

The physical constants as biosignature: an anthropic retrodiction of the Selfish Biocosm Hypothesis

James N. Gardner

Gardner & Gardner, 111 S. W. Columbia Street, Suite 810, Portland, Oregon 97201, USA
e-mail: jgardner@gardnerandgardner.com

Abstract: Goal 7 of the NASA Astrobiology Roadmap states: ‘Determine how to recognize signatures of life on other worlds and on early Earth. Identify biosignatures that can reveal and characterize past or present life in ancient samples from Earth, extraterrestrial samples measured *in situ*, samples returned to Earth, remotely measured planetary atmospheres and surfaces, and other cosmic phenomena.’ The cryptic reference to ‘other cosmic phenomena’ would appear to be broad enough to include the possible identification of biosignatures embedded in the dimensionless constants of physics. The existence of such a set of biosignatures – a life-friendly suite of physical constants – is a retrodiction of the Selfish Biocosm (SB) hypothesis. This hypothesis offers an alternative to the weak anthropic explanation of our indisputably life-friendly cosmos favoured by (1) an emerging alliance of M-theory-inspired cosmologists and advocates of eternal inflation like Linde and Weinberg, and (2) supporters of the quantum theory-inspired sum-over-histories cosmological model offered by Hartle and Hawking. According to the SB hypothesis, the laws and constants of physics function as the cosmic equivalent of DNA, guiding a cosmologically extended evolutionary process and providing a blueprint for the replication of new life-friendly progeny universes.

Received 1 September 2004, accepted 28 October 2004

Key words: anthropic principle, physical constants.

Introduction

The notion that we inhabit a universe whose laws and physical constants are fine-tuned in such a way as to make it hospitable to carbon-based life is an old idea (Gardner 2003). The so-called ‘anthropic’ principle comes in at least four principal versions (Barrow & Tipler 1988) that represent fundamentally different ontological perspectives. For instance, the ‘weak anthropic principle’ is merely a tautological statement that since we happen to inhabit this particular cosmos it must perforce be life-friendly or else we would not be here to observe it. As Vilenkin (2004) put it recently, ‘the “anthropic” principle, as stated above, hardly deserves to be called a principle: it is trivially true’. By contrast, the ‘participatory anthropic principle’ articulated by Wheeler and dubbed ‘it from bit’ (Wheeler 1996) is a radical extrapolation from the Copenhagen interpretation of quantum physics and a profoundly counterintuitive assertion that the very act of observing the universe summons it into existence.

All anthropic cosmological interpretations share a common theme: a recognition that key constants of physics (as well as other physical aspects of our cosmos such as its dimensionality) appear to exhibit a mysterious fine-tuning that optimizes their collective bio-friendliness. Rees (2000) noted that virtually every aspect of the evolution of the universe – from the birth of galaxies to the origin of life on

Earth – is sensitively dependent on the precise values of seemingly arbitrary constants of nature like the strength of gravity, the number of extended spatial dimensions in our universe (three of the ten posited by M-theory), and the initial expansion speed of the cosmos following the Big Bang. If any of these physical constants had been even slightly different, life as we know it would have been impossible:

The [cosmological] picture that emerges – a map in time as well as in space – is not what most of us expected. It offers a new perspective on a how a single ‘genesis event’ created billions of galaxies, black holes, stars and planets, and how atoms have been assembled – here on Earth, and perhaps on other worlds – into living beings intricate enough to ponder their origins. There are deep connections between stars and atoms, between the cosmos and the microworld. ... Our emergence and survival depend on very special ‘tuning’ of the cosmos – a cosmos that may be even vaster than the universe that we can actually see.

As stated recently by Smolin (2004), the challenge is to provide a genuinely scientific explanation for what he terms the ‘anthropic observation’:

The anthropic observation: *Our universe is much more complex than most universes with the same laws but different values of the parameters of those laws. In particular, it has a*

complex astrophysics, including galaxies and long lived stars, and a complex chemistry, including carbon chemistry. These necessary conditions for life are present in our universe as a consequence of the complexity which is made possible by the special values of the parameters.

There is good evidence that the anthropic observation is true. Why it is true is a puzzle that science must solve.

It is a daunting puzzle indeed. The strangely (and apparently arbitrarily) biophilic quality of the physical laws and constants poses, in Greene's view, the deepest question in all of science (Greene 2004). In the words of Davies (Gardner 2003), it represents 'the biggest of the Big Questions: why is the universe bio-friendly?'

Modern history of anthropic reasoning

Modern statements of the cosmological anthropic principle date from the publication of a landmark book by Henderson in 1913 entitled *The Fitness of the Environment* (Henderson 1913). Henderson's book was an extended reflection on the curious fact that there are particular substances present in the environment – pre-eminently water – whose peculiar qualities rendered the environment almost preternaturally suitable for the origin, maintenance, and evolution of organic life. Indeed, the strangely life-friendly qualities of these materials led Henderson to the view that 'we were obliged to regard this collocation of properties in some intelligible sense a preparation for the process of planetary evolution. ... Therefore the properties of the elements must for the present be regarded as possessing a teleological character'.

