. "Due process," *e.g.,* captures multiple and probably diverse meanings. Communication complexity theory helps us understand that short phrases can have multiple meanings. The individual inputs may not be recoverable, for the same reason that we cannot reconstruct historical information lost to entropy phase transitions.[121]

For the most part, written and oral communication follows the linguistic "maxim of quantity," a term created by philosopher of language Paul Grice to measure how effectively meaning is conveyed.[122] Under this maxim, information should be provided for a reason. Speakers or text writers should say enough, but not too much. Following the maxim, if the Second Amendment were intended to include individual gun ownership, then the first clause simultaneously says too little (by failure to directly mention it) and too much (by being unnecessary to the second clause). The maxim forces us to confront what function of meaning the first clause serves. Justice Scalia contends that it merely sets the context for the second clause, yet every one of the Bill of Rights could have had

---

[119] As noted *supra*, increased entropy results in information loss, as does any observation.

[120] Algorithmic complexity can be measured in ways other than string length, such as the character set used for inputs. I leave those explorations to others.

[121] *Supra*, n.97 and accompanying text.

[122] P. Grice, STUDIES IN THE WAY OF WORDS (1989). Grice's "system of maxims for interpreting language in conversational settings … appear remarkably similar in form to many of the leading maxims of statutory interpretation." Geoffrey P. Miller, *Pragmatics And The Maxims Of Interpretation*, 1990 Wis. L. Rev. 1179, 1182. The legal maxim "*expressio unius est exclusion alterius*" is an example.

an explanatory or justifying preface – there was certainly a rich history behind each – but did not, except for the Second Amendment.

Another of Grice's linguistic maxims – "manner" – is also relevant to the inquiry. This maxim holds that effective communication requires the speaker to be perspicuous: avoid obscurity of expression, avoid ambiguity, avoid unnecessary prolixity, and be orderly. Obscure and convoluted language, even in ancient texts, undermines understanding. One linguist explains how the maxims of quantity and manner interact. If too few or too many "units of information" are conveyed for the listener's effective understanding, then the maxim of quantity is violated. If the message is the right length but the information conveyed is too "curt or long-winded," then the maxim of manner is broken.[123] According to Grice, when his maxims are violated, the presumption of shared meaning in communication fails, resulting in an enhanced risk of misinterpretation.

C. Readability Scoring

Algorithmic complexity is one way to assess the comprehensibility of constitutional clauses. Another is to measure readability scores using one of the many algorithms developed to gauge how understandable is a body of text. If algorithmic information content and readability scores correlate well, it would lend credence to my thesis that long, incompressible clauses are harder to understand than short statements.

In this section, I apply two commercial readability algorithms to the Constitution: 1) Automated Readability Index (ARI), and 2) Flesch-Kincaid Grade Level. Both metrics calculate how difficult a written passage is to comprehend by determining the U.S. grade level needed to understand the text. ARI[124] relies on a factor of characters per word, while the Flesch-Kincaid Grade Level[125] test takes into account the total syllables per word. At surface level, determining readability by simply counting the number of characters or words and multiplying them by a particular ratio to determine its complexity is rather straightforward. However, a document as politically noteworthy as the Constitution necessitates a multitude of interpretations, particularly in the case of informationally complex clauses and amendments.

Applying these widely used algorithms to the Constitution strongly confirms the conclusion that more algorithmically complex clauses create

---

[123] See H. Schiffman, *Topics in Dravidian Linguistics*, https://www.sas.upenn.edu/~haroldfs/dravling/grice.html.

[124] Equation for ARI = 4.71 × (characters/words) + 0.5 × (words/sentences) - 21.43

[125] Equation for Flesch Kincaid Grade Level: 0.39 × (total words/total sentences) + 11.8 × (total syllables/total words) - 15.59

difficulty in interpretation. Focusing on the Constitution's highly complex clauses (i.e., those with binary values greater than 1024), their ARI scores have a mean value of 16.2, corresponding to post-graduate-level reading comprehension. Their average Flesch-Kincaid Grade Level is 15.7, or roughly senior year in college. Moreover, the correlation between calculated algorithmic content and ARI for complex clauses is .95; for Flesch-Kincaid, it is .92. These suggest nearly perfect correlations between binary length and readability, and tend to confirm my hypothesis of a difficult-to-understand Constitution, at least in many clauses.

Comparing the First and Second Amendments, as I have done for algorithmic information content, generates the following:[126]

Clause	**Binary Length**	**ARI**	**Flesch-Kincaid**
Congress shall make no law respecting an establishment of religion,	528	10.42	9.55
or prohibiting the free exercise thereof;	320	9.04	10.35
or abridging the freedom of speech,	272	4.34	4.46
or of the press;	120	-	-
or the right of the people peaceably to assemble,	384	4.00	3.65
and to petition the Government for a redress of grievances.	464	6.65	8.37
A well-regulated Militia, being necessary to the security of a free State, the right of the people to keep and bear Arms, shall not be infringed.	1,152	12.13	12.86

The ARI score for the Second Amendment, 12.13, corresponds to a 12th Grade reading level. Flesch-Kincaid rates it a bit higher – college freshman. The Free Exercise clause comes closest, presumably supported by the difficulty the Supreme Court has had in articulating its meaning. Other readability algorithms produce similar results. The Lexile test, developed by the education company Metametrics, uses sentence length and difficulty of vocabulary to gauge linguistic complexity.[127] According to the Lexile algorithm, the First Amendment as a whole has a reading comprehension score of 1000-1100, corresponding to average

[126] Readability scores calculated at https://www.online-utility.org

[127] *See* https://lexile.com

6th grade reading comprehension.[128] In contrast, the Second Amendment scores 1200-1300 on the Lexile scale, corresponding to a modern 12th grade level, a level attained by very few of the founding generation. The assertion that the Amendment was written so as to communicate meaning to the average 18th Century voter is a difficult one to sustain.

Originalism purports to discover the "original public meaning" of the Constitution and fix that meaning throughout time short of constitutional amendment. Complexity analysis undermines that heuristic through the mere realization that much of the Constitution was not readily understandable, either to the founding generation or currently.

V. Data Modeling of the Constitution

A. Law as Data

We tend to think of articulations of law, whether in positive law, decisional law, scholarly works, or private writings, in mostly analog terms. Even where we use digital tools, such as electronic research and e-discovery, the corpus of law consists mainly of doctrinal meanings expressed in symbolic language. However, recent advances in computational text analysis have allowed scholars to treat legal sources and content as data. Computational legal analysis is similar to quantitative empirical legal studies but takes advantage of the increasing power of computers and artificial intelligence to extract information and discover patterns in data that often escape semantic analysis. This domain treats legal text as data to which mathematical methods are employed to derive meaning and insights.

