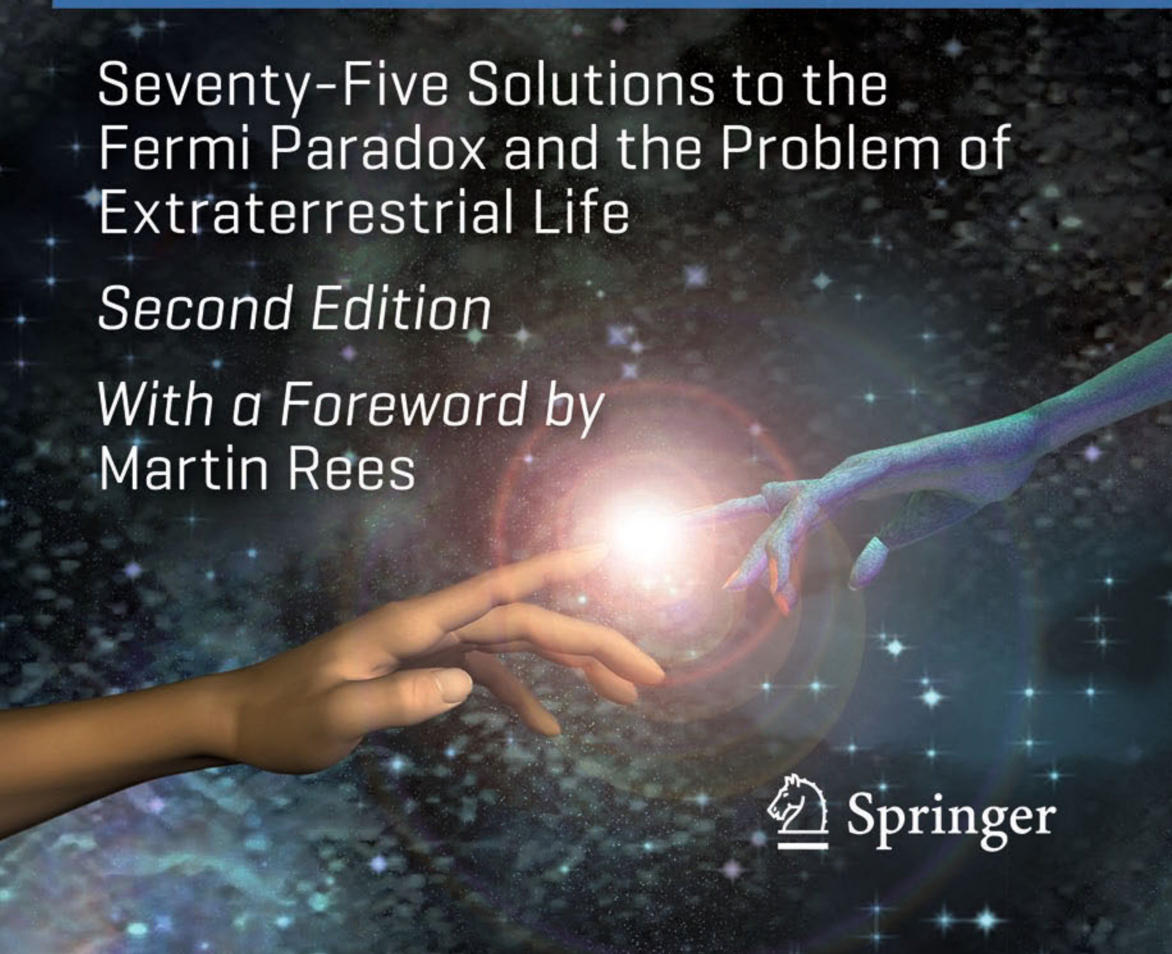
Stephen Webb

# If the Universe Is Teeming with Aliens ... Where Is Everybody?

Seventy-Five Solutions to the Fermi Paradox and the Problem of Extraterrestrial Life

*Second Edition*

*With a Foreword by*
Martin Rees



Springer

Science and Fiction



More information about this series at
http://www.springer.com/series/11657

## Science and Fiction – A Springer Series

This collection of entertaining and thought-provoking books will appeal equally to science buffs, scientists and science-fiction fans. It was born out of the recognition that scientific discovery and the creation of plausible fictional scenarios are often two sides of the same coin. Each relies on an understanding of the way the world works, coupled with the imaginative ability to invent new or alternative explanations—and even other worlds. Authored by practicing scientists as well as writers of hard science fiction, these books explore and exploit the borderlands between accepted science and its fictional counterpart. Uncovering mutual influences, promoting fruitful interaction, narrating and analyzing fictional scenarios, together they serve as a reaction vessel for inspired new ideas in science, technology, and beyond.

Whether fiction, fact, or forever undecidable: the Springer Series "Science and Fiction" intends to go where no one has gone before!

Its largely non-technical books take several different approaches. Journey with their authors as they

- Indulge in science speculation—describing intriguing, plausible yet unproven ideas;
- Exploit science fiction for educational purposes and as a means of promoting critical thinking;
- Explore the interplay of science and science fiction—throughout the history of the genre and looking ahead;
- Delve into related topics including, but not limited to: science as a creative process, the limits of science, interplay of literature and knowledge;
- Tell fictional short stories built around well-defined scientific ideas, with a supplement summarizing the science underlying the plot.

Readers can look forward to a broad range of topics, as intriguing as they are important. Here just a few by way of illustration:

- Time travel, superluminal travel, wormholes, teleportation
- Extraterrestrial intelligence and alien civilizations
- Artificial intelligence, planetary brains, the universe as a computer, simulated worlds
- Non-anthropocentric viewpoints
- Synthetic biology, genetic engineering, developing nanotechnologies
- Eco/infrastructure/meteorite-impact disaster scenarios
- Future scenarios, transhumanism, posthumanism, intelligence explosion
- Virtual worlds, cyberspace dramas
- Consciousness and mind manipulation

Stephen Webb

# If the Universe Is Teeming with Aliens . . . WHERE IS EVERYBODY?

## Seventy-Five Solutions to the Fermi Paradox and the Problem of Extraterrestrial Life

Second Edition

Springer

Stephen Webb
DCQE
University of Portsmouth
Portsmouth
United Kingdom

ISSN 2197-1188 ISSN 2197-1196 (electronic)
Science and Fiction
ISBN 978-3-319-13235-8 ISBN 978-3-319-13236-5 (eBook)
DOI 10.1007/978-3-319-13236-5

Library of Congress Control Number: 2015930256

Springer Cham Heidelberg New York Dordrecht London


Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

*To Heike and Jessica*

# Foreword

"Are we alone in the universe?" is one of the oldest and most universal questions. For a century or more it has stimulated brilliant science fiction—and it's now incentivizing real science and exploration. But we still lack evidence—indeed we know too little to say whether intelligent aliens are likely or unlikely to exist. That's why we need all the arguments that can be mustered. And that's why this book will be such a stimulus to all enquiring minds.

There may be simple organisms on Mars, or remnants of creatures that lived early in the planet's history; and there could be life, too, in the ice-covered oceans of Jupiter's moon Europa, or Saturn's moon Enceladus. But few would bet on this; and certainly nobody expects a complex biosphere in such locations. For that, we must look to the distant stars—far beyond the range of any probe we can now construct.

Prospects are here far brighter. In the last twenty years (and especially in the last five) the night sky has become far more interesting, and far more enticing to explorers, than it was to our forebears. Astronomers have discovered that many stars—perhaps even most—are orbited by retinues of planets, just like the Sun is. These planets aren't generally detectable directly. Instead, they reveal their presence by effects on their parent star that can be detected by precise measurements: small periodic motions in the star induced by an orbiting planet's gravity, and slight recurrent dimmings in a star's brightness when a planet transits in front of it, blocking out a small fraction of its light.

There is special interest in possible "twins" of our Earth—planets the same size as ours, orbiting other Sun-like stars, on orbits with temperatures such that water neither boils nor stays frozen. The Kepler spacecraft has identified many of these, and we can confidently infer that there are billions in our Galaxy.

Within twenty years the next generation of telescopes will image the nearest of these planets. Will there be life on them? We know too little about how life began on Earth to lay confident odds. What triggered the transition from complex molecules to entities that can metabolize and reproduce? It might have involved a fluke so rare that it happened only once in the entire Galaxy. On the other hand, this crucial transition might have been almost inevitable given the "right" environment. We just don't know—nor do we know if the DNA/RNA chemistry of terrestrial life is the only possibility, or just one chemical basis among many options that could be realized elsewhere.

Furthermore, even if simple life is widespread, we can't assess the odds that it evolves into a complex biosphere. And, even it did, the outcome might anyway be unrecognizably different. I won't hold my breath, but the SETI programme is a worthwhile gamble—because success in the search would carry the momentous message that concepts of logic and physics (if not consciousness) aren't limited to the hardware in human skulls.

Moreover, it might be too anthropocentric to limit attention to Earth-like planets. Science fiction writers have other ideas—balloon-like creatures floating in the dense atmospheres of Jupiter-like planets, swarms of intelligent insects, nanoscale robots etc. Perhaps life can flourish on planets flung into the frozen darkness of interstellar space, whose main warmth comes from internal radioactivity (the process that heats Earth's core). There could even be diffiuse living structures floating freely in interstellar clouds; such entities would live (and, if intelligent, think) in slow motion, but perhaps come into their own in the far future—like the "Black Cloud" envisaged by my Cambridge mentor Fred Hoyle.

No life would survive on a planet whose central Sun-like star became a giant and blew off its outer layers. Such considerations remind us of the transience of inhabited worlds (and life's imperative to escape their bonds eventually). We should also be mindful that seemingly artificial signals could come from super-intelligent (though not necessarily conscious) computers, created by a race of alien beings that had already died out.

Maybe we will one day find ET. On the other hand, this book offers 75 reasons why SETI searches may fail; Earth's intricate biosphere may be unique. That would disappoint the searchers, but it would have an upside: it would entitle us humans to be less "cosmically modest". Moreover, this outcome would not render life a cosmic sideshow. Evolution may still be nearer its beginning than its end. Our Solar System is barely middle aged and, if humans avoid self-destruction, the post-human era beckons. Life from Earth could spread through the Galaxy, evolving into a teeming complexity far beyond what we can even conceive. If so, our tiny planet—this pale blue dot floating in space—could be the most important place in the entire Galaxy, and the first interstellar voyagers from Earth would have a mission that would resonate through the entire Galaxy and perhaps beyond.

This debate will continue for decades. And Stephen Webb has condensed, within just one highly entertaining book, a fascinating cornucopia of arguments and speculation that will enrich the debate. We should be grateful to him.

Martin Rees, Astronomer Royal

# Preface to the Second Edition

I'd like to thank Chris Caron of Springer, both for his suggestion that I should update *Where Is Everybody?* and for his encouragement throughout the painful process of updating. I'm pleased that the second edition of the book will appear in Springer's Science & Fiction series, the brainchild of Chris and his colleague Angela Lahee, because any discussion of the Fermi paradox sits at that stimulating intersection between science and science fiction. A dozen years after the publication of the first edition I believe even more strongly that Fermi's question is one of the most pressing problems in science, but it remains the case that SF authors have contributed at least as much to the debate as professional scientists.

I've discussed the Fermi paradox with too many people over the years to mention them all by name, but I would particularly like to thank Milan Ćirković, Mike Lampton, Colin McInnes, Anders Sandberg, David Waltham and Willard Wells for sharing ideas, papers and manuscripts with me.

And of course I must thank Heike and Jessica, who make all this worthwhile.

Stephen Webb
Lee on the Solent, July 2014

# Preface to the First Edition

This book is about the Fermi paradox—the contradiction between the apparent absence of aliens, and the common expectation that we should see evidence of their existence. I was fascinated by the paradox when I first met it, some 17 years ago, and it fascinates me still. Over those years, many authors (too many to mention here, though their names appear in the reference list at the back of this book) have enthralled me with their writing about the paradox. Their influence upon this work will be clear. I have also discussed the paradox with many friends and colleagues; although they are too numerous to mention individually, I am indebted to them all.

Several people have contributed directly to the writing of this book, and I would like to take this chance to thank them. Clive Horwood of Praxis, and John Watson of Springer, have been very supportive of the project; the book would not have been completed had it not been for their advice and encouragement. (I would also like to thank John for sharing his favoured resolution of the paradox over an enjoyable working lunch.) Stuart Clark provided many useful comments on an early draft of the manuscript; Bob Marriott caught several errors and solecisms in a later draft (Bob also sent me a list of 101 resolutions of the paradox—75 of which I agree with); and I am extremely grateful to Steve Gillett for putting me right on many scientific points. (I am, of course, responsible for those errors that remain.) Several authors and organizations kindly gave permission to reproduce figures; I am particularly grateful to thank Lora Gordon, Geoffrey Landis, Ian Wall, Susan Lendroth, Reinhard Rachel, Heather Lindsay and Merrideth Miller for help in obtaining suitable figures. I would like to thank David Glasper, for sharing his recollections of a childhood incident that affected us both. Finally, of course, I would like to thank my family—Heike, Ron, Ronnie, Peter, Jackie, Emily and Abigail—for their patience. I spent time writing that I should instead have shared with them.

Stephen Webb
Milton Keynes, July 2002

# Contents

# 1
# Where *Is* Everybody?

There is something beguiling about paradox. The visual paradoxes of Maurits Escher's prints never fail to entice the eye. Poems such as Robert Graves' *Warning to Children*, which play with the paradox of infinite regress, make the head spin. Paradox lies at the heart of Joseph Heller's *Catch-22*, one of the 20th century's greatest novels. My favorite paradox, though, is that of Fermi.

I first came across the Fermi paradox in the summer of 1984. I had just graduated from Bristol University, and I should have spent the summer months studying Aitchison and Hey's *Gauge Theories in Particle Physics*—required reading before I started postgraduate studies at Manchester University. Instead, I spent my time enjoying the sunshine on the Bristol Downs, studying my favorite reading matter: *Isaac Asimov's Science Fiction Magazine*. As is the case with many people, SF sparked my interest in science. It was through reading the works of Isaac Asimov,[1] Arthur Clarke and Robert Heinlein and watching films such as *Forbidden Planet* that I became enamored with science. Two thought-provoking science-fact articles appeared in successive issues[2] of *Asimov's* that year. The first, by Stephen Gillett, was simply entitled *The Fermi Paradox*. The second, a forceful rebuttal by Robert Freitas, was entitled *Fermi's Paradox: A Real Howler*.

Gillett argued in the following way. Suppose, as the optimists believed, that the Galaxy is home to many extraterrestrial civilizations. (To save typing, I shall often refer to an extraterrestrial civilization as an ETC.) Then, since the Galaxy is extremely old, the chances are good that ETCs will be millions or even *billions* of years in advance of us. The Russian astrophysicist Nikolai Kardashev proposed a useful way of thinking about such civilizations. He argued that we could classify ETCs in terms of the technology they possessed, and he devised a 3-point scale for measuring the potency of that technology. A Kardashev type 1 civilization, or KI civilization, would be comparable to our own: it could employ the energy resources of a planet. A KII civilization would be far beyond our own: it could employ the energy resources of a star. A KIII civilization could employ the energy resources of an entire *galaxy*. According to Gillett, then, most ETCs in the Galaxy would be of a KII or KIII type. Now, everything we know about terrestrial life tells us that life has a natural



S. Webb, *If the Universe Is Teeming with Aliens … WHERE IS EVERYBODY?*, Science and Fiction, DOI 10.1007/978-3-319-13236-5_1

tendency to expand into all available space. Why should extraterrestrial life be any different? Surely ETCs would want to expand from their home world and out into the Galaxy. But—and this is the key point—a KII or KIII civilization should be able to colonize the Galaxy in a few million years. The Galaxy should be *swarming* with technologically advanced civilizations. They should already be here! And yet we see no evidence that ETCs exist. Gillett called this the Fermi paradox. (I learned why Fermi's name is attached to the paradox a few months later, when Eric Jones published a Los Alamos preprint describing the origins of the paradox; but more of this later.) For Gillett, the paradox pointed to a chilling conclusion: humankind is alone in the universe.

Freitas thought this was all hogwash. He compared Gillett's logic to the following argument. Lemmings breed quickly—about 3 litters per year, with each litter containing up to 8 offspring. In just a few years the total mass of lemmings will be equal to the mass of the entire terrestrial biosphere. The Earth must be swarming with lemmings. And yet most of us see no evidence that lemmings exist. Have *you* ever seen a lemming? The "Fermi paradox" line of reasoning would lead us to conclude that lemmings don't exist—yet, as Freitas pointed out, this would be absurd. More interestingly, he pointed out that the lack of evidence for ETCs isn't particularly strong: if small artificial probes were parked in the Asteroid Belt, say, or larger probes in the Oort Cloud, then we'd have essentially no chance of detecting them. Besides, he argued that the logic behind the so-called paradox is faulty. The first two steps in the argument are: (i) if aliens exist, then they should be here; (ii) if they are here, then we should observe them. The difficulty is those two "should"s. A "should" is not a "must" and therefore it's logically incorrect to reverse the arrow of implication. (In other words, the fact we haven't observed them doesn't allow us to conclude they aren't here, so we can't conclude they don't exist.)

Until we obtain some new information that can help us resolve a paradox, people are free to follow differing lines of reasoning. This is, after all, what makes paradox so interesting. In the case of the Fermi paradox, the stakes are so high (the existence or otherwise of alien intelligence) and the experimental input to the argument is so sparse (even now, we can't be sure ETCs aren't here) that arguments often become heated. In the Gillett–Freitas debate, I initially sided with Freitas. The main reason was sheer weight of numbers: there are perhaps as many as 400 billion stars in the Galaxy, and as many galaxies in the universe as there are stars in the Galaxy. Ever since the time of Copernicus, science has taught us there's nothing special about Earth. It followed, then, that Earth couldn't be the sole home to intelligent life. And yet . . .

Gillett's argument stuck in my mind. I'd been reading about cosmic wonders since I was a child. The Galaxy-spanning civilization of the *Foundation* trilogy, the astroengineering wonders of *Ringworld*, the enigma of the vessel in


**Fig. 1.1** The first image of Earth taken from the surface of another planet: this photograph was taken in March 2004 by the Mars Exploration Rover Spirit. Earth is just about visible on a computer screen; the limits of printing technology mean you might not be able to see it here. Earlier, in 1990, Voyager 1 sent back a photograph of Earth taken from much further away—a distance of about 6 billion kilometers. In Carl Sagan's words, Earth appeared as a pale blue dot. When contemplating the insignificance of the piece of rock we inhabit, and the billions of similar rocks that must be out there, it's difficult to believe we might be alone in the universe. (Credit: NASA)

*Rendezvous with Rama*—all these were part of my mental furniture. And yet where *were* these marvels? The imaginations of SF writers had shown me hundreds of possible universes, but my astronomy lecturers had made it clear that so far, whenever we look out into the real universe, we can explain everything we see in terms of the cold equations of physics. Put simply, the universe looks dead. The Fermi question: where *is* everybody? The more I thought about it, the more the paradox seemed to be significant.

It seemed to me the paradox was a competition between two large numbers: the plethora of potential sites for life versus the vast age of the universe.

The first number is simply the number of planets with suitable environments for the development of life. If we adopt the Principle of Mediocrity, and assume there's nothing special about Earth, it follows there are many millions of suitable environments for life in the Galaxy (and many billions of environments in the universe). Given so many potential seeding grounds, life should be common. This argument goes back at least as far as the 4$^{th}$ century BC when Metrodorus of Chios wrote that "a single ear of wheat in a large field is as strange as a single world in infinite space".

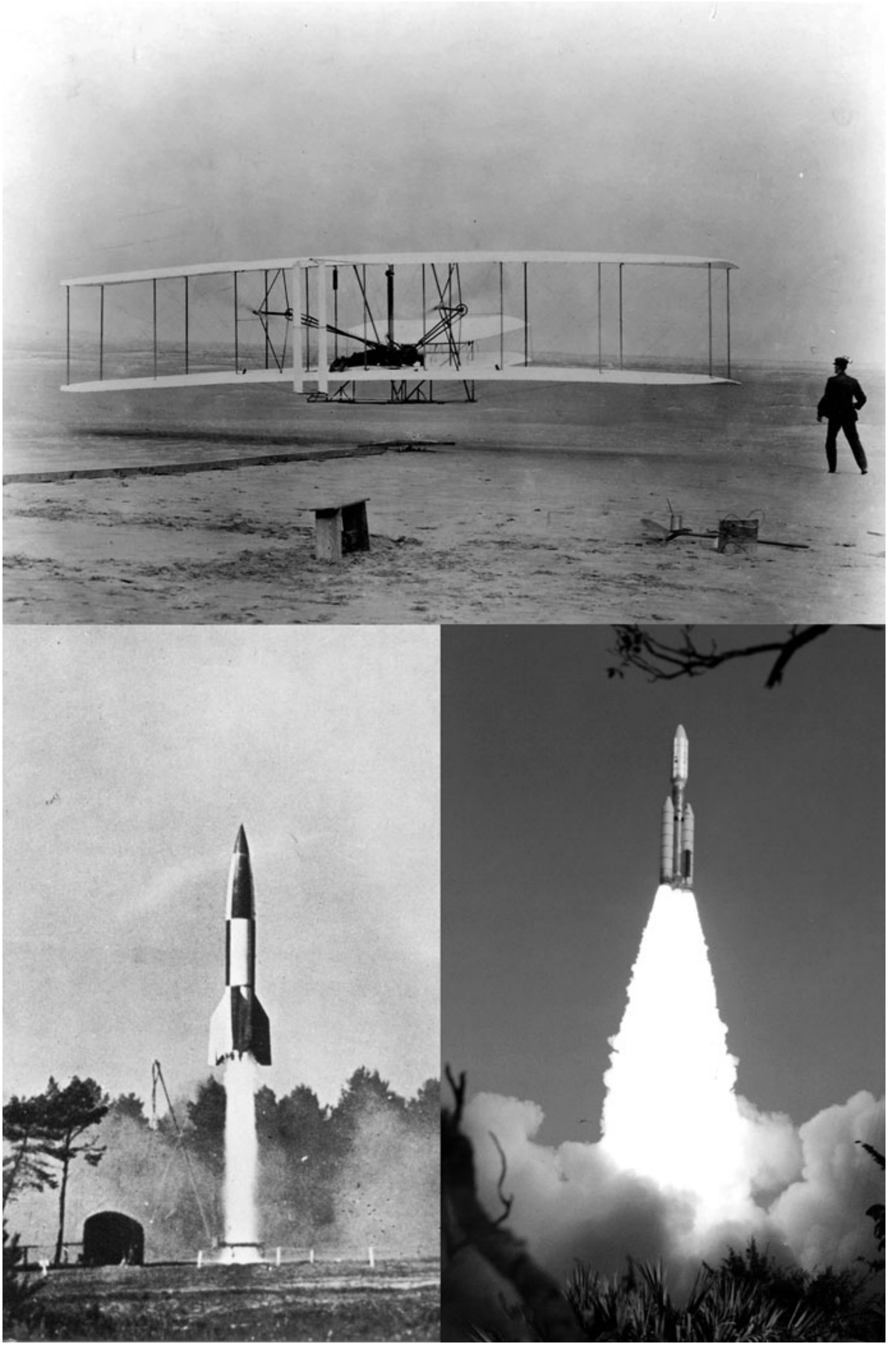


**Fig. 1.2** *Top*: Orville Wright at the controls in 1903. *Bottom left*: a rocket fired from a launchpad in Germany in 1945. *Bottom right*: the launch of Voyager 1 in 1977. Immense technological progress in less than a century. What will our craft look like in a thousand years? (Credit: *top*—USAF; *bottom left*—Crown Copyright 1946; *bottom right*—NASA)

**Table 1.1** In the Universal Year we compress 13.8 billion years into 365 days. On this timescale, an individual's lifespan is a fraction of a second. Jesus lived about 4.6 seconds before midnight on 31 December and the dinosaurs died out in the early hours of 30 December

| Real time | Time in a Universal Year |
|---|---|
| 70 yrs | 0.16 s |
| 100 yrs | 0.23 s |
| 437 yrs | 1 s |
| 1000 yrs | 2.3 s |
| 2000 yrs | 4.6 s |
| 10 000 yrs | 23 s |
| 100 000 yrs | 3 mins 50 s |
| 1 million yrs | 38 mins 20 s |
| 2 million yrs | 1 hr 16 mins 40 s |
| 10 million yrs | 6 hr 23 min 20 s |
| 100 million yrs | 2 days 15 hr 53 min 20 s |

The second number is now known with astonishing precision: the latest cosmological measurements[3] tell us the universe is 13.8 billion years old (give or take 37 million years). To evoke a feeling for such a large time span, it's usual in these discussions to compress the entire history of the universe into some standard interval. In this case, I'll compress the current age of the universe into a standard Earth year: in other words, the "Universal Year" compresses the entire history of the universe into 365 days. On this timescale, a second of real time corresponds to 437 years; in the Universal Year, western science begins about 1 second before midnight on 31 December. In 1903, the Wright brothers developed powered flight; less than four decades later a German V-2 rocket became the first object to achieve suborbital flight; about three decades after that, in 1977, Voyager 1 was launched on a Titan rocket and has now reached the edge of interstellar space. Within a typical person's lifespan, humans went from being an essentially Earthbound species to one capable of launching a craft that will eventually reach the stars. And yet that span of time represents just the final 0.16 seconds of the Universal Year. Even the entire history of our species takes up much less than 1 hour of the final day of the Universal Year. On this scale, however, the earliest ETCs could have originated in the early summer months. If the colonization of the Galaxy can take place in the equivalent of a few hours, then one would expect one or more of the advanced technological civilizations to have long since completed the job. Even if they all took some path other than colonization, wouldn't we at least expect to hear some evidence of their presence? But the universe is silent. The paradox might not logically *prove* aliens don't exist, but surely Fermi's question is worth some of our attention.

I wasn't the only one who found the Fermi paradox interesting. Over the years, many people have offered their resolutions to the paradox, and I developed the habit of collecting them. Although there's a fascinating range of answers to the question "where is everybody?", they all fall into one of three classes.

First, there are answers based around the idea that somehow extraterrestrials are (or have been) here. This is probably the most popular resolution of the paradox. Certainly, belief in intelligent extraterrestrial life is widespread. Polls repeatedly and consistently suggest that the majority of Americans believe flying saucers exist and are buzzing around Earth; the proportion of Europeans holding that belief seems to be smaller, but is nevertheless high.

Second, there are answers suggesting ETCs exist, but for some reason we've not yet found evidence of their existence. This is probably the most popular category of answer among practicing scientists.

Third, there are answers purporting to explain why humankind is alone in the universe, or at least in the Galaxy; we don't hear from extraterrestrial intelligence because there *is* no extraterrestrial intelligence.

In 2002 I published the first edition of this book. It contained discussions of 50 solutions to the Fermi paradox that I'd collected over the years, organized into the three classes mentioned above. Why, a dozen years later, do I feel there's a need for a second edition of the book? After all—and I don't believe this will come as a surprise to anyone—there's still no hard evidence for the existence of extraterrestrial intelligence. Well, although we have no definitive answer to the question "Where *is* everybody?", scientists have made tremendous progress in better understanding the relevant inputs to many of the proposed solutions. Over the past dozen years scientists have learned much about exoplanets, about planetary dynamics, about the limits to life . . . we've even learned more about the genesis of the first proposed solution to the paradox ("They are here and they call themselves Hungarian"). Thus many discussions in the first edition are now rather dated. There's also the fact that various new solutions have been proposed in recent years. A duodecennial update therefore seems appropriate.

The first edition of the book contained one or two light-hearted solutions. I've decided to keep these, and even add a couple more, but this is not to imply that the Fermi paradox need not be taken seriously. I believe that the Great Silence is becoming ever-more deafening. With each search that turns up negative, with each year that passes without scientists finding some evidential trace of extraterrestrial activity in the mountains of data captured by our telescopes, the paradox gains in strength. I believe that Fermi's question is becoming one of the most important in all of science—right up there with questions to do with the nature of consciousness and the unification of our physical theories.

**Exponential Notation** The book uses exponential notation. If you're unfamiliar with the notation, all you need to know is that it's a convenient method for handling very large and very small numbers.

In this book I always use 10 as a base and so, in essence, the exponent counts the number of zeros following the 1. Multiplying numbers together using this notation is simple: just add the exponents. For example:

$$100 = 10 \times 10 = 10^2$$

and

$$1000 = 10 \times 10 \times 10 = 10^3.$$

Division is just as easy: subtract one exponent from another. For example:

$$1000 \div 10 = 10^{3-1} = 10^2 = 100.$$

For numbers less than unity, the exponent is negative. A negative exponent gives the same value as the reciprocal of its matching positive exponent. Thus:

$$10^{-2} = \tfrac{1}{10^2} = \tfrac{1}{100} = 0.01$$

and

$$10^{-3} = \tfrac{1}{10^3} = \tfrac{1}{1000} = 0.001.$$

Using exponential notation we can write, for example, 1 million as $10^6$ and 1 billionth as $10^{-9}$. This is useful in science, where we routinely deal with very large and very small numbers. Using exponential notation we can discuss the number of stars in the universe (there are about $10^{22}$ of them) or the mass of an electron (which is about $10^{-36}$ kg) without resorting to unwieldy phrases such as "a thousand billion billion" or "a trillion trillion trillionth".

The purpose of this book, then, is to present and discuss 75 proposed solutions to Fermi's question. The list of solutions isn't intended to be exhaustive; rather, I've chosen them because they are representative or because I think they possess some feature that's particularly interesting. The solutions come from scientists working in several widely separated fields, but also from SF authors; in this topic, authors have been at least as industrious as academics, and in many cases they have anticipated the work of professional scientists.

The outline of the book is as follows.

Chapter 2 gives a brief biography of Fermi, focusing on his scientific achievements. I then discuss the notion of paradox and present a brief discussion of the history of the Fermi paradox.

Chapters 3–5 present 74 of my favorite solutions to the paradox. Not all of them are independent, and sometimes I revisit a solution in another guise, but all of them have been seriously proposed as answering Fermi's question. I arrange the answers according to the three classes mentioned above. Chapter 3 discusses 10 proposals based around the idea that ETCs are or were here. Chapter 4 discusses 30 answers based around the idea that ETCs exist, but we've yet to find evidence of them. Chapter 5 discusses 24 solutions of the paradox based around the idea that we are alone. There's a logic to the arrangement of the various discussions, but I hope the sections are all self-contained enough to allow readers to "dip into" the book and pick out solutions that particularly interest them. In the discussions I try to be as even-handed as possible, even if I disagree with the solution (which I often do).

Chapter 6 contains the 75th solution: my own view of the resolution of the paradox. It's not a particularly original suggestion, but it summarizes what I feel the Fermi paradox might be telling us about the universe in which we live.

This is followed by a chapter of notes and suggestions for further reading. The material discussed in this book covers various subjects, ranging from astronomy to zoology, so the references in the final chapter are necessarily wide in scope. They range from SF stories to popular science books to primary research articles published in scholarly journals. Many readers might encounter difficulties in accessing the more specialized references, but I hope they will at least find it possible to use this chapter to discover related information on the Web.

The book is specifically aimed at a popular audience. One of the beauties of the Fermi paradox is that it can be appreciated without the need for any mathematics beyond an understanding of exponential notation. It follows that anyone can present a resolution of the Fermi paradox; you don't need to have years of scientific and mathematical training behind you to contribute to the debate. I hope that a reader of this book might devise a solution that no-one else has thought of. If you do—please write to me and share it!

# 2
# Of Fermi and Paradox

Before considering the merits of the various proposed solutions to the Fermi paradox, this chapter presents some of the background. I first give a short biography of Enrico Fermi himself, focusing on just a few of his many and varied scientific accomplishments. I mention only those contributions to science that I refer to in later sections of the book. I ignore, for example, his contribution to cosmic ray physics: Fermi was the first to propose a realistic model for explaining the origin of the high-energy particles that bombard Earth from space. This work is honored by the naming of NASA's satellite mission for investigating cosmic rays—the *Fermi Gamma-ray Space Telescope*. Indeed, Fermi's scientific achievements were so numerous that the *Fermi Space Telescope* is only the latest in a variety of things named after him. Fermilab, in Batavia, IL, is one of the world's leading centers for particle physics; the element with atomic number 100, which was first synthesized in 1952 in a hydrogen bomb explosion, is called fermium (Fm); the typical length scale in nuclear physics, $10^{-15}$ m, is called the fermi; 8103 Fermi is a main-belt asteroid and Fermi is a large crater on the far side of the Moon; several members of the Enrico Fermi Institute at Chicago University have won Nobel prizes. For more details of Fermi's life, both inside and outside science, I recommend the interested reader to the biographies of Fermi listed in the References.

I then discuss the notion of paradox, and briefly look at a few examples from various fields. Paradox has played an important role in intellectual history, helping thinkers to widen their conceptual framework and sometimes forcing them to accept quite counterintuitive notions. It's interesting to compare the Fermi paradox with these more established paradoxes.

Finally, I discuss how the Fermi paradox itself—where *is* everybody?—came into being. It's worth noting that some people argue that this is neither a paradox nor is it Fermi's. Nevertheless, we shall see that Fermi's question can be cast into the shape of a formal paradox (if you feel the need to do so) and I explain how Fermi's name came to be attached to a paradox that is older than many people believe.



S. Webb, *If the Universe Is Teeming with Aliens . . . WHERE IS EVERYBODY?*, Science and Fiction, DOI 10.1007/978-3-319-13236-5_2

## The Physicist Enrico Fermi

*It is no good to try to stop knowledge from going forward. Ignorance is never better than knowledge.*

Enrico Fermi

Enrico Fermi was the most complete physicist of the last century—a world-class theoretician who carried out experimental work of the highest order. No other physicist since Fermi has switched between theory and experiment with such ease, and it's unlikely that anyone will do so again. The field has become too large to permit such crossover.

Fermi was born in Rome on 29 September 1901, the third child of Alberto Fermi, a civil servant, and Ida DeGattis, a schoolteacher. He showed precocious ability in mathematics,[4] and as an undergraduate student of physics at the Scuola Normale Superiore in Pisa he quickly outstripped his teachers.[5]

His first major contribution to physics was an analysis of the behavior of certain fundamental particles that make up matter. These particles—such as protons, neutrons and electrons—are now called *fermions* in his honor. Fermi showed how, when matter is compressed so that identical fermions are brought close together, a repulsive force comes into play that resists further compression. This fermionic repulsion plays an important role in our understanding of phenomena as diverse as the thermal conductivity of metals and the stability of white dwarf stars.

Soon after, Fermi's theory of beta decay (a type of radioactivity in which a massive nucleus emits an electron) cemented his international reputation. The theory demanded that a ghostly particle be emitted along with the electron, a particle he called the *neutrino*—"little neutral one". Not everyone believed in the existence of this hypothetical fermion, but Fermi was proved correct. Physicists finally detected the neutrino in 1956. Although the neutrino remains rather intangible in terms of its reluctance to react with normal matter, its properties play a profound role in present-day astronomical and cosmological theories.