Thoroughly modern in outlook, Henderson dismissed this apparent evidence that inanimate nature exhibited a teleological character as indicative of divine design or purpose. Indeed, he rejected the notion that nature's seemingly teleological quality was in any way inconsistent with Darwin's theory of evolution through natural selection. On the contrary, he viewed the bio-friendly character of the inanimate natural environment as essential to the optimal operation of the evolutionary forces in the biosphere. Absent the substrate of a superbly 'fit' inanimate environment, Henderson contended, Darwinian evolution could never have achieved what it has in terms of species multiplication and diversification.

The mystery of *why* the physical qualities of the inanimate universe happened to be so oddly conducive to life and biological evolution remained just that for Henderson – an impenetrable mystery. The best he could do to solve the puzzle was to speculate that the laws of chemistry were somehow fine-tuned in advance by some unknown cosmic evolutionary mechanism to meet the future needs of a living biosphere:

The properties of matter and the course of cosmic evolution are now seen to be intimately related to the structure of the living being and to its activities; they become, therefore, far more important in biology than has previously been suspected. For the whole evolutionary process, both cosmic and organic, is one, and the biologist may now rightly regard the Universe in its very essence as biocentric.

Henderson's iconoclastic vision was far ahead of its time. His potentially revolutionary book was largely ignored by his contemporaries or dismissed as a mere tautology. *Of course* there should be a close match-up between the physical requirements of life and the physical world that life inhabits, contemporary sceptics pointed out, since life evolved to survive the very challenges presented by that pre-organic world and to take advantage of the biochemical opportunities it offered.

While lacking broad influence at the time, Henderson's pioneering vision proved to be the precursor to modern formulations of the cosmological anthropic principle. One of the first such formulations was offered by British astronomer Fred Hoyle. A storied chapter in the history of the principle is the oft-told tale of Hoyle's prediction of the details of the triple-alpha process (Mitton 2005). This prediction, which seems to qualify as the first falsifiable implication to flow from an anthropic hypothesis, involves the details of the process by which the element carbon (widely viewed as the essential element of abiotic precursor polymers capable of autocatalysing the emergence of living entities) emerges through stellar nucleosynthesis. As noted by Livio (2003):

Carbon features in most anthropic arguments. In particular, it is often argued that the existence of an excited state of the carbon nucleus is a manifestation of fine-tuning of the constants of nature that allowed for the appearance of carbon-based life. Carbon is formed through the triple-alpha process in two steps. In the first, two alpha particles form an unstable (lifetime $\sim 10^{-16}$ s) ^8Be . In the second, a third alpha particle is captured, via $^8\text{Be}(\alpha, \gamma)^{12}\text{C}$. Hoyle argued that in order for the 3α reaction to proceed at a rate sufficient to produce the observed cosmic carbon, a resonant level must exist in ^{12}C , a few hundred keV about the $^8\text{Be} + ^4\text{He}$ threshold. Such a level was indeed found experimentally.

Other chapters in the modern history of the anthropic principle are treated comprehensively by Barrow & Tipler (1988) and will not be revisited here.

The new urgency of anthropic investigation

Two recent developments have imparted a renewed sense of urgency to investigations of the anthropic qualities of our cosmos. The first is the discovery that the value of dark energy density is exceedingly small but not quite zero – an apparent happenstance, unpredictable from first principles, with profound implications for the bio-friendly quality of our universe. As noted recently by Goldsmith (2004):

A relatively straightforward calculation [based on established principles of theoretical physics] does yield a theoretical value for the cosmological constant, but that value is greater than the measured one by a factor of about 10^{120} – probably the largest discrepancy between theory and observation science has ever had to bear.

If the cosmological constant had a smaller value than that suggested by recent observations, it would cause no trouble (just as one would expect, remembering the happy days

when the constant was thought to be zero). But if the constant were a few times larger than it is now, the universe would have expanded so rapidly that galaxies could not have endured for the billions of years necessary to bring forth complex forms of life.

The second development is the realization that M-theory – arguably the most promising contemporary candidate for a theory capable of yielding a deep synthesis of relativity and quantum physics – permits, in Bjorken's phrase (Bjorken 2004), 'a variety of string vacua, with different standard-model properties'.

M-theorists had initially hoped that their new paradigm would be 'brittle' in the sense of yielding a single mathematically unavoidable solution that uniquely explained the seemingly arbitrary parameters of the Standard Model. As Susskind (2003) has put it:

The world-view shared by most physicists is that the laws of nature are uniquely described by some special action principle that completely determines the vacuum, the spectrum of elementary particles, the forces and the symmetries. Experience with quantum electrodynamics and quantum chromodynamics suggests a world with a small number of parameters and a unique ground state. For the most part, string theorists bought into this paradigm. At first it was hoped that string theory would be unique and explain the various parameters that quantum field theory left unexplained.