One example of this is the use of "big data" to construct searches and derive meaning. Repositories of legal and other relevant sources can be explored using computational means to gain intelligence that may go unnoticed using conventional legal tools. Most commercial law search products now offer Natural Language Processing, a form of artificial intelligence, for both search queries and the extraction of legal argument and reasoning from data.

The sources of "big data" go way beyond customary legal databases. For instance, Mark Davies of Brigham Young University has collected more than a dozen corpora of English-language materials, some dating to the 1500s.[129] Dennis Baron used the Corpus of Founding Era American English and the Corpus of Early Modern English,[130] to find instances of

[128] This is Metametrics' internal scoring scale and doesn't correspond to other metrics.

[129] https://english-corpora.org

[130] *See* https://lawcorpus.byu.edu.

the term "bear arms" in documents from the founding era.[131] Of the 900 separate occurrences he found, after correcting for duplicates, nearly all "refer to war, soldiering, or other forms of armed action by a group rather than an individual."[132] None clearly applied the term to private uses or personal defense.[133] "This should constitute proof that the natural meaning of bear arms in the framers' day was military or quasi-military,"[134] flatly contradicting Justice Scalia's conclusion in *Heller*.

The BYU corpora contain over 200 billion words, making ordinary search cumbersome. To complicate matters, many words are auxiliary, such as the phrasal verb "bear" as in "bear arms," rendering interpretation even more difficult. Fortunately, artificial intelligence can help extract meaning from context. As Baron's essay shows, no originalist can avoid these data resources and tools if they want to be taken seriously.

There are other computational tools that are relevant to reading and understanding law. Many use text and underlying legal documents for quantitative statistical analysis. The Santa Fe Institute recently published a compendium of such works showing how data analytics can augment typical semantic analysis.[135] One work, for example, analyzed "intensifiers" and readability scores "to test the[] theory that justices broadcast weak legal position through the use of language."[136]

Another example of computational analysis that has appeared in legal literature is the frequency of words used in judicial opinions. Simple word frequency facilitates semantic analysis by comparing, say, Supreme Court opinions to that of other courts, opinions over time, or one justice's to others. One such study found that Supreme Court opinions are "a genre apart" from other courts, and "increasingly idiosyncratic."[137] Another found the "the court's language has become decidedly more 'grumpy' over the course of the past two centuries."[138]

---

[131] Dennis Baron, *Corpus Evidence Illuminates the Meaning of Bear Arms*, 46 Hastings Const. L.Q. 509 (2019).

[132] *Id*. at 510. One of the non-military uses was in a 1780 English translation from a French description of an orangutan.

[133] Seven instances were ambiguous or carried no military connotation. *Id*.

[134] *Id*.

[135] *Supra*, n.6.

[136] Keith Carlson, et al., *Style and Substance on the U.S. Supreme Court*, in LAW AS DATA, *id*, at 86.

[137] Michael A. Livermore, et al., *The Supreme Court And The Judicial Genre*, 59 Ariz. L. Rev. 837, 871 (2017).

[138] *Style and Substance, supra.* n.136 at 87.

In the next section, I apply a computational tool to measure the coherence and comprehensibility of the Constitution. If, as I conclude, the document deviated from linguistic norms, it would have been difficult for most readers (then or now) to derive meaning from it. A heuristic that relies on the supposed received public meaning to determine the contents of constitutional law fails if such public meaning is illusory.

B. Zipf's Law

Modern languages strive to maximize communication efficiency while minimizing the cost of communication. Demonstrating this principle, American linguist George Kingsley Zipf developed a theorem of communication efficiency that he called the *Principle of Least Effort*.[139] "Least effort" is a variant of "least work" or "least action," the minimal amount of energy an individual or group must expend in the pursuit of an objective. Zipf asserts "all physical process throughout the entire time-space continuum is governed by the one single superlative, least action."[140]

While there are many examples of the least action principle in physical processes, such as in Shannon Information Theory, discussed previously, here we are primarily concerned with its application to human language. Zipf argues that the principle controls here as well. This "empiric natural law" is the "primary principle that governs our entire individual and collective behavior of all sorts, including the behavior of our language and preconceptions."[141] In terms of communication, Zipf asserts that writers and speakers strive to convey meaning while economizing their speech. Words are tools for communication and using them economically reduces the effort required of the speaker. Lawyers are familiar with economies of speech in many areas, such as trademark and defamation (innuendo), as well as in our use of abbreviations, maxims, terms of art, and Latin expressions. Least effort also applies to our word choices.

One form of economy in speech (for the speaker) is found in using the same word repeatedly, even to express different meanings. We become familiar with our vocabulary, including our preconceptions of meaning, and thus develop preferential reuse of certain forms. This does not depend upon the education, eloquence or even the language of the speaker. For all speakers in all languages, it takes added effort to learn, differentiate and use different words to match all the various meanings we want to

---

[139] George Kingsley Zipf, HUMAN BEHAVIOR AND THE PRINCIPLE OF LEAST EFFORT: AN INTRODUCTION TO HUMAN ECOLOGY (Addison-Wesley Press, 1949).

[140] *Id*., location 231.

[141] *Id*., location 137. Zipf explains that an individual's preconceptions become part of her vocabulary of language. *Id*. at location 127.

express. An extreme and rare example of this phenomenon is a speaker who uses a single word to express all the separate meanings she wishes to convey.[142] Expanding one's vocabulary of usage, even when other words are at hand, takes effort that the speaker might prefer to expend elsewhere. Just as one might employ a multi-function tool, such as a Swiss Army Knife or Smart Phone, to perform tasks that a specialized device would perform better, we also tend to economize effort with our word choices.

In communication, there is a tradeoff between speaker and listener. Economy of speech for the former, such as use of idioms or fewer unique words to convey meaning, requires the latter to expend greater effort to uncover the conveyed meaning. If we were to instead emphasize least effort for the reader, in deconstructing intended meaning, then the speaker would need to use a different word for each separate meaning.