In 1938 Fermi was awarded the Nobel prize for physics, partly in recognition of a technique he developed to probe the atomic nucleus. His technique led him to the discovery of new radioactive elements; by bombarding the naturally occurring elements with neutrons, he produced more than 40 artificial radioisotopes. The award also recognized his discovery of how to make neutrons move slowly. This might seem a minor point but it has profound practical applications, since slow-moving neutrons are more effective than fast neutrons at inducing radioactivity. (A slow neutron spends more time in the

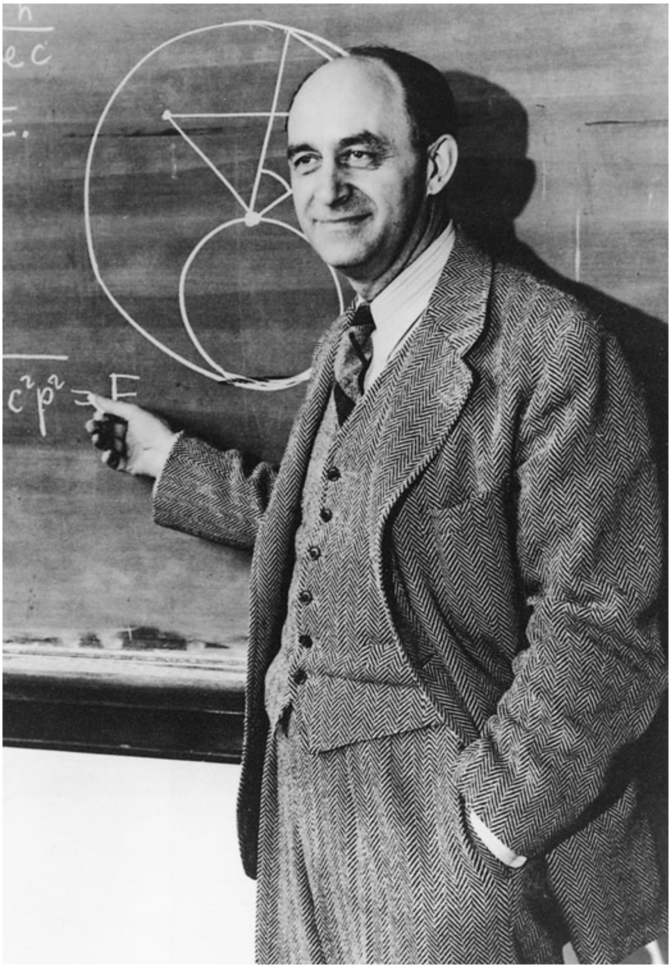



**Fig. 2.1** This photograph of Enrico Fermi lecturing on atomic theory appears on a stamp released by the US Postal Service on 29 September 2001 to commemorate the hundredth anniversary of Fermi's birth. (Credit: American Institute of Physics Emilio Segré Visual Archives)

neighborhood of a target nucleus, and so is more likely to interact with the nucleus. In a similar way, a well-aimed golf ball is more likely to sink into the hole if it's moving slowly: a fast-moving putt can roll by.) This principle is used in the operation of nuclear reactors.

News of the award was tempered by the worsening political situation in Italy. Mussolini, increasingly influenced by Hitler, initiated an anti-Semitic campaign. Italy's fascist government passed laws that were copied directly from the Nazi Nuremberg edicts. The laws didn't directly affect Fermi or his two children, who were considered to be Aryans, but Fermi's wife, Laura, was Jewish. They decided to leave Italy, and Fermi accepted a position in America.

Two weeks after arriving in New York, news reached Fermi that German and Austrian scientists had demonstrated nuclear fission. Einstein, after some prompting, wrote his historic letter to Roosevelt alerting the President to the probable consequences of nuclear fission. Citing work by Fermi and colleagues, Einstein warned that a nuclear chain reaction might be set up in a large mass of uranium—a reaction that could lead to the release of vast amounts of energy. Roosevelt was concerned enough to fund a program of research into the defense possibilities. Fermi was deeply involved in the program.

**Fermi Questions** Fermi's colleagues were in awe of him for his uncanny ability to see straight to the heart of a physical problem and describe it in simple terms. They called him the Pope because he seemed infallible. Almost as impressive was the way he estimated the magnitude of an answer (often by doing complex calculations in his head). Fermi tried to inculcate this facility in his students. He would demand of them, without warning, answers to seemingly unanswerable questions. How many grains of sand are there on the world's beaches? How far can a crow fly without stopping? How many atoms of Caesar's last breath do you inhale with each lungful of air? Such "Fermi questions" (as they are now known) required students to draw upon their understanding of the world and their everyday experience and make rough approximations, rather than rely on bookwork or prior knowledge.

The archetypal Fermi question is one he asked his American students: "How many piano tuners are there in Chicago?" We can derive an informed estimate, as opposed to an uninformed guess, by reasoning as follows.

First, suppose that Chicago has a population of 3 million people. (I haven't checked an almanac to see whether this is correct; but making explicit estimates in the absence of certain knowledge is the whole point of the exercise. Chicago is a big city, but not the biggest in America, so we can be confident that the estimate is unlikely to be in error by more than a factor of 2. Since we have explicitly stated our assumption we can revisit the calculation at a later date, and revise the answer in the light of improved data.) Second, assume that families, rather than individuals, own pianos and ignore those pianos belonging to institutions such as schools, universities and orchestras. Third, if we assume that a typical family contains 5 members, then our estimate is that there are 600,000 families in Chicago. We know that not every family owns a piano; our fourth assumption is that 1 family in 20 owns a piano. We thus estimate there are 30,000 pianos in Chicago. Now ask the question: how many tunings would 30,000 pianos require in 1 year? Our fifth assumption is that a typical piano will require tuning once per year—so 30,000 piano tunings take place in Chicago each year. Assumption six: a piano tuner can tune 2 pianos per day and works on 200 days in a year. An individual piano tuner therefore tunes 400 instruments in 1 year. In order to accommodate the total number of tunings required, Chicago must be home to 30,000/400 = 75 piano tuners. We want an estimate, not a precise figure, so finally we round this number up to an even 100.

As we shall see later, Fermi's ability to grasp the essentials of a problem manifested itself when he posed the question: "where *is* everybody?"

Physicists had many questions to answer before they could build a bomb, and it was Fermi who answered many of them. On 2 December 1942, in a makeshift laboratory constructed in a squash court under the West Stands of the University of Chicago stadium, Fermi's group successfully achieved the first self-sustaining nuclear reaction. The reactor, or pile, consisted of slugs of purified uranium—about 6 tons in all—arranged within a matrix of graphite.

The graphite slowed the neutrons, enabling them to cause further fission and maintain the chain reaction. Control rods made of cadmium, which is a strong neutron absorber, controlled the rate of the chain reaction. The pile went critical[6] at 2:20 PM, and the first test was run for 28 minutes.

Fermi, with his unmatched knowledge of nuclear physics, played an important role in the Manhattan Project. He was there in the Alamogordo desert on 15 July 1945, just 9 miles away from ground zero at the Trinity test. He lay on the ground facing in the direction opposite the bomb. When he saw the flash from the immense explosion he got to his feet and dropped small pieces of paper from his hand. In still air the pieces of paper would have fallen to his feet, but when the shock wave arrived, a few seconds after the flash, the paper moved horizontally due to the displacement of air. In typical fashion, he measured the displacement of the paper; since he knew the distance to the source, he could immediately estimate the energy of the explosion.

After the war, Fermi returned to academic life at the University of Chicago and became interested in the nature and origin of cosmic rays. In 1954, however, he was diagnosed with stomach cancer. Emilio Segré, Fermi's lifelong friend and colleague, visited him in hospital. Fermi was resting after an exploratory operation, and was being fed intravenously. Even at the end, according to Segré's touching account, Fermi retained his love of observation and calculation: he measured the flux of the nutrient by counting drops and timing them with a stopwatch.

Fermi died on 29 November 1954, at the early age of 53.

## Paradox

*These are old fond paradoxes, to make fools laugh i' the alehouse.*

William Shakespeare, *Othello*, Act II, Scene 1

Our word paradox comes from[7] two Greek words: *para* meaning "contrary to" and *doxa* meaning "opinion". It describes a situation in which, alongside one opinion or interpretation, there's another, mutually exclusive opinion. The word has taken on a variety of subtly different meanings, but at the core of each usage is the idea of a contradiction. Paradox is more than mere inconsistency, though. If you say "it's raining, it's not raining" then you've contradicted yourself, but paradox requires more than this. A paradox arises when you begin with a set of seemingly self-evident premises and then deduce a conclusion that undermines them. If your cast-iron argument proves it must be raining, but you look and see that it's dry outside, then you have a paradox to resolve.

**Fig. 2.2** A visual paradox. This impossible figure is a Penrose triangle. It's named after Roger Penrose, a British mathematician who devised it in the 1950s. (It was first created even earlier, in 1934, by the Swedish graphic artist Oscar Reutersvärd.) The illustration appears to show a three-dimensional triangular solid, but the triangle is impossible to construct. Each vertex of a Penrose triangle is in fact a perspective view of a right angle. Artists such as Escher and Reutersvärd delighted in presenting visual paradoxes. (Credit: Tobias R.)

A weak paradox or *fallacy* can often be clarified with a little thought. The contradiction usually arises because of a mistake in a chain of logic leading from premises to conclusion. For example, beginning students of algebra often construct "proofs" of obviously untrue statements such as $1+1 = 1$. Such "proofs" usually contain a step in which an equation is divided by zero. This is the source of the fallacy, since dividing by zero is inadmissible in arithmetic: if you divide by zero you can "prove" anything at all. In a strong paradox, however, the source of a contradiction is not immediately apparent; centuries can pass before matters are resolved. A strong paradox has the power to challenge our most cherished theories and beliefs. Indeed, as the mathematician Anatol Rapoport once remarked:[8] "Paradoxes have played a dramatic part in intellectual history, often foreshadowing revolutionary developments in science, mathematics and logic. Whenever, in any discipline, we discover a problem that cannot be solved within the conceptual framework that supposedly should apply, we experience shock. The shock may compel us to discard the old framework and adopt a new one."

Paradoxes abound in logic and mathematics and physics, and there's a type for every taste and interest.

## A Few Logical Paradoxes

An old paradox, contemplated by philosophers since the middle of the 4th century BC and still discussed, is that of the liar paradox. Its most ancient attribution is to Eubulides of Miletus, who asked: "A man says he is lying; is what he says true or false?" Whichever way one analyzes the sentence, there's a contradiction. The same paradox appears in the New Testament. St. Paul, in his letter to Titus, the first bishop of Crete, wrote: "One of themselves, even a prophet of their own, said the Cretans are always liars." It's not clear whether Paul was aware of the problem in his sentence, but when self-reference is allowed paradox is almost inevitable.

One of the most important tools of reasoning we possess is the sorites. In logicians' parlance, a sorites is a chain of linked syllogisms: the predicate of one statement becomes the subject of the following statement. The statements below form a typical example of a sorites:

all ravens are birds;
all birds are animals;
all animals require water to survive.

Following the chain we must logically conclude: all ravens need water to survive.

Sorites are important because they allow us to make conclusions without covering every eventuality in an experiment. In the example above, we don't need to deprive ravens of water to know that doing so would cause them to die of thirst. But sometimes the conclusion of a sorites can be absurd: we have a sorites paradox. For example, if we accept that adding one grain of sand to another grain of sand doesn't make a heap of sand, and given that a single grain doesn't itself constitute a heap, then we must conclude that no amount of sand can make a heap. And yet we see heaps of sand. The source of such paradoxes lies in the intentional vagueness[9] of a word such as "heap". Another paradox—Theseus' paradox—hangs on the vagueness of the word "same": if you restore a wooden ship by replacing each and every plank, is it the same ship? Politicians, of course, routinely take advantage of these linguistic tricks.

In addition to sorites, we all routinely employ induction—the drawing of generalizations from specific cases—when reasoning. For example, whenever we see something drop, it falls *down*: using induction we propose a general law, namely that when things drop they *always* fall down and never up. Induction is such a useful technique that anything casting doubt on it is troubling. Consider Hempel's raven paradox.[10] Suppose an ornithologist, after years of field observation, has observed hundreds of black ravens. The evidence is enough for her to suggest the hypothesis that "all ravens are black". This is the standard process of scientific induction. Every time the ornithologist sees a black raven it's a small piece of evidence in favor of her hypothesis. Now, the statement "all ravens are black" is logically equivalent to the statement "all non-black things are non-ravens". If the ornithologist sees a piece of white chalk, then the observation is a small piece of evidence in favor of the hypothesis that "all non-black things are non-ravens"—but therefore it must be evidence for her claim that ravens are black. Why should an observation regarding chalk be evidence for a hypothesis regarding birds? Does it mean that ornithologists can do valuable work whilst sat indoors watching television, without bothering to watch a bird in the bush?

Another paradox in logic is that of the unexpected hanging, wherein a judge tells a condemned man: “You will hang one day next week but, to spare you mental agony, the day that the sentence will be carried out will come as a surprise.” The prisoner reasons that the hangman can’t wait until Friday to carry out the judge’s order: so long a delay means everyone will know the execution takes place that day—the execution will not come as a surprise. So Friday is out. But if Friday is ruled out, Thursday is ruled out by the same logic. Ditto Wednesday, Tuesday and Monday. The prisoner, mightily relieved, reasons that the sentence can’t possibly take place. Nevertheless, he’s completely surprised when the executioner leads him to the gallows on Thursday! This argument—which also goes under the name of the “surprise examination paradox” and the “prediction paradox”—has generated a huge literature.[11]

## A Few Scientific Paradoxes

Although it’s often fun, and occasionally useful, to ponder liars, ravens and condemned men, arguments involving logical paradoxes too frequently—for my taste at least—degenerate into a discussion over the precise meaning and usage of words. Such discussions are fine if one is a philosopher, but for my money the really fascinating paradoxes are those that can be found in science.

The twin paradox, which involves the special relativistic phenomenon of time dilation, is perhaps one of the most famous. Suppose one twin stays at home while the other twin travels to a distant star at close to the speed of light. To the stay-at-home twin, his sibling’s clock runs slow: his twin ages more slowly than he does. Although this phenomenon is contrary to common sense, it’s an experimentally verified fact. But surely relativity tells us that the traveling twin can consider himself to be at rest? From *his* point of view, the clock of the earthbound twin runs slow; the stay-at-home twin should be the one who ages slowly. So what happens when the traveler returns? They can’t both be right. It’s impossible for both twins to be younger than each other! The resolution of this paradox is easy: the confusion arises from a misapplication of special relativity. The two scenarios aren’t interchangeable because it’s only the traveling twin who accelerates to light speed, decelerates at the half-way point of his journey, and does it all again on the trip back. Everyone can agree that the stay-at-home twin undergoes no such acceleration. So the traveler ages more slowly than the earthbound twin; he returns to find his brother aged, or even dead. An extraterrestrial visitor to Earth would observe the same phenomenon when it returned to its home planet: its stay-at-home siblings (if aliens have siblings) would be older or long-since dead. This behavior is certainly contrary to our experience, but it's not a paradox—rather, a sad fact of interstellar travel.[12]

The so-called firewall paradox is of much more recent vintage than the twin paradox. It was first proposed in 2012,[13] and since then a storm of papers have attempted to resolve the underlying riddle. As of the time of writing, no one has managed to douse the firewall; it remains a troubling issue for theoretical physics. The paradox arises because of an apparent contradiction between the predictions made by three fundamental theories of physics: quantum theory, general relativity and complementarity.

Quantum theory is our best theory of the physical processes that happen in nature. It's a probabilistic theory, which means that it doesn't predict what will definitely happen; rather, it gives the *probability* that some particular event will happen. Quantum theory thus only makes sense if the probabilities of all the different outcomes to an event add up to 1. If you add up the probabilities for all possible outcomes and find that the result is 0.8 or 1.3—or *any* value except 1—then the result is nonsensical. It follows that information in quantum theory can't be lost and it can't be cloned: if information somehow disappeared or could somehow be copied then probabilities wouldn't add up to 1 and the result would be nonsense.

General relativity, which is our best theory of gravity, is a classical rather than a quantum theory. In other words it gives a definite prediction for the outcome of an event rather than a range of probabilities for different possible outcomes. General relativity describes gravity in terms of the warping of spacetime, and one of its predictions is that when the warping of spacetime becomes intense enough a black hole can form. A black hole is a region of space where not even light itself travels fast enough to escape the grip of gravity. Surrounding a black hole is an event horizon, a "surface of no return". If you are outside the event horizon then it's always possible, if only in principle, to leave the vicinity of the black hole; fall over the event horizon, however, and any attempt to leave the black hole will inevitably end in failure. It's important to note that according to general relativity you wouldn't notice anything special as you passed the event horizon; there's no sign marking the boundary in space beyond which lies a black hole. The usual analogy is with a rowing boat on a river with an increasingly fast current that culminates in a weir. The river contains a point of no return, beyond which the muscle power of any rower will fail to overcome the current. If the boat passes the point of no return then its fate is sealed: it will be carried over the weir. But nothing in the river marks that point of no return, and the boat can drift quite peacefully past that point without noticing anything has changed. It's the same with the event horizon surrounding a black hole.

In the mid-1970s, Stephen Hawking introduced the black hole information paradox to physics. Hawking showed that black holes do in fact radiate: quantum effects close to the surface of the event horizon mean that particles can leave the vicinity of the horizon. Black holes emit so-called Hawking radiation, and this radiation carries information and energy with it. This effect causes the black hole to lose energy, which in turn means that it shrinks. Eventually, the black hole evaporates. The question is: what happens to the information that was inside the black hole? If the information was carried away by Hawking radiation then the information would have had to have been cloned: the information couldn't have escaped from inside the event horizon. But having two copies of information violates quantum theory because it would mean probabilities don't add up to one. So perhaps the information disappears when the black hole evaporates? But disappearing information violates quantum theory because it would mean probabilities don't add up to one. We have a paradox: quantum theory and general relativity appear to give conflicting accounts about what happens to any information that might fall into a black hole.

In the early 1990s, Leonard Susskind and co-workers proposed something called complementarity as a resolution of the black hole information paradox. Susskind's idea was that in a sense the problem is one of perspective: observers inside and outside the event horizon see different things. An observer outside the black hole sees information gather at the event horizon and then eventually flee the black hole in Hawking radiation. An observer inside the black hole sees information as being inside the event horizon. Since the two observers can't communicate, the paradox is avoided. Susskind's proposal in a sense allows the information to be both inside and outside the event horizon in a way that doesn't violate the requirements of quantum theory. His proposal was given a boost in 1997, when Juan Maldacena proposed an idea[14] called AdS/CFT correspondence. The idea says that string theory (which automatically contains gravity) is equivalent to a quantum theory without gravity in a space of fewer dimensions. Maldacena's paper has been hugely influential, because it allows physicists to attack problems that would otherwise be too difficult: if a problem is intractable in one regime simply switch to another regime where it might be tractable, do the work there, then switch back to the original regime. Crazy as it might sound, the AdS/CFT correspondence states that the three-dimensional interior of a black hole with gravity is equivalent to a quantum theory without gravity that sits just above the two-dimensional surface of the horizon. A lot of theoretical work based on this correspondence seemed to back up the complementarity proposal. It seemed that information isn't lost, quantum theory was saved and the information paradox was put to bed.

In 2012, however, four physicists (Ahmed Almheiri, Donald Marolf, Joseph Polchinski and James Sully, collectively known as AMPS) discovered something unsettling when they tried to describe the process of black hole evaporation in terms of complementarity. According to their analysis, when a black hole is about halfway through the evaporation process it has lost so much information through Hawking radiation that the remaining information at the two-dimensional horizon surface is insufficient to represent the three-dimensional interior of the black hole with gravity. This manifests itself in a phenomenon that AMPS called a firewall: an observer falling into a black hole gets burned to a cinder at a surface just above the event horizon. But this effect simply should not happen according to general relativity: nothing in space should mark the "surface of no return".

So the paradox is back, and worse than ever because now we have *three* elements competing for attention: quantum theory, general relativity and complementarity. At the time of writing the situation is one of confusion. Clarity will eventually return—perhaps with input from a separate area of science, such as information theory—and by resolving the firewall paradox we'll understand more about some of the fundamental concepts in physics.

Predating both the twin paradox and the firewall paradox is that named after Heinrich Olbers,[15] who considered a question asked by countless children—"Why is the night sky dark?"—and showed that the darkness of night is deeply mysterious. His reasoning was based upon two premises. First, the universe is infinite in extent. Second, the stars are scattered randomly throughout the universe. (Olbers was unaware of the existence of galaxies, which weren't recognized as stellar groupings until some 75 years after his death, but his reasoning is unaffected by this. His argument works in exactly the same way for galaxies as it does for stars.) From these premises he reached an uncomfortable conclusion: in whichever direction you look, your line-of-sight must eventually end on a star—therefore the night sky should be bright.

**Olbers' Paradox** Suppose all stars have the same intrinsic brightness. (The following argument is simpler under this assumption, but the conclusion in no way depends upon it.) Now consider a thin shell of stars (call it shell A) with Earth at its center, and another thin shell of stars (shell B), also centered on Earth, with a radius twice that of shell A. In other words, shell B is twice as distant from us as shell A.

A star in shell B will appear to be 1/4 as bright as a star in shell A. (This is the inverse-square law: if the distance to a light source *increases* by a factor of 2, the apparent brightness of the light source *decreases* by a factor of $2 \times 2 = 4$.) On the other hand, the surface area of shell B is 4 times larger than that of shell A, so it contains 4 times as many stars. Four times as many stars, each of which is 1/4 as bright: the total brightness of shell B is exactly the same as the total brightness of shell A! But this works for *any* two shells of stars. The contribution to the brightness of the night sky from a distant shell of stars is the same as from a nearby shell. If the universe is infinite in extent, then the night sky should be infinitely bright.

This argument is not quite correct: the light from an extremely distant star will be intercepted by an intervening star. Nevertheless, in an infinite universe with a uniform distribution of stars *any* line of sight will eventually run into a star. Far from being dark, the entire night sky should be as dazzling as the Sun. The night sky should blind us with its brightness!

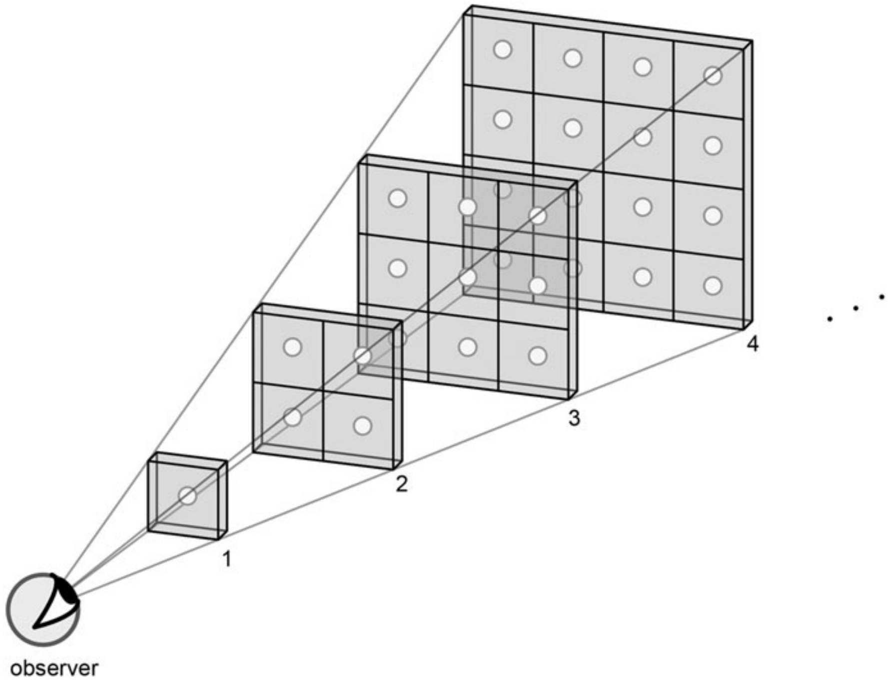



**Fig. 2.3** Suppose the stars are uniformly distributed throughout space. A star's brightness decreases as the square of its distance from an observer, but the number of stars increases as the square of the distance from the observer. These two effects cancel, and each grid shown above contributes the same amount of brightness. Since there is an infinite number of these grids, the night sky should be infinitely bright. Even allowing for nearby stars blocking the light from distant stars, the night sky should be blindingly bright. (Credit: Htykym)

How can we resolve the paradox? The first explanation that springs to mind is that clouds of gas or dust obscure light from distant stars. The universe does indeed contain such clouds, but they can't shade us from Olbers' paradox: if the clouds absorb light they'll heat up until they are at the same average temperature as the stars themselves. It turns out that the paradox is explained by one of the most dramatic discoveries ever made by astronomers: the universe has a finite age. Since the universe is only about 13.8 billion years old, the part we can see is finite in size. For the night sky to be as bright as the surface of the Sun, the observable universe would have to be almost 1 million times bigger than it is. (That the universe is expanding also helps to explain the paradox: light from distant objects is redshifted by the expansion, and so distant objects are less bright than one would expect from the inverse-square law. The principal explanation, though, comes from the finite age of the universe.)

It's fascinating that in pondering such a simple question—"Why is the night sky dark?"—one could infer that the universe is expanding and that it has a finite age. Perhaps the simple question that Fermi asked—"Where *is* everybody?"—leads to an even more important conclusion.

## The Fermi Paradox

*Sometimes I think we're alone. Sometimes I think we're not.*
*In either case, the thought is staggering.*

Buckminster Fuller

Thanks to detective work by the Los Alamos scientist Eric Jones, whose report I draw heavily upon[16] in this section, we know the genesis of the Fermi paradox.

The spring and summer of 1950 saw the New York newspapers exercised over a minor mystery: the disappearance of public trash cans. This year was also the height of flying saucer reports, another subject that filled the column inches. On 20 May 1950, *The New Yorker* published a cartoon by Alan Dunn that made amusing reference to both stories.

Fermi was at Los Alamos in the summer of 1950. One day, he was chatting to Edward Teller and Herbert York as they walked over to Fuller Lodge for lunch. Their topic was the recent spate of flying saucer observations. Emil Konopinski joined them and told them of the Dunn cartoon. Fermi remarked wryly that Dunn's was a reasonable theory because it accounted for two distinct phenomena: the disappearance of trash cans and the reports of flying saucers. After Fermi's joke, there followed a serious discussion about whether flying saucers could exceed the speed of light. Fermi asked Teller what he thought the probability might be of obtaining evidence for superluminal travel by 1960. Fermi said that Teller's estimate of one-in-a-million was too low; Fermi thought it was more like one-in-ten.

The four of them sat down to lunch, and their discussion turned to more mundane topics. Then, in the middle of the conversation and out of the clear blue, Fermi asked: "Where *is* everybody?" His lunch partners Teller, York and Konopinski immediately understood that he was talking about extraterrestrial visitors. And since this was Fermi, perhaps they realized it was a more troubling and profound question than it first appears. York recalls that Fermi made a series of rapid calculations and concluded that we should have been visited long ago and many times over.

Neither Fermi nor the others ever published these calculations, but we can make a reasonable guess at his thought processes. He must first have made an

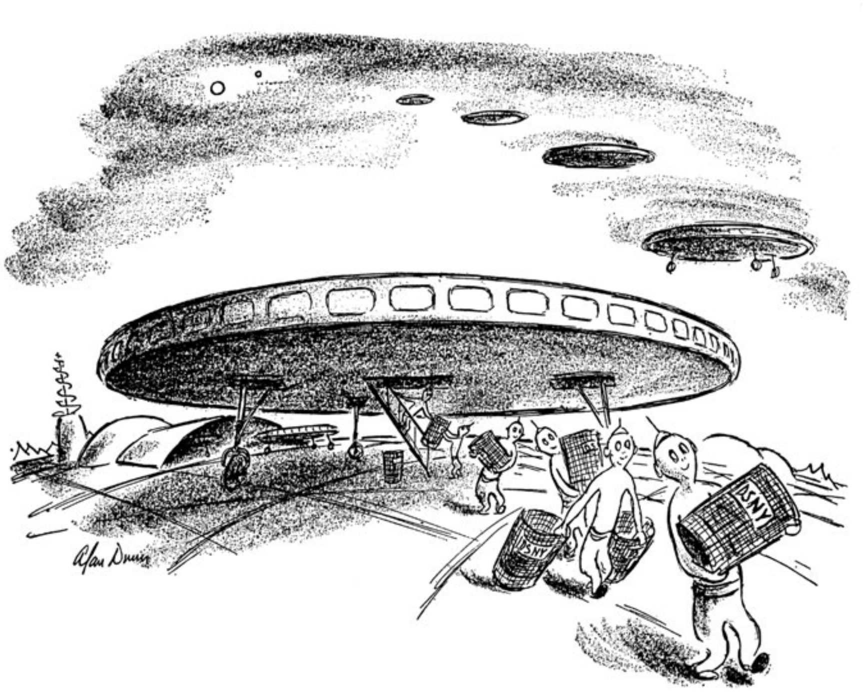



**Fig. 2.4** For reasons that make sense only to them, aliens are returning to their home planet with trash cans that are the property of New York's Department of Sanitation. (Credit: *The New Yorker* Collection 1950, drawn by Alan Dunn, from cartoonbank.com; all rights reserved)

estimate of the number of ETCs in the Galaxy, and this is something we can estimate ourselves. After all, the question "How many advanced communicating extraterrestrial civilizations are there in the Galaxy?" is a typical Fermi question!

**A Fermi Question: How Many Communicating Civilizations Exist?**

Represent the number of communicating ETCs in the Galaxy by the symbol $N$. To estimate $N$ we first need to know the yearly rate $R$ at which stars form in the Galaxy. We also need to know the fraction $f_p$ of stars that possess planets and, for planet-bearing stars, the number $n_e$ of planets with environments suitable for life. We also need the fraction $f_l$ of suitable planets on which life actually develops; the fraction $f_i$ of these planets on which life develops intelligence; and the fraction $f_c$ of intelligent life-forms that develop a culture capable of interstellar communication. Finally, we need to know the time $L$, in years, that such a culture will devote to communication. Multiplying all these factors together will provide us with an estimate for $N$. We can write it as a simple equation:

$$N = R \times f_p \times n_e \times f_l \times f_i \times f_c \times L.$$

**Fig. 2.5** Edward Teller (*left*) with Fermi in 1951, not long after Fermi first asked his question. (Credit: American Institute of Physics Emilio Segré Visual Archives)

Note that the equation shown in the previous box, namely

$$N = R \times f_p \times n_e \times f_l \times f_i \times f_c \times L$$

is no more a "proper" equation for the number of communicating ETCs than

$$N = p_c \times n_f \times f_p \times n_t \times R$$

is the equation for the number of piano tuners in Chicago. Nevertheless, if we assign reasonable values to the various factors in the equation—always with the understanding that such values can and will change as our knowledge increases—we will arrive at a rough estimate for the number of ETCs in the Galaxy. The difficulty we face is in our varying degrees of ignorance for the various terms in the equation. When asked to provide values for these terms, astronomers would provide responses ranging from "We're reasonably certain" (for the factor *R*) to "We'll pin this down over the next few decades" (for the factor $n_e$) to "How the hell should we know?" (for the factor *L*). At least when we try to estimate the number of Chicago-based piano tuners we can be reasonably confident that our various sub-estimates are not wildly in error; there can be no such confidence with our estimate for the number of communicating ETCs. Nevertheless, in the absence of any definite knowledge of ETCs, how else can we proceed? (The equation above, incidentally, has

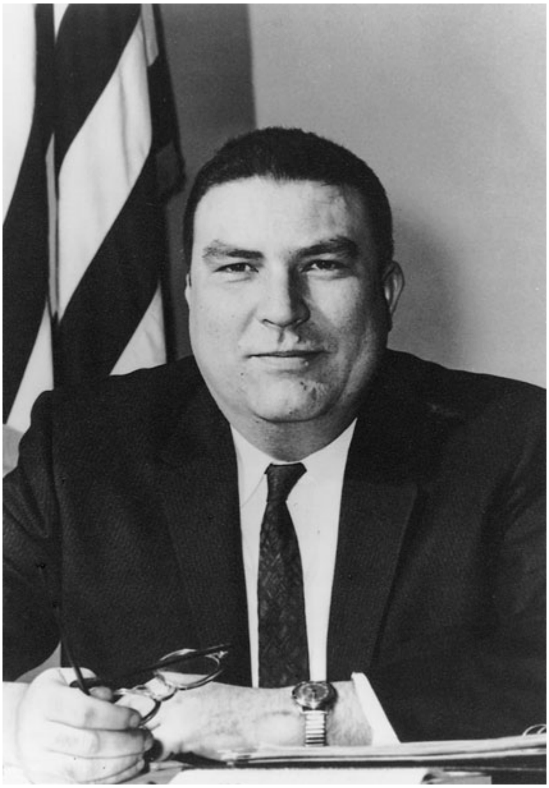


**Fig. 2.6** Herbert York, one of Fermi's lunchtime companions. (Credit: American Institute of Physics Emilio Segré Visual Archives)

reached a certain iconic status in science; it's known as the Drake equation, after the radio astronomer[17] Frank Drake who was the first to make explicit use of it. The Drake equation was the focal point of an extremely influential conference on the search for extraterrestrial intelligence, held at Green Bank in 1961—eleven years after Fermi's remark.)

In 1950, Fermi would have known far less about the various factors at play in the above "equation" than we do nowadays, but he could certainly have made some reasonable guesses—guided, as he would have been, by the Principle of Mediocrity: there's nothing special about Earth or our Solar System. If he guessed at a rate of star formation of 1 star per year he wouldn't have been too wrong. Values of $f_p = 0.5$ (half the stars have planets) and $n_e = 2$ (stars with planets on average each have 2 planets with environments conducive to life) seem to be not unreasonable. The other factors are much more subjective; if he were an optimist, Fermi might perhaps have chosen $f_l = 1$ (every planet

**Fig. 2.7** Emil Konopinski (*far left*), another one of Fermi's lunchtime companions. (Credit: American Institute of Physics Emilio Segré Visual Archives)

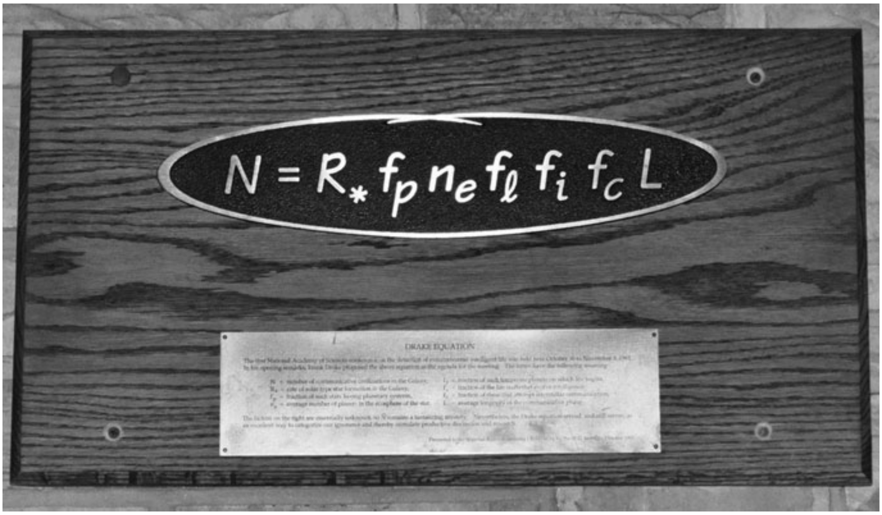



**Fig. 2.8** The Drake equation is a means of estimating the number of communicative civilizations in the Galaxy. Drake developed the equation so that it could form the agenda for the first ever SETI meeting—held in 1961 at NRAO Green Bank, WV. This commemorative plaque is on the same wall that held the blackboard where the equation was first written. (Credit: SETI League)

that can develop life *will* develop life), $f_i = 1$ (once life develops, intelligent life will certainly follow), $f_c = 0.1$ (1 in 10 intelligent life-forms will develop a civilization capable and willing to communicate) and $L = 10^6$ (civilizations remain in the communication phase for about 1 million years). Had Fermi argued in this way he would have arrived at the estimate $N = 10^6$. In other words, there could right now be a million civilizations trying to communicate with us. Some of them must be much more technologically advanced than us. So why don't we hear from them?