This hope has been dashed by the recent discovery that the number of different solutions permitted by M-theory (which correspond to different values of Standard Model parameters) is, in Susskind's words, 'astronomical, measured not in millions or billions but in googles or googleplexes'. This development seems to deprive our most promising new theory of fundamental physics of the power to uniquely predict the emergence of anything remotely resembling our universe. As Susskind puts it, the picture of the universe that is emerging from the deep mathematical recesses of M-theory is not an 'elegant universe' but rather a Rube Goldberg device, cobbled together by some unknown process in a supremely improbable manner that just happens to render the whole ensemble fit for life. In the words of University of California theoretical physicist Steve Giddings, 'No longer can we follow the dream of discovering the unique equations that predict everything we see, and writing them on a single page. Predicting the constants of nature becomes a messy environmental problem. It has the complications of biology'¹.

Two contemporary restatements of the weak anthropic principle: eternal inflation plus M-theory and many-worlds quantum cosmology

There have been two principal approaches to the task of enlisting the weak anthropic principle to explain the mysteriously small (and thus bio-friendly) value of the density

of dark energy and the apparent happenstance by which our bio-friendly universe was selected from the enormously large 'landscape' of possible solutions permitted by M-theory, only a tiny fraction of which correspond to anything resembling the Standard Model prevalent in our cosmos.

Eternal inflation meets M-theory

The first approach, favoured by Susskind (2003), Linde (2002), Weinberg (1999), and Vilenkin (2004) among others, overlays the model of eternal inflation with the key assumption that M-theory-permitted solutions (corresponding to different values of Standard Model parameters) and dark energy density values will vary randomly from bubble universe to bubble universe within an eternally expanding ensemble variously termed a multiverse or a meta-universe. Generating a life-friendly cosmos is simply a matter of randomly reshuffling the set of permissible parameters and values a sufficient number of times until a particular Big Bang yields, against odds of perhaps a googleplex-to-one, a permutation that just happens to possess the right mix of Standard Model parameters to be bio-friendly.

Sum-over-histories quantum cosmological model

The second approach invokes a quantum theory-derived sum-over-histories cosmological model inspired by Everett's 'many worlds' interpretation of quantum physics. This approach, which has been prominently embraced by Hawking & Hertog (2002), was summarized as follows by Hogan (2004):

In the original formulation of quantum mechanics, it was said that an observation collapsed a wavefunction to one of the eigenstates of the observed quantity. The modern view is that the cosmic wavefunction never collapses, but only appears to collapse from the point of view of observers who are part of the wavefunction. When Schrödinger's cat lives or dies, the branch of the wavefunction with the dead cat also contains observers who are dealing with a dead cat, and the branch with the live cat also contains observers who are petting a live one.

Although this is sometimes called the 'Many Worlds' interpretation of quantum mechanics, it is really about having just one world, one wavefunction, obeying the Schrödinger equation: the wavefunction evolves linearly from one time to the next based on its previous state.

Anthropic selection in this sense is built into physics at the most basic level of quantum mechanics. Selection of a wavefunction branch is what drives us into circumstances in which we thrive. Viewed from a disinterested perspective outside the universe, it looks like living beings swim like salmon up their favorite branches of the wavefunction, chasing their favorite places.

Hawking & Hertog (2002) have explicitly characterized this 'top down' cosmological model as a restatement of the weak anthropic principle:

We have argued that because our universe has a quantum origin, one must adopt a top down approach to the problem

¹ <http://www.edge.org/discourse/landscape.html>.

of initial conditions in cosmology, in which histories that contribute to the path integral, depend on the observable being measured. There is an amplitude for empty flat space, but it is not of much significance. Similarly, the other bubbles in an eternally inflating spacetime are irrelevant. They are to the future of our past light cone, so they don't contribute to the action for observables and should be excised by Ockham's razor. *Therefore, the top down approach is a mathematical formulation of the weak anthropic principle.* Instead of starting with a universe and asking what a typical observer would see, one specifies the amplitude of interest.

Critique of contemporary restatements of the weak anthropic principle

Apart from the objections on the part of those who oppose in principle any use of the anthropic principle in cosmology, there are at least three reasons why both the Hawking/Hogan and the Susskind/Linde/Weinberg restatements of the weak anthropic principle are objectionable.

First, both approaches appear to be resistant (at the very least) to experimental testing. Universes spawned by Big Bangs other than our own are inaccessible from our own universe, at least with the experimental techniques currently available to science. So too are quantum wavefunction branches that we cannot, in principle, observe. Accordingly, both approaches appear to be untestable – perhaps untestable in principle. For this reason, Smolin (2004) recently argued 'not only is the Anthropic Principle not science, its role may be negative. To the extent that the Anthropic Principle is espoused to justify continued interest in unfalsifiable theories, it may play a destructive role in the progress of science'.