> Indeed from the viewpoint of the auditor, who has the job of deciphering the speaker's meanings, the important internal economy of speech would be found rather in a vocabulary of such size that it possessed a distinctly different word for each different meaning to be verbalized. Thus if there were *m* different meanings, there would be *m* different words, with one meaning per word.[143]

But that doesn't typically happen in legal language. For example, take the modal verb "shall." As a modal verb, "shall" is used in conjunction with a base verb to indicate modality as well as tense, voice and syntax. Thus, it can be used in many different ways, such as: an *imperative command* (as in *mandatory*, *must*, *will, directed* or *obliged), permissive* (as in *can*, *may* or *could), hortatory (*as in *should, requested* or *exhorted), predictive (*as in *intended, planned, inevitable, or determined)*, *interrogative* (as in *offer* or *question*) and so on. The Oxford English Dictionary lists 6 pages of definitions and illustrative uses.[144] This word has obviously developed from its Old English origins to embody a vast array of different and sometimes conflicting meanings. For example, "shall" has

---

[142] A poignant example is provided by John Malkovich's character in *The New Pope*. As Pope John Paul III, Malkovich delivers a papal sermon consisting solely of the word "no," which he repeats over and over again. It is meant to convey a variety of reactions and emotions, including commentary on violence and cruelty, current events, treatment of sex, liberalized biblical interpretations, and so on. *The New Pope*, episode 1.5 (air date Jan. 24, 2020, HBO).

[143] *Id*, location 593.

[144] OXFORD ENGLISH DICTIONARY (Oxford, 1971), at "S" 607-613.

been used with 1,128 different base verbs in Supreme Court opinions.[145] As meanings expand, and the writer acclimates to them, it is often easier to use fewer words in different contexts. However, when a writer uses the word *shall* to convey multiple meanings in a text, the reader must expend greater effort to associate each usage with the author's intended meaning, by perhaps considering collocates and context, grammar, historical usages, and comparison to other uses in the same or related texts.

Consider how many different meanings are ascribed to *shall*, a word that appears 189 times in the Constitution. Here are a few examples of different usages, with their legal interpretations:

- "the judicial power *shall* extend to" (Art. III, Section 2), construed in *Martin v. Hunter's Lessee* as used in the imperative sense;[146]

- Congress has power "to make all laws which *shall* be necessary and proper" (Art. I, Section 8, par. 17), construed as discretionary in *McCulloch v. Maryland*;[147]

- "No state *shall* enter into any treaty …" (Art. I, Section 10), construed as prohibitory in *Providence Bank v. Billings*;[148]

- "A person charged in any State with treason, felony, or other crime, who *shall* flee from justice, and be found in another State, *shall*, on demand of the executive authority of the State from which he fled, be delivered up…" (Art. IV, Section 2, par. 2); first *shall* construed as "might" in *Prigg v. Pennsylvania*;[149] second *shall* determined to be obligatory in *Holmes v. Jennison*,[150]

- "The United States *shall* guarantee to every State in this Union a republican form of government, and *shall* protect each of them against invasion …" (Art. IV, Section 4); first clause held to be not judicially enforceable, and second clause held to be discretionary in *Luther v. Borden*.[151]

---

[145] Corpus of US Supreme Court Opinions (search term "shall" collocated with any verb within 1 word following), https://www.english-corpora.org/scotus.

[146] *Martin v. Hunter's Lessee*, 14 U.S. 304, 331 (1816).

[147] *McCulloch v. Maryland*,17 U.S. 316, 420 (1819).

[148] *Providence Bank v. Billings*, 29 U.S. 514, 559 (1830).

[149] *Prigg v. Pennsylvania*, 41 U.S. 539, 611 (1842).

[150] *Holmes v. Jennison*, 39 U.S. 540, 615 (1840).

[151] *Luther v. Borden*, 48 U.S. 1 (1849). *See also Groves v. Slaughter*, 40 U.S. 449, 496 (1841) (clauses are discretionary).

In many of the 189 instances in which the framers used the word *shall*, it has taken a Supreme Court decision to give it authoritative meaning.[152] It would have saved readers, constitutional lawyers and historians, considerable effort had the framers expended greater effort themselves to select instead from a menu of verbs, each with fewer or distinct meanings. As Bryan Garner, editor of Black's Law Dictionary, says: "*shall* violates the presumption of consistency: Words are presumed to have a consistent meaning in clause after clause, page after page. Which is why shall is among the most heavily litigated words in the English language (with hopelessly inconsistent court holdings)."[153] Unsurprisingly, the U.S. government's Plain Language Guidelines recommend that drafters of official documents "delete every *shall*,"[154] and yet, it is the third most common word found in the Constitution. Pity the poor ratifiers.

In examining the respective efforts required of writer and reader, Zipf asserts that "the two opposing economies are in extreme conflict," and that successful communication requires some balance in vocabularies used by the opposing forces. Thus, writing that is successful in conveying meaning will be a compromise between the speaker's and reader's internal economies. Optimized communication will be found somewhere between a single word for all of the writer's intended meanings and separate words for each.

Zipf posited that the success of any given communication can be measured empirically. This measurement of communicative efficiency is now known as Zipf's Law. Zipf found that patterns of speech evolve to promote this balance, and the balance can be evaluated statistically. The basic principle is that communicative efficiency is achieved at lowest cost when word frequency follows a particular power-law relationship.[155] In addition to measuring the balance of internal economies, Zipf's law also measures the connectedness of syntactically linked words, such as prepositions and conjunctions, that are a precondition for successful

[152] Not only is the word "shall" overused in the Constitution, it also seems to occupy a special place in Supreme Court opinions. Just in the first decade of the Court's existence, the word appeared 1,508 times. Treating it as a lemma (including "will") increases the count to 2,698 occurrences. These counts are per the Corpus of US Supreme Court Opinions, https://www.english-corpora.org/scotus.

[153] Bryan A. Garner, *Shall We Abandon Shall*, ABA Journal, Aug. 1, 2012.

[154] https://www.plainlanguage.gov/guidelines/conversational/shall-and-must

[155] R. Ferrer i Cancho and R.V. Solé, *Least effort and the origins of scaling in human language*. Proc. Natl Acad. Sci. 2003;100:788–791.

communication. Others have shown that Zipf distributions can arise from minimizing information-theoretic notions of cost.[156]

There are many power law relationships, both in nature and in social structures. Newton's inverse square law of gravitation is an example of the former. Pareto distribution is an example of the latter. Italian economist Vilfredo Pareto observed that income and wealth tend to follow a scalar distribution.[157] This is the field of Pareto power statistics, which describes relationships between rank and size, say of cities and their populations. A nation's second largest city is often roughly half the size of its largest, and so on. The related field of computational linguistics applies statistical methods to identify empirical rules, including least action, in human languages.

Zipf's law is a special case of Pareto distribution; it posits an exponential power relationship between ordinal rank and frequency of word use in any large corpus. In the standard version, the second most used word in a sample will generally appear half as often as the most frequently used; the third will appear one third as often as the first, and so on. Thus, relative frequency $f \approx 1/r$ where r is the rank number of a word's occurrence and f is its relative frequency (compared to the most frequent word in the sample). If "the" is the most frequent word in a corpus, found 100 times, and "to" is the second most frequent, Zipf's law would predict 50 occurrences of "to."