Indeed, pressing on with this line of reasoning, why aren't they already here? If some civilizations are extremely long-lived then we might expect them to colonize the Galaxy—and have done so before multicellular life even developed on Earth. The Galaxy should be swarming with extraterrestrial civilizations. Yet we see no sign of them. We should already know of their existence, but we don't. *Where are they?* It's tempting to see this as being nothing more than a throwaway question about interstellar travel, but it's possible to explicitly formulate the argument as a paradox[18] and we can be fairly sure that Fermi would have appreciated the paradoxical aspect of his question. Where *is* everybody? This is the Fermi paradox.

Note that the paradox isn't that extraterrestrial civilizations don't exist. (I have no idea whether Fermi believed in the existence of extraterrestrial intelligence but I suspect that, as with many physicists, he did.) Rather, the paradox—or at least the expanded version of the paradox, which notes that not only aren't they here but that we haven't heard from them or seen any evidence of their activity out in the Galaxy—is that we observe no signs of them when we might expect to. One explanation of the paradox is indeed that we are the only advanced civilization—but it's only one of several explanations.

* * *

We can appreciate the strength of the Fermi paradox when we realize that it has been independently discovered *four* times: it might more properly be called the Tsiolkovsky–Fermi–Viewing–Hart paradox.

Konstantin Tsiolkovsky, a scientific visionary[19] who worked out the theoretical basis of spaceflight as long ago as 1903, believed deeply in the monistic doctrine that ultimate reality is entirely of one substance. If all parts of the universe were the same, it followed that there must be other planetary systems similar to our own, and that some of those planets would possess life. However, not unnaturally given his interest in the details of spaceflight, Tsiolkovsky also firmly believed that mankind would construct habitats in the Solar System and then move out into space. His feelings were revealed in his famous phrase: "Earth is the cradle of intelligence, but it is impossible to live forever in the cradle." The monist in him impelled him to argue that if *we* expand into space then all those *other* species must do the same. The logic is inescapable, and

Tsiolkovsky was aware that this led to a paradox when maintaining both that mankind will expand into space *and* that the universe is brimful with intelligent life. In 1933, long before Fermi asked his question, Tsiolkovsky pointed out that people deny the existence of ETCs because (i) if such civilizations existed, then their representatives would have visited Earth, and (ii) if such civilizations existed, then they would have given us some sign of their existence. Not only is this is a clear statement of the paradox, Tsiolkovsky offered a solution: he believed that advanced intelligences—"perfect heavenly beings"—consider mankind to be not yet ready for a visitation.

Tsiolkovsky's technical works on rocketry and spaceflight were widely discussed, but the rest of his copious output was generally ignored in the Soviet era. An appreciation of his discussion of the paradox therefore came only recently. (Fermi's own contribution didn't fare much better. Sagan mentioned Fermi and his question in a footnote of a paper published in 1963, but gave no reference except to say that the Los Alamos discussion was "now rather well known". In their influential 1966 book *Intelligent Life in the Universe*, Sagan and Shklovsky introduce a chapter with the quote "Where are they?"; they attribute the quote to Fermi, but they incorrectly state that it was uttered in 1943. In a later paper, Sagan says that Fermi's quote was "possibly apocryphal".)

In 1975, English engineer David Viewing clearly stated the dilemma.[20] A quote from his paper encapsulates it nicely: "This, then, is the paradox: all our logic, all our anti-isocentrism, assures us that we are not unique—that they *must* be there. And yet we do not see them." Viewing acknowledges that Fermi was first to ask the important question—"Where are they?"—and that this question leads to a paradox. To my knowledge, then, this paper is the first that refers directly to the Fermi paradox.

However, it was a 1975 paper[21] by Michael Hart in the *Quarterly Journal of the Royal Astronomical Society* that sparked an explosion of interest in the paradox. Hart demanded an explanation for one key fact: there are no intelligent beings from outer space on Earth at the present time. He argued that there are four categories of explanation for this fact. First, "physical explanations"—these are based on some difficulty that makes space travel unfeasible. Second, "sociological explanations"—in essence these explanations propose that extraterrestrials have chosen not to visit Earth. Third, "temporal explanations"—these suggest that ETCs have not had time to reach us. Fourth, there are explanations arguing that perhaps extraterrestrials *have* visited Earth but we don't see them now. These categories were meant to exhaust the possibilities. Hart then forcefully showed how none of these four categories provide a convincing account of the key fact, which led him to offer his own explanation: *we are the first civilization in our Galaxy*.

Hart's paper led to a vigorous debate, much of it appearing in the pages of the *Quarterly Journal.* It was a debate anyone could enter—one of the earliest contributions came from the House of Lords[22] at Westminster! Perhaps the most controversial offering came from Frank Tipler, in a paper with the uncompromising title "Extraterrestrial intelligent beings do not exist". Tipler reasoned[23] that advanced ETCs could use self-replicating probes to explore or colonize the Galaxy cheaply and in a relatively short time. The abstract to Tipler's paper sums it up: "It is argued that if extraterrestrial intelligent beings exist, then their spaceships must already be present in our Solar System." Tipler contended that the SETI program had no chance of success, and was therefore a waste of time and money. His argument poured oil on the fires of the debate and led to a further round of argument. The coolest and best summary[24] of the arguments came from David Brin, who called the paradox the "Great Silence".

In 1979, Ben Zuckerman and Michael Hart organized a conference to discuss the Fermi paradox. The proceedings were published[25] in book form, and although the volume contains a variety of views it's difficult to read it without concluding that ETCs have the means, motive and opportunity to colonize the Galaxy. The means: interstellar travel seems to be possible, if not easy. The motive: Zuckerman showed how some ETCs would be forced into interstellar travel by the death of their star, and in any case it seems a wise idea for a species to expand into space to guard against the possibility of planetary disaster. The opportunity: the Galaxy is 13 billion years old, but colonization can take place over a period of only a few million years. Yet we don't see them. If this were a murder mystery, we would have a suspect but no body.

Not everyone was struck by the force of the argument. The mathematician Amir Aczel argued that the probability of extraterrestrial life is 1.[26] The physicist Lee Smolin wrote that[27] "the argument for the non-existence of intelligent life is one of the most curious I have ever encountered; it seems a bit like a ten-year-old child deciding that sex is a myth because he has yet to encounter it." Referring to Tipler's contention that ETCs would deploy probe technology to colonize the Galaxy, the late Stephen Jay Gould wrote that[28] "I must confess that I simply don't know how to react to such arguments. I have enough trouble predicting the plans and reactions of people closest to me. I am usually baffled by the thoughts and accomplishments of humans in different cultures. I'll be damned if I can state with certainty what some extraterrestrial source of intelligence might do."

It's easy to sympathize with this outlook. When considering the type of reasoning employed with the Fermi paradox, I can't help but think of the old joke about the engineer and the economist[29] who are walking down a street. The engineer spots a banknote lying on the pavement, points to it, and says, "Look! There's a hundred-dollar bill on the pavement." The economist walks on, not bothering to look down. "You must be wrong", he says. "If there

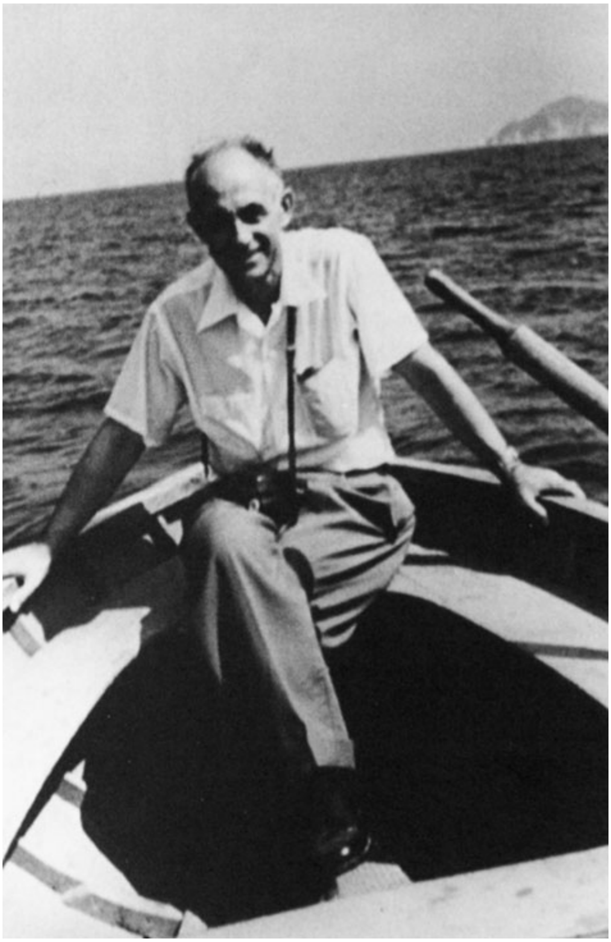


**Fig. 2.9** Enrico Fermi, sailing off the island of Elba. The photograph was taken shortly before his death. (Credit: American Institute of Physics Emilio Segré Visual Archives)

were money there, someone would already have picked it up." In science it's important to observe and experiment; we can't know what is out there unless we look. All the theorizing in the world achieves nothing unless it passes the acid test of experiment.[30]

Nevertheless, surely Hart's key fact *does* require an explanation. Astronomers have been searching for ETCs for more than half a century. And the continuing silence, despite intensive searches, is beginning to worry even some of the most enthusiastic proponents of SETI. We observe a natural universe when we could so easily observe an artificial universe. Why? Where *is* everybody? Fermi's question still demands an answer.

# 3
# They Are (or Were) Here

The simplest resolution of the Fermi paradox is that "they"—in other words, intelligent representatives from extraterrestrial civilizations—are already here (or, if they aren't here now, they were at least here some time in the past). Of the three classes of solution to the paradox, this one is by far the most popular amongst the general public. Opinion polls consistently demonstrate that a large percentage of people accept the notion that the UFO phenomenon is best explained in terms of alien spacecraft. The proportion of the population believing it was alien rather than human engineers who built assorted ancient structures around the globe is smaller; nevertheless the idea that the Egyptian pyramids, for example, have an extraterrestrial origin is hardly an underground opinion. (I've just performed an internet search using the words "pyramids extraterrestrial" and it yielded 332,000 hits.) And a surprising number of people even claim to have been in contact, either willingly or unwillingly, with beings from another planet. For many people, then, Fermi's question—where is everybody—is quite easily answered.

Scientists are much more skeptical about these various propositions, not only because of the inherent unlikelihood of the claims but also because of the poor quality of the supporting evidence. Nevertheless, it's worth at least considering these proposals as potential resolutions of the paradox. Although some of the proposed solutions are frankly risible, we shouldn't dismiss all related ideas without at least examining them with an open mind. Indeed, some serious scientists would argue that, until we've explored our neighborhood much more thoroughly and can definitely rule out the presence of alien artifacts, there really *is* no Fermi paradox.

Note that I interpret the title of this chapter rather loosely: I consider "here" to be not just Earth but the whole Solar System—and even, in Solutions 9 and 10 of this chapter, our entire universe. To begin, however, I discuss a localized resolution of the paradox that predates Fermi's original question.



S. Webb, *If the Universe Is Teeming with Aliens . . . WHERE IS EVERYBODY?*, Science and Fiction, DOI 10.1007/978-3-319-13236-5_3

# Solution 1 They Are Here and They Call Themselves Hungarians

*. . . the cleverest man I ever knew, without exception.*
Jacob Bronowski on John von Neumann, *The Ascent of Man*

Fermi would surely have been aware of one solution to the paradox even before he posed his question: it was a joke that used to do the rounds at Los Alamos.

The joke originated[31] at Los Alamos in 1945 or '46, when the American physicist Phil Morrison made up a tale of how Martians were planning—should the need ever arise—to occupy Earth. Morrison realized that a Martian invasion of Earth would be an even more difficult task than the recent Allied invasion of Normandy. So how would they go about it? Morrison argued that Martians would take the long view and spend a millennium or two getting to know the place, and he presented a number of reasons why Hungary would be their beachhead. To be successful in their long-term surveillance activities the Martians would have to pass themselves off as human and, clearly, they had been extremely successful in hiding their evolutionary differences—except for three traits. The first trait was wanderlust: this found its outlet in the Hungarian gypsy. The second was language: Hungarian is unrelated to any of the Indo-European languages spoken in the neighboring countries of Austria, Croatia, Romania, Serbia, Slovakia, Slovenia and Ukraine. The third was intelligence: their brainpower was beyond that of mere humans. A few years later, by the time Fermi asked his question, Morrison's tale had become a whimsical story often repeated within the Theoretical Division at Los Alamos. As the joke went: "they are among us and they call themselves Hungarians".

Unfortunately for the theory, many peoples have exhibited wanderlust at some point in their history; and the Hungarian language is hardly unique, related as it is to Finnish, Estonian and some languages spoken in Russia. But that third trait was in evidence during the Manhattan Project: Fermi's colleagues included Leo Szilard, Edward Teller, Eugene Wigner and John von Neumann. All four had been born in Budapest within ten years of each other. Another native of Budapest who made a major contribution to the war effort was Theodore von Kármán, but he was born slightly before the others. These "Martians" certainly constituted a formidable array of intellect:[32] the physicist Szilard made contributions in several fields; Teller went on to be the prime mover behind the development of thermonuclear weapons; Wigner won the 1963 Nobel prize in physics for his work in quantum theory; and von Kármán's aerodynamics research led to the design of the first supersonic aircraft. Easily the most brilliant of the Martians, though, was von Neumann.

John von Neumann, whom we shall meet again later in the book, was one of the outstanding mathematicians of the 20th century. He developed the discipline of game theory, made fundamental contributions to quantum theory, ergodic theory, set theory, statistics and numerical analysis, and gained fame when he helped develop the first flexible stored-program digital computer. Toward the end of his career he was a consultant to big business and the military, allotting time to various projects as if his brain were a time-share mainframe computer. His ability to calculate in his head the answers to mathematical problems was legendary—he routinely beat Fermi whenever the pair had a calculating contest—and his near-photographic memory just added to an aura of unearthly intelligence. He possessed other talents that chimed nicely with the "Hungarians are aliens" story. "Good-Time Johnny" imbibed large amounts of alcohol at Princeton parties with seemingly no detriment to his mental faculties. He was involved in road traffic accidents at alarming rate—one junction in Princeton was known as "von Neumann Corner" after the number of accidents he caused there—yet he always walked away unscathed. (The natural conclusion is that alcohol impaired his driving, but there's no clear evidence that this was the case; he seems just to have been a bad driver.)

Nevertheless, even the "cleverest man in the world" sometimes got it wrong. Although von Neumann played a pivotal role in the development of the digital computer, and has thus affected our lives in a way that few other mathematicians have ever done, he apparently thought that computers would always be huge devices, useful only for building thermonuclear bombs and controlling the weather. He failed completely to foresee a day when manufacturers would embed computers in everything from toasters to tumble dryers. Surely a real Martian would have known better.

## Solution 2 They Are Here and They Call Themselves Politicians

*What one man can fantasize, another man will believe.*

William K. Hartmann

Many of us, at one time or another, must have expressed the opinion that our political leaders aren't quite normal. Some of them, indeed, we've probably condemned as being downright weird. In the case of a certain type of English politician, I've always held that their weirdness must be the product of overweening ambition crossed with an eccentric public school system (and for the benefit of readers who aren't based in the UK it's perhaps worth pointing out that "public" schools are private). In other countries there are doubtless other explanations for the abnormal behavior of politicians. But would you say that any of them are *alien*? That's precisely what David Icke—an ex-soccer player and one-time sports presenter on the BBC—argues. According to Icke,[33] a race of alien, extradimensional lizard-people project their identities onto key US and UK politicians. (It's not just politicians: Queen Elizabeth, Prince Philip and Prince Charles are all shape-shifting reptilians. Although Princess Anne is a reptilian, she apparently has never been seen to shape-shift.)

Icke (the name is pronounced "ike" not "icky") isn't alone in his belief that some of those in power are not human. Paul Hellyer, a respected Canadian public figure who became his country's Minister of Defence in the early 1960s and who served in Pierre Trudeau's administration as the Senior Minister in the cabinet, believes that extraterrestrials currently walk the Earth. In particular, in testimony given to the Citizen Hearing on Disclosure in May 2013, Hellyer stated that two members of President Obama's administration[34] are aliens. One politician has even confessed to having repeated intimate encounters with aliens: Simon Parkes, who serves on the Whitby Town Council, claims to have fathered a child with an alien he calls the Cat Queen. (I have to concede that Parkes' political career is not at the same level as those mentioned by Icke and Hellyer. Parkes represents a small community in the north east of England; his 2012 local election success[35] was in a ward with an electorate of 2758, of whom 648 bothered to vote.)

The "Hungarians are extraterrestrials" story was always intended as a joke; Icke, Hellyer and Parkes are serious. To people such as these, then, there clearly is no Fermi paradox: extraterrestrials are here and they are our overlords or lovers or something. It's easy to dismiss these as crackpot ideas—so I shall: these are crackpot ideas—but it's not purely for the sake of completeness that I'm presenting this as a solution to the paradox. It's quite likely that of all the solutions in this book (with the probable exception of Solution 4) *this* would be accepted by the largest number of people. Certainly, more people will read Icke's books than will read mine, and a remarkable number of online reviewers see Icke's meanderings as anything but crackpottery. Hundreds of thousands of people have watched Hellyer's testimony, and much of the feedback on various YouTube recordings of the Disclosure Hearing are supportive. When Parkes appeared as a guest on breakfast TV the follow-up phone calls were generally encouraging and sympathetic. The notion that aliens have abducted certain unfortunates and subjected them to a bodily examination seems to be treated seriously by a significant section of the community.

Now, I can understand how an individual might come to believe that the Queen is a lizard shape-shifter, or that a government minister is an extraterrestrial in disguise, or that aliens visit them for regular sex sessions: ultimately the only experiences that any of us can truly know are those that go on inside our heads and, for people such as Icke, the thoughts that bubble up are perhaps taken to represent an external reality. (Far subtler minds than Icke have followed the same path. John Nash, an outstanding mathematician who built on von Neumann's work in game theory, suffered the debilitating disease of paranoid schizophrenia. Someone asked him how he, a mathematician, could believe that extraterrestrials were sending him messages. He replied that those ideas came to him in the same way as his creative mathematical ideas—so he was forced to take them seriously.)[36] What I can't understand is why so many other people would *choose* to believe the statements of Icke, Hellyer and Parkes. Although the notion that politicians are extraterrestrials might be a popular hypothesis (and admittedly it does have the virtue of explaining Tony Blair) we surely need to look for a more plausible solution to the paradox.

## Solution 3 They Are Throwing Stones at Radivoje Lajic

> *Magicians have calculated that million-to-one chances crop up nine times out of ten.*
>
> Terry Pratchett, *Mort*

I encountered an entirely new solution to the Fermi paradox during my holiday reading in the summer of 2013. An excellent popular science book on materials science,[37] by a distinguished and reputable researcher, contained a throwaway yet seemingly serious comment about Radivoje Lajic—a Bosnian man who claims that his house has been struck by meteorites on six separate occasions. As the book quite rightly stated, the chance of the same house being hit this many times is so tiny that Lajic's own explanation seemed more plausible: he (or at least his house) is being targeted by extraterrestrials.

**The Chance of Being Hit by a Meteorite** What's the chance that a house will be hit on six separate occasions by a meteorite? Well, that's a classic Fermi question. I'll leave the estimate to you, but here are a few relevant numbers about the Lajic case to start you off.

First, Earth's surface area is approximately 500,000,000,000,000 $m^2$.

Second, Mr Lajic's dwelling, in a north Bosnian village near Prijedor, is rather humble—for the purposes of the estimate, his roof has a surface area of about 10 $m^2$.

Third, I'd guess that about 100,000 meteorites[38] with a diameter larger than 5 cm reach Earth's surface each year.

Combine these numbers appropriately and, if your estimate is anything like mine, you'll conclude either that (i) Mr Lajic has even worse luck than the couple who won the UK lottery but lost their ticket[39] or (ii) extraterrestrials are indeed targeting his house. Or, of course, that there's something fishy about the whole story.

The throwaway comment so exercised me (Lajic's "claim" does, after all, provide a quite definite answer to Fermi's question) that as soon as I got back from holiday I dug a little deeper into the story. A quick internet search shows that Lajic's story appeared in a clutch of newspapers and websites in the month of April 2008 (at which time his house had apparently been hit by meteorites five times) and then again in the month of July 2010 (after the sixth hit). Unfortunately it's not easy to track down the original source of these accounts. The 2008 report might have originated in an online publication that appeared, perhaps significantly, on the first of the month. The initial publication of the 2010 story seems to have appeared on 19 July, with one of the first to break the story being the UK-based tabloid newspaper *Metro*.

At this point it's worth highlighting the work of the English comedian and rationalist Dave Gorman, who was intrigued by the large number of "face of Jesus Christ" stories that appear in the *Metro*. (According to the *Metro*, the face of Jesus has in recent years appeared in a tree stump, chicken feathers, a tea towel and a dozen other places besides.) Gorman decided to manufacture an image of Jesus using fabric conditioner on an old t-shirt; he then sent *Metro* a photograph of the stained garment along with a short, jokey story purportedly written by a student called Martin Andrews[40]—and, of course, the newspaper printed it. What's interesting is that, within hours, Gorman's fabricated story got picked up by other publications and websites. Even more interesting is the way in which the propagation of Gorman's story seems to have mirrored the propagation of Lajic's story, even down to the twisting of language. In Gorman's source material a student jokes that the stain looks like "Jesus posing as the Fonz from Happy Days"; after a Chinese whisper or two the student is "convinced" that this is the face of Christ. In the 2010 *Metro* report Lajic merely "says" he is being targeted; the language quickly changes so that Lajic "insists" that his house is being "bombarded".

A moment's reflection, then, suggests that Radivoje Lajic's story bears more resemblance to an April Fool's joke than a resolution to the Fermi paradox. It's therefore rather depressing to see so many news outlets report a story such as this straight and choose to file it under "bizarre" without even a hint that there could be mundane and terrestrial factors at play; it's sad to read the comments section of blog posts about the story and see that, for every cynic who suggests neighborhood children might be flinging stones at Mr Lajic's house there's a believer who takes this as proof of "something going on"; and it's a shame to see the story reach even serious outlets for science communication without critical comment. Many fine SF stories tell the tales of unfortunate individuals being targeted by extraterrestrial beings. But those are stories. There's no evidence that such beings are targeting Radivoje Lajic—or, indeed, any of those other people who claim to have had their lives adversely affected by aliens. As with the previous two solutions, we can't seriously accept that something such as this could resolve the paradox.

## Solution 4 They Are Watching Us from UFOs

*You can observe a lot by watching.*

Yogi Berra

Shakespeare has Juliet ask: "What's in a name?" In certain situations the answer is: everything. For example, for thousands of years people have seen strange lights in the sky.[41] No great attention was paid to the phenomenon until the lights acquired a catchy name. Call them "flying saucers" and suddenly everyone is interested.

We can date the precise moment when a person first saw a flying saucer. On 24 June 1947, Kenneth Arnold was flying his private plane[42] over the Cascade Mountain range in Washington State. From his cockpit he saw several airborne objects; when he landed he reported his sighting, describing the objects as skipping "like saucers across a pond". The name stuck. The press was hungry for gossip about these "flying saucers" and the term found resonance with an American public nervously entering the Cold War. Many people took it for granted that the flying saucers were crewed by aliens—either Russian or extraterrestrial.

If flying saucers are real and if they are indeed spacecraft crewed by extraterrestrials then the Fermi paradox is instantly resolved. Of all the proposed resolutions of the paradox this one has perhaps most support amongst members of the public. As surveys consistently show,[43] over a third of Americans believe flying saucers are visiting Earth right now; the proportion of Europeans holding that belief is smaller but is still significant. Many people even believe that a flying saucer crashed at Roswell, New Mexico, in late June/early July of 1947 (suspiciously close to the time of Arnold's sighting), and that the US military recovered alien bodies from the wreckage. Nevertheless, science is not a democratic process. Hypotheses are not proven right or wrong through a ballot. No matter how many people believe in the truth of a particular hypothesis, scientists will accept the hypothesis (and then just provisionally) only if it explains many facts with a minimum of assumptions, if it can withstand vigorous criticism, and if it does not run counter to what is already known. So the question is: how well does this hypothesis—that flying saucers are evidence of ETCs—stand up to scrutiny?

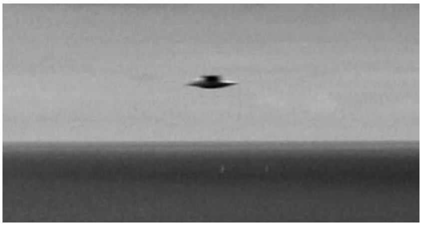


**Fig. 3.1** A UFO—or is it an IFO? This photo was taken by a holidaymaker near St Austell in Cornwall in August 2011. (By coincidence I was holidaying in St Austell during this period, but I can confirm that this isn't my handwork.) The object is entirely identifiable: it's a flying seagull depositing a sticky white dielectric substance. The only mystery is why GCHQ, a British intelligence and security agency, should choose to include it in a presentation about UFOs and internet use. (Credit: Initial creator unknown; presentation created by GCHQ and leaked by whistleblower Edward Snowden)

Before discussing this, I believe it's best to employ the neutral term unidentified flying object, or UFO, when examining claims about strange lights or objects in the sky. The term was coined by Edward Ruppelt,[44] who undertook an investigation of UFOs for the USAF. The terms UFO and flying saucer are often used interchangeably, which is unfortunate, but if used correctly a UFO is precisely what its name implies: an aerial phenomenon that is *unidentified*. Everything we see in the atmosphere is either a UFO or an IFO (an identified flying object). Upon investigation, a UFO can become an IFO; and an IFO *might* turn out to be a flying saucer—but we can only reasonably make that determination after careful research.

Under this definition it's undeniable that UFOs exist. Indeed, it's tempting to say that if you haven't seen a UFO then you haven't been looking hard enough. The sky is host to a myriad of interesting phenomena, both natural and artificial. Upon even a cursory examination, however, most UFOs are explicable—they quickly become IFOs. For example, people often mistake Venus for an artifact; aircraft can create unusual visual effects; each day, 4000 tons of extraterrestrial rock and dust burn up in the Earth's atmosphere and produce the occasional light show; and so on. A few UFOs will be the result of unusual but mundane events; for example, one mysterious light turned out to be the result of a golfball thrown onto a bonfire. The files of UFO researchers must be filled with the observations of startled people noticing one-off events such as this. Some other UFOs would undoubtedly require a thorough and detailed investigation before being classified as IFOs. For example, the novaya zemlya, fata morgana and fata bromosa mirages, which have fooled people for hundreds of years, are caused by relatively rare atmospheric conditions; perhaps the same mechanism can explain some UFOs? Perhaps some of those strange lights in the sky are the beams of car headlights refracted through abnormal air conditions? The explanation of some UFOs might even require advances in science. For example, the phenomenon of ball lightning is still poorly understood and not particularly well researched—ironically for the same reasons that many scientists feel uncomfortable with the idea of UFOs. Finally, some UFOs turn out to be the result of deliberate hoaxes.

Upon investigation, then, most UFOs become IFOs. But each year has a tiny residue of cases in which no rational account is forthcoming. We shouldn't find this too surprising. After all, as the noted skeptic Robert Sheaffer [45] points out, police don't achieve a 100% success rate for solving murders. But while it's generally accepted that not all murder cases will be solved many people find it unacceptable that a UFO might remain unidentified; they want an explanation for *all* sightings. How, then, should we try to account for the inexplicable UFOs?

If a reported UFO were simply a light in the sky then one could reasonably argue that, no matter how strange the light appeared to be, we don't *have* to explain it. Life is too short for scientists to explain every instance of every phenomenon. A scientist no more has to explain the detailed circumstances that produced a particular light in the sky than he has to explain the shape of, say, the strange Pooh Bear-like cloud formation I can see out of my window as I write this. There are more important things to study. But what if an explanation is *demanded*?

My feeling is that we need no new hypothesis to explain the residue of UFO sightings: the reasons that account for most UFOs would account for *all* UFOs if we were clever enough and had enough resources and patience to carry through the necessary investigations. Sheaffer highlights the interesting finding that the percentage of apparently inexplicable UFOs varies little within the overall number of sightings. In other words, whether it's a busy year or a quiet year for UFO sightings, the IFO/UFO ratio is about the same—which is hardly what one would expect if those inexplicable UFO sightings represented alien craft. The simplest explanation of this finding is that, in Sheaffer's words, "the apparently unexplainable residue is due to the essentially random nature of gross misperception and misreporting".

None of this proves that we are *not* receiving visits from etcs. (Nor does it prove that when we see UFOs we aren't watching fairy craft, manifestations of ghosts or the sporadic intersection of higher-dimensional beings with our own spacetime.) But neither does the observation of UFOs prove that we *are* receiving visits. The cast-iron, unimpeachable sightings of lights in the sky are just that: sightings of lights in the sky. If you see a strange light in the sky and you can't explain it then you have to leave matters there: you've seen a UFO. If you call that light in the sky a flying saucer then you've identified it, but without any grounds for making that identification. The existence of

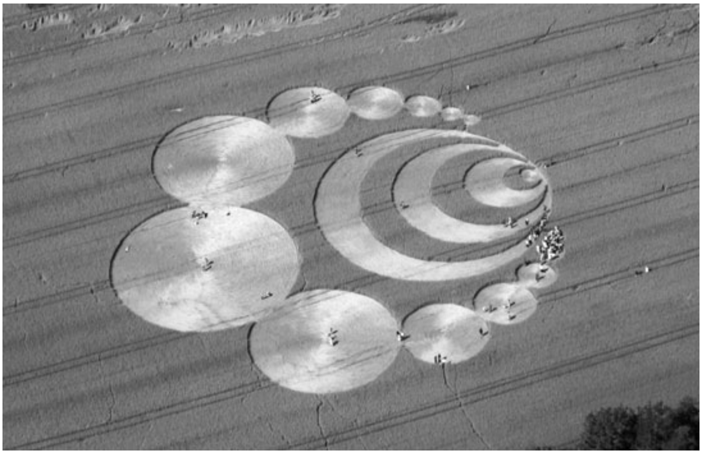


**Fig. 3.2** The majority of crop circles occur in southern England, but this is a Swiss version. Such a beautiful pattern couldn't be made by natural phenomena such as wind or rain. Surely the conclusion, therefore, should be that it was made by people—not that it was made by a flying saucer! Just visible in the photograph are sightseers. In England, at least, it's possible to earn money by providing guided tours of a particularly intricate crop circle. (Credit: Jabberocky)

unidentified aerial phenomena simply doesn't provide any evidence for the existence of extraterrestrial visitations.

Of course, some flying saucer claims go far beyond mere lights in the sky.

For example, some enthusiasts claim that alien craft have crashed; the Roswell incident, mentioned above, is the most widely publicized instance. Regardless of whether it's likely that a craft could successfully travel interstellar distances yet fail to negotiate a planetary atmosphere, the *evidence* for all these reports is shoddy. An item of advanced equipment or a piece of an unknown alloy or a sample of alien tissue would prove the case. Instead, we get claims of military cover-ups and government conspiracies and, in the case of Roswell, a video of an alien autopsy—a video that turned out to be a (profitable) hoax. One occasionally sees claims from people saying that a flying saucer landed and aliens disembarked—and proceeded to anally probe them, medically examine them or, bizarrely, mutilate their cattle. (Some even claim, as we saw in Solution 2, that the aliens get jobs in Downing Street or the White House). It's hardly worth stating that the *evidence* required to support such claims is poor.

A more restrained claim is that alien craft occasionally land without fuss and without attempting to make contact. Consider, for example, the crop circle phenomenon. (Crop circles actually have a variety of shapes. There are crop hexagons, crop fractals, crop Nike swooshes . . . but they are generally referred to as circles.) Since it's difficult to understand how a complex design could be imprinted in a field of wheat through a natural process this was proof, according to some cereology experts, that at least some crop circles were caused by flying saucers. Matthew Williams, a self-confessed maker of crop circles, took issue with this conclusion; he wanted to demonstrate that it's quite easy for people to make elaborate crop circles. In 2000 he proved his point by making a 7-pointed shape—something that one leading cereologist claimed was impossible to fabricate. Armed only with some planks, bamboo poles and a torch, Williams proceeded to create his 7-point shape over three nights in a farmer's field of ripening wheat. Personally, I admire his devotion to rationality, but the farmer wasn't impressed; neither was a judge, who issued a £100 fine for criminal damage and £40 costs. Williams continued to produce crop circles, only ceasing his activity in 2013 due to worsening hay fever.

Unfortunately, even though people have confessed to making crop circles and have shown others how to do it, there remain those who are convinced that the crop circle phenomenon is an unexplained and perhaps inexplicable mystery. How can one argue with people who are so wedded to a particular idea, except to say that when thinking about phenomena we should use Occam's razor?[46] One formulation of the razor is that explanations of unknown phenomena should first be sought in terms of known quantities. We can explain crop circles, cattle mutilations and other fringe phenomena in terms of known quantities. We simply don't need unknown hypotheses to explain them.

Whenever an extraordinary claim is made for flying saucers, no extraordinary evidence is presented to support the claim. Instead, we get lies, evasions and hoaxes. The flying saucer hypothesis might be the most popular explanation of the Fermi paradox, but surely there are better explanations.

* * *

Incidentally, I should state here that I have seen a UFO, and it remains one of my most vivid memories. While playing soccer in the street as a child—this was before the increasing number of cars stopped children playing in the street—I looked up and saw a pure white circle about the size of the full moon. Protuberances on either side of the circle made it look rather like Saturn showing its rings edge-on. Whatever it was, it seemed to hover for a few seconds before moving off at tremendous speed. I was with a friend, who also saw it and remembers it still. Interestingly, we differ in our recollections: I remember it shooting away to our left as we watched; my friend says that it moved away to our right. (People are poor observers, and I know from experience that I am a *very* poor observer. But I am adamant that it moved to the left!) We definitely saw *something* in the sky that day and I have absolutely no idea what. But no, it wasn't a flying saucer. It was just a light in the sky.

## Solution 5 They Were Here and Left Evidence of Their Presence

*Tell them that I came, and no one answered.*

Walter de la Mare, *The Listeners*

The evidence that extraterrestrials are currently visiting Earth is essentially non-existent. But perhaps they visited Earth, or at least our Solar System, some time in the past—long ago, maybe, at a stage in human development when no one could recognize them for what they were? If that happened then they might have left behind evidence of their technology, either here on Earth or at least in Earth's neighborhood. Is there any evidence for this? This is an important question because it has the potential to widen the scope of the search for extraterrestrial intelligence: in addition to looking for signals (an activity that's discussed later in the book) we could look for footprints of alien technology.[47] Let's work through the Solar System, beginning at home.

### Earth

Suppose extraterrestrials visited Earth in the distant past—tens of millions of years ago, say. Could they have left behind physical traces that might yet survive?[48] Well, it's exceedingly unlikely. Earth is an active planet and tens of millions of years of glaciation, tectonic activity and weathering would erase most types of evidence. Nevertheless, one can imagine a couple of activities whose aftereffects we might possibly detect. For example, some radionuclides possess half lives measured in millions of years, so if extraterrestrial visitors dumped nuclear waste on the Cretaceous landscape it might leave a trace that we could detect today. (There was a natural deposit of uranium in Oklo, Gabon that went critical when Earth was about two thirds of its present age; the Oklo reactor[49] left behind a range of radionuclides that we can detect 1.7 billion years later.) If we found plutonium traces, for example, we would be forced to explain the find in terms of a technological civilization—either our own or an ETC: the half life of plutonium is such that there are no natural sources of the element. A second activity that might leave its mark over geological timescales would be large-scale quarrying: if the extraterrestrials came here to engage in the industrial mining of minerals then current geological surveys would in principle be able to detect the quarries (in the same way that meteor impact craters created millions of years ago are detectable, even if they are buried under further strata).