Second, both approaches violate the mediocrity principle. The mediocrity principle, a mainstay of scientific theorizing since Copernicus, is a statistically based rule of thumb that, in the absence of contrary evidence, a particular sample (Earth, for instance, or our particular universe) should be assumed to be a typical example of the ensemble of which it is a part. The Susskind/Linde/Weinberg approach, in particular, flouts this principle. Their approach simply takes refuge in a brute, unfathomable mystery – the conjectured lucky roll of the dice in a crap game of eternal inflation – and declines to probe seriously into the possibility of a naturalistic cosmic evolutionary process that has the capacity to yield a life-friendly set of physical laws and constants on a non-random basis.

Third, both approaches extravagantly inflate the probabilistic resources required to explain the phenomenon of a life-friendly cosmos. (Think of a googleplex of monkeys typing away randomly until one of them, by pure chance, accidentally composes a set of equations that correspond to the Standard Model.) This should be a hint that something fundamental is being overlooked and that there may exist an unknown natural process, perhaps functionally akin in some manner to terrestrial evolution, capable of effecting the emergence and prolongation of physical states of nature that are, in the abstract, vanishingly improbable.

The Darwinian precedent

Hogan (2004) has analogized the quantum theory-inspired sum-over-histories version of the weak anthropic principle to Darwinian theory:

This blending of empirical cosmology and fundamental physics is reminiscent of our Darwinian understanding of the tree of life. The double helix, the four-base codon alphabet and the triplet genetic code for amino acids, any particular gene for a protein in a particular organism – all are frozen accidents of evolutionary history. It is futile to try to understand or explain these aspects of life, or indeed any relationships in biology, without referring to the way the history of life unfolded. In the same way that (in Dobzhansky's phrase), 'nothing in biology makes sense except in the light of evolution,' physics in these models only makes sense in the light of cosmology.

Ironically, Hogan misses the key point that neither the branching wavefunction nor the eternal inflation-plus-M-theory versions of the weak anthropic principle hypothesize the existence of anything corresponding to the main action principle of Darwin's theory: natural selection. Both restatements of the weak anthropic principle are analogous, not to Darwin's approach, but rather to a mythical alternative history in which Darwin, contemplating the storied tangled bank (the arresting visual image with which he concludes *The Origin of Species*), had confessed not a magnificent obsession with gaining an understanding of the mysterious natural processes that had yielded 'endless forms most beautiful and most wonderful', but rather a smug satisfaction that of course the Earthly biosphere must have somehow evolved in a just-so manner mysteriously friendly to humans and other currently living species, or else Darwin and other humans would not be around to contemplate it.

Indeed, the situation that confronts cosmologists today is reminiscent of that which faced biologists before Darwin propounded his revolutionary theory of evolution through natural selection. Darwin confronted the seemingly miraculous phenomenon of a fine-tuned natural order in which every creature and plant appeared to occupy a unique and well-designed niche. Refusing to surrender to the brute mystery posed by the appearance of nature's design, Darwin masterfully deployed the art of metaphor² to elucidate a radical hypothesis – the origin of species through natural selection – that explained the apparent miracle as a natural phenomenon.

A significant lesson drawn from Darwin's experience is important to note at this point. Answering the question of why the most eminent geologists and naturalists had, until

² The metaphor furnished by the familiar process of artificial selection was Darwin's crucial stepping stone. Indeed, the practice of artificial selection through plant and animal breeding was the primary intellectual model that guided Darwin in his quest to solve the mystery of the origin of species and to demonstrate in principle the plausibility of his theory that variation and natural selection were the prime movers responsible for the phenomenon of speciation.

shortly before publication of *The Origin of Species*, disbelieved in the mutability of species, Darwin responded that this false conclusion was ‘almost inevitable as long as the history of the world was thought to be of short duration’. It was geologist Charles Lyell’s speculations on the immense age of Earth that provided the essential conceptual framework for Darwin’s new theory. Lyell’s vastly expanded stretch of geological time provided an ample temporal arena in which the forces of natural selection could sculpt and reshape the species of Earth and achieve nearly limitless variation.

The central point for the purposes of this paper is that collateral advances in sciences seemingly far removed from cosmology (complexity theory and evolutionary theory among them) can help dissipate the intellectual limitations imposed by common sense and naïve human intuition. And, in an uncanny reprise of the Lyell/Darwin intellectual synergy, it is a realization of the vastness of time and history that gives rise to the novel theoretical possibility to be discussed subsequently. Only, in this instance, it is the vastness of future time and future history that is of crucial importance. In particular, sharp attention must be paid to the key conclusion of Wheeler: most of the time available for life and intelligence to achieve their ultimate capabilities lies in the distant cosmic future, not in the cosmic past. As Tipler (1994) has stated, ‘Almost all of space and time lies in the future. By focusing attention only on the past and present, science has ignored almost all of reality. Since the domain of scientific study is the whole of reality, it is about time science decided to study the future evolution of the universe’. The next section of this paper describes an attempt to heed these admonitions.