Zipf derived the equation $r \times f = C$ as a measure of vocabulary balance. In the ideal case, C would be a constant; all words in a given corpus would have the same C value. He tested his theory by examining James Joyce's *Ulysses* and found a remarkably uniform value for C.[158] Moreover, when plotted on a double logarithmic chart, the idealized slope of rank vs. frequency is -1; i.e., a straight line at 45 degrees downward to the right. That appears to hold for texts that have a strong vocabulary balance; *i.e.*, are seen as successful communications.

---

[156] Benoit Mandelbrot, *On the theory of word frequencies and on related markovian models of discourse*, Structure of Language and its Mathematical Aspects. 1962:190–219

[157] Pareto Distribution is an empirical observation describing fairly consistent ratios in distributions of wealth (e.g., 20% of a society's population owns 80% of its wealth), sizes of cities, traffic patterns, insurance casualty losses and other quantities. See Vilfredo Pareto, *Manuale di economia politica con una introduzione alla scienza sociale* (Milano 1906). A refined form of this is known as a *Lorenz curve*, named after Max Lorenz, that quantifies in a canonical manner the distribution of income and wealth in most societies.

[158] Zipf, *supra*, n. 139, at location 633, 701-761.

With this experiment, Zipf posited a "fundamental regularity of some sort of an underlying governing principle that is not inconsistent with our theoretical expectations of vocabulary balance."[159] This holds true "regardless of whether or not the speakers and auditors are aware of the existence of the principle."

Let us now apply Zipf's law to the original, un-amended, Constitution. Excluding signatures, it contains 4,427 total words and 853 unique words. This table shows the ten most common words in historical English and in the Constitution.

Rank	Early Modern English[160]	Founding Era English[161]	Constitution	Frequency (Constitution)
1	the	the	the	413
2	of	of	of	284
3	and	to	shall	189
4	to	and	and	188
5	in	in	be	128
6	that	a	to	114
7	a	that	in	88
8	is	be	states	81
9	it	i	or	79
10	his	it	united	54

There are few surprises here, except for the word *shall*, which is the third most frequently used word in the Constitution.[162] In contrast, *shall* ranks a lowly 84th in Early Modern English,[163] and 2,217th in current usage.[164] Another way to see this anomaly is that "shall" appears 48 times more frequently in the Constitution, as a percentage of the corpus, than in

[159] *Id*. at location 777.

[160] This is based on a wildcard search in the BYU Corpus of Early Modern English (COEME), which is based on over 1 billion words in usage between 1475 and 1800, https://lawcorpus.byu.edu/byucoeme. For modern word frequency, *see* https://en.wikipedia.org/wiki/Most_common_words_in_English.

[161] This is based on a wildcard search in Corpus of Founding Era American English (COFEA), based on 136 million words in usage between 1760 and 1799, https://lawcorpus.byu.edu/cofea. Many of these BYU sources are related to the Constitution, so may not be independent of the latter's lexicon.

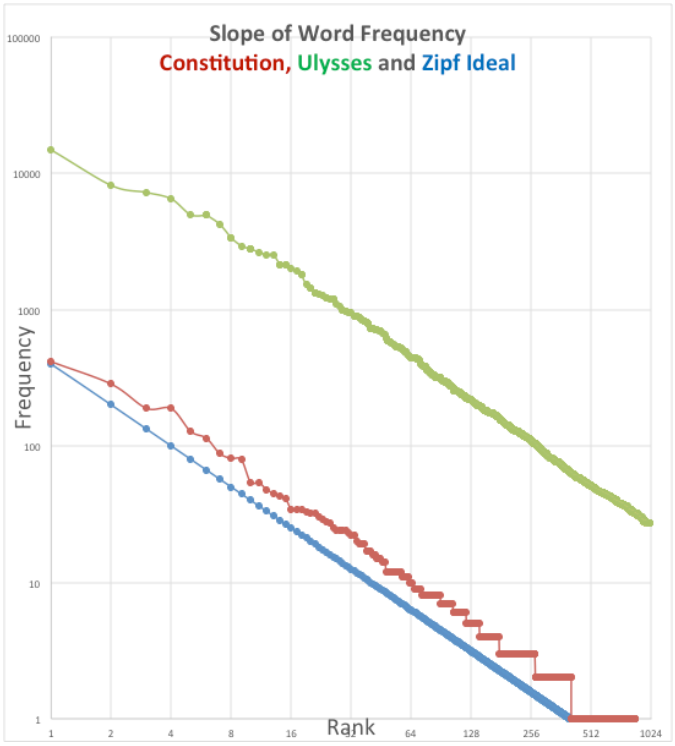
[162] "United" and "states" also appear frequently, which is to be expected in the *United States Constitution*.

[163] See n.160.

[164] This is the rank found in https://www.wordfrequency.info/free.asp?s=y..

common usage at the time.[165] More specific verbs were available and typically used in early modern English, but the framers seemed fixated on *shall*. Perhaps it conveyed formality or urgency, along with its many disparate actual meanings. It was surely a workhorse in the Constitution. Because of its deviant word frequency, the document lacked a proper balance of vocabularies between writer and reader. And because communication efficiency was lost, one would expect reduced comprehension for many or most readers. Readers could not actuate the principle of least effort. Rather, they would have to employ significant effort to understand the particular meaning of each of the 189 uses of the word *shall*.

Also of interest is that the slope of word frequency deviates from the Zipf ideal, -1. The following chart plots *rank* (x axis) against *frequency* (y axis), both for the Constitution and for a hypothetical source that perfectly conforms to Zipf's law. For good measure, I also show the slope of James Joyce's *Ulysses*, from which Zipf's Law is derived. Each axis is logarithmic, as provided by Zipf, to indicate the *C* product slopes.[166]