It would cost next to nothing to look for evidence of anomalous radionuclides or ancient quarries—geologists are performing the surveys in any

**Fig. 3.3** This is the result of a meteor not a spacecraft. The meteor entered Earth's atmosphere on the morning of 15 February 2013 at a speed 60 times that of sound. It exploded as an airburst, about 23 km above Chelybinsk. This image is a simulation by Mark Boslough, rendered by Brad Carvey, using a Sandia National Laboratory supercomputer. The original photograph of the meteor was by Olga Kruglova. (Credit: Sandia Labs)

case—so even if the chance of finding something is exceedingly low there surely can be no harm in keeping an eye open for signs of past extraterrestrial visitations. If you don't look, you can't find. However, even if you think it likely that extraterrestrial industrialists visited Earth in the past (and personally I find it extremely improbable: I can't imagine why intelligent creatures would travel light years just to pan for gold) finding evidence of the visit would require a *huge* dollop of luck. Perhaps there have been signs of more recent visits?

The famous Tunguska event of 1908, an explosion that felled tens of millions of trees across the Siberian taiga, packed quite a punch—the equivalent of about 15 million tons of TNT. However, the first people to reach this desolate place found none of the debris that one would expect from the likeliest cause of such an event—an asteroid impact—so for many years the event was something of a mystery. Once the immense power of nuclear explosions became apparent, soon after World War II, the notion circulated that the Tunguska event had been a nuclear blast—the crash-landing of an alien nuclear-powered spacecraft. The idea was taken semi-seriously, and there was a simple means of testing it: go to Tunguska and search for traces of radioactivity. This was done, but researchers found no traces of the radioactivity that would have come from a nuclear engine; they also ruled out an antimatter engine. Scientists now believe that the Tunguska event was the result of a stony meteoroid, or possibly a comet, that exploded in the atmosphere.

There have been Tunguska-like events in the past and there will be similar events in the future. Indeed, around the time I began thinking about a second edition of this book, the Russian city of Chelyabinsk suffered an airburst. The Chelyabinsk event was minuscule in comparison to Tunguska, but because it occurred over an inhabited region it caused injuries to more than 1200 people. When accounting for the Tunguska and Chelyabinsk events there's simply no need to invoke the hypothesis of a downed spaceship; it's Nature that supplies us with these spectacular displays. If a spaceship ever did crash-land in the past, we haven't found the evidence—Roswell notwithstanding.

We might lack evidence for the existence of extraterrestrial quarries or crashed craft, but some would argue that these are the wrong things to be looking for. In the 1970s, Erich von Däniken became famous for a series of books[50] in which he claimed that extraterrestrial visitors built many of the enigmatic structures we see dotted around the world—Stonehenge, the lines on the Nazca Plain in Peru, the Easter Island statues, and so on. None of the books contained proof to back up his claims but his large reading public were nevertheless happy enough to stick with him during his lengthy spell in prison for fraud; they supported him after his claims were painstakingly and thoroughly debunked; only when they became bored, and taste and style moved on, did they drop him. Now, like several pop groups from that era, von Däniken and his ideas are back in fashion even though, in the forty-odd years since the books were first published, he's still produced no proof to support his speculations—something that von Däniken himself cheerfully admits and seems to find irrelevant. Since the supporters of von Däniken are unlikely to be swayed by rational argument, we might as well move on—and accept that there's no evidence that members of an ETC have ever been on Earth. As always, this is not to say definitely that they have *not* been here. But in the absence of any evidence to the contrary, we might as well assume that Earth has been untouched.

## Moon

The Moon is a much less active place than Earth. If extraterrestrials visited the Moon tens of millions years ago and left behind traces of technological activity—such as large-scale quarrying or the dumping of radioactive waste, as mentioned above—then there's a good chance that those traces would still be visible. Structures would not have to withstand wind or rain or glaciation; a radioactive dumping site would not be buried by tectonic processes. Meteorites occasionally strike the Moon's surface, throwing up dust and regolith, but if items of interest are more than a few meters in size then it would be hundreds of millions of years before this "gardening" process covered them.[51] Furthermore, the Moon is close enough for us to search for evidence of past extraterrestrial visitations.[52] Indeed, we could already perform a search: NASA's Lunar Reconnaissance Orbiter began mapping the Moon at high resolution in 2009. The Orbiter's cameras have a maximum resolution of 50 cm per pixel, which is sufficient for them to detect evidence of visitations. (The cameras have already detected evidence of activity on the Moon, though of course these were human visitations made during the Apollo program.) Nowadays there are a number of highly effective "citizen science" projects, whereby members of the public donate their time to various scientific endeavors: there are occasions in which the pattern-recognition capabilities of the human brain far outperform computers. Well, a citizen science project devoted to the Lunar Reconnaissance Orbiter would enable us to search the entire lunar surface for evidence of past extraterrestrial activity; it would be a low-cost addition to the SETI program.

Although there's no evidence that ETCs have ever visited our satellite it's worth mentioning that, until fairly recently, some people claimed to have seen signs of extraterrestrial activity on the Moon. In 1953, for example, the astronomer Percy Wilkins discovered what appeared to be an artificial structure—a bridge.[53] When other observers failed to see the structure through more powerful telescopes the astronomical community decided, quite reasonably, that the bridge was a trick of the light. This didn't dampen the enthusiasm of those who believed in the Moon as an abode of alien life. Enthusiasts pointed out that the Moon shows only one side to Earth (to be precise, due to the phenomenon of libration we see only 59% of the Moon's surface). If we never see 41% of the lunar surface, who knows what might be hiding on the far side of the Moon? It wasn't until the late 1970s, well after the many landers and orbiters had observed the entire surface of the Moon, that "life" enthusiasts finally stopped promoting the idea of bridges and other artifacts.

## Earth–Moon Lagrangian Points

As we shall see later, one can reasonably argue that an ETC wishing to explore our Solar System would send small unmanned (unaliened?) probes rather than a fleet of crewed spacecraft. Where might we find such probes? There are three cases to consider. First, the probes could be programmed to attract our attention. Since we see no evidence for beacons, it's safe to assume that such probes aren't here. Second, the probes could be programmed to hide from us. Since we are unlikely ever to find such probes, we needn't spend time discussing the best strategy to observe them. Third, an ETC might send probes and not care whether humans observe them. If *that* is the case, where might we find them?[54]

We can reasonably argue that of all the planets in the Solar System, ours is most worthy of study. Earth is an interesting planet for a variety of reasons—most importantly it is, as far as we know, the only planet to harbor life. So probes would most likely be programmed to investigate Earth. (This argument of course reeks of anthropocentrism. Who knows what an alien mind might want to investigate? Who knows what technology it might employ? But such logic is all we have, so we lose nothing if we continue the argument and see where it leads us.) Earth's surface would be a poor site for long-term studies of our planet. It would make more sense to view the entire planet from space,[55] where solar energy is more readily available, and where there's no need for the probe to protect itself against the effects of the Earth's geological activity.

Several types of orbit are suitable for parking an observational probe, but perhaps the best known are the Lagrangian points.[56] If a small mass is near two much larger orbiting masses, then there are five points at which the small mass can orbit at a fixed distance from the larger masses. These five Lagrangian points mark the positions where the gravitational pull of the two larger masses exactly balances the centripetal force required to rotate with them. Three of the Lagrangian points—L1, L2 and L3—are unstable: nudge the small mass and it will move away from the L point. But L4 and L5 are stable: nudge the small mass and it will return to the L point. (To be precise, L4 and L5 are stable only if the most massive of the three bodies is at least 24.96 times as massive as the intermediate body. This condition is satisfied in the Sun–Earth system, since the Sun is much more massive than Earth. The condition is also satisfied in the Earth–Moon system, since Earth is 81 times as massive as the Moon. The Sun's gravitational influence tends to destabilize the L4 and L5 points of the Earth–Moon system; however, it smears the stable *points* into *volumes* of space in which stable orbits exist.)

**Satellites at the Sun–Earth Lagrangian Points** The L1 point is home to the satellites ACE, SOHO and WIND—from this vantage point the satellites have an uninterrupted view of the Sun. The L2 point has been home to several high-profile astronomical missions including WMAP, the *Planck* Space Observatory and the *Herschel* Space Telescope—these satellites have observed the universe in unprecedented detail, and many more L2 missions are planned. Note that, although the L1 and L2 points are unstable, it's possible to find orbits around those points that enable a spacecraft to stay in place while expending only a small amount of energy. The space agencies are unlikely to find a use for the L3 point of the Sun–Earth system since it's on the opposite side of the Sun from Earth. The regions around L4 and L5 contain interplanetary dust and at least one asteroid.

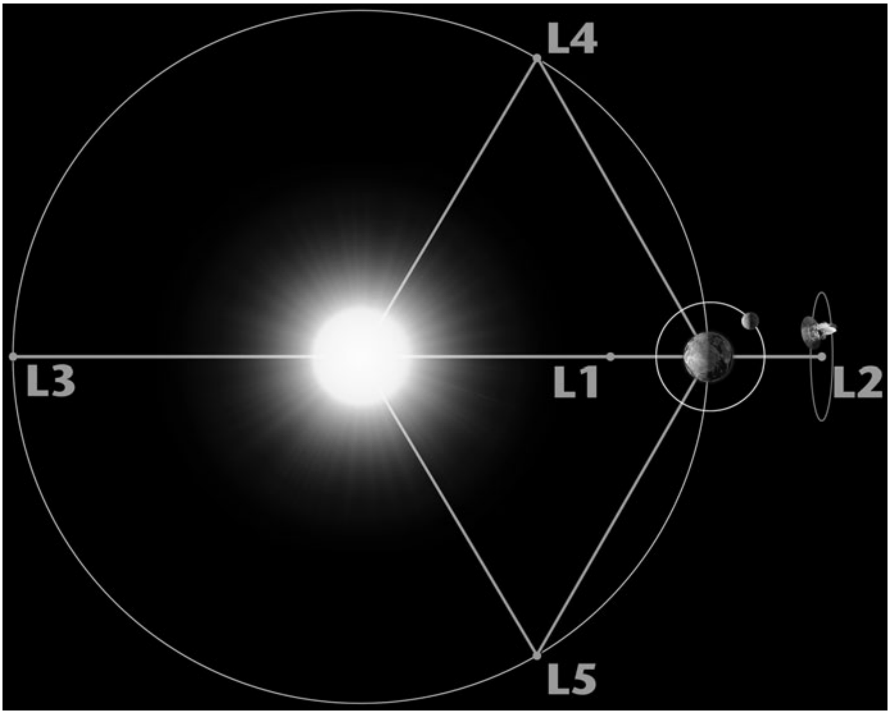



**Fig. 3.4** The five Lagrangian points of the Sun–Earth system (not drawn to scale). In general, Lagrangian points are places in the vicinity of two orbiting masses where a third smaller body can maintain a fixed distance from the larger masses. The points L1, L2 and L3, which lie on a line connecting the two large masses, are unstable: after a perturbation, the small body will move away from the Lagrangian point. Under certain circumstances, the points L4 and L5 are stable: after a perturbation, the small body will return to the Lagrangian point. A similar configuration exists for the Earth–Moon system: might there be extraterrestrial probes parked at these Lagrangian points to observe Earth, just as we park probes at Sun–Earth Lagrangian points to observe the universe? (Credit: NASA)

The space agencies NASA and ESA already make heavy use of parking facilities offered by the Sun–Earth Lagrangian points. If NASA and ESA find it convenient to use those points then perhaps ETCs would do so too. Perhaps we might find probes at Lagrangian points in the Earth–Moon system? In particular, perhaps we might find probes at the L4 and L5 points since at those locations the probes could in principle observe over long periods of time without expending too much energy. Well, at least one dedicated search has been made. Furthermore, astronomers have already studied the L4 and L5 points of the Earth–Moon system, since the points are interesting from a general astronomical viewpoint. In neither the dedicated search nor the general scans was any evidence of probes found. Furthermore, a recent study has shown that the L4 and L5 points in the Earth–Moon system might not provide the stable vantage point[57] that was once thought. If it were just the Earth and Moon that had to be considered then L4 and L5 would indeed be stable. But the Solar System contains other objects. It turns out that the gentle gravitational tug of the other planets perturbs the stability of the Lagrangian points, and any probe placed at L4 or L5 would eventually gently drift away. It would be strange if we found evidence for extraterrestrial probes there.

Increasingly, other near-Earth orbits are being scanned—this time by astronomers searching for potentially lethal asteroids. As a by-product of this research we might hope to find artifacts; so far, though, none have been found. Probes would give off heat, but no anomalous infrared signals have been observed; probes might be expected to transmit messages back to their creators, but no such transmissions have been detected.

Some people have claimed that the phenomenon of long-delayed radio echoes (LDEs)—echoes that appear between 3 and 15 seconds after the transmission of a radio signal—is best explained in terms of electromagnetic transmissions from ETC probes. The LDE phenomenon has been observed since the dawn of radio, and it remains somewhat mysterious. Radio echoes from the Moon are common, but this can't explain the LDE phenomenon because the echoes appear 2.7 seconds after transmission of the main signal—this being the time it takes light to travel to the Moon and back. Similarly Venus, the nearest planet, can't be the culprit: echoes only appear 4 minutes after the main signal. One explanation is that the echoes are radio returns from ETC probes lying beyond the distance of the Moon. A more prosaic explanation[58] is that they are a natural phenomenon caused by plasma and dust in the Earth's upper atmosphere.

Although the search for near-Earth probes is not complete—indeed, the search has hardly begun, as Earth could be bathed in signals at certain frequencies and we wouldn't necessarily know about them—all observations to date have given a negative result. Our telescopes *have* occasionally detected transmissions from probes in the depths of our Solar System—but these are transmissions from our own spacecraft.

## Mars

As we shall see later, there are good reasons for supposing Mars might have played a role in the development of life on Earth. But could it be home to its own life—and its own technological civilization?

Mars has indeed long been thought to be home to life,[59] but much of the fuss stemmed from a mistranslation. Giovanni Schiaparelli, in a series of observations beginning in 1877,[60] saw features on Mars that he called *canali*—an Italian word meaning "channels" or "canals". It's clear from his writings that when he named these features Schiaparelli thought *natural* processes had formed them. English-speaking astronomers, however, translated the word as "canals"—*artificial* structures connecting two bodies of water.

Percival Lowell also saw[61] the surface features recorded by Schiaparelli, and he finally counted 437 of them. However, Lowell failed to acknowledge he was working at the limits of observation; he failed to appreciate that evolution has primed the human visual system to look for familiar features in random patterns. He became convinced he was seeing artificially constructed linear canals, and he speculated that the canals supplied water from the polar caps to a desert world. The notion of canals was in the public consciousness anyway—the Suez Canal, a modern wonder of the world, had opened to navigation in 1869—and the general public was gripped by the possibility that intelligent beings had constructed the Martian canals. Science fiction writers were quick to use it as a source of stories. It was a popular and romantic notion, and even as late as 1960 some maps of the planet showed oases and canals. Several astronomers continued to believe that seasonal changes in the Martian surface markings might be due to changing vegetation patterns.

Meanwhile, in the early 1960s,[62] Shklovsky discussed a peculiarity in the orbit of Phobos, the larger of Mars' two moons, and offered an ingenious explanation.

The orbit of Phobos is decaying. The peculiarity was that, according to observations made by the astronomer Bevan Sharpless in the 1940s, the *rate* of decay was difficult to explain. Several mechanisms were suggested—the effect of a hypothetical large Martian magnetic field, tidal interaction with Mars, a possible solar influence—but none of them were feasible. Neither was the obvious explanation, namely that Phobos was passing through the thin outer regions of the Martian atmosphere, because atmospheric drag wouldn't affect a rock the size of Phobos to the extent observed by Sharpless. The audacious Shklovsky wondered whether Phobos were *hollow*. A hollow Phobos would be less massive than its size would suggest, so its orbit would be much more affected by the Martian atmosphere. If Phobos really were hollow, then it couldn't be natural: Shklovsky therefore suggested that the satellite was artificial—the product of a Martian civilization. (It was a suggestion more imaginative than anything in the books of von Däniken, yet it was based on the best available observational data.) Shklovsky thought the satellite would have been launched millions of years ago, but other scientists thought the launch could have been more recent. Frank Salisbury pointed out[63] that the two Martian moons were discovered in 1877 by Asaph Hall, who used a 26-inch telescope. Fifteen years earlier, when Heinrich d'Arrest trained a larger telescope on the red planet, the

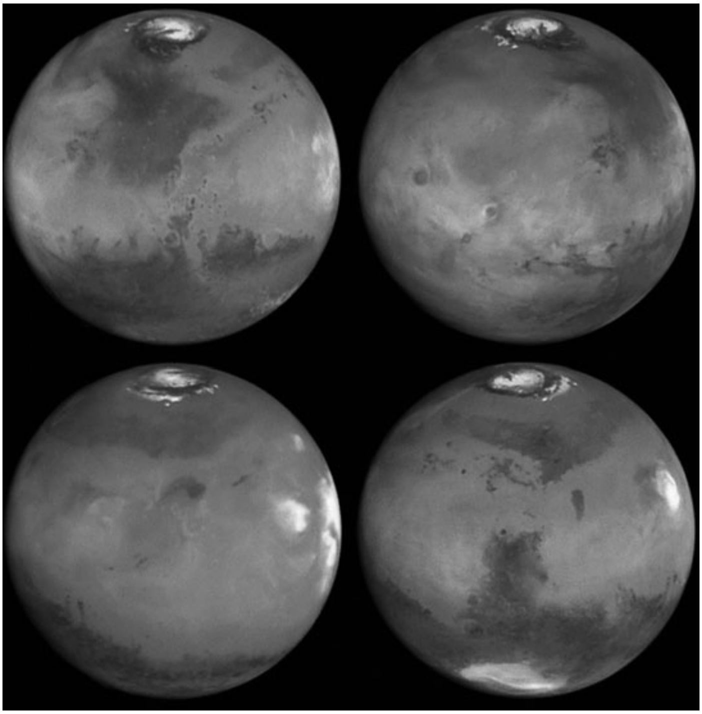


**Fig. 3.5** Four faces of Mars, as photographed by the Hubble Space Telescope on 30 March 1997. There are no signs of canals. (Credit: Phil James (University of Toledo), Todd Clancy (Space Science Institute), Steve Lee (University of Colorado) and NASA/ESA)

conditions for viewing Mars had been better. How could d'Arrest have missed seeing the moons in 1862? Was it possible, Salisbury asked, that the moons were artificial satellites launched some time between 1862 and 1877?

The romantic notion of an advanced Martian civilization capable of building canals and launching satellites didn't survive the 1960s. It was laid to rest when the early Mariner spacecraft flew by at close range, returning photographs that showed none of the canals seen by Lowell. The Viking landers of 1976 and the Pathfinder and Mars Global Surveyor missions of 1997 also failed to find canals. Similarly, the flyby missions saw nothing at all artificial about Phobos. It's a small pockmarked piece of rock—perhaps a captured asteroid (although the origin of the two Martian moons remains a matter for research). Furthermore, although its orbit is indeed decaying, recent measurements indicate that the rate of decay is only half that calculated by Sharpless. Armed with this improved measurement, theorists can now explain the origin of the drag on Phobos: it's the result of tidal interaction with Mars. (Phobos draws closer to Mars by about 1 inch every year. The satellite will hit Mars some time within the next

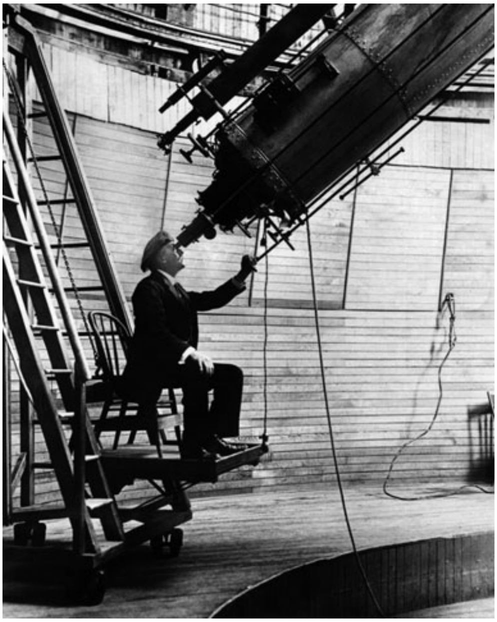


**Fig. 3.6** Percival Lowell in 1914, using the 24-inch refractor at the Lowell Observatory. (Photographer unknown)

40 million years, leaving a basin the size of Belgium. Although 40 million years is a short time on the astronomical scale, it's a long time on the human scale. A pity—it would be a spectacular event.)

The evidence from the various flyby, orbiting and lander missions almost killed the belief in an ancient Martian civilization. Almost, but not quite. In 1976, Viking photographed the Cydonia region on Mars, and nasa released the photographs soon afterward. Almost immediately, enthusiasts pointed out that one of the low-resolution photographs appeared to show a human face. You could make out an eye, a mouth, and a nostril (though the enthusiasts often failed to point out that the "nostril" was actually an artifact of the way the image had been processed, and corresponded to no physical structure on Mars). The face was large, roughly a square of 1 km, and seemingly carved out of stone. nasa scientists emphasized that this was a natural formation; the image was simply the result of sunlight falling on a hill one Martian afternoon. Others argued that the formation was an artificial structure; the stone "face" was proof that Mars was once home to an ancient civilization.

If you search through a large collection of random data long enough and hard enough, conveniently ignoring arrangements of the data that are of no interest and not defining beforehand what you are looking for, then eventually you

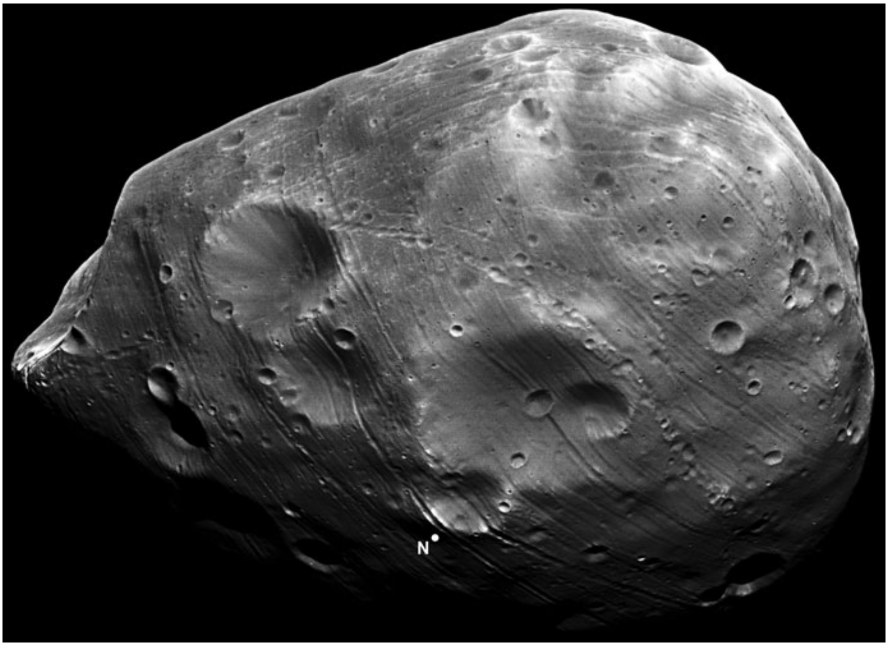



**Fig. 3.7** Phobos, the larger of Mars' two moons, is a potato-shaped rock about 16 miles by 10 miles in size. It is quite possibly a captured asteroid. The letter N on the image marks the north pole. (Credit: G. Neukum (FU Berlin) et al., Mars Explorer, DLR, ESA)

**Fig. 3.8** The "face" on Mars. This low-resolution image contains many black dots, which are artifacts of the image-processing techniques employed by the Jet Propulsion Laboratory, and do not correspond to any Martian feature. (Credit: NASA)

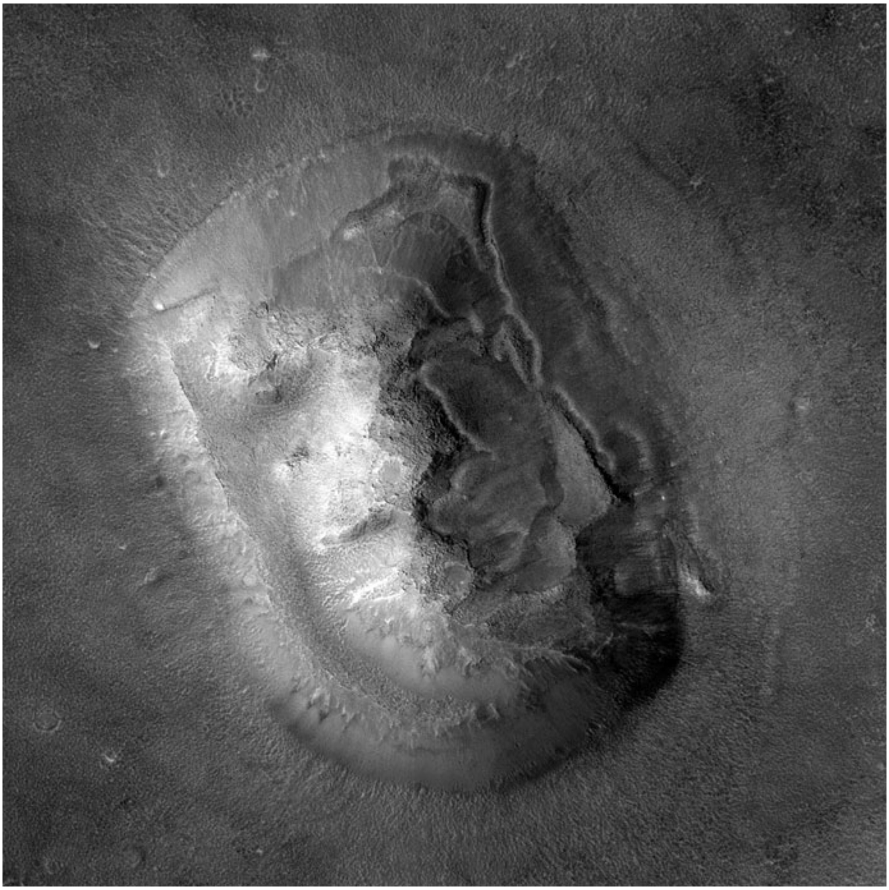


**Fig. 3.9** Is it a shield? Is it a footprint? Is it Chewbacca? A high-resolution image of the Cydonia region, this time taken by Mars Global Surveyor in 1998, shows no evidence for a face. (Credit: NASA/JPL)

*will* find something remarkable. The surface of Mars covers 150 million $km^2$; it would be strange if one of those square kilometers did *not* vaguely resemble something familiar. Planetary scientists argued that the Martian “face” has as much significance as the patterns you see in the coals on a fire. It was another instance of an observer imposing meaning on a meaningless pattern.

Mars Global Surveyor revisited the Cydonia region and took a more detailed photograph. The evidence for the face, of course, evaporated. (It’s only fair to point out that the illumination is different in the two photographs. Nevertheless, modern computer imaging techniques can retain the detail of the Global Surveyor photograph while simulating the feature in the same afternoon light that Viking saw. If I scrunch my eyes, then I can just about make out Chewbacca from *Star Wars*—but no human face.)[64]

## Asteroids

Michael Papagiannis argued[65] that we need to rule out the possibility of ETCs being in the Asteroid Belt before we can conclude they aren't here in the Solar System. The Asteroid Belt would be an ideal place for ETCs to set up space colonies. They could mine the asteroids for natural resources,[66] and they would have plentiful supplies of solar energy. Who knows—perhaps the fragmentation of the Asteroid Belt components is the result of large-scale mining projects by ETCs? If space colonies *were* in the Asteroid Belt, we wouldn't necessarily know about them: craft that were, say, 1 km or less in size would be difficult to distinguish from natural asteroids.

On the other hand, if they really are in the Asteroid Belt, there are questions to ask. Why have we detected no leakage of electromagnetic radiation? Why have we not observed a single object that possesses an effective temperature higher than is justified by its distance from the Sun? And why, if they are there, have they chosen to remain silent for so long?

## Outer Solar System

Beyond the asteroids we see numerous "anomalies"—such as the axial tilt of Uranus or the retrograde orbit of Triton—that could, if we were so inclined, be taken as evidence for tampering by ETCs. And beyond the orbit of Neptune, from a distance extending between about 30–50 astronomical units from the Sun, lies the Kuiper Belt. More than one thousand objects have been found in the Kuiper Belt and it's believed that the Belt contains more than one hundred thousand objects. Most of them are small; the largest is Pluto, which in 2006 was ignominiously demoted from its status as a planet to a mere trans-Neptunian object. Before its reclassification, David Stephenson suggested that Pluto's unusual orbit might be the result of an astroengineering project.[67] However, all these "anomalies"—Uranus' tilt, Triton's retrograde motion, Pluto's eccentric and inclined orbit—can be explained more prosaically as the result of collisions and interactions that took place in the early history of the Solar System. There is simply no need to invoke other explanations. Nevertheless, Kuiper Belt objects might play a role in the search for ETCs. In 2012, the Harvard astronomer Abraham Loeb and the Princeton astronomer Edwin Turner published a paper in which they pointed out that biological creatures, once they possess a certain level of technological civilization, are likely to artificially illuminate their home planet during the dark phase of its diurnal cycle (they'll light up at night, in other words). Our own civilization employs two types of night-time illumination—quantum (from devices such as LEDs and fluorescent lamps) and thermal (incandescent light bulbs)—and both of them have a spectral signature that is quite different from the natural emission one gets from a naturally warm object such as a planet. If we could detect leakage from artificial illumination then we could infer the presence of an ETC. Imagine a city the size of Tokyo on a Kuiper Belt object, and suppose that its night-time illumination was at the same level as present-day Tokyo. Loeb and Turner showed[68] that existing telescopes could detect that artificial illumination. We could use this technique to search for civilizations in the Kuiper Belt right now. (A negative result would, as always, prove nothing: the aliens might shield the radiation from us, or be adapted to low light levels, or use technology that we can't even imagine . . . )

Ten times further out than the outer edge of the Kuiper Belt lies a zone that offers, one could argue, a rather logical place to look for probes in our Solar System. The argument begins with the observed fact that the path of a light ray bends if it passes close to a large mass. Einstein's theory of general relativity explains why this should happen—the mass causes space to curve, and light rays just follow the curvature. The path of a light ray is also bent if it passes through an optical lens. The "bending mechanism" involved in the two cases is quite different, of course, but in principle it's possible for a large enough mass to bring light to a focus in the same way that a lens can bring light to a focus—the mass is then acting as a gravitational lens. In 1979, von Eshleman, a professor of electrical engineering[69] at Stanford University, applied the theory of gravitational lensing to the case of the Sun. He showed that if a telescope could be placed at a distance of 548 AU from the Sun—almost 14 times the distance between the Sun and Pluto—then it would be able to take advantage of the magnification afforded by the Sun's gravitational lens. (The distance of 548 AU calculated by Eshleman is the minimum distance at which the Sun generates a gravitational lens. As one moves out past this minimum distance one finds an infinite number of focal points in all directions. Indeed, a telescope might be better placed at, say, 1000 AU since at such a large distance there would be less need to compensate for the complicating effects of the Sun's corona. But these are just details.)

**Optical and Gravitational Lenses** When light hits a boundary between regions in which it has different speeds of propagation it tends to bend towards the region in which it moves more slowly. (It's like driving a car and finding the nearside wheels hitting a patch of snow. The wheels on the road turn faster than the wheels on the snow, and the car turns—it starts to skid.) Since light travels much more slowly in glass than it does in air, a light ray bends when it passes from air to glass. The amount of bending depends on the angle at which the light hits the glass, but if you shape a lens correctly you can arrange matters such that all light rays hitting the glass will bend so that they converge on a single point: the focal point. The bending mechanism is different when it comes to gravity: light rays are bent in the vicinity of a large mass because space itself is warped by the presence of mass. A light ray follows the shortest path through space but, near a large mass, the shortest path is bent. However, although the mechanism is different, the end result can be the same.

A telescope based at the Sun's focal point would be an astronomer's dream: it could study distant planets, stars and galaxies with incredible detail. It could also be used as a powerful tool in the search for extraterrestrial intelligence, as pointed out by the Italian astronomer Claudio Maccone, who perhaps more than anyone[70] has championed the importance of the Sun's focal point for future astronomy missions. Maccone has also shown that the vast transmission gains in stellar gravitational lensing systems allow a means of communicating between nearby stars using only moderate transmission powers; the gain is really quite astonishing.

What does all this have to do with searching for evidence of extraterrestrials? Well, suppose an ETC sets out to explore the Galaxy by probe (we'll look at particular models of exploration later in the book). Communication between a probe and the parent civilization would presumably take place, but a sensible communication strategy would have the probe keep in touch with neighboring stellar systems rather than the original system. (The structure of the Milky Way, combined with its large size when compared with the limiting speed of light, mean that it would be difficult to maintain direct communication with the home system. Not only that, a communication strategy based on having the original system as a central hub means that the entire probe network would be at risk should the original civilization collapse, migrate or simply lose interest.) And the easiest way in which a relatively small object could communicate over interstellar distances would be if it used the gravitational lenses kindly provided by Nature. In other words, if *exploration* probes are or were here in the Solar System then we'd be likely to find *communication* probes at those focal points of the Sun—1000 AU would seem to be a sensible distance—that permit information exchange with nearby stellar systems. This suggestion, from the Belgian astrophysicist Michaël Gillon,[71] provides a simple means of directing the search for probes because, for any particular nearby star, we can easily calculate the location of the relevant points in space.

Unfortunately, as Gillon himself points out, it would be difficult to find the probes even if we knew where to look. Suppose the probe used a solar sail to provide its motive power. (The probe would need to compensate for the Sun's tiny, but not completely negligible, gravitational pull. An ETC might have access to energy sources we can only dream about, of course, but let's suppose the probe uses a large sail to capture solar power. This would be the best-case scenario in terms of our chances of observing them.) It turns out that, for a probe as massive as the Voyager spacecraft, a circular solar sail with a radius of about 500 m would be required. The question then becomes: can we detect a solar sail of that size from a distance of 1000 AU? Unfortunately, even with the spectacular observatories[72] that are being planned (such as the European Extremely Large Telescope), it won't be feasible to directly image such an object.

The probe would simply be too faint. A second possibility would be to use the occultation method, and search for the dip in brightness of a distant star that would occur when the probe moved in front of it. Turns out this isn't feasible either: the change in brightness would be too small, too fleeting. So having deduced that the solar focal regions of nearby stars are a good place to look for probes, do we now have to concede that we can't search for them? Well, Gillon has three more suggestions. First, we could send our own probes out there and take a look around. However, the two Voyager missions set off in 1977 and, at the time of writing, one of them is at a distance of 127 AU from the Sun and the other is at a distance of 104 AU. It will be a *long* time before one of our spacecraft reaches 1000 AU. Second, we could search for leakage radiation from the probes. That's possible in principle, highly unlikely in practice. Third, we could take the initiative and attempt to contact the probes directly—prod them with a radio blast and listen for a reaction. To my mind this third option is the only realistic search option for such probes, at least with the technology we'll have available over the next couple of decades. Let's send them a message and see if they respond. If they do, the world changes. If we hear only silence, which I strongly suspect would be the case ... well, we're back where we started.