The Selfish Biocosm hypothesis

In a paper published in *Complexity* (Gardner 2000), I first advanced the hypothesis that the anthropic qualities which our universe exhibits might be explained as incidental consequences of a cosmic replication cycle in which the emergence of a cosmologically extended biosphere could conceivably supply two of the logically essential elements of self-replication identified by von Neumann (1948): a controller and a duplicating device. The hypothesis proposed in that paper was an attempt to extend and refine Smolin’s (1977) conjecture that the majority of the anthropic qualities of the universe can be explained as incidental consequences of a process of cosmological replication and natural selection (CNS) whose utility function is black hole maximization. Smolin’s conjecture differs crucially from the concept of eternal inflation advanced by Linde (1998) in that it proposes a cosmological evolutionary process with a specific and discernible utility function – black hole maximization. It is this aspect of Smolin’s conjecture rather than the specific utility function he advocates that renders his theoretical approach genuinely novel.

As demonstrated previously (Rees 1997; Baez 1998), Smolin’s conjecture suffers from two evident defects: (1) the fundamental physical laws and constants do not, in fact,

appear to be fine-tuned to favour black hole maximization; and (2) no mechanism is proposed corresponding to two logically required elements of any von Neumann self-replicating automaton: a controller and a duplicator³. The latter are essential elements of any replicator system capable of Darwinian evolution, as noted by Dawkins (Gardner 2000) in a critique of Smolin’s conjecture:

Note that any Darwinian theory depends on the prior existence of the strong phenomenon of heredity. There have to be self-replicating entities (in a population of such entities) that spawn daughter entities more like themselves than the general population.

Theories of cosmological eschatology previously articulated (Dyson 1988; Wheeler 1996; Kurzweil 1999) predict that the ongoing process of biological and technological evolution is sufficiently robust and unbounded that, in the far distant future, a cosmologically extended biosphere could conceivably exert a global influence on the physical state of the cosmos. A related set of insights from complexity theory (Gardner 2000) indicates that the process of emergence resulting from such evolution is essentially unbounded.

A synthesis of these two sets of insights yielded the two key elements of the Selfish Biocosm (SB) hypothesis. The essence of that synthesis is that the ongoing process of biological and technological evolution and emergence could conceivably

³ Both defects were emphasized by Susskind in a recent on-line exchange with Smolin which appears at www.edge.org. Smolin has argued that his CNS hypothesis has not been falsified on the first ground (Smolin 2004) but conceded that his conjecture lacks any hypothesized mechanism that would endow the putative process of proliferation of black-hole-prone universes with a heredity function:

The hypothesis that the parameters p change, on average, by small random amounts, should be ultimately grounded in fundamental physics. We note that this is compatible with string theory, in the sense that there are a great many string vacua, which likely populate the space of low energy parameters well. It is plausible that when a region of the universe is squeezed to Planck densities and heated to Planck temperatures, phase transitions may occur leading to a transition from one string vacua to another. But there have so far been no detailed studies of these processes which would check the hypothesis that the change in each generation is small.

As Smolin noted in the same paper, it is crucial that such a mechanism exist in order to avoid the conclusion that each new universe’s set of physical laws and constants would constitute a merely random sample of the vast parameter space permitted by the extraordinarily large ‘landscape’ of M-theory-allowed solutions:

It is important to emphasize that the process of natural selection is very different from a random sprinkling of universes on the parameter space P . This would produce only a uniform distribution $p_{\text{random}}(p)$. To achieve a distribution peaked around the local maxima of a fitness function requires the two conditions specified. The change in each generation must be small so that the distribution can ‘climb the hills’ in $F(p)$ rather than jump around randomly, and so it can stay in the small volume of P where $F(p)$ is large, and not diffuse away. This requires many steps to reach local maxima from random starts, which implies that long chains of descendants are needed.

function as a von Neumann controller and that a cosmologically extended biosphere could, in the very distant future, function as a von Neumann duplicator in a hypothesized process of cosmological replication.

In a paper published in *Acta Astronautica* (Gardner 2001) I suggested that a falsifiable implication of the SB hypothesis is that the process of the progression of the cosmos through critical epigenetic thresholds in its life cycle, while perhaps not strictly inevitable, is relatively robust. One such critical threshold is the emergence of human-level and higher intelligence, which is essential to the eventual scaling up of biological and technological processes to the stage at which those processes could conceivably exert a global influence on the state of the cosmos. Four specific tests of the robustness of the emergence of human-level and higher intelligence were proposed.

In a subsequent paper published in the *Journal of the British Interplanetary Society* (Gardner 2002) I proposed that an additional falsifiable implication of the SB hypothesis is that there exists a plausible final state of the cosmos that exhibits maximal computational potential. This predicted final state appears to be consistent with both the modified ekpyrotic cyclic universe scenario (Khoury *et al.* 2001; Steinhardt & Turok 2001) and with Lloyd's (2000) description of the physical attributes of the ultimate computational device: a computer as powerful as the laws of physics will allow.