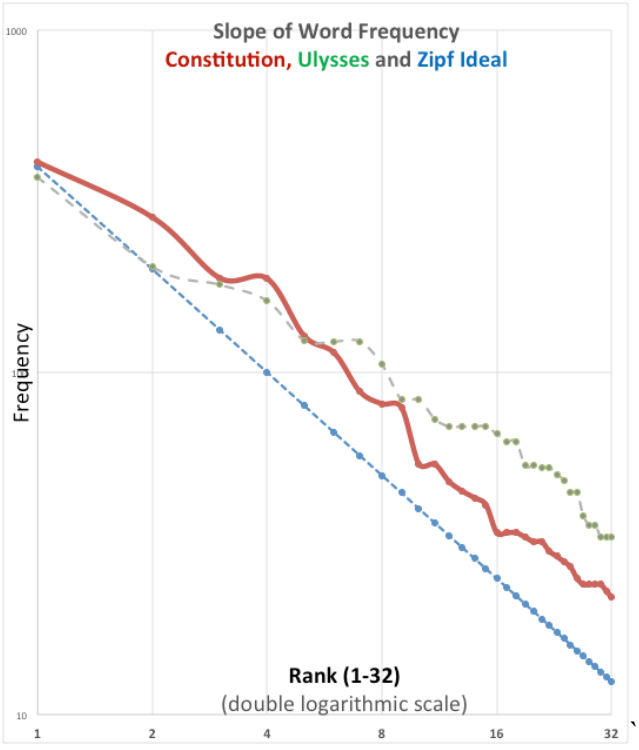


[165] "Shall" comprises 6.2% of all words in the un-amended constitution, compared to .13% of all words in the COEME corpus.

[166] The chart shows the product of rank and frequency: $r \times f = C$ for each unique word. Zipf's law predicts that *C* will remain constant throughout a text. The raw data used to generate the charts is available at ___. Rank/frequency data for Ulysses was produced by A. Corral, et al, and can be found at https://figshare.com/articles/data_for_Corral_et_al_PLoS_ONE_2015/1430465.

From this chart, it appears that the Constitution's word frequency generally conforms to Zipf's law with an overall slope very nearly equal to -1. The data points and slope for *Ulysses* also appear close to the ideal.[167] The tail of that slope is not plotted because there are nearly 30,000 unique words in *Ulysses*, thirty-five times as many as in the Constitution. Including them would make a visual comparison less effective. The "steps" at the lower right on the Constitution's slope are to be expected because word frequency is always a function of whole integers. As word count attenuates into the single digits, the plot will show them as steps.

At this resolution, the Zipf-based slope of the Constitution appears to indicate balance in vocabularies and efficiency in communication. However, conforming slopes for words with low frequency (say, below 10 occurrences) may simply be an artifact of a dependent relationship,[168] making it is harder to estimate curve reliability. It may be more useful to focus instead on the Zipf relationship of the most frequently used words. Zooming in on the 32 most common words in the Constitution we see significant deviation from proper balance. Again, *Ulysses* and the idealized Zipf slope are shown for comparison.



[167] The Zipf values are manufactured to show an ideal slope, rather than empirical for the English language as a whole

[168] See, e.g., Steven T. Piantadosi, *Zipf's word frequency law in natural language: A critical review and future directions*, Psychonomic Bulletin & Review. 2014 Oct; 21(5): 1112.

At this resolution,[169] we see some irregular slopes, which may indicate imbalance between the internal economies of writer and reader. The 5th ranked word (*be*) would be expected to appear 1/5th (or .2 times) as often as the 1st ranked (*the*). Instead, its ratio is .31. The same is found for the 10th ranked word (*united*, .13 vs. .1) and the 20th ranked word (*may*, .08 vs. .05). The ratios for the 50th ranked word (*thereof*) and the 100th most frequent word (*between*), although not shown on the chart, are also different than predicted by Zipf (.03 vs. .02. and 02 vs. .01, respectively). The average deviation from standard Zipf's law expectancy for the 100 most common words in the Constitution is 69.5%. This lack of balance might suggest that a reader would need to expend greater effort in deciphering meaning, at least for frequently used words, while the authors enjoyed least effort. The latter makes sense when we consider that authoring the Constitution was a group effort, with a final output the result of compromise in Philadelphia. But that's not much consolation to the average reader, either in 1787 or now.

Even the slope of the most frequent words in *Ulysses*, at this scale, track Zipf's law slightly better than does the Constitution. Yes, it also has irregularities. But if the best we can say about the comprehensibility of the Constitution is that it is no less understandable than *Ulysses*, one of the most difficult English literary works, we should not be impressed.[170]

Power-law distributions have been noted for many physical and social phenomena, ranging from earthquakes to genetics to law. They generally seem to be based on some foundational statistical power-structure of complex systems.[171] For instance, the frequency of an event is often inversely proportional to its rank in severity, showing a negative slope or "thick tail distribution." Zipf's law has also been noted to apply to federal court filings and the number of district court judgeships.[172] Close ad-

---

[169] The chart uses a base 2 log scale; other bases produce the same slope, but a different x (rank) axis. To make the Ulysses comparison more visually useful, its word rank is adjusted by a factor of 1/40 so as to cross the y (frequency) axis near the Constitution's value.

[170] Zipf calculated that for *Ulysses*, the product of rank and frequency, *C,* remained remarkably consistent (24,800 – 28,680) at key locations in the first 1,000 words and throughout the book as a whole. This lead Zipf to find "clear evidence of the existence of a vocabulary balance." Zipf, supra, at location 745. In contrast, the value of *C* for the Constitution ranges from 406 to 853, again showing imbalance in the vocabulary used.

[171] *See* generally Iddo Eliazar, POWER LAWS: A STATISTICAL TREK (Springer, 2020).

[172] Thomas Bak, *Power-Law Distributions and the Federal Judiciary*, 46 Jurimetrics J. 139 (2006) (ranking judicial districts according to numbers of case filings).

herence was found to the nominal Zipf value of -1.[173] For textual analysis, examining word rank and frequency can reveal new insights. For example, a related statistical analysis was used to identify James Madison as the pseudonymous author of certain essays in the *Federalist*.[174]

A fair amount of work has been done in statistical linguistics in the 70 years since Zipf developed his empirical formula.[175] Some have questioned whether Zipf's power relationship is just an artifact of ordering, rather than indicative of communicative efficiency. I tend to agree with Zipf's defenders who find his law to be fundamental in natural languages.[176] Others have improved on the precise formula, such as by adding a constant to the rank variable, or adjusting the power exponent.[177] But, the basic rank-frequency relationship has been confirmed, even if its theoretical foundations remain elusive. What Zipf's law and its variants demonstrate is that while human languages are highly complex, they tend to have reliable structures that follow predictable patterns. Among those patterns, word frequency relationships "are one of the most basic properties of humans' communicative system and play a critical role in language processing and acquisition."[178]

Other relationships have been seen, such as between relative and absolute frequencies. Zipf found that the 1st most frequent word in a corpus typically appears approximately every 10 words; the 2nd most frequent, every 20 words; and so on. This harmonic series can be expressed as: $r_n \approx 1/10n$, "as an illustration of the high degree of orderliness with which linguistic forces operate."[179] However, this particular function may not

---

[173] *Id*. at 149-150 (finding a standardized regression coefficient of -0.953). These are not independent variables since they closely track district populations.