When we begin to discuss the Kuiper Belt and the solar focal regions we begin to realize just how big the Solar System is. There are 50 billion billion billion cubic miles of space within a sphere that encloses the orbit of Pluto; and the Solar System extends to the Oort Cloud of comets, almost a light year from the Sun. The chances of finding a small alien artifact by accident are essentially zero. Only if an artifact draws attention to itself—by signaling us, perhaps, or by being in a visible location—will we detect it. We therefore can't rule out the possibility[73] that observational probes were once in the Solar System nor, indeed, that they are still here. Some would argue that until we *can* rule out that possibility, there is no Fermi paradox.

What we can say with confidence, however, is that no evidence for alien artifacts has yet been uncovered.[74] It certainly makes sense to search for them: as I've already mentioned the search would be cheap, and although the chance of success is exceedingly low the payoff for success is exceedingly high. But until such times as we observe them, why should we assume they are here?

## Back to Earth

Perhaps we are looking in entirely the wrong place. The discussion has revolved around alien artifacts—evidence of engineered objects. Perhaps an ETC has been here and left *information* rather than *things*?

An entertaining science fiction story from the 1950s suggested that the reason so many people dislike spiders is that the class Arachnida consists of alien creatures. They were carried here on some spacecraft, and then escaped; humans, instinctively recognizing the spiders' alien heritage, recoil from them. (Needless to say, spiders aren't aliens. As we'll see later, in Solution 64, all life on this planet is related. However much you might dislike spiders, you share a large part of your DNA with them.) In the 1970s, some scientists finally caught up with SF writers and made the suggestion that biological material might carry a coded message from an ETC. In theory, this would be possible: after all, the whole point of DNA is that it encodes information. Indeed, the genetic code can stay unchanged for billions of years and yet it could readily be modified if one wanted to embed some sort of signal.[75]

A message encoded in DNA seems an unlikely communication channel. For one thing, the sender could convey a message only to a planet that possessed the same biochemistry. (In our case, the sender's biochemistry would have to be based on L-amino acids, have protein synthesis based on the same genetic code as ours, and so on.) Even if it were possible for the recipient to distinguish between a natural and an artificial sequence, the message might become garbled through random mutations and the sender would have no way of preventing that. The vagaries of evolution might erase the message altogether. Nevertheless, genomic DNA is already being used here on Earth to store information so it's not impossible that others might have embedded a message. A few investigations have been performed[76] to test the idea, and analysis of certain types of viral DNA has found nothing resembling an artificial pattern. Now that biologists have sequenced the entire genome of several creatures, including man, more detailed searches could be performed for coded messages. Such searches must be low on the list of priorities for geneticists, but eventually someone will sift through the genome data looking for patterns. My guess is that patterns *will* be found, but they'll have the same source as the Martian canals and the Cydonian face. Such patterns are evidence of intelligence—but at the observer's end of the telescope or microscope.

## Solution 6 They Exist and They Are Us—We Are All Aliens!

*I should have known what fruit would spring from such a seed.*

Lord Byron, *Childe Harold*

In the discussion of Solution 5 we considered the idea that ETCs might have encoded a message in the DNA of terrestrial organisms. Although this is a remote possibility, a broader version of the idea is, paradoxically, more plausible. With each breakthrough in the study of genetics it becomes increasingly apparent that all life on this planet is deeply related. Perhaps individual species are not alien, but we can't discount the possibility that *every* species came from the same extraterrestrial source. Perhaps life itself is the message. Perhaps we are *all* aliens.

The idea that life originated elsewhere and was somehow transported to Earth is an old one. The notion of *panspermia* —literally "seeds everywhere"—probably dates back to Anaxagoras.[77] It wasn't until the 19$^{th}$ century, however, with work by Berzelius, Richter, Helmholtz and others, that the panspermia hypothesis took a modern form. Scientists of the time discussed various forms of panspermia. Lord Kelvin, for example, in an address to the British Society for the Advancement of Science in 1871, pondered whether life might be distributed through space on meteoric rocks—*lithopanspermia*. However, it was a book by Arrhenius[78] in 1908 that popularized the panspermia idea. Arrhenius proposed that the universe is full of living spores that are driven through space by the pressure of starlight—*radiopanspermia*. Some spores fell on the early Earth, flourished, and evolved into the life we see today.

As we shall discuss in more detail later (see Solution 64), one of the deep mysteries of the origin of life is the almost indecent haste with which it arose on Earth. There scarcely seems to have been enough time for random physical and chemical processes to generate life from lumps of inanimate matter. The panspermia idea is attractive, since it removes the problem of timescales: life dropped "ready-made" onto Earth. Nevertheless, the Arrhenius hypothesis quickly fell from favor for several reasons. One reason why the idea was shelved was the difficulty of imagining spores hardy enough to withstand the rigors of an aeons-long journey through space; in particular, cosmic radiation would surely prove deadly to spores. Another reason was that it merely removed the problem of the ultimate origin of life from Earth to somewhere in space (it would of course be nice to know where life originated, if only to settle a fact of history).

The idea that there might be microbial life in space did not entirely disappear. For example, Hoyle and Wickramasinghe championed the idea that microbes travel to Earth on comets, causing occasional mass outbreaks of disease.[79]

The claim was lent some credence by the discovery that bacteria traveled to the Moon on unmanned lunar landers, and were still alive when brought back to Earth by Apollo astronauts. More recently, researchers have investigated the ability of some extremophiles[80]—microorganisms that can thrive in extremely harsh terrestrial environments—to withstand the conditions found in space. Experiments have shown that extremophilic microorganisms, when protected by microsized carbonaceous grains, can survive hours of intense radiation from a synchrotron source—the equivalent of an accumulated radiation dose from millions of years of solar radiation. Thus *microlithopanspermia*—the transfer of microbial life on small dust grains rather than large boulders—would seem to be another possibility. Even if the process of panspermia is too destructive for life to move from one planet to another (after all, the organisms would have to contend with more than just the harsh conditions of space; they'd also have to survive the shocks associated with being ejected from one home planet, landing on another, and passing through both planetary atmospheres) perhaps it's possible that at least genetic information from inactivated virus-like organisms[81] or dead bacterial fragments might be sufficient to get life on Earth "going"—*necropanspermia*.

Although panspermia is perhaps not in the mainstream of biological thought, the hypothesis has certainly not been ruled out. If it turns out to be true, then the chances of life being a frequent occurrence in the universe are greatly increased (though it doesn't necessarily say anything about the existence or otherwise of *intelligent* life and ETCs). In 1973, however, Crick and Orgel published the idea of *directed* panspermia:[82] panspermia plus intelligence, as Dyson put it. Crick and Orgel felt that the chance of viable microorganisms landing on Earth after an interstellar journey measured in light years was small. But *deliberate* seeding is different. Directed panspermia is the suggestion that an ancient ETC deliberately aimed spores toward planets with conditions favorable to the survival of life. Maybe primitive life didn't arrive here haphazardly inside a meteorite; maybe it was *sent* here via a probe. (Why would an ETC seed planets in this way? Perhaps they were preparing planets for subsequent colonization, but somehow failed to get round to colonizing Earth. Perhaps they were performing grand astrobiological experiments. Perhaps they faced a global catastrophe, and wanted to ensure the survival of their genetic material. Who can tell?)

It's difficult to know how to test the hypothesis of directed panspermia. Billions of years after the event, how can we distinguish between primitive life emerging from the primordial ooze, primitive life arriving inside a meteorite, or primitive life arriving by space probe? In their paper, Crick and Orgel argued that directed panspermia could resolve certain puzzles. For example, why is there only one genetic code on Earth? A universal code follows naturally if all life on Earth represents a clone derived from a single set of microorganisms. Another argument offered in support of the idea relates to the dependence of many enzymes on molybdenum. This metal is rather uncommon—it ranks 56th in order of abundance of the elements in the crust of the Earth—and yet it plays an important biochemical role. This slightly odd state of affairs would be less surprising if life on Earth derived from a system in which molybdenum was much more abundant. Of course biochemists have more orthodox answers to these puzzles, and so the evidence in favor of directed panspermia is weak.

If biologists develop a convincing theory of how life originated naturally from the materials available on the primordial Earth, then panspermia—directed or otherwise—would be unnecessary. Or Crick and Orgel might some day be proved right: we might even meet the ETC that seeded our part of the Galaxy. The hypothesis of directed panspermia remains as a possible resolution of the Fermi paradox. Where are they? They are here, because *we* are aliens.

## Solution 7 The Zoo Scenario

*Someone told me it's all happening at the zoo.*
*I do believe it, I do believe it's true.*

Paul Simon, *At The Zoo*

The zoo scenario was proposed[83] by John Ball in 1973 as a means of resolving the Fermi paradox. In fact, Ball called it the "zoo hypothesis"; variants of the idea, some of which are described below, also call themselves "hypotheses", and they appear as such in the literature. I prefer to call them scenarios, because in science an hypothesis usually implies a speculation framed in such a way that it can be tested. As we shall see, Ball's speculation in its basic form can't be tested. This is not to say the zoo scenario is untrue, illogical or somehow more unlikely than other explanations. We've already met ideas that are far more improbable than Ball's speculation. The problem is that we can't readily falsify it.

Ball proposed that ETCs are ubiquitous; many technological civilizations will stagnate or face destruction (from within or without) but some will develop their level of technology over time. Arguing in analogy with terrestrial civilizations, he reasoned that we need only consider the most technologically advanced civilizations. Those ETCs will, in some sense, be in control of the universe[84] because the less advanced will be destroyed, tamed or assimilated.

The important question becomes: how will highly developed ETCs choose to exert their power? Arguing in analogy with how mankind exerts its power over the natural world, wherein we set aside wilderness areas, wildlife sanctuaries and zoos so that other species can develop naturally, Ball speculated that Earth is in a wilderness area set aside for us by ETCs. The reason there seems to be no interaction between them and us is that they don't want to be found—and they have the technological ability to ensure we don't find them. The zoo scenario involved the idea that advanced ETCs are simply observing us. (Variants on the idea were less appealing; the laboratory scenario would have us as the subjects of laboratory experiments.)

This general idea has a long history in science fiction, predating Ball's publication. For example, *Star Trek* had its "Prime Directive", which stated that the Federation should not interfere with the natural development of a planet. (The Directive was more honored in the breach than the observance, of course, since the writers had to generate plots.) And before that the leading SF magazine of the 1950s, *Astounding*, under the strong but quixotic editorship of John Campbell[85] established the trope of Earth under quarantine—either because ETCs were protecting us or, more commonly, because mankind was a threat to them. One could also argue that Tsiolkovsky's solution to the paradox, namely that ETCs have set Earth aside in order to let mankind evolve to a state of perfection, contains the seeds of the zoo scenario.

Believers in flying saucers tend to favor the zoo scenario as if it legitimizes their belief. Yet the zoo scenario specifically predicts that we should *not* see flying saucers or any other manifestation of superior technology. If flying saucers are spacecraft then the zoo scenario is wrong. (James Deardorff proposed a variant of Ball's idea, known as the leaky embargo scenario, which is compatible with observations of flying saucers. The idea is that advanced and benevolent ETCs have put in place an embargo on official contact with mankind. But the embargo is not total: aliens contact those citizens whose stories are unlikely to be credible to scientists and the government. The aliens want to slowly prepare us[86] for the shock that might come later when they reveal themselves. Deardorff's proposal is so unscientific—though again not necessarily untrue—that it probably does not merit even the term "scenario".)

The zoo scenario has been criticized on several grounds. A major drawback to my mind is that it leads us nowhere: it's not a testable hypothesis. A good hypothesis generates ideas for observations that might confirm or falsify it, and in doing so generates new hypotheses. It's difficult to think of any observation that could test the validity of the speculation. Its one prediction is that we won't find ETCs, but the failure to find them hardly confirms the initial statement. There is something unsatisfying about an approach in which, no matter how hard we look, no matter how thoroughly we search, the absence of ETCs is explained simply by saying they don't want us to see them. (I can explain the lack of observational evidence for fairies at the bottom of my garden by saying they become invisible whenever people look their way. Irrespective of whether fairies exist, this is a poor sort of explanation from a scientific standpoint.)

Others have criticized the scenario on the grounds that it is anthropocentric. Why should an ETC have any interest at all in a species such as us? (Assuming, of course, that it is *us* they are interested in and not dolphins or monkeys or bees . . . ) Since we have no conception of what alien minds might find diverting, I guess we can't rule out the possibility that Earth—for whatever reason—has been set aside as the galactic equivalent of a national park. However, a further weakness is that the zoo scenario fails to explain why aliens didn't colonize Earth long before complex life-forms appeared: the scenario might describe the reaction of ethically advanced ETCs to the discovery of intelligent life on Earth, but would that reaction be the same if only primitive single-celled organisms were involved?

A more serious criticism is that it takes only *one* ETC to break the embargo, just one immature civilization that decides to poke its fingers through the bars of the cage, for us to see them here on Earth. Furthermore, it fails to explain why we observe no evidence of them out there in the Galaxy. The proposition here is that intelligent life is ubiquitous, so where are their astroengineering projects? Where are their communications? It's one thing for them to keep Earth free from development, but quite another for them to stop all activity on our account.

Finally, the scenario suffers in a way common to all solutions to the Fermi paradox that depend upon the motivations of alien intelligences. It supposes that *all* ETCs at *all* times behave in the same way with regard to us.

An expanded version of the idea, known as the interdict scenario, attempts to generalize Ball's idea and address some of the weaknesses.

## Solution 8 The Interdict Scenario

*Ever absent, ever near.*

Francis Kazinczy, *Separation*

The interdict scenario—an expanded form of the zoo scenario[87] that provides reasons why *all* life-bearing planets, not just Earth, are off limits—was proposed in 1987 by Martyn Fogg.

Fogg presented the results of a simple model of the origin, expansion and interaction of early galactic civilizations. As with many authors before him he found that, using seemingly plausible values for model parameters, the Galaxy fills quickly with intelligent species. Depending upon the parameters, either a few species dominate with large "empires" or there are many different smaller "empires". The conclusion of Fogg's model is that, whatever the value of the parameters, ETCs would colonize the Galaxy even before our Solar System forms.

Fogg argues that once the colonization phase is over and nearly every star supports intelligent life-forms, the Galaxy enters a new "steady-state" era. The expansionist urge withers, and the problems of aggression, territoriality and population growth are solved. The distribution of intelligence becomes increasingly well-mixed and homogeneous, and the steady-state era becomes an age of communication. According to the model, we are billions of years into this (wonderful sounding) era.

If the scenario Fogg describes is true, then Earth is located within a sphere of influence of one or more advanced ETCs. So why haven't they taken over? He argues that, in a steady-state era, knowledge will be the most valuable resource. Advanced ETCs would have a reason to leave a life-bearing planet well alone, if only because the planet will provide a non-renewable source of information. And the sacrifice of *lebensraum* need not be great. As Asimov pointed out,[88] ETCs might move beyond the need for planet-dwelling. If ETCs can travel between the stars in space arks, then they need not visit Sun-like stars; any star will do, and bright O-type stars might be best. Such space arks might therefore, on principle, avoid Sun-like stars with habitable planets. Fogg suggests the number of stars that ETCs must avoid might be small: he gives a figure of 0.6% for the fraction of stars possessing a life-bearing planet. (This figure is, of course, debatable.) Leaving a small number of systems untouched is a small price to pay for the information content their life-bearing planets will eventually possess.

In the steady-state era, then, an era in which ETCs communicate with each other and common approaches are agreed upon, the "Galactic Club" agrees not to interfere with already populated planets. In the words of Newman and Sagan, a *Codex Galactica* is established.[89] Fogg's suggestion is that the Solar System was placed under interdict when, billions of years ago, an ETC visited the Earth and discovered primitive organisms. Since then, organisms on Earth have lived in a zoo—studied for the complex patterns of information they generate.

To my mind, some of the premises that underly the interdict scenario are unconvincing. To take just one, I believe that the cultural homogeneity that Fogg suggests is unlikely to come to pass. I find it implausible that truly alien intelligences, if they exist, can communicate so efficiently that they reach "an enhanced level of understanding [and] mutual agreement". The problems

**Fig. 3.10** A galaxy such as our Milky Way is typically 100,000 light years in diameter. The galaxy shown here, NGC 2841, is even larger—150,000 light years across. The interdict scenario requires a "Galactic Club" to be able to enforce its rules and traditions from one end of the galaxy to the other. In a relativistic universe, this is extremely difficult to achieve. (Credit: NASA/ESA/Hubble Heritage Collaboration)

in establishing a transgalactic communication system go way beyond mere translation difficulties. For example, the differential rotation of the Galaxy causes a star like the Sun to move relative to other stars. Fifty million years ago, Earth might have been in a region of the Galaxy in which the zoo keepers were punctilious; right now, though, we could be entering a region where the zoo keepers have evolved and decided to take some time off. If they did that, who else would know? And what could the other members of the Galactic Club do to stop it? We live in a universe that possesses a speed limit for information flow, and it makes galactic cultural homogeneity extremely difficult to achieve. McDonald's might have conquered the world, but it won't conquer the Galaxy.

So even without questioning the detailed parameters and assumptions underpinning Fogg's computer model, the conclusions are open to debate. Putting those reservations to one side, the interdict scenario suffers from some of the criticisms leveled at the original zoo scenario. In particular, there seems to be no way of discovering whether we are under interdict (until, perhaps, we advance enough as a species to be elected as members of the Galactic Club) so there are no testable predictions. The scenario also supposes that advanced ETCs, at all stages in their own evolution, can hide their activities from us. Well, perhaps they can. But if the Galaxy really is teeming with ancient ETCs, as is suggested, would we not see the occasional grand astroengineered structure or catch the occasional piece of interstellar gossip? Putting a planet under interdict is one thing; hiding all evidence of their existence is something else. Finally, as discussed above, even if deep communication were established in the steady-state era of the Galaxy, would a uniformity of motive regarding life-bearing planets really arise? The existence of just *one* advanced ETC that fails to share the values discussed above could be enough to invalidate the scenario.

## Solution 9 The Planetarium Hypothesis

*Real are the dreams of Gods.*

John Keats, *Lamia, I*

Stephen Baxter has proposed[90] an interesting variant on the zoo scenario. He calls it the planetarium hypothesis. The speculation is far wilder than Ball's idea, but it merits the term "hypothesis" rather than "scenario" because it offers testable predictions. Is it possible, Baxter asks, that the world we live in is a simulation—a virtual-reality "planetarium" engineered to present us with the *illusion* that the universe is devoid of intelligent life?

The physics behind such an idea has a modern feel to it. Indeed, the planetarium hypothesis could only reasonably have been proposed in recent years—times that have seen an incredible increase in the power of computers. And yet the "things are not what they seem" concept that underlies the planetarium hypothesis is an established trope of science fiction. In Heinlein's novella *Universe*, the inhabitants of a generation ship (see page 79) find a universe beyond the confines of their vessel. In a light-hearted short story by Asimov, written two years before Soviet satellites photographed the far side of the Moon, the first astronauts to orbit the Moon find not a cratered surface but a huge canvas propped up by two-by-fours: the "trip" was a simulation that enabled psychologists to study the effects of a lunar mission on the crew. The protagonist of *The News from D-Street*, a much more somber story by Andrew Weiner, discovers that the totality of his familiar yet strangely restricted world is the product of a computer program. More recently, mainstream media have explored the concept of people interacting with various engineered realities. Several episodes of the TV show *Star Trek: The Next Generation*, for example, were set on the "holodeck"—a technology that emulated material objects with which users could interact. The movie *The Matrix* had humans forcibly immersed in a virtual reality, this time through a technology in which brains

**Fig. 3.11** In a well-designed planetarium we can lose ourselves in a realistic representation of the universe. (Credit: courtesy of Carl Zeiss)

were stimulated directly by implants. The protagonist of the movie *The Truman Show* was the unwitting star of a TV show that had him living inside an engineered reality; in this case it was a "low-tech" reality, a fake town[91] below a painted dome designed by the show's producers.

Many of these stories and movies have a haunting quality, perhaps because they touch upon matters of deep philosophical concern. After all, questions about the nature of reality, and about how each of us perceives the external universe, have kept philosophers in business for millennia. The planetarium hypothesis suggests that our commonly accepted understanding of the external universe might be wrong. Exactly how wrong depends on the type of planetarium the ETC has provided for us ("low-tech" as in *Truman* or "high-tech" as in *Matrix*) and also its scope—the position of the boundary between human consciousness and external "reality".

The planetarium hypothesis taken to extreme is similar to solipsism. The true solipsist believes that everything he experiences—people, events, objects—is part of the content of his consciousness, rather than an external reality in which we all share. It's not just that his is the only mind that exists. (The sole survivor of some planet-wide catastrophe might be correct if he believed his was the only mind, and yet he wouldn't necessarily be a solipsist.) Rather, the true solipsist in principle can attach no meaning to the idea that other minds experience thoughts and emotions. It's an egocentric view of the universe. The most extreme planetarium, therefore, would have an ETC generate an artificial universe directly into *my* consciousness. The universe appears to me to be empty because an ETC, for some reason, wants to fool *me* into so thinking.

Solipsism seems to lead nowhere and is rarely defended directly. (The true solipsist when defending his philosophy presumably has to inform his opponents they don't exist, which seems a rather ludicrous thing to do.) Less extreme planetaria still have a solipsistic flavor but are slightly less outrageous. For example, perhaps we humans are real but some or all of the objects we see around us are simulations—like the holodeck in *Star Trek*. Or perhaps reality consists of everything on Earth plus those places in the Solar System we have visited, but the stars and galaxies are simulated—like a large-scale version of *The Truman Show* dome.

Occam's razor gives us a reason for rejecting all these planetaria. Suppose you throw a ball and watch its parabolic path: you'll conclude the ball is an autonomous object obeying Newton's law of gravity. The alternative—that some system (whether an individual consciousness or a sophisticated virtual-reality generator) contains laws that simulate the properties of the ball and its motion under gravity—is a more complex explanation of the same phenomenon. Both explanations fit the observations. But Occam's razor tells us to use the simplest explanation, which in this case is that the ball is "real". It has an autonomous existence. We can make the same argument regarding our observations of the universe.

On the other hand, if we are willing to put Occam's razor to one side for the moment and take the planetarium hypothesis seriously, Baxter shows how we can test whether we are living in certain types of engineered reality. This is an advance on the original zoo and interdict scenarios, neither of which make hard predictions.

Baxter points out that a fundamental requirement of a planetarium is that scientific experiments should always yield consistent results. (At this point, we don't ask why an ETC would bother simulating a universe for our benefit. It's enough to note that a *perfect* simulation of a system—in other words, a simulation that can't be distinguished from the original physical system by any conceivable test—can in theory be generated.) If an experiment highlights inconsistencies in the fabric of reality, then we might be led to postulate the existence of an "outside".

Physicists can calculate the information and energy demands required to create a perfect simulation of any given size. We can therefore ask whether an ETC has the capacity to meet the energy demands for the construction of any particular planetarium. (We have to assume that the planetarium designers are subject to the same laws of physics as us. If they aren't restricted by physics—if, for example, they can alter the value of the Boltzmann constant—then we can't take the argument further.)

**The Bekenstein Bound** Jacob Bekenstein showed[92] how quantum physics places a limit to the amount of information a physical system can code. The uncertainty relations show that the amount of information inside a system of radius $R$ (in meters) and mass $M$ (in kilograms) can never be greater than the mass multiplied by the radius multiplied by a constant (which has a value of about $2.5 \times 10^{43}$ bits per meter per kilogram). Nature permits a surprising amount of information to be encoded before the *Bekenstein bound* is reached. For example, a hydrogen atom can encode about 1 Mb of information. A typical human can code about $10^{39}$ Mb of information—far more information than can be handled by any hard disk in existence.

Natural physical systems seem to encode much less information than Nature permits. But the Bekenstein bound gives planetarium designers plenty of opportunity to engineer perfect simulations of varying size and scope. Standard thermodynamic calculations give us the energy required to construct a perfect simulation of any particular size and mass.

It turns out that a KI civilization could generate a *perfect* simulation of about 10,000 km$^2$ of Earth's surface and to a height of about 1 km. In other words, a KI civilization could not have generated a perfect simulation of the ancient Sumerian empire, much less our present world. A planetarium designer *could* have fooled the people of Sumer with a less-than-perfect simulation; it would be unnecessary to emulate material 200 m below Earth's surface, for example, since humans of that time were unlikely to dig so deep. Various tricks and short-cuts would also be available to the planetarium programmer—but note the resulting simulation would not be *perfect*, and in principle an inconsistency might be revealed. The protagonist in Weiner's *The News From D-Street* finds himself in exactly this situation.

A KII civilization could have generated a simulation to fool Columbus. But the voyages of Captain Cook might have uncovered inconsistencies in their planetarium design.

A KIII civilization could generate a perfect simulation of a volume with a radius of about 100 AU. This is a large distance, and when I wrote the first edition of this book our civilization was unable to test whether our universe is "real" or the result of a simulation developed by a KIII civilization. But the situation has changed. Voyager 1 is already 127 AU from home and it didn't bump into a metal wall painted black! We know that we don't live in a perfect simulation. We might yet live in a simulation that is less than perfect; after all, only the two Voyager spacecraft have travelled further than 100 AU. The planetarium builders might have scrimped on simulating some aspects of reality in order to extend the boundary of their simulation. But it can't be a *perfect* simulation; our instruments can in principle detect the inconsistencies in such a lower-quality simulation.

The planetarium hypothesis defies both Occam's razor and our basic intuition about how the universe works. It verges on paranoia to suppose that a KIII civilization would go to such effort simply to persuade us that our universe is empty. Baxter himself advanced it only as a possibility to be eliminated (and I'm sure he doesn't believe it to be true). But at least we *can* eventually eliminate it. In the decades to come, as we explore more of the universe and test the fabric of reality at ever-larger distance scales, we will either find an inconsistency in the simulation or be forced to accept that the universe is "real". And if it turns out the universe is "real"—which I am sure most readers would wager is the case[93]—then we'll have to look elsewhere for a resolution of the Fermi paradox.

## Solution 10 God Exists

*Chance is perhaps God's pseudonym when he does not want to sign.*
Anatole France, *Le Jardin d'Epicure*

Some have suggested that SETI scientists are engaged in a theological pursuit: since ETCs are likely to be far in advance of us, they'll be almost omniscient, omnipotent beings. We'd think of them as gods. Many SETI scientists would disagree: an ETC's technology might indeed be so far advanced that it is, to use Clarke's phrase, indistinguishable from magic, but surely we know enough to consider these beings as master engineers. At worst, we'd look on them as thaumaturgists. We know enough not to think of them as gods.[94]

Others have argued that God—the creator of our universe—exists. And that, since God is everywhere, our search for extraterrestrial intelligence would be satisfied if we found God. I'm hopelessly unqualified to argue these points. However, there is a speculation from the realms of theoretical physics that might, if proved true, demonstrate the existence of many other universes that are conducive to the development of ETCs; an even more speculative suggestion is that one of those civilizations created our own universe. They would, in a sense, be God. The work is *highly* speculative, but the theory makes a definite prediction that can be tested. The argument is as follows.

A "theory of everything", which physicists have been pursuing for decades, is a physical theory that unifies gravity with the other forces and that explains the observed relationships between the various forces. A theory of everything would answer *fundamental* physics questions; every type of question a physicist might ask could *in principle* be answered in terms of the theory. In practice, most questions would *not* be explained in terms of ultimate principles, any more than present problems in protein synthesis require a knowledge of quantum chromodynamics for their answers. And a theory of everything certainly doesn't have to explain love or truth or beauty. But the theory should explain the workings of black holes and elementary particles ... and the birth of the universe.

The present best candidate for a final theory is called M-theory. (As far back as the 19$^{th}$ century physicists thought they were on the verge of producing a theory of everything, so it's always best to take these things with a pinch of salt.) The mathematics of M-theory is exceedingly difficult; indeed, much of the mathematical machinery needed to develop the theory has yet to be invented. However, suppose in the next few decades M-theory is developed to a high degree of sophistication. Will it explain "everything"? Perhaps it will; that is the hope of most workers in the field. Nevertheless, there are indications that the theory—whatever it turns out to be—will have a number of parameters, such as the masses of the fundamental particles and the relative strengths of the fundamental forces, whose values must be put into the theory "by hand". The equations of our final theory might say, for example, that the electron mass should be non-zero but it's unclear whether they'll say anything about why its mass should be so tiny: $10^{-22}$ in natural units. It might turn out that the electron mass, and the various other parameters in the theory, could have taken *any* value.

If a theory of everything fails to explain why fundamental quantities take the values we observe, if the theory is self-consistent no matter what numbers we insert for the various free parameters, then we'd have a final theory that describes a multitude of possible universes. Each universe would have different values for the various fundamental parameters. Indeed, for a variety of reasons, the notion of the "multiverse" is being taken increasingly seriously by physicists. How, though, can physicists begin to answer a perfectly reasonable question, such as: "Why is the mass associated with the cosmological constant $10^{-60}$ in natural units when we would naively expect its mass to be about 1?" How can we proceed?

One approach is to say the parameter values were set by chance. How, though, can we explain the fact that the observed values of these parameters seem to be necessary for life? You can tinker with the parameters a little, but not much: life requires chemistry, chemistry requires stars, stars require galaxies ... and all of these require the parameters to lie within a narrow range of values. Decrease the strength of the strong interaction by a factor of four, say, and no stable nuclei can exist: we would not have stars. Change the cosmological constant by a factor of 10, say, and you end up with a universe totally unlike the one we inhabit. The physicist Lee Smolin estimates a probability of 1-in-$10^{229}$ for picking a random set of parameters that generate a universe favorable to life. If Smolin's estimate is correct, then we simply can't appeal to good luck.

**A 1-in-$10^{229}$ Chance** It's difficult to convey just how *fantastically* unlikely a 1-in-$10^{229}$ chance has of occurring. For example, imagine you have a single ticket in a cosmic lottery that has roughly the same odds as the UK National Lottery: about a 13 million-to-1 winning chance. You might think it worth entering: you're not likely to win but, hey, someone has to. Now, suppose the commissioners of this cosmic lottery are miserly beings. Their lottery has been drawn once a second, every second, since the start of the universe some 13 billion years ago—so there have been roughly $10^{17}$ draws. But they pay out on only *one* of those draws; all other draws are void, and they keep the money. So there's only one chance in a hundred million billion that your ticket is eligible for the prize draw; and even if it is eligible, there's only a 13 million-to-1 chance that it will win. With these odds even the most optimistic gambler surely wouldn't bother to enter. But the chance of winning such a lottery does not even *begin* to convey the sheer improbability of a 1-in-$10^{229}$ chance coming up. In fact, only an economist might think such an event is credible: explaining the poor performance of a hedge fund during the financial crisis of 2007 the Chief Financial Officer of Goldman Sachs said that "we were seeing things that were 25 standard deviation moves, several days in a row". Forget about several days in a row—you'd expect to see a 25 standard deviation move on one trading day out of $3.1 \times 10^{136}$.

A second approach is to invoke some form of anthropic principle (for more discussion of the principle, see page 214). In other words, we could argue that the parameters are tuned to these unlikely values in order for rational creatures to exist. Perhaps God explicitly set the parameters to create a universe with life; or, taking a less theological view, perhaps the multiverse contains a vast number of universes, each of which has different laws and constants of physics. We then must find ourselves in a universe where the parameters are conducive to life—after all, we can hardly find ourselves in a universe where physics does not allow life to exist. Many scientists feel vaguely uneasy with such arguments, since anything can be explained this way; to argue like this is almost an abdication of scientific responsibility. Furthermore, a persistent criticism of the anthropic approach is that, with a couple of debatable exceptions, it fails to make predictions that can be tested by observation.

A third approach, promoted by Smolin, is to apply Darwin's evolutionary ideas to cosmology.[95] Equations can't explain why physical parameters have fine-tuned values such as $10^{-60}$, *but evolutionary processes can.* Smolin suggests that the physical constants, and perhaps even the laws of physics, have evolved to their present form through a process similar to mutation and natural selection.

How can this be? Smolin's key assumption is that the formation of a black hole in one universe gives birth to another, different expanding universe. He further assumes that the fundamental parameters of the child universe are slightly different from those of the parent universe. This process is thus rather like mutation in biology: the child has a similar genotype to the parent, but there can be a slight variation. In this picture, then, the universe we live in was generated through the formation of a black hole in a parent universe with

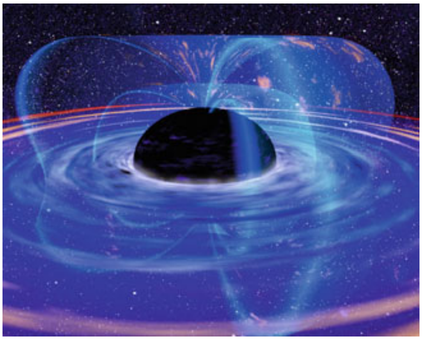


**Fig. 3.12** An artist's impression of the black hole in galaxy MCG-6-30-15. The cores of most galaxies contain supermassive black holes. Could each of these black holes create a universe with physical parameters like our own? If so, our universe might have given rise to billions of similar universes. Black holes formed in stellar collapse are even more common than supermassive black holes. If these objects create new universes, then our own universe might have a billion billion offspring! (Credit: NASA)

similar physical constants to our own. A universe with parameters that permit the formation of black holes has offspring that will in turn produce black holes. A universe with parameters that lead to little or no black-hole formation will produce little or no offspring. Very quickly, no matter how fine-tuned the parameters need to be, universes with parameters that lead to black-hole formation will come to dominate: pick a universe at random and the chances are overwhelming that you pick a universe in which many black holes form.

Now, so far as we know, the most efficient way for a universe to produce black holes naturally is through the gravitational collapse of stars. For example, our own universe will create as many as $10^{18}$ black holes—and thus, in Smolin's picture, child universes—through stellar collapse. So, no matter how "improbable" the values of the fundamental physical parameters that allow stars to form, we expect cosmic evolution to generate a preponderance of universes in which there are innumerable stars. And a universe with physical parameters that gives rise to stars is a universe that inevitably has heavy nuclei, and chemistry, and long enough timescales for complex phenomena to emerge. In other words, it is a universe that might possess life. Note that the fine-tuning of the constants is for the benefit of black-hole production rather than the production of life. In Smolin's picture, life is simply an incidental consequence of a universe that has sufficient complexity to allow the formation of black holes.

This might sound like pure speculation, and it is. There's no evidence (and perhaps there never can be) that the formation of a black hole creates a different expanding universe. Even if a new universe *does* form, we can't answer many of the questions we'd like to ask. (Exactly how do the physical parameters change at the birth of each child universe? Does a single black hole always give rise to a single universe? Does the mass of the black hole play a role? What about its spin? What happens if several black holes merge? And so on, and so on.) Until we have a quantum theory of gravity, we can't even begin to attack such questions. Nevertheless, Smolin's idea has a certain attraction: it links key scientific ideas—evolution, relativity and quantum theory—to explain the long-standing puzzle of the values of the fundamental parameters of physics. Moreover, it makes a specific forecast,[96] a prediction against which the theory can be tested. The prediction is that, since we live in a universe that creates many black holes and can therefore assume that the fundamental parameters are close to optimum for black-hole formation, a change in any of the fundamental parameters would lead to a universe with fewer black holes.

In a few cases, physicists have been able to calculate what would happen if a fundamental parameter differed from its observed value. In each case, it would indeed lead to a reduction in the number of black holes formed by stellar collapse. At present, though, we don't understand enough about astrophysics to calculate the effects of varying all the parameters. Smolin's idea is neither ruled in nor ruled out; it remains an intriguing speculation.