Key retrodiction of the SB hypothesis: a life-friendly cosmos

The central assertions of the SB hypothesis are: (1) that highly evolved life and intelligence play an essential role in a hypothesized process of cosmic replication; and (2) that the peculiarly life-friendly laws and physical constants that prevail in our universe – an extraordinarily improbable ensemble that Pagels (1983) dubbed the cosmic code – play a cosmological role functionally equivalent to that of DNA in an Earthly organism: they provide a recipe for cosmic ontogeny and a blueprint for cosmic reproduction. Thus, a key retrodiction of the SB hypothesis is that the suite of physical laws and constants that prevail in our cosmos will, in fact, be life-friendly. Moreover – and alone among the various cosmological scenarios offered to explain the phenomenon of a bio-friendly universe – the SB hypothesis implies that this suite of laws and constants comprise a robust program that will reliably generate life and advanced intelligence just as the DNA of a particular species constitutes a robust program that will reliably generate individual organisms that are members of that particular species. Indeed, because the hypothesis asserts that sufficiently evolved intelligent life serves as a von Neumann duplicator in a putative process of cosmological replication, the biophilic quality of the suite emerges as a retrodicted biosignature of the putative duplicator and duplication process within the meaning of Goal 7 of the NASA Astrobiology Roadmap, which provides in pertinent part:

Determine how to recognize signatures of life on other worlds and on early Earth. Identify biosignatures that can

reveal and characterize past or present life in ancient samples from Earth, extraterrestrial samples measured *in situ*, samples returned to Earth, remotely measured planetary atmospheres and surfaces, and other cosmic phenomena.

Does this retrodiction qualify as a valid scientific test of the validity of the SB hypothesis? I propose that it may, provided two additional qualifying criteria are satisfied:

- the underlying hypothesis must enjoy consilience⁴ with mainstream scientific paradigms and conjectural frameworks (in particular, complexity theory, evolutionary theory, M-theory, and theoretically acceptable conjectures by mainstream cosmologists concerning the feasibility, at least in principle, of 'baby universe' fabrication); and
- the retrodiction must be augmented by falsifiable predictions of phenomena implied by the SB hypothesis but not yet observed.

Retrodiction as a tool for testing scientific hypotheses

There is a lively literature debating the propriety of employing retrodiction as a tool for testing scientific hypotheses (Oldershaw 1988; Gee 1999; Cleland 2001, 2002). Oldershaw (1988) has discussed the use of falsifiable retrodiction (as opposed to falsifiable prediction) as a tool of scientific investigation:

A second type of prediction is actually not a prediction at all, but rather a 'retrodiction.' For example, the anomalous advance of the perihelion of Mercury had been a tiny thorn in the side of Newtonian gravitation long before general relativity came upon the scene. Einstein found that his theory correctly 'predicted,' actually retrodicted, the numerical value of the perihelion advance. The explanation of the unexpected result of the Michelson-Morley experiment (constancy of the velocity of light) in terms of special relativity is another example.

As he went on to note, 'Retrodictions usually represent falsification tests; the theory is probably wrong if it fails the test, but should not necessarily be considered right if it passes the test since it does not involve a definitive prediction'. Despite their legitimacy as falsification tests of hypotheses, falsifiable retrodictions are qualitatively inferior to falsifiable predictions, in Oldershaw's view:

But, in the final analysis, only true definitive predictions can justify the promotion of a theory from being viewed as one of many plausible hypotheses to being recognized as the best available approximation of how nature actually works. A theory that cannot generate definitive predictions, or whose definitive predictions are impossible to test, can be regarded as inherently untestable.'

⁴ Wilson has identified consilience as one of the 'diagnostic features of science that distinguishes it from pseudoscience' (Wilson 1998):

The explanations of different phenomena most likely to survive are those that can be connected and proved consistent with one another.

A less sympathetic view concerning the validity of retrodiction as a scientific tool was offered by Gee (1999), who dismissed the legitimacy of all historical hypotheses on the ground that ‘they can never be tested by experiment, and so they are unscientific ... No science can ever be historical’. This viewpoint, in turn, has been challenged by Cleland who contends that ‘when it comes to testing hypotheses, historical science is not inferior to classical experimental science’ but simply exploits the available evidence in a different way:

There [are] fundamental differences in the methodology used by historical and experimental scientists. Experimental scientists focus on a single (sometimes complex) hypothesis, and the main research activity consists in repeatedly bringing about the test conditions specified by the hypothesis, and controlling for extraneous factors that might produce false positives and false negatives. Historical scientists, in contrast, usually concentrate on formulating multiple competing hypotheses about particular past events. Their main research efforts are directed at searching for a smoking gun, a trace that sets apart one hypothesis as providing a better causal explanation (for the observed traces) than do the others. These differences in methodology do not, however, support the claim that historical science is methodologically inferior, because they reflect an objective difference in the evidential relations at the disposal of historical and experimental researchers for evaluating their hypotheses.