[174] *See* Frederick Mosteller, David L. Wallace, INFERENCE AND DISPUTED AUTHORSHIP: THE FEDERALIST (U. Chicago, 1964).

[175] *See, e.g.,* Steven T. Piantadosi, *Zipf's word frequency law in natural language: A critical review and future directions*, Psychonomic Bulletin & Review. 2014 Oct; 21(5): 1112, and authorities collected there.

[176] *See, e.g.*, Ramon Ferrer-i-Cancho and Brita Elveva, *Rndom Texts Do Not Exhibit the Real Zipf's Law-Like Rank Distribution*, PLoS ONE 5(3): e9411. doi:10.1371/journal.pone.0009411 (2010).

[177] Piantadosi, supra, n. 175, describing Mandelbrot's refinement, $f(r) \propto \frac{1}{(r+\beta)^{\alpha}}$, where β is the constant and α is the exponent (in Zipf's law, $\alpha = 1$ and no constant was added to the rank).

[178] *Id*.

[179] George K. Zipf, THE PSYCHO-BIOLOGY OF LANGUAGE: AN INTRODUCTION TO DYNAMIC PHILOLOGY (MIT Press, 1965).

be independent of Zipf's (first) law, and it is unclear what we should make of conformities and deviations, including anomalies in the Constitution.[180] My point is only that statistical analysis of the Constitution can provide insights that are hidden in traditional searches for semantic meaning.

In sum, significant deviations from the nominal Zipf's law are rare in nature and in social structures. The fact that the slope for word frequency in the Constitution differs from ordinary English usage may suggest reduced comprehensibility. If most Americans in 1787 found the draft constitutional text discordant, confusing, or requiring great effort to understand, that realization might tell us something about its "original public meaning," or lack thereof.

VI. Gödel's Incompleteness Theorems

Our final mathematical analysis of the Constitution is based on Kurt Gödel's *Incompleteness Theorems*.[181] Gödel was a colleague and close friend of Einstein at the Institute for Advanced Study in Princeton. Mathematician Jim Holt calls him "the greatest logician since Aristotle."[182] His work was popularized by Douglas Hofstadter in his award-winning book, *Gödel, Escher, Bach*, about knowledge representation, meaning, and cognition.[183]

Gödel developed two theorems regarding axioms or "proofs" for formal systems. A system of knowledge is "formal" if it contains a full rule set for the resolution of problems that may arise within the system. Thus, if a system's axioms (self-evident proofs) and inferences claim to be sufficient to solve any arbitrary problem, without the need to resort to external factors, that system is a formal one. Gödel's theorems concern the limits of provability and "truths" in formal systems. In this section I show how the theorems lead to constitutional indeterminacy under originalism.

---

[180] The 10 highest ranked words in the Constitution appear every (x) words: 1(11), 2(16), 3(23), 4(24), 5(35), 6(39), 7(50), 8(55), 9(56), 10(82).

[181] Kurt Gödel, ON FORMALLY UNDECIDABLE PROPOSITIONS OF PRINCIPIA MATHEMATICA AND RELATED SYSTEMS (1931).

[182] Jim Holt, WHEN EINSTEIN WALKED WITH GÖDEL: EXCURSIONS TO THE EDGE OF THOUGHT (Farrar, 2018), 4.

[183] Douglas Hofstadter, GÖDEL, ESCHER, BACH: AN ETERNAL GOLDEN BRAID (Basic Books, 1979). There is an international society named for Gödel to promote research in logic, philosophy and mathematics. *See* https://kgs.logic.at.

Gödel's first theorem asserts that if a formal system is internally *consistent*, it cannot be *complete*. Or, conversely, any complete system cannot be internally consistent. "Consistent" means that every axiom of the system is true. "Complete" means that the system contains all the axioms needed in order to resolve problems within the system. According to Gödel's first theorem, every formal system will confront one or more problems that have no solutions. A simple example is the sentence "This statement is false." The sentence and its truth-value are contradictory, negating any possibility of proof. Gödel generalized this, establishing that all formal systems are either incomplete or inconsistent. This was rather disturbing to logicians who thought they could always construct a system (even if only theoretical) that could prove any proposition, even its own validity. Gödel showed this was not logically feasible.

Gödel's second theorem builds on the first stating that the consistency of a system cannot be proven using only the system's axioms. In other words, one cannot prove the validity of a system axiom from within the system; resort to external factors is always necessary. A system that asserts its own consistency is inconsistent. Thus, if proposition *G* can be demonstrated as true within a supposedly complete system, then its converse, *not-G*, is also demonstrable within the system. Thus, any formal system will contain at least one statement that may be true but is unprovable. As Hofstadter describes it, "Gödel showed that provability is a weaker notion than truth, no matter what axiom system is involved."[184] Others have seen in the incompleteness theorems a rejection of the notion of "objective truth." What one investigator finds to be true in a system does not necessarily hold for other investigators. Accordingly,

> there is no nonshifting foundation on which any system rests. All truths—even those that had seemed so certain as to be immune to the very possibility of revision—are essentially manufactured. Indeed the very notion of the objectively true is a socially constructed myth.[185]

While Gödel developed his theorems for mathematical numbering systems, they have been extended to other aspects of human knowledge and constructed systems.[186] Outside of arithmetic, the theorems were first applied to computing, with Alan Turing theorizing his "Universal Turing Machine" (general purpose computer) around the same time as Gödel

---

[184] Hofstadter, *supra*, n. 183.

[185] Rebecca Goldstein, INCOMPLETENESS: THE PROOF AND PARADOX OF KURT GÖDEL (W. W. Norton, 2005) (Kindle Locations 194-196). See also *Law's Complexity, supra*, n. 88 at location 328.

[186] *See* Sigmund, 2017 (any string of symbols such as a language can be written as mathematical statements).

developed his theorems.[187] In this context, the incompleteness theorem tells us that we cannot prove *ex ante* whether a problem can be solved on a computer, irrespective of whatever algorithm that can be devised.[188] An algorithm consists of a series of steps for solving a problem. It can be expressed in any symbolic language; it need not be only mathematics. Thus, Gödel's theorems have been generalized to any problem that can be expressed in human language, not just problems involving numbers and equations,[189] including the formal system of law.[190] As shown next, Gödel's theorems undermine the notion of originalism as a complete or consistent theory. Even if it were, it could not be proved, and thus cannot serve as a basis for finding constitutional truths.