And the relationship of all this to the problem of extraterrestrial intelligence is what? Well, Edward Harrison takes the speculation one step further.[97] He too highlights the long-standing puzzle of why the physical constants seem to be *just* right for the development and maintenance of organic life. Smolin's theory goes part of the way to explaining the puzzle, but Harrison argues that the link between black-hole formation and the conditions necessary for life is too tenuous. Suppose, though, some time in the future, Smolin's idea transmutes into established cosmological theory. Then, Harrison suggests, we might come to believe we should make as many black holes as possible, for in doing so we would increase the probability that other universes might contain intelligent life. Furthermore, a technological civilization doesn't have to bother with stellar collapse in order to create a black hole. It's possible that by building the Large Hadron Collider humans *already* possess a machine for generating black holes; these would be tiny black holes, but that presumably doesn't matter. We might already possess the technology to create universes by the bucketload; a more technologically advanced civilization than ours will surely be able to generate vast numbers of black holes. If in the future *we* might create child universes, perhaps our *own* universe was created by intelligent life. Perhaps God didn't labor for six days; perhaps it was an ETC, in a universe with fundamental physical parameters much like our own, that labored to create a black hole—a black hole that led to the formation of our universe and, eventually, us.

I'm not sure whether Harrison's suggestion could ever resolve the Fermi paradox to everyone's satisfaction. Could the ETC squeeze information through the bounce that creates another universe? If not, how could we ever know whether our universe was artificially produced in a laboratory inside some other universe? The notion that they could squeeze through a message is, however, intriguing. If we found such a message then we'd know that even if *our* universe was devoid of any other intelligent life, we weren't alone in the multiverse.

# 4
# They Exist, But We Have Yet to See or Hear from Them

The position many scientists take on the question of extraterrestrial life is the following. The Galaxy contains billions of habitable, Earth-like planets. Some of those planets, perhaps tens of thousands of them, are home to life. And on some of those planets ETCs exist that are technologically far in advance of our own. This conclusion seems to follow from the Principle of Mediocrity—the notion that Earth is a typical planet orbiting a common type of star in an ordinary part of the Galaxy. The principle has served science well since the time of Copernicus. Scientists who take this position, however, have to answer Fermi's question. If ETCs exist, then why aren't they here? At the very least, why haven't we heard from them?

There are a variety of answers, ranging from the technological (interstellar travel is impossible to achieve, for example) to the practical (communication over interstellar distances is inherently difficult, for example) to the sociological (all societies sufficiently advanced to develop interstellar travel or communicate over interstellar distances inevitably destroy themselves, for example). This chapter discusses 40 resolutions of the paradox that argue "they" exist but that there are technological, practical, sociological or other reasons why to date we have no evidence for the existence of extraterrestrial civilizations.

One weakness of some of these resolutions of the paradox, particularly the sociologically-based arguments, is that to answer Fermi's question they must apply to *every* ETC. I leave the reader to decide whether such answers can resolve the paradox, either singly or in combination.

Several suggestions in this chapter are based on the observation that, here on Earth, the desire and ability to perform computation is steadily increasing. If the trend continues, who knows where it will take us? If advanced ETCs are motivated by the wish to maximize computation, who knows where that desire will take them? As an example of how computing can address the Fermi paradox, consider the aestivation hypothesis[98] of Anders Sandberg, Stuart Armstrong and Milan Ćirković. Aestivation (don't feel bad; I had to look it up too) is a period of prolonged torpor into which some creatures slip. Unlike hibernation, which is a response to the cold of winter, aestivation is a response



S. Webb, *If the Universe Is Teeming with Aliens … WHERE IS EVERYBODY?*,
Science and Fiction, DOI 10.1007/978-3-319-13236-5_4

to heat or drought and is thus most often observed during summer months. Sandberg and his colleagues point out that the cost of a given amount of (irreversible) computation is proportional to temperature. One joule of energy will buy you a certain amount of computation today. If you wait, however, the universe will expand and as it does so it will cool and that joule of energy will be worth more in terms of the computation it can buy. In terms of the amount of computation that can be performed there is *huge* value in waiting before deciding to use your energy endowment; if you wait a trillion years, you gain by a factor of about $10^{30}$.

So here's an idea: civilizations, which are motivated by the desire to maximize computation, colonize a certain amount of the universe in order to gain access to sufficient raw materials, and then aestivate until it becomes rational for them to use those resources for computation. We don't see ETCs now because they are "sleeping", sheltering from the unbearable heat of our present-day universe.

A variety of other elements must be added to this argument in order for the aestivation hypothesis to fully address Fermi's question. Sandberg and his colleagues are currently developing those elements as I write this, so I'm unable to present it here as a separate solution. Nevertheless, as our own society becomes increasingly digital, I'm sure that more people will wonder whether computation plays a role in resolving the Fermi paradox—through aestivation, through a headlong rush to the Singularity, or more likely through some other mechanism that we have yet to dream of.

## Solution 11 The Stars Are Far Away

*. . . between stars, what distances.*

Rainer Maria Rilke, *Sonnets to Orpheus*, Part 2, XX

Perhaps the most straightforward solution to the Fermi paradox is that the distances between stars are too great to permit interstellar travel. Perhaps, no matter how technologically advanced a species becomes, it can't overcome the barrier of interstellar distance. (This would explain why ETCs haven't *visited* us, but not necessarily why we haven't *heard* from them. But let's put this criticism to one side for the moment.)

That the stars are far away doesn't in itself make interstellar travel unattainable. It's certainly possible to build a vessel that can leave a planetary system and then travel through interstellar space. Take our Solar System as an example: its escape velocity, starting at Earth's distance from the Sun, is 42 km/s. In other words, if we launch a vessel traveling at 42 km/s relative to the Sun, then it can escape the grip of the Sun's gravitational influence. It can become a starship. No problem: NASA has already built several such vessels. With our present technology we have to cheat a little and use the gravity assist offered by the planets: the so-called "slingshot effect" is needed to boost a slow-moving craft to escape velocity. But, however we get there, the fact is that with our current level of technology we can reach interstellar space.

Voyager 1, launched in September 1977, toured the outer planets before heading out into space. In February 1998 it became the most distant artificial object and at the time of writing, in June 2014, it is just over 127 AU from the Sun—four times farther out than the outermost planet, Neptune. Unless alien probes pick it up, as happened to the fictional Voyager 6 in *Star Trek: The Motion Picture*, it will eventually make its closest approach to a star—it will drift within 1.6 light years of an unprepossessing M4 star called AC +79 3888. The trouble is, Voyager will take[99] tens of thousands of years to reach its closest encounter with the star. And *that* is the difficulty with interstellar travel: unless you travel fast, the transit times are long.

The best way to rate a starship's speed is in terms of $c$, the speed of light,[100] since $c$ is a universal speed limit. The speed of light in a vacuum is 299,792.458 km/s. So Voyager 1, which as I write is traveling at 17.26 km/s away from the Sun, travels at a mere $0.000058c$. Now, the stars are so widely separated that a favored method of presenting interstellar distances is to use the light year: the distance light travels in one year. For example, the nearest star to our Sun[101] is Proxima Centauri, which is 4.22 light years distant. So the fastest possible "craft"—photons of light—take more than 4 years to reach the nearest star; Voyager 1, were it traveling in that direction, would take almost 73,000 years to complete the same journey. Another way of appreciating these numbers is to recognize that after decades of traveling Voyager 1 is only 17.6 light hours away; much less than a light day. It's the huge travel time involved when traveling at sub-light speed that leads many commentators to conclude that interstellar travel, while perhaps not theoretically impossible, is impracticable.

But perhaps exploration of the Galaxy, even at Voyager speeds, is possible. As long ago as 1929, John Bernal proposed the idea[102] of the "generation ship" or "space ark": a slow-moving self-contained craft that would effectively constitute the whole world for its passengers. After setting off from the home planet, many generations of passengers would live and die before the craft arrived at its destination. Bernal's idea was wonderfully dramatized in Heinlein's story *Universe*.[103] Another possibility would be to put the passengers into suspended animation, as in the film *Alien*, and revive them upon arrival. It has even been suggested that frozen embryos could be transported on slow-moving craft and then grown in artificial wombs at journey's end. And the notion of directed panspermia (page 60) doesn't presuppose the use of relativistic spacecraft; the Galaxy could be seeded using slow-moving probes.

However, it seems clear that we need to build craft that can travel at a substantial fraction of the speed of light if we wish to reach the stars in a reasonable time. Even then, the travel times involved would be long on an individual human scale. For example, ignoring the acceleration and deceleration times at either end of a journey, a craft traveling at the enormous speed of $0.1c$ would take 105 years to reach Epsilon Eridani, which is one of the nearest Sun-like stars. Few crew members seeing their new star for the first time would remember the star they left behind. Is that necessarily a problem, though? When talking about travel times, we tend to assume that people will choose *not* to spend so many years of their life away from home. But we base this assumption in terms of the present human lifespan. After gaining their degrees, several of my more adventurous contemporaries chose to spend a year—which is roughly 2% of their adult life—simply traveling around the world. If human lifespans were increased by a factor of ten, say, and truly relativistic speeds could be reached, then perhaps an adventurous soul would be quite willing to spend a mere decade of his life traveling to the stars. There would be things that we can learn,[104] things that we can experience, only by going out there and studying parts of the universe *in situ*; that fact alone might be enough to tempt people to voyage. Perhaps even a century-long journey would not be uncommon. Who knows? As always, it's difficult to argue about future activities based on present technology.

The journey time mentioned above—105 years to reach Epsilon Eridani, at $0.1c$—is the time that Earthbound observers would measure. People on the ship would measure a slightly smaller interval due to the special relativistic effect of time dilation. Time dilation is another of the unusual consequences of special relativity. Just as moving objects increase in mass, so moving clocks go slow. The faster a clock moves relative to an observer here on Earth, say, the slower that clock seems to tick compared to a clock carried by the Earthbound observer. We are justified in ignoring time dilation effects for on-board observers traveling at $0.1c$, since the effect is only about 0.5%. The closer the speed is to $c$, however, the more noticeable the effect. A craft traveling to Epsilon Eridani at $0.999c$ would take 10.5 years to complete the journey as measured by Earthbound observers, but to a crew member the journey would take only 171 days! If it were possible to travel at speeds infinitesimally smaller than $c$, then *for the traveler* the journey would take a mere fraction of a second. A trip to the farthest galaxies would be possible within a human lifetime[105]—though to Earthbound observers the trip would take so long that Earth itself would be consumed in the Sun's death throes.

What is the likelihood that an intelligent species could develop techniques for interstellar travel at reasonable speeds? (By "reasonable" I mean any speed

**Fig. 4.1** The 110-m-tall Apollo 11 spacecraft was launched from Pad A, Launch Complex 39, Kennedy Space Center, at 09:32, 16 July 1969. On board were astronauts Armstrong, Aldrin and Collins. This vehicle, the first to land men on another world, would be impractical for interstellar travel. (Credit: NASA)

that enables a mission to reach nearby stars on a timescale of hundreds rather than tens of thousands of years. Highly relativistic speeds would be preferable, of course, since they would put the stars within reach of individuals living a human lifespan. But a craft leaving the Solar System traveling at $0.01c$ will reach the nearest star in about 430 years, which puts the stars within range of generation ships.) To answer this, we need to consider the various space-travel technologies that have been suggested. I give only a brief overview here; the notes in a later chapter point to further resources. (Note that if technologically advanced ETCs currently have spaceships moving at relativistic speeds then we might be able to detect them[106] from the way in which light is reflected off the ships. Clumps of matter generally don't move at speeds of 0.1–0.5*c*, so if we spotted a Doppler shift associated with reflection from such a fast-moving object then we might well conclude that it had an artificial origin.)

Although I concentrate here on propulsion methods, it's worth bearing in mind that there are other factors to consider. For example, a starship traveling at high speeds would suffer a ferocious bombardment: tiny dust particles from the interstellar medium would deposit large amounts of energy into the starship structure. Protecting the structure against such erosion, and protecting the crew from the more insidious problem of cosmic-ray bombardment, would require sophisticated shielding. There's also a navigation problem:[107] the stars move with different velocities in three dimensions, making it difficult for a slow-speed mission to rendezvous with a particular star. Nevertheless, these problems are moot if no systems exist that can propel a ship to the stars. If interstellar travel remains forever impractical then perhaps we have a solution to the Fermi paradox.

## Rockets

Most people's initial idea for a starship propulsion mechanism is the self-contained rocket. The familiar chemical rockets employed by NASA and ESA to launch satellites obtain all their energy and expellant mass from on-board reserves. Consider the Apollo missions, for example. The multi-stage Saturn V rockets burned liquid propellants: a mixture of kerosene with liquid oxygen for the first stage, and liquid hydrogen with liquid oxygen for the second stage. The exhaust from these chemical reactions was sufficient for reaching the Moon, but this approach is simply not feasible for interstellar travel. Proxima Centauri is 100 million times more distant than the Moon: the kerosene tanks needed to reach it would be enormous!

Nevertheless, it might be possible to employ variations on this theme. For decades, scientists have considered various alternatives to chemical rockets. An ion rocket, for example, would expel charged atoms to generate thrust; a nuclear fusion rocket would generate high-speed particle exhaust by means of controlled thermonuclear reactions. Perhaps the boldest possibility is the antimatter rocket, first suggested in 1953 by Eugen Sänger.[108] When a particle of matter comes into contact with its antiparticle, both particle and antiparticle mutually annihilate and produce energy. Choose the initial particles correctly and it might be possible to channel the annihilation products into a directed exhaust. Although further analysis showed that Sänger's initial design couldn't succeed, advances in antimatter physics made in recent decades have stimulated proposals that *might* one day lead to an antimatter rocket.

## Fusion Ramjets

The whole concept of using a self-contained rocket—which has to carry the energy source *and* the payload—might be inappropriate when thinking about interstellar travel. It would be much more efficient to use a propulsion system that doesn't require the ship to carry its own fuel. In 1960, Robert Bussard suggested that a *fusion ramjet*[109] might power its way to the stars.

The space between stars is far from being empty: there exists an interstellar medium, comprised chiefly of hydrogen. A ramjet would use an electromagnetic field to scoop up this hydrogen and funnel it to an on-board fusion reactor, which in turn would "burn" the hydrogen in thermonuclear reactions to produce thrust. As with Sänger's antimatter rocket design, Bussard's fusion ramjet proposal suffers from a host of practical difficulties and it's unlikely that the initial idea could be made to work. Nevertheless, several studies have proposed methods to improve the design. Perhaps one of these designs could eventually form the basis of a working starship. Enthusiasts remain enticed by the possibility of the ramjet because in theory it could attain speeds close to $c$ after just a few months.

## Laser Sails

In the 1970s, the American physicist Robert Forward began to consider[110] the possibility of using the *laser sail* as a means of reaching the nearest stars. Imagine a vast "sail" attached to a spaceship and imagine a giant solar-powered laser aiming a narrow beam of radiation toward the ship. Photons from the beam would cause a tiny pressure on the sail and the ship would be gently pushed toward the stars. A laser sail could accelerate to extremely high velocities; hitting the brakes would be more difficult, although deceleration mechanisms have been proposed. Forward's idea has been refined over recent decades, and enthusiasts have designed schemes[111] to use laser sails for both a one-way colonization mission and a round-trip to the stars. A sail would be expensive,[112] at least with our current level of technology, but it does seem to be technically feasible and it would allow speeds of $0.3c$.

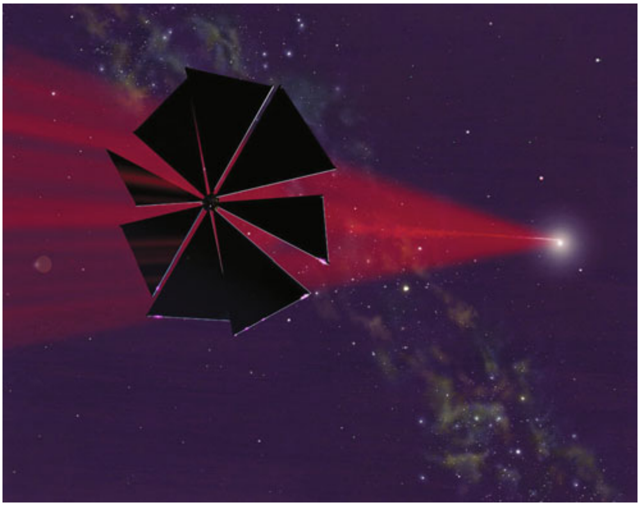


**Fig. 4.2** This beautiful painting shows a solar-powered space-based laser focusing a beam on the huge lightweight sails of a spacecraft. (Credit: Michael Carroll, Planetary Society)

It's worth mentioning here a variant of the sail idea. It has nothing to do with laser power, and it could be developed only by KII civilizations or above, but it does highlight the power of sails. A *Shkadov thruster*, or stellar engine, is a gigantic, massive mirror[113] that reflects a large fraction of a star's radiation pressure. Since more radiation would be emitted from the star in one direction compared to another there would be a tiny net thrust. A Shkadov thruster wouldn't attract the boy-racers among extraterrestrial civilizations: a thruster using a Sun-like star as the source would accelerate from 0 to 20 km/s in one billion years. But if a civilization faced an existential risk, or simply fancied a move, then (if certain dynamical stability issues can be solved) the thruster might work for them: they could move a star 34,000 light years in a billion years.

## Gravity Assists

In 1958, Stanislaw Ulam considered the possibility of accelerating a ship to high velocity using its gravitational interaction with a system of two much larger astronomical bodies in orbit around each other. It was a trick similar to the gravity-assist trajectories that gave Voyager 1 sufficient velocity to leave the Solar System. A few years later, Freeman Dyson considered more realistic—though still, of course, speculative—scenarios. Using Dyson's approach, an advanced technological civilization might employ two orbiting neutron stars to accelerate spaceships to near light speed.[114]

## Fancy Physics

The technologies mentioned above are all based on established physics. The construction of starships using these ideas are, of course, way beyond our present capabilities; indeed, engineering considerations may make it impossible *in practice* to construct starships. But there seems to be nothing wrong with these ideas *in theory*. They break no physical laws.

For many years, people have wondered whether it could ever be possible to travel *really* fast. If we could travel at speeds greater than *c*, then the stars would no longer be grindingly distant. Faster-than-light (FTL) travel would bring the ends of the Galaxy within reach. Nearly all ideas for FTL travel can immediately be discounted, since they clearly violate established physical principles. A few suggestions, however, are still sometimes discussed.

## Tachyons

The special theory of relativity doesn't absolutely forbid superluminal travel. It states, rather, that massive particles can't be accelerated to light speed, while massless particles (such as photons) always travel at the speed of light. Particles with *imaginary* mass must always travel *faster* than the speed of light. Such imaginary-mass particles are called tachyons.

There's nothing particularly unusual about imaginary quantities: we represent several physical quantities by imaginary numbers. But it's difficult to understand what an imaginary *mass* represents. We have no problem understanding the idea of a positive mass; nor is there any difficulty with the idea of a zero mass; we can even ascribe meaning to negative mass[115] (and note that, if negative mass existed, we might be able to use it in a propulsion device). But *imaginary* mass? Whatever it might mean, physicists have searched for signs of it. So far, the tachyon remains hypothetical. There's no evidence such particles exist,[116] and our theories work fine without them. Even if we found tachyons, how could we harness them for FTL travel? We are clueless, here, and it seems reasonable to strike tachyon drives from the list of propulsion possibilities.

## Wormholes and Warp Drives

Most of us are familiar with the Newtonian picture of gravity. We are taught in school that massive objects attract one another by exerting a mysterious influence through empty space. Einstein's general theory of relativity presents a very different picture of gravity. In this view space—or rather, spacetime—plays an active part in the gravitational interaction. In the words of John

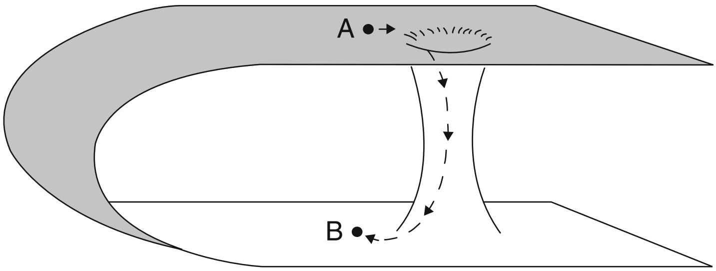



**Fig. 4.3** If space folds over on itself, then a wormhole linking A to B might allow travelers to move between these points without having to traverse the "normal" spacetime between the points

Wheeler: "mass tells spacetime how to curve, and curved spacetime tells mass how to move".

We can think of special relativity as a particular case of general relativity. It applies locally to any region of spacetime small enough that its curvature can reasonably be neglected. The interesting point to consider here is that general relativity permits FTL travel—so long as the *local* restrictions of special relativity are obeyed. The speed of light is a local speed limit, but general relativity permits ways to circumvent this limit. Although this might seem peculiar, there are well-established examples of FTL phenomena in general relativity. For example, standard cosmological models suggest that, due to the expansion of the universe, distant regions of space recede from us at FTL speeds. Only if the expansion were to slow would those regions appear over the light speed horizon and become visible to us. In fact, what appears to be happening is that the expansion is *accelerating*, so in the future more parts of the universe will disappear from view over the light speed horizon: the far future universe will be a lonely place for our descendants.

So far, general relativity has passed every experimental test. It correctly predicts the bending of light rays near the limb of the Sun, the orbits of binary pulsars, and the arrival time of signals in GPS systems. However, most tests of the theory occur in situations where spacetime curvature is small. Sometimes, the distribution of matter can cause a large curvature of spacetime. At the singularity of a black hole, for example, the density of matter is infinite; the very fabric of spacetime is punctured.

It's difficult to interpret the results of general relativity in the extreme situations that occur near the singularity of a black hole. Perhaps the theory can't be applied in such situations; we probably require a quantum theory of gravity to describe what happens there. But in an attempt to understand these extreme regions of spacetime, physicists have pushed the theory. One speculation is

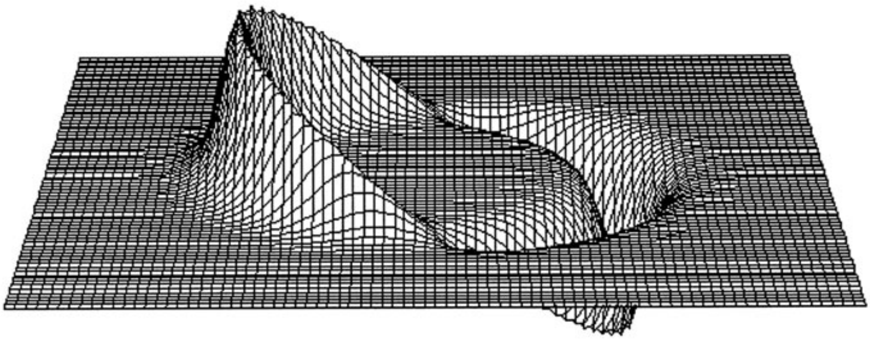


**Fig. 4.4** The figure shows the curvature of space in the region of Alcubierre's warp. Space expands at the rear of the warp and contracts at the front; the flat region is pushed forward

that the formation of a black hole can lead to the formation of a *wormhole*—a "bridge" that links two separate black holes. The two holes might link two quite separate points of spacetime or two different regions of the universe. Enter one black hole and you might emerge from the other hole moments later, thousands of light years from your starting point. As you traveled through the bridge you would have observed the local speed limit and moved slower than *c*; yet your effective speed could be millions of times greater than *c*. Sagan used this idea in his SF novel *Contact*.[117]

Although based on solid work, the wormhole remains a hypothetical creature in the theoretical physicist's bestiary. Wormholes might not exist. Even if they *do* exist, we might be unable to travel through them: calculations suggest that they are likely to be small and wildly unstable. Nevertheless, there remains a tantalizing possibility that an ETC in possession of "exotic" matter (matter with a negative mass–energy) could take a microscopic wormhole, stabilize it, inflate it to a large size—and then use it to traverse huge distances. Alternatively, the engineers of an advanced technological civilization might be able to make use of a solution to general relativity first pointed out by the Russian physicist Sergei Krasnikov. Krasnikov showed that a certain class of wormhole[118] has the property that, no matter how far you travel, you can get home soon after you left. Perhaps a KIII civilization could use a Krasnikov tube for interstellar travel?

There's another way in which general relativity might permit superluminal travel (and in the style to which *Star Trek* has accustomed us). Imagine a spaceship, one as large and luxurious a cruise liner, inside a flat region of spacetime. Everything on board the ship would behave as it does in the flat region of spacetime we inhabit here on Earth. Now imagine that, at the rear of the volume, space expands—in the same way that the universe itself expands. And at the front of the volume, space contracts—as would happen if the universe were to collapse into a Big Crunch (which it won't). The result of this particular warp in space is that the flat-space volume, containing the ship, would move forward—propelled by the expansion of space at the rear and the contraction of space at the front. The ship effectively surfs a spacetime wave.[119]

The warp can travel at arbitrarily large speeds, perhaps many times faster than *c*, and it carries the ship with it. With respect to the local volume of flat space, however, the ship is at rest. There's no relativistic mass increase and no time dilation. For the crew, everything is as normal. As they speed toward the stars at a speed of 100*c*, the passengers are free to enjoy the hospitality of the Spaceship QE2.

The properties of this peculiar solution to Einstein's equations were first analyzed by Miguel Alcubierre while he was at Cardiff University. I have a soft spot for the Alcubierre warp drive, since I was wasting time in the office opposite Miguel while he was working on his idea. Nevertheless, the Alcubierre drive, at least as first proposed, is unlikely to work. First, we've no practical idea how to produce the required curvature of space. Second, the energy density within the warped region is *very* large, and negative. Some theorists would argue this second problem kills the whole idea of a working Alcubierre drive. Quantum theory provides circumstances in which a negative energy density can occur, so if we ever advance to the stage where we can produce large quantities of exotic matter then *perhaps* we could make some form of Alcubierre drive, but a warp large enough to carry the Spaceship QE2 would require a total negative energy ten times larger than the positive energy of the entire visible universe. It seems unlikely.

The Belgian physicist Chris Van Den Broeck might have found a way around some of the problems of the Alcubierre drive. The construction of a microscopically small warp bubble would require just small amounts of exotic matter; combine this with some topological gymnastics, which are allowable in general relativity, and you can end up with an interior volume of the warp bubble that's large enough to hold a spaceship. It would be rather like the Tardis in *Dr Who*: microscopically small on the outside, but roomy enough for passengers on the inside. When we finally have a full quantum theory of gravity we might find that the Van Den Broeck drive is ruled out. In any case, it's worth emphasizing that his drive possesses unrealistic features[120]—unreasonably large energy densities are required, for example. Perhaps, then, wormhole and warp-drive transportation will never be practical. But they haven't yet been shown to be impossible. Maybe one day . . .

## Zero-Point Energy

The quantum uncertainty principle tells us we can't know simultaneously both the position and the momentum of a particle. Therefore even at absolute zero a particle must jitter, since if it were at a perfect standstill we would know both its position and momentum. Energy and time also obey the uncertainty principle; similarly, then, a volume of empty space must contain energy (since to establish that the energy was zero we would have to take measurements for eternity). The Casimir effect[121]—a small attractive force that acts between two uncharged parallel conducting plates brought into close proximity—is the clearest example of the existence of *zero-point energy* (ZPE). The effect can only be explained in terms of quantum fluctuations of the electromagnetic field.

Some writers suggest there's an infinite supply of energy in the vacuum and that some day we'll be able to tap into this ZPE: perhaps we can use ZPE for a propulsion system. Indeed, NASA even sponsored a meeting on innovative propulsion systems in which ZPE was identified as a potential breakthrough technology. If it works, then we'll have limitless cheap energy. Personally, I remain highly skeptical of the idea; we never get something for nothing. But it's yet another suggestion of how a technologically advanced ETC might use the possibilities inherent in the laws of physics to develop technologies that seem almost magical to beings at our level of development.

* * *

I've only touched on the various proposals for interstellar propulsion systems. At present, we couldn't build *one* of the devices mentioned above and use it to reach the stars. With our present level of technology, we'd find it almost impossible to send people safely to Saturn and back,[122] let alone Sirius. There is a host of problems—economic, political, scientific and technical—that we (and presumably an ETC) would have to overcome in order to travel to the stars. What is remarkable, though, is the number of methods that reputable scientists have proposed for starflight. The methods range from the slow to the essentially instantaneous; from the tried-and-tested to the exotic. Although the human race can't build a starship in 2014, what about in 2114? What about in 3014? Other civilizations might be millions, even billions, of years older than our own. Is it likely that *none* of them have the requisite technological skill (or, if relativistic travel is impossible, simply patience) for space travel?

The stars are indeed distant. This fact alone might explain why we haven't been visited (though it does not necessarily explain the "Great Silence"—the absence of signals from ETCs—nor why we see no other evidence of advanced civilizations). However, for those who are optimistic about the reach of science and technology, the distance barrier can be overcome. For those people, the size of the Galaxy alone does not explain the Fermi paradox.

## Solution 12 They Have Not Had Time to Reach Us

*Had we but world enough, and time.*

Andrew Marvell, *To His Coy Mistress*

A common reaction when people first hear of the Fermi paradox is: "Oh, they haven't had time to reach us." Hart, in his influential paper on the absence of ETCs, called this the temporal explanation of the paradox.[123]

As we saw on page 27, Hart argued that this explanation is not tenable if one assumes that interstellar travel is possible. To recap, Hart reasoned that if an ETC sends colonization ships to nearby stars at a speed $0.1c$, and if the colonies in turn send out their own colonization ships, then that ETC would quickly colonize the Galaxy. If the ships didn't pause between trips, then a colonization wavefront would sweep through the Galaxy at a speed of $0.1c$. If the time *between* voyages was about the same as the voyage time itself (travelers have to rest, after all), then the colonization wavefront would move at $0.05c$, so it could travel from one end of the Galaxy to the other in 0.6 to 1.2 million years. For ease of use, we can say that under these assumptions the galactic colonization time is 1 million years.

One million years is an immensely long time at the level of individuals; it's a long time even at the level of an entire mammalian species. But it's extremely short compared to the total time available for colonization. Consider the various timescales involved in terms of the Universal Year. The galactic colonization time corresponds to just 38 minutes 20 seconds—less than one half of a soccer match. On this timescale, civilizations might have been popping into existence since the late spring months, and there seems to be no compelling reason why the first ETC couldn't have arisen by about May Day. So although the first species with the inclination and ability to engage in interstellar travel might have arisen at any time in the 8 months between May and December, according to Hart the temporal explanation asks us to accept that this species started traveling no earlier than 11:21 PM on 31 December. It would be a remarkable coincidence if our civilization emerged so soon after the emergence of the first civilization that decided to reach for the stars.

Hart's argument is compelling, but one can dispute a number of his assumptions. An obvious problem is the speed of the colonization wavefront, which Hart assumes to be a large fraction of the speed of individual spacecraft. As Sagan pointed out: "Rome was not built in a day—although one can cross it on foot in a few hours." In other words, for the city of Rome, the speed of the "colonization wavefront" was an infinitesimal fraction of the speed of the craft used to "colonize" it. More explicitly, throughout all of human history there has never been a colonization wavefront that moved anything like as fast as the speed of individual craft. Why should it be any different for a civilization busy colonizing the Galaxy?

Hart calculated his galactic colonization time simply by dividing the diameter of the Galaxy by an assumed travel speed. Since the publication of Hart's paper, several authors have developed[124] more sophisticated computer models of galactic colonization and thereby arrived at more plausible colonization times. Eric Jones analyzed a model in which colonization was driven by population growth. He assumed a population growth rate of 0.03 per year and an emigration rate of 0.0003 per year (which was the emigration rate from Europe during the colonization of North America in the $18^{th}$ century). His model showed that, under these assumptions, a single space-faring ETC could colonize the Galaxy in 5 million years. In subsequent analyses he offered a preferred colonization time of 60 million years, though this time can be made larger with different assumptions for the rates of emigration and population growth. A colonization time of 60 million years is much longer than Hart's figure, but it's still too short to permit a temporal explanation of the Fermi paradox. On the human scale a process that takes 60 million years is not even glacially slow, but on the cosmic scale a colonization wave moves like a flash flood through the Galaxy.

However, Jones himself made assumptions that can be disputed. For example, Newman and Sagan argued that galactic colonization can't possibly be driven by the demands of population growth.[125] Look at humankind. In the last century, the world population more than tripled in size. If the population were to continue to grow at that rate, and if we wished to maintain Earth's present population density, then in a few hundred years a colonization wavefront would be moving at light speed. Once we reached that point, the population growth rate would *have* to decline! This is an extreme example, but it demonstrates that ETCs will not establish colonies simply as a means of avoiding overcrowding on the home planet. In the long run they simply can't travel fast enough—one can't outrun an exponential increase. A civilization has to curb its population growth regardless of whether it develops space travel. Newman and Sagan therefore modeled galactic colonization as a *diffusion process*, and applied the well known mathematics of diffusion to a particular colonization model. Their results seemed to show that if ETCs practice zero population growth, then the *nearest* civilization would reach Earth only if it had a lifetime of 13 billion years. This *is* long enough to provide a temporal explanation of why extraterrestrials are not here—though it doesn't address the question of why we haven't heard from them.

**Diffusion Processes** In physics, diffusion is a random molecular process, whereby energy or matter flows from a higher concentration to a lower concentration until a uniform distribution is attained. For example, if you heat one end of a rod, then the heat diffuses from the hot end to the cool end. The rate of the diffusion process depends upon the rod's material; in a metal rod, the diffusion is quick; in an asbestos rod, the diffusion is slow. Another example of a diffusion process occurs when you put a sugar lump in a cup of tea; unless you stir the tea, sugar molecules diffuse only slowly through the liquid. A solid can even diffuse into another solid: if gold is plated on copper, the gold diffuses into the surface of the copper—though it takes thousands of years for gold atoms to penetrate more than a tiny distance.

The Newman–Sagan model was in turn subject to criticism. In their model, it turns out that the galactic colonization time is rather insensitive to the speed of interstellar travel. What matters is the time taken to establish a planetary colony, which in turn depends upon the population growth rate. Newman and Sagan assumed *very* low population growth rates—rates that many people find too conservative. Even if one accepts their rates for population growth, there's a problem with their conclusion. The differential rotation of the Galaxy turns the expansion zone into a spiral, rather like the path of a drop of thick cream when you slowly stir it into a cup of coffee. Take this factor into account and the galactic colonization time shortens dramatically. A final criticism: even if advanced ETCs are not driven to expansion by population pressure, would they not explore the Galaxy out of curiosity?

Nevertheless, an extremely detailed model of galactic exploration[126] made by Bjørk strengthened the Newman–Sagan result. Suppose an ETC decides to explore the Galaxy in the following way. It sends out 8 "host" probes, each of which has 8 smaller probes to use at its discretion, and tells the probes to explore a region of the Galaxy containing 40,000 stars. So a host probe travels to a destination star and dispatches its 8 smaller probes to visit those stars that haven't already been explored. The small probes travel at $0.1c$ and perform fly-by investigations. If a probe detects intelligent life it lets the home planet know; if it detects nothing it moves on to the next unexplored star. Once the probes have visited all 40,000 stars on the list they return to the host, which moves to a new destination star and the exploration process starts afresh. Using this method of exploration, Bjørk found that it would take about 300 million years to explore just 4% of the Galaxy. This is a *painfully* slow method of exploration, and at first glance adds strong support to the notion of a temporal explanation for the paradox. Cotta and Morales expanded on Bjørk's model[127] and reached a similar conclusion. As with all such models it's possible to take issue with some of the underlying assumptions (and the authors themselves supply criticisms of their models). There are two aspects of these models that I believe cast doubt on the conclusion. First, they exclude colonization: if ETCs engage in colonization then this would alter the nature of the exploration strategy. Second, they explicitly exclude the possibility of self-replicating probes—but I leave a discussion of this point until Solution 22.