Cleland’s approach has the merit of preserving as ‘scientific’ some of the most important hypotheses advanced in such historical fields of inquiry as geology, evolutionary biology, cosmology, palaeontology, and archaeology. As Cleland (2002) has noted:

Experimental research is commonly held up at the paradigm of successful (a.k.a. good) science. The role classically attributed to experiment is that of testing hypotheses in controlled laboratory settings. Not all scientific hypotheses can be tested in this manner, however. Historical hypotheses about the remote past provide good examples. Although fields such as paleontology and archaeology provide the familiar examples, historical hypotheses are also common in geology, biology, planetary science, astronomy, and astrophysics. The focus of historical research is on explaining existing natural phenomena in terms of long past causes. Two salient examples are the asteroid-impact hypothesis for the extinction of the dinosaurs, which explains the fossil record of the dinosaurs in terms of the impact of a large asteroid, and the ‘big-bang’ theory of the origin of the universe, which explains the puzzling isotropic three-degree background radiation in terms of a primordial explosion. Such work is significantly different from making a prediction and then artificially creating a phenomenon in a laboratory.

In a paper presented to the *2004 Astrobiology Science Conference*, Cleland (2004) extended this analytic framework to the consideration of putative biosignatures as evidence of the past or present existence of extraterrestrial life. Acknowledging that ‘because biosignatures represent indirect

traces (effects) of life, much of the research will be historical (vs. experimental) in character even in cases where the traces represent recent effects of putative extant organisms’, Cleland concluded that it was appropriate to employ the methodology that characterizes successful historical research:

Successful historical research is characterized by (1) the proliferation of alternative competing hypotheses in the face of puzzling evidence and (2) the search for more evidence (a ‘smoking gun’) to discriminate among them.

From the perspective of the evidentiary standards applicable to historical science in general and astrobiology in particular, the key retrodiction of the SB hypothesis – that the fundamental constants of nature that comprise the Standard Model as well as other physical features of our cosmos (included the number of extended physical dimensions and the extremely low value of dark energy) will be collectively bio-friendly – appears to constitute a legitimate scientific test of the hypothesis. Moreover, within the framework of Goal 7 of the NASA Astrobiology Roadmap, the retrodicted biophilic quality of our universe appears, under the SB hypothesis, to constitute a possible biosignature.

Caution regarding the use of retrodiction to test the SB hypothesis

Because the SB hypothesis is radically novel and because the use of falsifiable retrodiction as a tool to test such an hypothesis creates at least the appearance of a ‘confirmatory argument[t] resemble[ing] just-so stories (Rudyard Kipling’s fanciful stories, e.g. how leopards got their spots)’ (Cleland 2001), it is important (as noted previously) that two additional criteria be satisfied before this retrodiction can be considered a legitimate test of the hypothesis:

- the SB hypothesis must generate falsifiable predictions as well as falsifiable retrodictions; and
- the SB hypothesis must be consilient with key theoretical constructs in such ‘adjoining’ area of scientific investigation as M-theory, cosmogenesis, complexity theory, and evolutionary theory.

As argued at length elsewhere (Gardner 2003), the SB hypothesis is both consilient with central concepts in these ‘adjoining’ fields and fully capable of generating falsifiable predictions.

Concluding remarks

In his book *The Fifth Miracle* Davies (1999) offered this interpretation of NASA’s view that the presence of liquid water on an alien world was a reliable marker of a life-friendly environment:

In claiming that water means life, NASA scientists are ... making – tacitly – a huge and profound assumption about the nature of nature. They are saying, in effect, that the laws of the universe are cunningly contrived to coax life into being against the raw odds; that the mathematical principles of physics, in their elegant simplicity, somehow know in

advance about life and its vast complexity. If life follows from [primordial] soup with causal dependability, the laws of nature encode a hidden subtext, a cosmic imperative, which tells them: 'Make life!' And, through life, its by-products: mind, knowledge, understanding. It means that the laws of the universe have engineered their own comprehension. This is a breathtaking vision of nature, magnificent and uplifting in its majestic sweep. I hope it is correct. It would be wonderful if it were correct. But if it is, it represents a shift in the scientific world-view as profound as that initiated by Copernicus and Darwin put together.

An emerging consensus among mainstream physicists and cosmologists is that the particular universe we inhabit appears to confirm what Smolin calls the 'anthropic observation': the laws and constants of nature seem to be fine-tuned, with extraordinary precision and against enormous odds, to favour the emergence of life and its by-product, intelligence. As Dyson put it eloquently more than two decades ago (Dyson 1979):

The more I examine the universe and study the details of its architecture, the more evidence I find that the universe in some sense must have known that we were coming. There are some striking examples in the laws of nuclear physics of numerical accidents that seem to conspire to make the universe habitable.

Why this should be so remains a profound mystery. Indeed, the mystery has deepened considerably with the recent discovery of the inexplicably tiny value of dark energy density and the realization that M-theory encompasses an unfathomably vast landscape of possible solutions, only a minute fraction of which correspond to anything resembling the universe that we inhabit.