The essential claim of originalism is that the Constitution is a formal system that is both complete and consistent. It is *complete* in that all questions of constitutional law can be answered by resort to the axioms of originalism – *fixation* and *constraint* – without more. It is *consistent* in that the axioms are necessarily valid, and that all interpretations faithful to the axioms will produce the same "true" outcomes.

The fixation axiom tells us that all constitutional "truths" can be found within the document's text as it was understood at formation. A necessary postulate to that axiom is that there was in fact a single original public meaning that is now discoverable. Every interpretation is either true or false. The constraint axiom requires that all official acts conform to that original meaning.

The fixation axiom can be expressed mathematically: $I(x)_{tn} = I(x)_{t0}$ where $I(x)$ stands for the interpretation of term x, t_0 stands for time at the origin, and t_n represents any time t arbitrarily later than t_0. This axiom requires that $I(x)_{t0}$ have a fixed and objective truth value that can be identified and measured. All truth-seekers will compute the same values if properly using the equation, and be bound by it (the constraint axiom). Thus, originalists view constitutional law as a formal system, one whose outcomes are both certain and verifiable without resort to external factors such as subjective human values or other theories of interpretation.

---

[187] Turing reached the same conclusion as Gödel in 1936.

[188] According to the "Halting Problem," it is not possible to know whether a computer program will ever halt (finish running) other than by running it. No proof of halting can be devised in any programming language.

[189] *See* Gödel's Incompleteness Theorems, *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/goedel-incompleteness.

[190] *See* POLITICAL NUMERACY, *supra*. n.3, ch. 11; Mark R. Brown and Andrew C. Greenberg, *On Formally Undecidable Propositions of Law: Legal Indeterminacy and the Implications of Metamathematics*, 43 Hastings L.J. 1439 (1992).

Barnett and Bernick acknowledge that not all contemporary issues were anticipated by the precise *letter* (text) of the Constitution. When "the text runs out,"[191] the constraint principle requires judges to faithfully implement the original *spirit* of the text. That "spirit" can never be used to override original meaning; only to amplify it. Since original meaning controls in all cases, the system supposedly remains *complete* and *consistent* even in cases of ambiguity.

Gödel's first theorem tells us that the formal system of constitutional law described by originalism cannot be both complete and consistent. It is incomplete because one needs more than fixation and constraint in order to resolve problems within the system. First of all, these are merely arguments, rather than axioms in the sense of agreed proofs. Even among originalists, there are different interpretations of originalism, hence of the axioms. Second, even taking the tenets of "new originalism" as valid, they are still insufficient to resolve constitutional problems. By *themselves*, they do not tell us what truths were fixed, how to find them, how to resolve linguistic conflicts[192] or the solutions to constitutional holes.[193]

This is especially true where interpretation is bound, not by a formal statement in the Constitution (the *letter*), but by fidelity to its *spirit.* This, according to Barnett and Bernick, means the Constitution's "functions, purposes, goals or aims implicit in its individual clauses and structural design."[194] Finding those requires "good-faith constitutional construction."[195] Indeed, some form of construction is unavoidable whenever a court is called upon to make a constitutional judgment. But, "good-faith" and "conscience" are not stated in the Constitution's text, or even implicit; they are external values. They are perhaps desired, and certainly politically sound, but measurable only by external metrics.

To the extent that originalism is complete, it cannot be consistent because fixation and constraint are not self-evidently true. Historical facts themselves are matters of interpretation, rather than axiomatic. Moreover, given the complexity of many constitutional clauses, and the noise

---

[191] *Letter and Spirit*, supra, n.14 at 10.

[192] Examples of conflict include the various usages of "shall" and the different meanings of "the people," noted in *Heller*, at 554 U.S. 579-80, among others.

[193] For example, the Constitution fails to specify what happens if a victorious presidential candidate dies after the election in early November but before there is a "President-Elect" upon certification of electoral votes in early January. The problem is compounded by the decision in *Chiafalo v. Washington*, 140 S.Ct. 2316 (2020), allowing states to disable "faithless electors."

[194] *Id*. at 3.

[195] *Id*.

created by 32,000 Supreme Court opinions since the founding,[196] it is impossible to prove consistency in interpretation, even as an ideal. The word "overruled" appears 8,514 times in decisions of the Supreme Court,[197] hardly what one would expect in a consistent system.

Consider again the injunction to abide by the *spirit* of the constitution when the *letter* is undefined. This is not the type of limiting factor that produces consistent constitutional "truths," even when confined to a "narrow construction zone." I suspect there are quite a few interpretations of the meaning of "good faith" in the 7,483 instances that the term appears in Supreme Court opinions,[198] or the nearly one million state and federal cases in which it appears.[199] A judge's obligation of good faith construction is not even as formal as the explicit command in Article II that the President "take care that the laws be *faithfully executed*."[200] Yet, given the number and scope of Executive Orders over the years, that too hasn't proven to be much of a consistent "constraint."

Gödel's second theorem reinforces these conclusions. We cannot prove the validity of originalism simply by resorting to constitutional sources. Some external values, whether other historical sources or one's own ideological agenda, are necessary to justify the interpretive methodology or to supply additional axioms. Originalism is itself an ontology external to the Constitution. Some argue that the very notion of "judicial power" in Article III was understood by the founders and ratifiers to require interpretations faithful to the original public meaning of law.[201] Whether this is true or not, it still requires resort to external sources for validation. It is surely not within the *letter* of the text, and *spirit* is not a self-contained rule. The axioms of originalism – fixation and constraint – do not prove the theory; rather they are outputs of the theory, if otherwise proven.

Not only are the axioms of originalism not provable within the formal system of constitutional law, neither are its outcomes. Thus, originalist constructions necessarily rely on heuristics,[202] contextual enrichment,[203]

---

[196] https://www.english-corpora.org/scotus/

[197] *Id*. A 4-gram collocate of "overruled" and "decision" suggests that up to 182 cases have been overruled by the Court. The remainder of the 8,514 overruled matters are likely lower court decisions and particular rulings.

[198] https://www.english-corpora.org/scotus/

[199] Using the search term "good faith" on plus.lexis.com

[200] U.S. Constitution, Art. II, s. 3.

[201] *See, e.g.,* THE GOOD CONSTITUTION*, supra*, n.31 at 751.

[202] *Letter and Spirit, supra,* n.14 at 16.

[203] *Id*. at 17.

fiduciary rules,[204] contract law,[205] empirical discovery,[206] founding era political and moral philosophies,[207] and other historical sources.[208] The incompleteness of constitutional text and the spirit of individual clauses is manifest. Originalists prove this point every time they cite Blackstone, the *Federalist Papers* or other Eighteenth Century documents for their preferred constitutional outcomes. If the originalists' constitution were a complete and consistent system, external sources would be irrelevant.