Yet other models have been analyzed.[128] For example, a calculation by Ian Crawford suggests that the Galaxy can be colonized in as little as 3.75 million years. The biggest uncertainty in Crawford's figure is not the speed of interstellar spaceships, but the time it takes for colonies to establish themselves and then send out their own spaceships. And Fogg, in developing his interdict scenario, analyzed the results of a model in which ETCs arise at the rate of 1 every 1000 years, and 1 in 100 of these ETCs attempts to colonize the Galaxy. His model provided the time to "fill" the Galaxy for different speeds of colonization wavefront. Even under the most pessimistic assumptions, he found that ETCs filled the Galaxy in 500 million years, which is short compared to the age of the Galaxy and makes it difficult to support a temporal explanation of the paradox. Nikos Prantzos reinforced this conclusion[129] in an analysis of the Drake equation and the Fermi paradox.

These different models—and we'll look at more in forthcoming Solutions—show we can argue either way. We can argue that interstellar travel is slow and expensive, that ETCs haven't reached us because there hasn't been enough time for them to reach us. Equally, we can argue that interstellar travel is fast and cheap for a civilization with a sufficiently advanced technology. Personally, I'd like to think that our descendants would think of ways to explore the Galaxy under reasonable timescales. And if we were able to do that, so could others in the past. They've had billions of years to reach us. That's time enough.

## Solution 13 A Percolation Theory Approach

*All things flow; nothing abides.*

Heraclitus

The colonization models mentioned in Solution 12 address the paradox in terms of the time it might take one or more ETCs to spread throughout the Galaxy. One can imagine various types of colonization model, however, and they can offer quite different insights. A model proposed by Geoffrey Landis presents an interesting solution to Fermi's question.

Landis bases his model[130] on three key assumptions. First, he assumes that interstellar travel is possible but difficult. No dilithium crystals, no warp engines, no Starship *Enterprise* boldly going; just a long, slow haul to the closest stars. As we've seen, this is a reasonable assumption: to the best of our certain knowledge, the laws of physics don't forbid interstellar travel, but they don't make it easy. Landis thus argues that there exists a maximum distance over which an ETC can establish a colony directly. For example, humankind might one day establish a colony directly around Tau Ceti (just under 12 light years distant from Earth) but might find it impossible to directly colonize any of the stars in the Hyades cluster (150 light years distant from Earth). Any given ETC will have only a small number of stars both suitable for colonization and within the maximum travel distance from its home planet. Therefore any given ETC will establish only a small number of direct colonies. More distant outposts can be settled only as secondary colonies.

Second, since interstellar travel is difficult, Landis assumes a parent civilization will possess only weak—possibly non-existent—control of its colonies. If the timescale over which a colony develops its own colonization capability is long, then every colony will possess its own culture—a culture independent of the colonizing civilization.

Third, he assumes a civilization will be unable to establish a colony on an already colonized world. This is tantamount to saying that invasion is unlikely over interstellar distances, which seems reasonable. If interstellar travel is difficult and costly then invasion must be even more difficult and more costly. There goes the plot of several Hollywood blockbusters.

Finally, he proposes a rule. A culture either has a drive to colonization or it does not. An ETC possessing such a drive will definitely establish colonies around all suitable stars within reach. An ETC having no uncolonized stars within reach will, of necessity, develop a culture lacking the colonization drive. Therefore any given colony will have some probability $p$ of developing into a colonizing civilization, and a probability $1 - p$ of developing into a non-colonizing civilization.

**Probabilities** A probability $p$ must, by definition, lie in the range between 0 and 1. A probability of $p = 0$ corresponds to an event that is impossible; a probability of $p = 1$ corresponds to an event that is certain to happen. If an event has only two outcomes—either the event happens or it does not—then the probability of the outcomes must add up to 1: it's certain that *something* occurs! So if the probability of the event happening is $p$, the probability of it not happening is $1 - p$.

These three assumptions, plus the rule, generate a *percolation* problem. The key task in a percolation problem[131] is to calculate, for a specific system, the probability that there's a continuous path from one end of the system to the other. The word percolation comes from the Latin phrase meaning "to flow through", and those who developed percolation theory perhaps had in mind coffee percolation when they named it: to make a drink, water must find a path through the ground coffee and into the pot. Coffee-making is a particular example of the general problem of the diffusion of liquid through a porous solid; but percolation models have also been used to study phenomena as diverse as the propagation of forest fires, the spread of contagious disease in a population, and the behavior of quarks in nuclear matter.

In essence, percolation is merely a way of filling a large array of empty spaces with objects. (Strictly, percolation theory is valid only for arrays that are

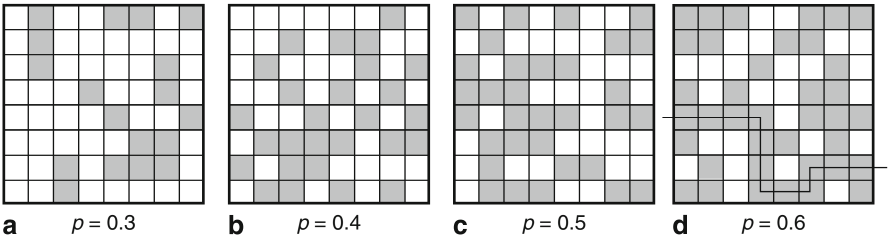



**Fig. 4.5** The cells in each of these four arrays have been shaded (occupied) at random. In (**a**), each cell has a 30% chance of being occupied. In (**d**), each cell has a 60% chance of being occupied. Even in (**a**) there are "clusters"—cases where two or more nearest-neighbor cells are occupied. (The nearest neighbor of a cell is one that is directly above, below, left or right of the cell.) In (**d**) we can see a "spanning cluster": a path through nearest neighbors from one end of the array to the other

infinitely large, so the systems of interest must be large for percolation theory to apply.) The array need not be rectangular, nor need it be two-dimensional: some phenomena are best modeled with a one-dimensional array, others with a three-dimensional array, and still others with higher-dimensional arrays. To fix ideas, though, it's easiest to imagine a large two-dimensional array of $N$ cells, rather like an extended chessboard.

**Percolation Theory** Suppose each cell of an array has a probability $p$ of being populated. Each cell is independent of the others; just because a particular cell happens to be populated doesn't mean that its neighboring cells are more or less likely to be populated. Clearly, $p \times N$ of the cells will be populated and $(1 - p) \times N$ will be empty. If the probability $p$ is large, then the array will contain lots of filled cells; if $p$ is small, then the array will be sparsely populated. Figure 4.5 shows four computer-generated 8 × 8 arrays. In (a) the probability of occupancy for a cell is 30%; in (b) it is 40%; in (c) it is 50% and in (d) it is 60%. (Physicists deal with much larger simulations than this, of course, but an 8 × 8 grid is fine for the purposes of illustration.) Two occupied cells that are next to each other are called *neighbors*, and groups of neighbors are called *clusters*. For the two-dimensional array shown in the illustration, each cell, except those on the edges, can have four neighbors: the cells directly above and below, and to the left and right. Percolation theory deals mainly with how these neighbors and clusters interact with each other, and how their density affects the particular phenomenon being studied. A cluster that spans the length or width (or both) of an array is particularly important in percolation theory. It's called the *spanning cluster* or *percolation cluster*. For an infinite lattice, a spanning cluster occurs only when the probability $p$ is above a critical value $p_c$.

In general, the value of $p_c$ can't be derived analytically. Instead, we must use computer simulations to estimate $p_c$ for a given system. An infinite square lattice, for example, has a value of $p_c$ at about 0.592,75. A simple example should make clear the importance of a spanning cluster. Imagine a large chunk of some electrical insulating material, in which we embed a certain fraction, by volume, of identical electrically conducting spheres. Below the critical value $p_c$, no spanning cluster exists and the material remains an insulator. Above the critical value $p_c$, a spanning cluster exists and the material can conduct electricity. The same considerations tell us the density of people at which a disease will spread or the density of trees at which a fire will consume an entire forest.

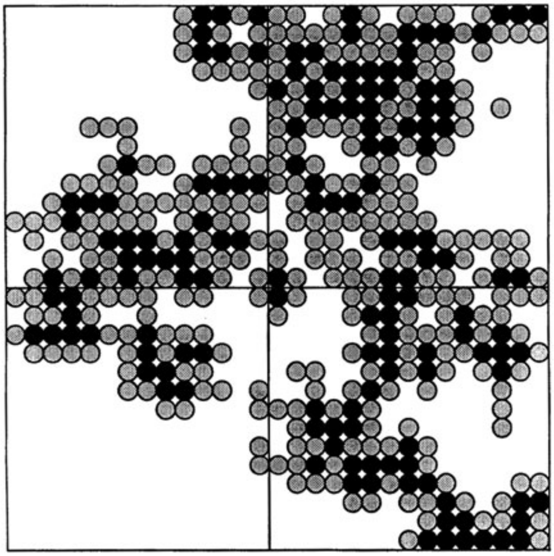


**Fig. 4.6** A slice from a typical percolation simulation on a simple cubic lattice in three dimensions. For this array the critical value is 0.311, while the simulation is for $p = 0.333$. The black circles denote "colonizing" sites; the gray circles denote "non-colonizing" sites. The absence of circles denotes sites that have not been visited. Note the irregular shape of the boundary and the large voids. Does Earth perhaps lie in one of the voids? (Credit: Geoffrey Landis)

What has this to do with the Fermi paradox? Well, if Landis is right, we can use the well-honed techniques of percolation theory to simulate the flow of ETCs through the Galaxy. Although percolation problems are difficult to study analytically, they can be easily simulated on computer. Readers with some programming expertise can set up the Landis model and study for themselves the distribution of ETCs under different model parameters. Figure 4.6 shows a typical result.

As in any percolation problem, the final lattice depends upon the relative values of $p$ and $p_c$. In the Landis model, if $p < p_c$, then colonization will always end after a finite number of colonies. Growth will occur in clusters, and the boundary of each cluster will consist of non-colonizing civilizations. If $p = p_c$, then the clusters will show a fractal structure, with both empty and filled volumes of space existing at all scales. If $p > p_c$, then clusters of colonization will grow indefinitely, but small voids will exist—volumes of space that are bounded by non-colonizing civilizations. We produce a Swiss-cheese model of colonization: civilizations span the Galaxy, but there are holes.

The percolation approach thus suggests that colonizing extraterrestrials have not reached Earth for one of three reasons. First, $p < p_c$, and any colonization that has taken place stopped before it reached us. Second, $p = p_c$, and Earth happens to be in one of the large uncolonized volumes of space that inevitably occur. Third, $p > p_c$, and Earth is in one of the many small unoccupied voids. Which of the three suggestions is most probable? To answer this we need to know the value of the colonizing probability $p$ and also the typical number of stars available for colonization. Of course, we have absolutely no idea of what a reasonable value for $p$ might be; Landis takes $p = 1/3$, which is as good as any other estimate. As for colonization sites, Landis argues that suitable candidates exist only around stars sufficiently similar to the Sun—in other words, single main-sequence stars within a restricted spectral range. Within a distance of 30 light years of Earth there are only five candidate stars, so a reasonable guess for this number is 5. These values produce a model that is close to critical: there are large colonized volumes of space and equally large empty volumes of space. According to the Landis model, then, the reason we haven't been visited by the many ETCs that exist in the Galaxy because we inhabit one of the voids.

This conclusion is similar[132] to one reached later by Osame Kinouchi, who points out that when one observes the nocturnal Earth from space the non-uniform distribution of human colonies—to say nothing about the perverse distribution of global wealth—becomes clear. One sees many human colonies—cities, in other words—but also vast uninhabited areas. Human civilization can travel by air at 1000 km/h, and so clearly has had time to colonize the globe, and yet there are areas that remain persistently unvisited. A member of an Amazonian tribe that has not yet been contacted by global civilization would be wrong to conclude that no such global civilization exists. Kinouchi's "persistence solution" to the Fermi paradox suggests that Earth is in a huge, poorly inhabited domain of the Galaxy: we are a "persistent site" unvisited by the colonization process. Similarly, Robin Hanson, who approached the problem of galactic colonization from an economist's point of view,[133] concluded on the basis of his model that Earth might exist as an oasis in a quiet region of the Galaxy. Colonization, exploration and furious consumption of resources could be taking place outside of that region. However, even Hanson found it puzzling that we see no signs activity as waves of colonization spread through the Galaxy nor any signs of past waves of colonization; we observe no "burning of the cosmic commons".

The percolation approach addresses the Fermi paradox in an attractive way. Rather than attributing uniformity of motive or circumstance to ETCs, it assumes civilizations will have a variety of drives, abilities and situations. The resolution of the paradox arises naturally as one possible consequence of the model. Of course, one could quibble about the details of the model; Landis himself discusses various points in his paper. For example, the model ignores the peculiar motion of stars. Stars aren't fixed as the squares on a chessboard are fixed, but instead move relative to each other. Although the relative movement of stars is slow, it might affect the percolation model. It's also possible to suggest ways to improve the analysis. For example, we could develop more complex models, taking into account galactic boundaries, habitable zones and the actual distribution of stars. One can also challenge the basic assumptions of the percolation approach. For example, is it realistic to assume the existence of a distance horizon, beyond which no civilization will ever colonize? After all, if a civilization can travel 50 light years, would a trip of 100 light years really be so much more difficult? And what of the assumption that only a few suitable stars will lie within the horizon? A suitably advanced civilization might well find it possible—indeed preferable—to construct habitats around a variety of stellar types. Furthermore, simple extensions of the model[134] can radically alter the conclusions. For example, perhaps individual colonies—as opposed to entire civilizations—can die. If a colony dies it opens up a path through which otherwise trapped colonizing civilizations can travel: this small change modifies the model quite substantially. And perhaps colonies can change culturally over time; perhaps non-colonizing cultures on the frontier occasionally get bitten by the urge to explore. Modify the percolation model by adding these two elements—colony death and colony mutation—and it turns out that there are no longer any voids in space. Eventually, the Galaxy becomes fully saturated.

Even if one accepts the original percolation model as an explanation for our lack of visitors, can the model explain why we haven't heard from an ETC or seen signs of their activity? This question is particularly troublesome if one of the $p \geq p_c$ cases is true, and we inhabit a void surrounded on all sides by advanced civilizations: even if daughter civilizations become independent of their parents, surely they'd occasionally want to communicate with each other. Keeping in contact using radio or optical channels would be trivial compared to the problem of physically traveling between stars. It's hard to believe that all these civilizations would travel, and then adopt and maintain a policy of silence. So why haven't we overheard just one of these conversations? Why haven't we seen a single "we are here" beacon? (In the Landis model, ETCs should have nothing to fear in revealing their position: one of the inputs to the model is that colonization of an inhabited system is so difficult it never takes place.) Why haven't we seen even one astroengineering project, of the kind an advanced ETC might undertake? The answer to all these questions might simply be that space is big and we haven't looked hard enough nor listened long enough. Nevertheless, although a percolation model provides an elegant explanation of why we haven't been visited, I personally find it ultimately unconvincing.

## Solution 14 Wait a Moment

*Everything comes if a man will only wait.*

Benjamin Disraeli, *Tancred*

One of the advantages of the Landis percolation theory approach was that it addressed the Fermi paradox by making explicit some simple assumptions about galactic colonization and then using a computer to explore the consequences of those assumptions. Nowadays many of us have access to computers powerful enough to explore our own pet theories of galactic colonization, and one way of doing this is to use models based on *cellular automata*. This is what the astronomers Bezsudnov and Snarski did, and thereby provided a related, yet slightly different, insight into the paradox.

**Cellular Automata** Cellular automata were first studied by Stanislaw Ulam and John von Neumann in the 1940s, but only came to prominence in the 1970s when Martin Gardner popularized[135] John Conway's *Game of Life*.

It's easy to create cellular automata. Take a board and divide it, chessboard-like, into a number of squares: each square is called a cell, and a cell can be one of a finite number of distinct colors. You need two more elements. First, you need a clock. Second, you need to define a transition rule that applies at every tick of the clock: a cell checks its own color, and those of its neighbors, and decides its color depending on the rule. When the clock ticks, all cells change at the same time. So to run a cellular automaton you color the cells in some pattern, start the clock, and then watch as the pattern develops at each clock tick.

In Conway's *Game of Life* each cell can be one two colors—black or white, say, which correspond to the states dead or alive—and the transition rule is simple: if a white (live) cell has either two or three white neighbors then it stays white; otherwise it turns black (it dies); a black (dead) cell turns white if it has exactly three white neighbors. These simple deterministic rules generate complex, unpredictable outcomes. If you haven't played *Game of Life*, do try out one of the freely available simulation systems—it can be quite hypnotic to watch a pattern evolve: some patterns blink; others oscillate; yet others glide across the space, or eat the gliders. In November 2013, 43 years after the game came to public attention, the first self-reproducing pattern was announced.

Bezsudnov and Snarski generate a cellular automaton model of galactic civilization in the following way. They slice the Galaxy into a set of square cells (to simplify things the Galaxy exists in a two-dimensional space, so after slicing it looks like the grid in a game of *Battleship*) and then make some assumptions about how ETCs will come into existence, expand and expire. First assumption: a civilization can appear at any point in unoccupied space with some small probability. Second, and crucially, all civilizations have the same natural lifetime, $T_0$, after which they start to die. (The authors believe the universal cause of the death of civilizations would be the loss of "basic functions—knowledge functions". In other words, having learned all it can about itself and its environment a civilization has no desire to continue. It withers, dies.) Third: if one civilization makes contact with another then the lifetime of both increases by a time $T_b$: the contact generates new things to learn, new conversations to be had, a spur to further development. (The authors call this the "bonus stimulated" model, but this generates a rather unfortunate acronym: the BS model. I'll instead spell it out in full.)

In the bonus stimulated model a civilization is represented by a square of cells, with the central square being the birthplace of the civilization. The model can be defined as a cellular automaton by casting the assumptions in the form of transition rules. The first rule is that a new civilization can be born in any empty cell; the probability of birth is $n$ and the civilization begins as a single cell. The second rule is that, with each tick of the clock, the civilization changes size by one layer of cells on each side: if the civilization is younger than $T_0$ then it increases in size; if the civilization is older than $T_0$ then it decreases in size; and when the size becomes zero, in other words when there are no more cells, then the civilization is dead. Third rule: if a growing civilization meets another civilization—which in this model means that a cell has to belong to both civilizations—then the lifetime of both civilizations is increased by the bonus time $T_b$; if there are several civilizations in a cluster then they all get the bonus time. The subsequent development of each civilization is as in the second rule.

**Monte Carlo Methods** Cellular automata aren't the only available computational method for exploring questions relating to extraterrestrial intelligence. For example, one can use Monte Carlo methods.

In a Monte Carlo approach you run a simulation many times over and then repeatedly sample at random in order to obtain the distribution of some unknown quantity. Perhaps the earliest Monte Carlo method came in the 18$^{th}$ century, when Georges-Louis Leclerc, Comte de Buffon showed how a random, probabilistic approach could be used to estimate $\pi$. Suppose you have a wooden board consisting of parallel strips, each of which has the same width $d$. Now drop a needle, whose length $l$ is less than the width of the strips, onto the floor. Buffon asked: what is the probability, $P$, that the needle will cross a line between two strips? He showed that $P = 2l/d\pi$, which you can rearrange to get an expression for $\pi$. But you can measure $P$ experimentally, by doing the experiment lots of times, and measuring how many times the needle crosses a line. Suppose it happens $h$ times out of $n$ attempts, so that $P = h/n$. In that case $\pi = 2ln/hd$. If you have enough patience you can do this Monte Carlo experiment for yourself, and estimate $\pi$.

According to Nicholas Metropolis, the man who came up with the name "Monte Carlo" method for the modern version of this approach, Fermi was the first to dabble[136] with this technique. Fermi used it when studying neutron diffusion in the 1930s, but didn't publish this research. It was the work of Ulam and von Neumann, mentioned above, that began the widespread use of Monte Carlo techniques. Ulam and von Neumann used Monte Carlo methods in nuclear weapons research at Los Alamos; nowadays the technique is used wherever there's a need to model some phenomenon that has a lot of uncertainty in its inputs—so you'll find it used throughout physics, engineering, weather forecasting, business ... it really is ubiquitous. The Scottish astrophysicist Duncan Forgan has pioneered the application of Monte Carlo techniques to the study of the Fermi paradox.[137]

The "Galaxy" that Bezsudnov and Snarski simulated was a square grid with ten thousand cells on each side—a hundred million cells in total—and the system evolved over 320,000 ticks of the clock. Each time they ran the simulation they used a slightly different combination of the variables $n$, $T_0$ and $T_b$, since these three numbers are clearly going to govern the behavior of the system. Of course, we have no idea what these three numbers might actually be—but it's easy to run the simulation and see what happens in different cases.

It turns out that if there is no bonus time (in other words, if $T_b = 0$) then the volume of the Galaxy occupied by civilizations is directly proportional to the birth probability $n$ and to the cube of the natural lifetime $T_0$. In this case, for the values of the variables considered by Bezsudnov and Snarski, contact between civilizations is highly unlikely: even if ETCs exist their paths won't cross. The situation changes if the bonus time $T_b$ is non-zero. Small values of $T_b$ don't make much of a difference, but at some threshold level civilizations have enough time to develop into a large cluster that fills the Galaxy. The conclusion drawn by Bezsudnov and Snarski is that, if the actual values of $n$, $T_0$ and $T_b$ are in the right range, we simply have to wait: right now there is a process taking place in which civilizations are expanding throughout the Galaxy.

If one is of a mind to criticize the bonus stimulation model it's easy to identify weaknesses. For example, in this model the younger colonies die first; isn't it just as likely, or indeed more probable, that a civilization would die "from the inside out"? Put this transition rule into the model and the results change because the colonies have more time to reap the reward of contact. Furthermore, the bonus stimulated model is unrealistic in that the benefits of contact are assumed to propagate essentially instantaneously throughout the interacting civilizations: even if civilizations that are spread over large volumes of the Galaxy are able to maintain the cultural homogeneity required for the benefit to be universal, such propagation would violate special relativity. The model is therefore far too simplistic for it to provide a convincing argument that an expanding wave of colonization is heading our way—but in fairness that's not really what Bezsudnov and Snarski were aiming for. Rather, they wanted to show how one can build upon the Landis approach and use cellular automata to investigate a particular scenario. The Serbian astronomers Branislav Vukotić and Milan Ćirković have developed much more sophisticated cellular automata,[138] based on knowledge that takes into account the knowledge gained by astrobiologists in recent years. The "Landis solution", whereby Earth finds itself an unexplored void, occupies one corner of their much larger landscape of parameters.

The beauty of cellular automata is their simplicity. If you are reasonably competent with a computer then it's straightforward to set up a model and then watch it evolve. If you have an idea about how extraterrestrial civilizations come into being and subsequently evolve, why not try and describe them with suitable transition rules? You can model their development, watch their fate unfold and perhaps come up with a new solution to the Fermi paradox. The solutions proffered by this approach so far, however, are to my mind unconvincing.

## Solution 15 The Light Cage Limit

*You can cage the singer but not the song.*

Harry Belafonte

Models of galactic colonization based on diffusion (such as the Newman–Sagan proposal), percolation (the Landis proposal) or cellular automata (Bezsudnov and Snarski) make statements about the migratory behavior of species that are assumed to hold true over timescales measured in hundreds of thousands or even millions of years. Colin McInnes developed a model of migration[139] that only needs to hold true over a period of a few millennia in order to account for the lack of extraterrestrial visitors here on Earth. It's a rather bleak resolution of the Fermi paradox; unfortunately, when one considers the behavior of the human species, it seems rather plausible.

McInnes ponders the likely characteristics of a young civilization that has just found itself with the technological capacity to succeed at interstellar migration. He argues that if a species has the drive and motivation to develop the necessary technologies then the species is likely to be highly competitive, since it would have had to outcompete other species during its early evolutionary development. If a species realizes that it can engage in interstellar travel on a large and economic scale, and in doing so exploit new material resources, then it isn't going to hold back. Indeed, any subgroup of that species will find it can gain a competitive advantage by colonizing space and acquiring new resources: there will be a race to get out there and take advantage of the opportunities. Wealth, activity and population will continue to increase, and the species will experience a wave of economic expansion. For a while, the species will have never had it so good. They aren't likely to stop.

The colonization process will presumably take place star by star, but for the purposes of a model we can think of it as spherical wave of expansion with the centre of the sphere being the home star. Now, the total population of the species will increase, but let's assume that the species will wish to keep the population within the colonized sphere at a constant density: after all, the average population density will be limited by the carrying capacity of their environment. Suppose then that the population grows at an annual rate of just 1%—modest, but perhaps not unreasonable. By introducing that population growth rate we have sown the seeds of disaster. The growth will be exponential and, as stated earlier, you can't outrun an exponential.

McInnes shows that in order to maintain a constant average population density the migration speed must increase linearly with radial distance from the home star. But at some point that migration speed will be equal to the speed of light. Beyond that radius it becomes impossible to maintain a constant population density. Stephen Baxter calls this radius[140] the "light cage". What happens according to this model is that the sphere of colonization increases ever more quickly until it reaches the light cage limit; after that time this still-young and vibrant civilization finds that it can't disperse inhabitants quickly enough to maintain a constant average population density. The population density increases unsustainably at the outer edge of the sphere, just inside the light cage limit, and the carrying capacity of the environment is exceeded. Resource limits are exceeded. The civilization collapses. It's inevitable—if the population grows at an annual rate of just 1%.

One might think that if the population growth rate is just 1% then the light cage must be at a large distance from the home star. If the light cage were, say, at a distance of 50,000 light years then there would be plenty of "wiggle room" for a species; they could colonize a significant proportion of the Galaxy. But if you think that way it's because you don't have an intuitive feel for the power of exponential growth; few of us do. A growth rate of 1% per year implies a light cage limit of just 300 light years. Furthermore, a civilization will reach the light cage limit in just a few thousand years—an eyeblink on cosmic scales. (If the maximum expansion velocity is less than the speed of light then the boundaries of the cage contract: a maximum velocity of $0.05c$ gives a cage limit of just 15 light years. There are only about 50 stars within 15 light years of Earth, and most of those would be unsuitable for colonization. In this picture, colonization hardly seems worth bothering with.)

So here is a scenario that explains why we haven't been visited. Any civilization that develops the ability to embark on large-scale, economic colonization of its stellar neighborhood inevitably collapses within a few thousand years because its rate of migration can't possibly match its growth. After it crashed, a civilization would be so starved of resources that it would be unable to mount a second attempt at colonization. Civilizations pop in and out of existence, and they aren't here because they never get beyond the light cage limit.

It's a gloomy scenario—but is it inevitable? In fact, the trap is so obvious that one could hope that at least *one* technologically advanced civilization would see it and take steps to avoid it. One way of avoiding the trap would be to keep net population growth very low, although there might then perhaps be dangers associated with the stagnation of the civilization. Another would be to constrain growth once the limits of resource are reached, but to allow rapid growth at the frontier. Surely there would be *one* technologically advanced civilization that could see the existential peril of unconstrained growth and have the wisdom to act accordingly. Wouldn't there?

## Solution 16 They Change Their Mind

*Never, never, never give up.*

Winston Churchill

When one investigates galactic colonization using a computer model, it's quite easy to forget the reality that the model is supposed to represent. In a cellular automaton model, for example, civilizations somehow pop into existence and with each "tick" they expand or contract or die; the action takes place on a chessboard and it all looks so simple. Each tick of the clock, though, corresponds to vast reaches of real time. If we accept that faster-than-light travel is impossible then even the most technologically optimistic of us must accept that colonization is a long-term project. Travel is only part of that project. After reaching their destination, ETCs might well find that planets must be re-engineered[141] to provide appropriate living space—that's inevitably going to take more time. Establishing a human colony on an inhospitable exoplanet would presumably take at least a thousand years; a more realistic estimate is that it would take a hundred thousand years. This is an eye blink in cosmic terms, and hardly worthy of mention in computer models where one "tick" represents a dozen millennia. But in terms of the real world it raises a question: is it likely that any culture will be stable over timescales of tens to hundreds of thousands of years and, moreover, will continue to commit substantial resource to a program of expansion? Even a long-lived and fixed culture might at some time reorder its priorities. Faced with these sorts of timescales, they might simply give up on the notion of large-scale colonization.

Claudius Gros has developed some simple equations[142] that govern the population dynamics of technologically advanced civilizations, given that those civilizations might change in character. In Gros' formulation new civilizations will be born at some rate, existing civilizations will go extinct at some rate, and there will be some growth rate. Gros, however, drops the assumption that the colonies of an expanding civilization automatically inherit the characteristics of that parent civilization. In other words colonies, for whatever reason, might simply give up on colonization; conversely, stagnant civilizations might, for whatever reason, decide to resume expansion. Putting all these various rates together, Gros shows that it's possible for a large population of expansionary—but mutable—ETCs to settle into a stable equilibrium.

As Gros himself admits, we know nothing of the various numbers that appear in his equations—the birth and extinction rates, the rate at which a formerly expanding civilization changes character and stagnates, and so on—but the work provokes a thought. Is it likely that human society can be stable over millennia while being continuously dedicated to exploration and expansion? If we doubt it of ourselves we can doubt it of others. Could this explain why they aren't here?

## Solution 17 We Are Solar Chauvinists

> *... the suns of home.*
>
> Rupert Brooke, *The Soldier*

Models of galactic colonization implicitly assume that the important objects out in space are stable, middle-aged, G2-type stars such as the Sun and watery planets such as Earth. But who knows *where* a civilization much older than ours would choose to live? Even if Earth-like conditions are necessary for the genesis and early evolution of life, once a civilization is technologically advanced and can construct a habitat for itself it might not want to remain on the surface of a planet orbiting a commonplace star. We tend to think ETCs would love to get their hands (tentacles, feelers, whatever) on the prime piece of real estate that is our Solar System, but that might simply be a reflection of our solar and planetary chauvinism. Perhaps the various galactic colonization models aren't wrong; perhaps they are they simply inapplicable.[143]

For example, Dyson once suggested that a KII civilization might choose to tear apart some of the planets in its system and use the material to create a sphere that encloses the star.[144] By doing this, all the star's energy output could be utilized; compare that with the situation on Earth, where we intercept only a billionth of the energy emitted by the Sun. If that civilization was also capable of interstellar travel, then presumably it could construct a Dyson sphere around any star it visited. (The Shkadov thruster, which we discussed briefly in Solution 11, is essentially half of a Dyson sphere.) If that were the case why would it bother with our Sun, when so much more energy is available from stars of spectral class O? An O5 star is about 800,000 times more luminous than the Sun; a Dyson sphere around such a star could harvest almost $10^{18}$

times more solar energy than we can here on Earth. Perhaps, then, advanced ETCs are nomads, traveling from O-type star to O-type star in generation ships. They could arrive, enjoy a copious energy supply for the few million years of the star's life, then leave before the star went supernova. Brilliant O-type stars provide unsuitable environments for life to evolve, because they die so quickly, but they might be the stars of choice for KII civilizations.

Would KII civilizations require stars at all? Perhaps they mine energy from the quantum vacuum or extract energy from black holes. They might live in their generation ships, never feeling the need to set foot (or alien pedal equivalent) on a planetary surface. Most models of galactic colonization have been based on an analogy: the colonization of America by Europeans or of the Pacific islands by Polynesians. Perhaps a better analogy for colonization[145] would be life's migration from water to land. Just as fish don't meet fowl perhaps ETCs won't meet us. Perhaps they colonize space, but they don't bother to colonize our particular piece of real estate; other places are more attractive to them.

## Solution 18 Aliens Are Green

*Live at home.*

George Washington Carver

Perhaps Fermi's question—"Where *is* everybody?"—draws its power from our gut feeling of how galactic colonization should progress: a civilization develops on one planet and then colonizes a second world; that world colonizes two further planets; each of those planets colonizes two more worlds . . . and rather soon the Galaxy is teeming with intelligent life. This is a picture of colonization that derives from our own history: our forebears stepped tentatively out of Africa, managed to establish viable colonies in other continents, and then spread across the globe. Our reach now is impressive: Dubai is a stunning metropolis rising from desert sands; the largest cruise ships carry as many passengers across the oceans as inhabit the small town where I reside; there are even permanently staffed research stations in Antarctica, the most inhospitable place on Earth. Won't ETCs colonize our Galaxy just as humans have colonized Earth?

The human success story is perhaps best illustrated by the growth in our population. As is clear from Figure 4.7, the growth follows an exponential curve. I've yet to meet anyone who can look at this curve and not conclude that something has to change: if this keeps on then in a century or two there'll be standing

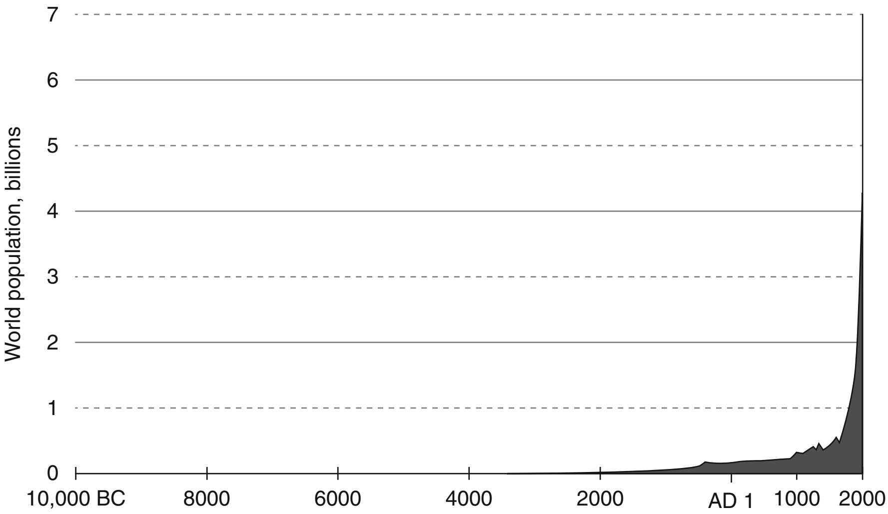



**Fig. 4.7** The growth in human population over history has a few minor bumps; the dip at around 1350, for example, is due to the Black Death—a pandemic that killed somewhere in the region of 100 million people. Essentially, though, human population has grown exponentially. It can't do so indefinitely. (Public domain)

room only. Of course, we won't reach such an extreme of population; the question is: which factor will cut in to curb the population increase? Some writers argue that an environmental catastrophe will provide the limit to population increase, and if they are correct then the future won't be pleasant for the next few generations. The problem is exacerbated because not only is our population growing but each individual is consuming more. As more and more people aspire to Western-style levels of consumption, and as Western-style consumption itself becomes increasingly extravagant, the stress on the planet's finite resources must surely increase. Thanks to scientists such as Norman Borlaug, modern agricultural practices have managed to feed many more mouths than was once thought possible—but can we feed ten billion people if they all want an American-style diet? Thanks to modern engineering techniques we can manage water resources much more efficiently than in the past—but if a drought strikes a region would we be able to slake people's thirst? Thanks to advances in technology we can generate more power now than ever before—but can we generate enough power to satisfy future demand? We consume more energy, more water, more food . . . we are *Homo consumens*. And that can't continue.

**Fig. 4.8** The Rapa Nui people carved 887 moai in the years between 1250 to 1500. The average height of a moai is 4 m and a typical weight is 12.5 tonnes. It's not known for certain why the Rapa Nui people made these carvings, but their creation and transportation must have taken a significant amount of the civilization's resource and focus. The islanders toppled the moai after their civilization collapsed. (Credit: Arian Zwegers)

**The Curious Case of Easter Island** Easter Island, or Rapa Nui, is extremely remote: the nearest continental point lies more than 3500 km away in Chile. Nevertheless, Polynesian voyagers settled there in the first millennium AD and developed a thriving culture. When the first Polynesian people arrived, the island was covered with a thick forest; by 1650 the forest was gone, cut down by the islanders. In 1650 the population of Easter Island was beginning to decline from a peak of 15,000; when the first Europeans arrived in 1722, the population had dropped to perhaps as low as 2000. The cause of the collapse of the Rapa Nui is still a subject of debate, but it seems likely that deforestation played a large role: the islanders were no longer able to build decent boats and so their ability to fish was compromised. Furthermore, the loss of trees caused a collapse in the population of land birds and sea birds. The islanders' number and variety of food sources dwindled. Did the Rapa Nui civilization grow in an unsustainable way?