Confronted with such a deep mystery, the scientific community ought to be willing to entertain plausible explanatory hypotheses that may appear to be unconventional or even radical. However, such hypotheses, to be taken seriously, must:

- be consistent with the key paradigms of 'adjoining' scientific fields;
- generate falsifiable predictions; and
- generate falsifiable retrodictions.

The SB hypothesis satisfies these criteria. In particular, it generates a falsifiable retrodiction that the physical laws and constants that prevail in our cosmos will be biophilic – which they are.

References

- Baez, J. (1998). On-line commentary on *The Life of the Cosmos* (available at <http://www.aleph.se/Trans/Global/Omega/smolin.txt>).
- Barrow, J. & Tipler, F. (1988). *The Anthropic Cosmological Principle*. Oxford University Press.
- Bjorken, J. (2004). The classification of universes, astro-ph/0404233.
- Cleland, C. (2001). Historical science, experimental science, and the scientific method. *Geology* **29**, 978–990.
- Cleland, C. (2002). Methodological and epistemic differences between historical science and experimental science. *Phil. Sci.* **69**, 474–496.
- Cleland, C. (2004). Historical science and the use of biosignatures. Unpublished summary of presentation abstracted in *Int. J. Astrobiol. Supplement* 2004, 119.
- Davies, P. (1999). *The Fifth Miracle*. Simon and Schuster.
- Dyson, F. (1979). *Disturbing the Universe*. Harper and Row.
- Dyson, F. (1988). *Infinite in All Directions*. Harper and Row.
- Gardner, J. (2000). The Selfish Biocosm: complexity as cosmology. *Complexity* **5**(3), 34–45.
- Gardner, J. (2001). Assessing the robustness of the emergence of intelligence: testing the Selfish Biocosm hypothesis. *Acta Astronautica* **48** (5–12), 951–955.
- Gardner, J. (2002). Assessing the computational potential of the eschaton: testing the Selfish Biocosm hypothesis. *J. British Interplanet. Soc.* **55**(7/8), 285–288.
- Gardner, J. (2003). *Biocosm*. Inner Ocean Publishing.
- Gee, H. (1999). *In Search of Deep Time*. The Free Press.
- Goldsmith, D. (2004). The best of all possible worlds. *Natural History* **5**(6), 44–49.
- Greene, B. (2004). *The Fabric of the Cosmos*. Knopf.
- Hawking, S. & Hertog, T. (2002). Why does inflation start at the top of the hill?. hep-th/0204212.
- Henderson, L. (1913). *The Fitness of the Environment*. Harvard University Press.
- Hogan, C. (2004). Quarks, electrons, and atoms in closely related universes, astro-ph/0407086.
- Khoury, J., Ovrut, B.A., Seiberg, N., Steinhardt, P. & Turok, N. (2001). From Big Crunch to Big Bang, hep-th/0108187.
- Kurzweil, R. (1999). *The Age of Spiritual Machines*. Viking.
- Linde, A. (2002). Inflation, quantum cosmology and the anthropic principle, hep-th/0211048.
- Linde, A. (1998). The self-reproducing inflationary universe. *Scientific American* **9**(20), 98–104.
- Livio, M. (2003). Cosmology and life, astro-ph/0301615.
- Lloyd, S. (2000). Ultimate physical limits to computation. *Nature* **406**, 1047–1054.
- Mitton, S. (2005). *Conflict in the Cosmos: Fred Hoyle's Life in Science*. Joseph Henry Press.
- Oldershaw, R. (1988). The new physics: physical or mathematical science? *Amer. J. Phys.* **56**(12), 1075–1081.
- Pagels, H. (1983). *The Cosmic Code*. Bantam.
- Rees, M. (1997). *Before the Beginning*. Addison-Wesley.
- Rees, M. (2000). *Just Six Numbers*. Basic Books.
- Smolin, L. (1997). *The Life of the Cosmos*. Oxford University Press.
- Smolin, L. (2004). Scientific alternatives to the Anthropic Principle, hep-th/0407213.
- Steinhardt, P. & Turok, N. (2001). 'Cosmic evolution in a cyclic universe, hep-th/0111098.
- Susskind, L. (2003). The anthropic landscape of string theory, hep-th/0302219.
- Tipler, F. (1994). *The Physics of Immortality*. Doubleday.
- Vilenkin, A. (2004). Anthropic predictions: the case of the cosmological constant, astro-ph/0407586.
- von Neumann, J. (1951). *On the General and Logical Theory of Automata*. In *Cerebral Mechanisms in Behavior*. The Hixon Symposium (L.A. Jeffren ed.) New York: John Wiley and Sons, pp 1–41.
- Weinberg, S. (1999). A designer universe? *New York Review of Books*, 21 October.
- Wheeler, J. (1996). *At Home in the Universe*. AIP Press.
- Wilson, E.O. (1998). Scientists, scholars, knaves and fools. *American Scientist* **86**, 6–7.