For example, if within a system of constitutional originalism we try to establish proposition *G* – "the Second Amendment guarantees the right of individual gun ownership," the same system also establishes *not-G* – "the Second Amendment does not guarantee the right of individual gun ownership."[209] The majority and dissenting opinions in *Heller* establish how treating the Constitution as a formal system generates strongly opposite conclusions. This is confirmed by Barnett and Bernick's criticism of Scalia's *Heller* opinion, which "illustrates the incapacity of originalist interpretation supported by extensive research, standing alone."[210]

Gödel proposed to prove the applicability of his theorems to constitutional law, but at the most inopportune time. In 1947 he decided to become an American citizen. At his naturalization hearing, the judge asked him "do you think a dictatorship like that in Germany could ever arise in the United States?"[211] Gödel answered yes and "began explaining to the judge how the U.S. Constitution contained a loophole that would allow a dictatorship to come into existence."[212] Einstein and Oskar Morgenstern, who had accompanied Gödel to the hearing, interrupted to save him from antagonizing the judge, and he never did identify the loophole.

One does not need to resort to Gödel's naturalization challenge to see the incompleteness of originalism. Take the recent case of *Juliana v. United States*,[213] in which a number of young Americans sued to enjoin federal

---

[204] *Id*. at 19.

[205] *Id*. at 26-30 (describing obligations of contract interpretation)

[206] *Id*. at 33-34.

[207] *Id*. at 24, *et seq*.

[208] *Id*. at 35.

[209] It would do this by treating the first clause, not simply as prefatory, but as containing the application condition for the second clause.

[210] *Letter and Spirit, supra*, n.14 at 38.

[211] Jørgen Veisdal, *Kurt Gödel's Brilliant Madness*, https://medium.com/cantors-paradise/kurt-gödels-brilliant-madness-84288dd96eda

[212] Holt, *supra*, n.182 at 11.

[213] 9th Cir. No. 18-36082 (Jan. 17, 2020).

subsidies to the fossil fuel industry on the ground that promoting the use of fossil fuels contributes to climate change which endangers the planet and "a climate system capable of sustaining human life." The Ninth Circuit found the basic facts undeniable:

> A substantial evidentiary record documents that the federal government has long promoted fossil fuel use despite knowing that it can cause catastrophic climate change, and that failure to change existing policy may hasten an environmental apocalypse.[214]

Nonetheless, despite the fact that plaintiffs had alleged concrete and particularized injury resulting from federal policy, they lacked Article III standing to sue. That was because the Constitution provided no standards that a court could invoke to protect the planet, or even assure "the country's perpetuity."[215] And even though the government's actions may "irreparably devastate our Nation," the Constitution provides no relief. Of course, that would have come as a surprise to John Jay and James Madison, who argued that the principal purpose of the Constitution was to "preserve and perpetuate" the Union.[216] Still, Article III is now read, by originalist standards, as to remove courts from that equation and render the document incapable of its own preservation.[217] It seems Gödel, in his naturalization hearing, underestimated the incompleteness of the Constitution when it comes to the executive branch putting us on a path to the "Nation's willful destruction."[218]

In summary, while Gödel's theorems formally apply only to mathematics and logic, their basic premise has been extended to many other fields, including law.[219] The Constitution, as modeled by the axioms of originalism, is both *incomplete* and *inconsistent*. Originalists must resort to outside sources to support both their underlying theory and their axioms of fixation and constraint. Like any Gödel system, one must import external axioms and values in order to solve system problems.

Constitutional axioms come in many flavors. As an interpretive theory, one could just as easily insist on the ability of courts or polities to con-

---

[214] Slip op. at 11.

[215] *Id*. at 26-29.

[216] *Id* at 38 (Staton, J., dissenting) (quoting the Federalist Papers).

[217] While the Ninth Circuit did not resort to original understanding on its own, it did rely on several cases that did. See, e.g., *Trump v. Hawaii*, 138 S. Ct. 2392, 2426 (2018), *Ariz. State Legis. v. Ariz. Indep. Redistricting Com*m'n, 135 S. Ct. 2652, 2672 (2015), *Steel Co. v. Citizens for a Better Env't*, 523 U.S. 83, 128 (1998), *Pacific States Tel. & Tel. v. Oregon*, 223 U.S. 118, 138 (1912).

[218] *Id*. at 33.

[219] *Id*. at 10 (referring to physicist Roger Penrose); Goldstein, supra, n.185 at 27.

form the Constitution to contemporary values (say, to preserve the nation) without being stymied by an incomplete and defective amendment process. Otherwise, like the Articles of Confederation before it, the system may need to be dismantled in order to be saved.

VII. Conclusion

Originalism is based on the theory that the Constitution had a commonly accepted *public meaning* when received by voters in 1787, whose meaning became "fixed" upon adoption of the document. There are many criticisms of the theory, including that whatever meaning was shared by the free white males, mostly of wealth or prestige, and mostly Protestant, who drafted and adopted the Constitution[220] is not relevant to a far different country of the 21st Century.

This article presents a different critique. It uses select mathematical analyses, mostly drawn from recent developments in the field, to challenge the underlying premise of originalism. Much of the Constitution is written in terms that defy linguistic balance and regularity, rendering it less comprehensible than would be necessary to establish a common public meaning. While the document as a whole may have conveyed strong sentiments to its ratifiers, individual clauses did not. Accordingly, their meaning could not have been "fixed" in 1787, or at any time since.

More importantly, an originalist model of the Constitution is unprovable. The formal axioms of originalism are neither complete nor consistent, as shown by Kurt Gödel's incompleteness theorems. The moment we leave the four-corners of the Constitution in order to add originalist interpretation as a constitutional command, the theory becomes self-defeating. One cannot use external sources to argue for originalism, but then deny their use for other interpretations and values.

Fixation and constraint may be useful theories of interpretation, but they are not axiomatic within our constitutional system. Rather, they compete with other heuristics, such as pragmatic contemporary interpretation or "living constitutionalism." Which theory and outcomes one adopts is a matter of policy preference, not of constitutional command. As Rebecca Goldstein says, "[m]etaquestions about… law are not normally questions that are contained in the field itself; they are not … legal [questions]."[221]

---

[220] *See* U.S. NATIONAL ARCHIVES AND RECORDS ADMINISTRATION, *America's Founding Fathers*, http://www.archives.gov/exhibits/charters/constitution_founding_fathers_overview.html (last visited July 19, 2016).

[221] Goldstein, *supra*, n.185, at 27.