We've already seen in the discussion of the light cage that fast growth is unsustainable. Haqq-Misra and Baum propose[146] that the Fermi paradox isn't telling us there are no ETCs—rather, it tells us there are no *fast-growing* ETCs. If an ETC grows exponentially then it hits the light cage limit as proposed by McInnes; it lives fast, dies young, and therefore we don't see it. (Although if a spacefaring civilization knew it was about to die, wouldn't it at least try to make a graveyard broadcast or deploy some signal to warn others and let the universe know that they were, for a brief while, great? We've heard nothing.) The alternative is very slow, sustainable growth and we won't have seen civilizations that follow this path because they won't have had time to reach us—if they do move out into the Galaxy it will be by a diffusion process which, as Newman and Sagan showed (see page 91), involves achingly long timescales.

The Haqq-Misra and Baum argument resonates because it speaks to issues that are increasingly important to human society. The exponentially increasing aspect of our consumption patterns might well cause our civilization to collapse and then it won't be humans that colonize the Galaxy. But can issues of sustainability really explain the Fermi paradox? I'm not convinced.

The next few decades will undoubtedly have problems in store for us. Nevertheless, with some luck, plenty of goodwill, and perhaps a few more scientists such as Norman Borlaug, human civilization can avoid a collapse caused by overpopulation and overconsumption. We can already have some confidence that the exponential rise in population *won't* continue: birth rates have been decreasing in many countries for several years and, although a quirk of demographics means that the total number of people on Earth will continue to rise for some time, there's every chance the world's population will stabilize some time in the middle of this century and thereafter begin to fall. Indeed, the most pressing difficulty might not be an exponentially growing population but a declining birthrate. This is already recognized as a problem in Germany, which has a birthrate of just 1.36 children per woman: how will a small number of young people generate the resources to look after a large number of old people? Combine a stable or declining world population with advances in science, technology and computing—advances that seem to follow their own exponentially increasing curve—and we have a recipe for a sustainable civilization. If that came to pass, the expansionist Western-style civilizations will have colonized a large fraction of the planet without imploding or stagnating; the slow-growing human civilizations (Haqq-Misra and Baum mention the Kung San people of the Kalahari desert in this regard) will in a sense have lost out. Earth's own history will have provided a counterexample to the sustainability solution of the Fermi paradox. Furthermore, it's important to recognize that as a civilization expands the resources available to it also expand; with sufficient wisdom and ingenuity (resources that unfortunately might always be in short supply) a civilization can avoid the light cage limit. In particular, as we shall see in Solution 22, a civilization need only survive long enough to produce a single self-replicating probe before it could in principle explore the Galaxy in an entirely sustainable way.

If human civilization can survive and reach a stage where interstellar exploration becomes possible then we have a chance of exploring the Galaxy—and doing so in a "green" fashion. And if we believe *we* might do so in the future then we have to believe that other, much older, civilizations could have done so in the past. Back to the paradox.

## Solution 19 They Stay at Home...

*There's no place like home.*

J. H. Payne

One of the most thrilling events of my childhood happened on 20 July 1969.[147] My father woke me to watch Neil Armstrong and Buzz Aldrin land on the Moon. I guess most people of my age felt the same awe when they saw Apollo 11 touch down. Decades later, we lack the ready capability—and motivation—to repeat the venture. Since Gene Cernan shook the lunar dust from his boots in 1972 no one has set foot on the Moon, and there are no definite plans for anyone to do so. Some space enthusiasts continue to do valuable work on establishing the factors needed for a manned trip to Mars, but such a trip is unlikely to happen soon. An assumption shared by many, including myself, is that intelligent species such as ours will inevitably expand into space—so why aren't we out there? Perhaps the assumption is wrong. Perhaps an unfortunate mixture of factors—apathy, perhaps; economics, almost certainly; and maybe the increasing capacity to gather information from the cosmos without the need for traveling—means that ETCs stay at home. It's not that they are "green" or lacking in technology or worried about the consequences of unconstrained growth. It's simply that they never get round to extensive space exploration. Perhaps that is the rather sad solution to the Fermi paradox.

There's reason to hope the suspension of manned space exploration is simply a pause. As technology improves, the journey into space will become cheaper, safer and happen more frequently. And it won't necessarily be government agencies that are the sole transport providers. Robert Heinlein imagined the possibilities of entrepreneurial space exploration a long time ago, and we have already seen the first space vacationist: in 2001, Dennis Tito paid $20 million to the Russian space program for the privilege of spending eight days in orbit on the International Space Station. This is not to say that private space travel is easy. For some time to come, getting beyond the grip of Earth's pull is likely to remain an activity for nation states rather than companies or charitable foundations. In February 2013, for example, Tito's Inspiration Mars Foundation announced that it hoped to launch a manned flyby mission to Mars in 2018; by December 2013 it was clear that such a mission would be impossible without significant NASA involvement. And Richard Branson promised in 2004 that his *Virgin Galactic* would soon be the world's first commercial spaceline; ten years and several false starts later, the prototype has managed to ascend to about half the altitude that Felix Baumgartner reached in a balloon. Nevertheless, it's quite plausible that the interests of tourism could join the interests of science and high-tech industry in driving manned space travel in the years ahead.

**Not All Cultures are Expansionist** The most frequently cited example of an isolationist civilization is that of China under the Ming dynasty.

The Ming dynasty was founded in 1368 by Zhu Yuanzhang, who became the Hongwu emperor (which in translation means Extremely Martial). Under his rule, and later that of the Yongle emperor, China expanded her empire.[148] The Yongle emperor and his successor, the Xuande emperor, sent the great admiral and explorer Zheng He on seven remarkable voyages. The voyages took him as far as India, the Persian Gulf and the coast of East Africa. Zheng He commanded one of the greatest armadas in history—on his first voyage, 60 of the 317 ships were 400-foot "Treasure Ships"; it must have been an awe-inspiring sight—and undoubtedly China was the leading maritime power of the day. Indeed, China was probably the most technologically advanced nation on Earth. But after the deaths of Zheng He and the Xuande emperor, and for reasons that are still debated, China ceased its expansionist policies, forbade foreign trade, and embarked on an inward-looking path.

In the longer run, there's a compelling reason why we should establish viable independent colonies on Mars, say, or in O'Neill habitats: doing so would help ensure the survival of humanity should disaster strike Earth. In recent years we've come to appreciate just how dangerous a place our planet can be. A large meteor strike would wipe us out just as efficiently as the Chicxulub strike wiped out the dinosaurs. Following the eruption of a super-volcano our technological civilization would crumble. Climate change, whatever the cause, could destroy our way of life. Things have been relatively peaceful here on Earth over the span of recorded human history, but our history corresponds to just a few seconds of the Universal Year. Believing Earth to be a calm and placid place because we've never seen it otherwise is taking the attitude of a man who jumps off the top of a tall building and figures that, since 29 of the 30 floors have passed without incident, he's going to be okay.

In the even longer run, it makes sense to establish colonies around other stars in case something happens to the Sun. A coronal mass ejection only a few times more powerful than the most intense solar flare on record would cause us serious problems.[149] Ultimately, if our species survived long enough, it would witness the Sun move off the main sequence on its way to becoming a red giant—and that really would force a move of home. (Zuckerman has shown[150] that if the Galaxy contains between 10 and 100 long-lived civilizations, then almost certainly at least one of them would have been forced to migrate due to the death of its star. If there are 100,000 such civilizations, then the Galaxy should have been completely colonized by civilizations whose home stars have evolved off the main sequence.)

Humankind has not exactly rushed headlong into space, but it's surely too early to say we will *never* attempt space travel. Our civilization has had the capacity to launch space vehicles for only a few decades; in the context of the Fermi paradox we must think in terms of thousands or millions of years. And although it's probably fruitless to speculate upon the motives of putative extraterrestrials, there seems to be a universal logic to the establishment of, if not interstellar colonies, at least off-world colonies. A species with all its eggs in one planetary basket risks becoming an omelette. Surely technologically advanced ETCs will move, however hesitantly, into space?

The idea that *all* ETCs stay at home seems (to me, at least) unlikely—unless there's a good reason why they should stay at home . . .

## Solution 20 . . . and Surf the Net

*Human kind*
*Cannot bear very much reality.*

T. S. Eliot, "Burnt Norton," *Four Quartets*

The planetarium hypothesis (see page 66) is Baxter's suggestion that we exist in a virtual reality; the universe appears to be devoid of life because advanced ETCs have engineered our reality to make it appear that way. We can invert the planetarium hypothesis to provide a less paranoid resolution to the Fermi paradox: perhaps ETCs generate virtual realities for their *own* use. Perhaps we don't hear from them because they stay at home and engage with an engineered reality more interesting and fulfilling than "real" reality.

It's easy to dream up scenarios in which an ETC might choose to disengage with the real world and instead inhabit a virtual one. For example, suppose their physicists discover a theory of everything, their biologists trace the origin of life back to its chemical underpinnings, their astronomers amass a wealth of observational data that fits their cosmological model, their economists finally understand something meaningful, and their philosophers combine the whole shebang into a consilient theory of knowledge. In short, suppose they conclude that their science is complete. Furthermore, suppose the computing power available to this ETC is far in excess of our own: their simulations can provide satisfying sensory-rich experiences. Finally, what if such a civilization decided interstellar travel, although possible, is too difficult or costly or boring to be worth the effort? Perhaps, under those circumstances, they would cease from exploration. They might instead choose to investigate the almost limitless possibility of artificial realities.

We've no idea whether such a scenario is probable. Some would argue that there'll never be an end to the process of science, that there'll always be some new knowledge for a civilization to discover and some new intellectual vistas to explore. But it's possible that the universe obeys a small set of laws and that the phenomena emerging from those laws are relatively small in number. It's possible that a long-lived technological society might eventually find its science to be essentially concluded. In that case, would they choose to explore inner space rather than outer space? Others would argue that it's impossible to generate virtual realities as convincing as the reality we inhabit. Recall our discussion of the planetarium hypothesis: to generate a virtual reality capable of fooling an advanced civilization requires impossible amounts of computing power. But that misses the point. We aren't talking here about a Baxter planetarium. The computing power required to satisfy *knowing* participants is much less than is required to *fool* humankind. In other words, the simulation designers could take shortcuts. There'd be no need for them to calculate the trillions of interactions in a particle physics experiment; no need to simulate the outputs of protein-folding calculations; no need to present the results of gravitational microlensing observations. Their scientists would already have generated that knowledge in the "real" universe. Since the participants of a virtual reality wouldn't be "kicking the walls" of the simulation, the designers could concentrate upon generating satisfying, compelling and imaginative realities

My guess is that, if our own technology permitted it, a large fraction of us would prefer to live in a virtual reality. Wouldn't you want to spend your time there if a simulation could provide you with a safe yet perfect sensory experience of walking on the surface of Mars, or being chased by dinosaurs, or scoring the winning goal in a Cup Final? It would be infinitely better than TV—and consider how much time we waste on that.

The scenario of a stay-at-home surf-the-Net civilization seems to me to be an uncomfortably plausible future for humankind,[151] but it doesn't alone solve the Fermi paradox. It's an example of a sociological condition that has to apply to *every* technological species for it to work. *We* might come to prefer virtual reality but why should couch-potatohood be a universal characteristic of intelligent species? And even for a society that believes its understanding of the physical universe is complete, there's still the possibility of learning new things by engaging with the universe: discovering the art and the history and the philosophy of an alien civilization isn't something that could be achieved by staying at home. To do that, an ETCs must explore, either directly or by probe. Or, at the very least, it must try and engage in conversation. Surely at least one civilization would try?

## Solution 21 Against the Empire

*Man who man would be,*
*Must rule the empire of himself.*

Percy Bysshe Shelley, *Political Greatness*

The Serbian astronomer Milan Ćirković is one of the most thoughtful and prolific writers on the Fermi paradox. Ćirković points out[152] that the entire premise behind galactic colonization might be flawed. Perhaps an advanced technological civilization will be motivated by something quite different than the need to expand.

Ćirković presents a handful of reasons why it's wrong to think of the development of ETCs in imperial terms. First, as we shall see on page 194, there are grounds for supposing that a sophisticated technological civilization is highly likely to make a transition to a postbiological stage. There are various forms this transition could take—perhaps minds are "uploaded" to silicon; carbon-based bodies might merge with metal-based robots; there are many possibilities, as several SF writers have investigated—but, however the transition occurred, various biological imperatives would be called into question. In a postbiological future will creatures have a genetic heritage to transmit, a biologically determined sex to maintain, infants to protect? And if expansion and colonization have their roots in phenomena such as this, the removal of biological pressures will remove the drive to colonize space. (It's not clear to me that all selection pressures are indeed removed in this scenario, but we're at the limits of speculation here.) Ćirković argues that postbiological civilizations will be motivated by a different definition of success: it won't be measured in terms of the extent to which they control space but rather the extent to which they control the substrate of their environment. In particular, success will probably be measured in terms of the amount and quality of digital computation they can access.

Second comes a more down-to-Earth argument: cost. If our present understanding of physics is correct, and there are no shortcuts to interstellar travel, then colonization is unlikely to be cheap. Furthermore, we've already seen how several authors have pointed out the potential pitfalls of unconstrained growth. If we can see this now, other ETCs will also see those risks early in their development. Ćirković argues that technological civilizations will deliberately choose a different developmental path—perhaps one based on a "city-state" rather than imperial model.

Third is an ethical argument: would a species or its postbiological descendants feel they had the right to influence the biosphere of an alien world? The dilemma of exploration versus potential planetary contamination is already being debated.[153] Wouldn't an ethical civilization choose a developmental pathway that avoided this dilemma? Wouldn't "city-state" trump "empire" on ethical grounds?

Fourth is a political argument: imperialism can become tyranny. Some futurologists believe that global totalitarianism poses one of the most serious existential risks to a technologically sophisticated society. In particular, an artificial entity that has undergone some runaway process of increasing intelligence has the potential to form a *singleton*, to use Nick Bostrom's term.[154] A "good" singleton could be beneficial. A "bad" singleton would be utterly disastrous: a global, stable and long-lived totalitarian regime would be nightmarish. (Recall that chilling line from Orwell's *1984*: "If you want a picture of the future, imagine a boot stamping on a human face—for ever.") This is surely a fate that all civilizations will want to avoid. Ćirković argues that since politics based on a city-state model are on balance more likely to avoid this fate than those based on the empire model, ETCs will tend to opt for the former. (Again, this isn't clear to me. I can imagine situations in which the risks are reversed.)

Fifth is a conclusion drawn from our own experience (assuming we can usefully apply any lessons from human history to the problem of extraterrestrial intelligence): colonial expansion has not been the rule for human societies. Successful civilizations based on the city-state paradigm appeared in the Indus Valley, Babylonia, ancient Greece, Mayan Mexico, medieval Italy, Germany . . . humans aren't driven to empire. And just as city-state civilizations on Earth engage in commerce, communicate with neighbors and are interested in the wider world we might expect ETCs to be open, inquisitive and thirsty for knowledge. They might want to share knowledge with their neighbors rather than conquer them.

We've seen no signs of an interstellar empire. That could be for one or more of the reasons we've discussed in earlier Solutions. But perhaps it's because—for one or more of the reasons mentioned above—ETCs don't build empires. Could the Fermi paradox be telling us that civilizations instead build "city-states"?

## Solution 22 Bracewell–von Neumann Probes

*. . . I looked to these very skies,*
*And probing their immensities . . . .*

Robert Browning, *Christmas Eve*

Of the many contributions to science made by von Neumann (a partial list can be found on page 33), perhaps the most important were in the theory of computing. He became interested in computing at Los Alamos, where he was in charge of the calculations needed for the design of the bomb. Crude calculating machines had been developed to help von Neumann's team in its tasks; after the War, von Neumann turned his mind to what was required of more general-purpose computing machines. His considerations led to many of the fundamental principles of practical computing, and most of today's computers—which are based on the general logical design and mode of operation he championed—are known as von Neumann machines.

The questions involved in the design of a general-purpose computing machine led von Neumann to ask an even bigger question: What is life? As a step toward answering this, he developed the idea of a *self-reproducing automaton*, a

**Fig. 4.9** John von Neumann (*right*) in conversation with Stanislaw Ulam (*left*) and Richard Feynman. All three men played an important role in developing the computers used at Los Alamos. (Credit: American Institute of Physics Emilio Segré Visual Archives)

device that could (a) function in the world and (b) make copies of itself. (Such a device is also sometimes called a "von Neumann machine," but this leads to confusion with *the* von Neumann machine—the architecture at the heart of modern computers. I'll use the term "self-reproducing automaton" when I refer to this hypothetical device.) In von Neumann's scheme, the automaton has two logically distinct parts. First, it has a *constructor*, which manipulates matter in its environment to carry out tasks. A universal constructor has the capacity to make *anything*—including the construction of units it can then use to assemble a copy of itself—as long as it has suitable instructions. Second, it has a *program*, stored in some sort of memory bank, which contains the instructions needed by the constructor.

An automaton can reproduce itself as follows. The program first tells the constructor to make a copy of the program's instructions and place the copy in a holder. It then tells the constructor to make a copy of itself with a clear memory bank. Finally, it tells the constructor to move the copy of the program from the holder to the memory bank. The result is a reproduction of the original device. The reproduction can function in the same environment as the original and is itself capable of self-reproduction.

Of course, von Neumann could not give explicit details of how to *build* a self-reproducing automaton. (Even today, we are far from being able to build such a device, although the seeming convergence of several technologies suggests that we may be able to do so in a few decades. When I wrote the first edition of this book the notion of a "3D printer" was fanciful. Just over a decade later, 3D print shops are on the high street. Other relevant technologies are progressing just as quickly.) But von Neumann wasn't interested in the precise engineering details behind particular mechanisms; rather, he was interested the logical underpinnings of self-reproducing systems. In a lecture first given in 1948, he discussed the relevance of self-reproducing automata to the question of life. He argued that a living cell, when it reproduces, must follow the same basic operations as a self-reproducing automaton. Within living cells, there must be a constructor and there must be a program. He was right. We now know that nucleic acids play the role of the program and proteins play the role of the constructor. All of us are self-reproducing automata. (We discuss the function of nucleic acids and proteins later; see page 271.) What concerns us here, however, is not what von Neumann's self-reproducing automata might tell us about life. Rather, it's whether ETCs could use such automata to spread throughout the Galaxy.

As long ago as 1980,[155] Robert Freitas sketched the outlines of a self-reproducing interstellar probe and Frank Tipler discussed the relevance of self-reproducing automata to galactic exploration. The basic idea is that an ETC can spread throughout the Galaxy by launching self-reproducing

*Bracewell–von Neumann probes*. (These devices are usually called simply von Neumann probes in the literature. However, to the best of my knowledge, von Neumann never considered probes in the context of interstellar travel. The first person to suggest[156] that probes could be useful for interstellar communication was Ronald Bracewell. Although a Bracewell probe need not be a self-reproducing automaton, adding the capacity for self-reproduction to such a probe greatly increases its effectiveness. It seems reasonable to refer to these devices as Bracewell–von Neumann probes.)

The probe need not be a single device (indeed, it's better thought of as a collection of different devices, which taken together have the overall capacity for reproduction) but it need not be a vast machine. In the scenario sketched by Tipler, a Bracewell–von Neumann probe can be small: the payload need be nothing more than a self-reproducing automaton—one with a universal constructor and an intelligent program—and a basic propulsion system for use within the target system. After arriving at the target star, the program instructs the probe to find suitable material with which it can reproduce itself and make copies of the propulsion system. If the planetary system resembled our own, then there would be plenty of raw material available for the constructor: asteroids, comets, planets and dust could all be broken down and utilized. If necessary, radio signals from the home planet could send revisions to the program, so that the probe's software need never be out of date. Soon after arrival there would be a host of probes, each undertaking some pre-programmed task. Some might explore the planetary system, sending back scientific data to the home world. Some might construct a suitable habitat for later colonization by the home species. Some might even raise members of the original species from frozen embryos stored as part of the payload (or, as we shall see on page 121, they might be able to restore an entire biosystem from programs beamed from the home world). And some would move on to another star, where the process would be repeated until every star in the Galaxy had been visited.

Tipler argued that if probes traveled between stars at the rather stately speed of $c/40$, and if the propagation of the probes was directed rather than random, then a colonization wave could surge through the Galaxy in roughly 4 million years—a period that equates to just 2 hours 33 minutes of the Universal Year. As is to be expected, this time is much shorter than the colonization times in the Newman–Sagan, Fogg, Bjørk and Cotta–Morales models. Probes need not stay in a planetary system and wait for colonists to give them instructions on how to proceed: they already *have* their instructions. The galactic colonization time is short because the process is *planned* to be efficient. Tipler's initial analysis was perhaps overly optimistic, but a variety of more recent investigations[157] seem to confirm the basic result: self-reproducing probes traveling at quite a small fraction of the speed of light can colonize the Galaxy on a timescale of 5–10 million years.

Not only is colonization by probe quick, it's cheap. Implicit in most of the models considered earlier was the idea that planetary systems would be investigated and colonized by living beings—an expensive activity since the payload would have to contain food, water, life support and so on. Probes don't have this problem. An ETC must construct the first few probes and send them out, but after that Nature picks up the tab in terms of providing raw material for the continuing process.

Can such a probe ever be built? Well, it's possible in principle. A spaceship containing a sufficient number of human couples, the appropriate life support systems, stored knowledge in the form of large databases and a sophisticated onboard factory would constitute a Bracewell–von Neumann probe. It would be impractical, of course: the cost benefits mentioned above would vanish because of the need to feed, shelter and entertain the human passengers. In principle, however, it would work: the system could reproduce itself and continue the process of exploration. The trick to creating a more practical Bracewell–von Neumann probe would be to replace humans with some form of artificial intelligence. Certainly there are significant technical and engineering hurdles to overcome, but this is the sort of technology humankind will have to develop if we want to explore and exploit the Asteroid Belt, for example, or the Oort Cloud. And if in the next few centuries *we* might use probes for interplanetary exploration and exploitation,[158] surely a technological civilization in advance of us by thousands or millions of years might develop interstellar probes. There seems to be no fundamental reason why they couldn't.

Colonization of the Galaxy by probe is technologically possible; it's quick; and it's cheap. Even if the aim is contact rather than colonization, Bracewell showed there are circumstances in which probes are more effective than radio signals. So as Fermi would ask: where are the probes?

We touched on this question in chapter 3, when we discussed the possible use of probes in directed panspermia and when we considered places where a monitoring probe could hide. But such probes aren't of the Bracewell–von Neumann type, the sort of probe that can dismantle planets, undertake astroengineering projects and colonize the Galaxy in the cosmological blink of an eye. While we can't rule out the existence of monitoring probes being in the Solar System right now, we'd surely have noticed if a self-reproducing probe bent on colonization had visited the Solar System. There's no evidence for such activity elsewhere in the Galaxy, either.

Even if an ETC had the ability to construct Bracewell–von Neumann probes would it necessarily choose to deploy the technology? It's not exactly a risk-free technology,[159] after all. The probes reproduce as if they were living beings, rather than replicate as if they were crystals, so inevitably there'll be reproductive errors. There'll be mutations. Probes would evolve, just as biological creatures evolve. The Galaxy could soon be home to different probe "species", each with its own interpretation of its goals. There'd be a risk, for example, of a probe returning to the home system and failing to recognize it—not good news for the ETC if the probe's orders are to dismantle planets and use the material to construct something else. But is it a risk *every* ETC declines to take, a problem *every* ETC fails to solve? My gut feeling is that any civilization advanced enough to construct an efficient Bracewell-von Neumann probe would have the technological nous to put in place the necessary safeguards.

Since colonization of the Galaxy by probe seems straightforward, at least on paper, some authors argue that there's an inevitable motivation for an ETC to engage in colonization: if we don't do it, some other species will. Stake your claim early, in other words. (This sort of argument might have appealed to von Neumann, who was a strong proponent of the nuclear first strike. In an interview with a *Time* magazine reporter, von Neumann said: "If you say why not bomb them tomorrow, I say, why not today? If you say five o'clock, I say at one o'clock." We must be grateful that, in the 1950s and 1960s, wiser counsel than von Neumann's prevailed.) Perhaps Ćirković is right and we can hope intelligent species develop to the stage where they have no urge to own every star, inhabit every planet, and populate their galaxy with beings just like themselves. Nevertheless it takes only *one* ETC to reason that it shouldn't take the risk of losing out on all that real estate . . . Indeed, probe technology seems so straightforward, at least on paper, that we don't have to think in terms of entire civilizations engaging in colonization. Perhaps *subgroups* of a truly advanced ETC would have the capacity to colonize a galaxy. Why hasn't one of those subgroups tried to colonize our Galaxy, or promulgate their particular religion, or simply spread out so as to minimize existential risks?

A discussion of Bracewell–von Neumann probes is relevant to any discussion of the Fermi paradox, but you might ask why I present it in a part of the book devoted to *solutions* of the paradox. Well, a surprising number of people seem to believe that probe technology *does* resolve the paradox. They argue that we don't see aliens because they'd send probes rather than travel interstellar distances themselves. Of course, this entirely misses the point. Fermi's question refers either to aliens or the product of alien technology. After all, if we detected an object in space that was clearly artificial yet not made by us then we could deduce the existence of a civilization that constructed the object. We see no evidence of aliens *nor of their probes*. Far from resolving the paradox, the possibility of Bracewell–von Neumann probes makes Fermi's question all the more intriguing. Indeed, recent work has significantly sharpened the paradox:[160] Stuart Armstrong and Anders Sandberg, researchers at Oxford University, have shown that an ETC capable of colonizing a galaxy using Bracewell–von Neumann probes possesses the capacity to colonize the reachable universe! If intelligent, technologically advanced civilizations arose a billion years ago, and if they developed the ability to send out probes traveling at 0.8*c*, then representatives from over a million galaxies could have reached us by now. It's not just the Milky Way galaxy we must take into account when considering Fermi's question, it's all our neighbors too. Bracewell–von Neumann probes give the paradox real bite.

## Solution 23 Information Panspermia

*There are no foreign lands.*
*It is the traveller only who is foreign.*

Robert Louis Stevenson, *The Silverado Squatters*

The Armenian mathematical physicist Vahe Gurzadyan has posited an interesting hypothesis:[161] we might inhabit a Galaxy "full of traveling life streams"—strings of bits beamed throughout space. The argument goes as follows.

We know that strings of characters can contain information. Consider two strings, each containing a trillion characters. The first string starts "101010 . . ." and continues in that way until the trillionth character is reached; the second string starts "x9Y$m& . . ." and carries on in a seemingly randomly pattern. The *Kolmogorov complexity*[162] of strings such as these is defined as the minimum length in bits of a binary-coded program that describes the string. The Kolmogorov complexity of the first string is small because one requires only a short program to describe it: in words, the program could be something along the lines of "Print alternating sequence of 1s and 0s, starting with 1 and ending after the trillionth digit". The Kolmogorov complexity of the second string is large because there's no obvious way of compressing the information it contains; any program describing the string would likely be as long as the string itself. Gurzadyan argued that the Kolmogorov complexity of the human genome—indeed, of the totality of terrestrial life—is relatively low. There's a vast amount of genetic information contained in the millions of species on Earth, but the program that describes that information might be much smaller.

**Genome Sequencing: of Mice and Men** The human genome was sequenced and published in draft form in 2001, and in complete form in 2006. Many other animals have had their genomes sequenced. For example, in 2013, and picking a few creatures at random, the African lion, the greater false vampire bat and the peregrine falcon all had their genomes sequenced. So did our extinct cousins, the Neanderthal.[163] Complete genome sequencing has become routine.

As more species are sequenced, it's becoming clear that there's a great deal of similarity between genomes. For example, the mouse genome has been studied in detail and so we know that mice and humans share virtually the same set of genes. That's not surprising: human and mouse shared a common ancestor that lived about 80 million years ago. Indeed, all mammals share a common ancestor and so will exhibit similarities in their genomes; if one goes further back, all life shares a common ancestor. So once we know the complexity of the human genome, the additional complexity of terrestrial species will be small.

Suppose we wanted to communicate all the genomic information contained in terrestrial life. Communication takes energy: the more bits we have to transmit, the greater the energy requirements. If we wanted to send a file containing all of Earth's genetic data then the cost in energy would be prohibitive; if instead we sent a *program* that could recover that information then the energy cost would be small. This is the same argument that says transmitting a trillion digits is much more costly than transmitting the string "Print a trillion alternating 1s and 0s". Gurzadyan showed that with an Arecibo-like antenna it would be possible to transmit the genomes of terrestrial organisms throughout the Milky Way galaxy.

Gurzadyan, then, imagines a type of what might be called "information panspermia". He describes the possibility of a Galaxy in which ETCs establish a network of self-replicating Bracewell–von Neumann probes and life is propagated not by sending the genomes themselves but by sending the programs that can recover the genomic information. In other words the probes, which could be many light years away from their home planet, would receive coded strings and from those strings reconstitute the full panoply of that planet's life. Even now, life might be raining down on us. But it would be a strangely desiccated form of life: not living creatures, but rather ghostly strings of information that have the potential to become living.

Gurzadyan says this idea can eventually approach a solution to Fermi's paradox. However, I'm not entirely sure how this is so. The hypothesis certainly has implications for SETI: perhaps we should be analyzing radiation for evidence of bit strings? However, if ETCs are indeed spreading their form of life via a Galaxy-wide network of Bracewell–von Neumann probes then why, as we've already argued, aren't they already here? They've had plenty of time to reach us but, unless you believe that terrestrial life is the result of the unpacking of a transmitted bit string, we see no signs of them. To my mind, Gurzadyan's hypothesis, rather than being a solution to the Fermi paradox, is a particular example of how directed panspermia might be made to work. (What perhaps could resolve the paradox is if ETCs either won't or can't make self-replicating probes. In that case, might they send out life streams anyway, like dandelion "clocks" in the wind, hoping that occasionally someone somewhere will catch one and reconstitute the life they contain?)

## Solution 24 Berserkers

*In the long run, we are all dead.*

J. M. Keynes

During the 1950s, Cold War strategists toyed with the idea of a "Doomsday weapon". Such a weapon was terrible, uncontrollable, capable of destroying all human life on Earth—including the owners of the weapon. If your enemy knew you were willing to deploy a Doomsday device then—so the Cold War logic went—they would not dare attack you. I suspect that Fred Saberhagen had the Doomsday weapon in mind when he wrote his famous berserker stories.[164]

Berserkers are sentient, self-reproducing machines that are savagely inimical to organic life. Think of them as paranoid Bracewell–von Neumann probes with a mean streak. The relevance to the Fermi paradox is clear: the originators of berserkers are either dead or in hiding; all other ETCs have either been prevented from arising by berserkers, wiped out by berserkers, or else are keeping quiet for fear of attracting berserkers. It's an elegant solution to the Fermi paradox. But could berserkers exist outside the pages of science fiction?

If an ETC could build probes capable of colonizing the Galaxy, then unfortunately berserker construction would presumably not be beyond them technically. It's hard to imagine any intelligent species actually *wanting* to develop berserkers, since the technology is so dangerous to the creators as well as to all other life. Besides, what would be their motivation for constructing berserkers? If their aim was to colonize the Galaxy for itself then, if Bracewell–von Neumann probes are indeed possible to build, it could fulfill its aim simply by being the first to colonize: remember that the colonization time for the Galaxy is much less than the age of the Galaxy. (If the probes are impossible to build then so are berserkers.) However, perhaps we shouldn't be overly sanguine about the prospect of berserkers. Suppose the programming of a "well-adjusted" probe mutates; perhaps a collision with a stray cosmic ray changes the line of code in its core module from "seek out new life and new civilizations" to "seek out new life and new civilizations, *and kill them*." Self-reproducing probes will inevitably evolve, and so berserker-type devices *might* develop.

The berserker solution has been criticized on several grounds. Even if berserkers exist, would they be an inevitable Nemesis? Could not ETCs "inoculate" themselves, much as they would inoculate themselves against a virulent disease? Most tellingly, the berserker scenario suffers from a Fermi paradox of its own: if berserkers exist, then how come *we* are here? Berserkers should already have sterilized our planet. Instead, as we shall see in later sections, the geological record indicates life has been present on Earth for billions of years. To be sure, Earth has seen several mass extinctions, but there are natural explanations for these events—the universe is dangerous enough without berserkers. So why have berserkers silenced all other civilizations but left us alone? We could argue that berserkers destroy only technological life-forms and need a "trigger"—presumably the detection of radio waves—before they begin work. But that extra step in the argument spoils what is potentially an elegant resolution of the Fermi paradox. Besides, we've been using radio for a century and may soon go radio-quiet despite our burgeoning level of technology. If berserkers are all they are cracked up to be, then where are they?

## Solution 25 They Are Signaling but We Don't Know How to Listen

*The world should listen then—as I am listening now!*

Percy Bysshe Shelley, *To a Skylark*

Perhaps large-scale interstellar travel is unattainable, either for crewed starships or for probes. This would explain why we haven't been *visited*, but not necessarily why we haven't *heard* from them. I've never been to Australia because it's too time-consuming and expensive to get there, but I keep in touch with antipodean goings-on through a variety of telecommunication tools. Fermi asked "Where *is* everybody?" in the context of a discussion about spacecraft, but surely his question should refer to more than the mere absence of visitors. It must surely refer also to the absence of *any* evidence that technologically advanced extraterrestrial civilizations exist.

A civilization will presumably quickly discover whether interstellar travel is attainable. If it concludes that stellar journeys are impossible then why should it hide? After all, it needn't fear invasion by an aggressive neighbor, since any neighbors would be too distant to pose a threat. Given that's the case, one can think of a number of reasons why it might choose to signal its presence. It might shout for help—perhaps it's facing a long-term existential threat that it hopes other civilizations have surmounted and could advise upon—or at least announce its existence if it knows its end is nigh. It might want to brag about its cultural achievements and high points. It might want to convert others to its religion, or sell information, or simply shout out to try and end a sense of loneliness. There are numerous possibilities. Such an ETC has nothing to lose by signaling, and the potential reward is huge: mutually satisfying dialogs with equally advanced civilizations. But if advanced civilizations are out there, educating each other, gossiping, holding conversations that are the interstellar equivalent of the Algonquin Round Table, then why haven't we been asked for our opinions? At the very least, why haven't we heard the babble of discussion?

One plausible answer is that we don't know how an ETC would choose to send a signal. We therefore don't know how to listen.

It's certainly true that we can't know what communication technology ETCs might possess. As my editor once pointed out, if a radio engineer from 1939 were somehow transported to modern-day New York, he could build a radio receiver and conclude there were almost no useful radio broadcasts being made: he wouldn't know about FM. Similarly, he would be blissfully unaware of communication devices employing lasers, fiber optics or geosynchronous satellites. So it's conceited of us to suppose we could understand what communication channels might be available to technical cultures that are millions of years in advance of our own. If they wanted to talk to each other secretly (perhaps they don't want to influence the development of young species such as our own?) then presumably they could maintain secrecy without difficulty. But the situation is different if they *want* to be heard, and heard widely. We can assume that every civilization must obey the laws of physics; moreover, any ETC will know that other ETCs must obey those same laws. Since we all have to pay our energy bills, the number and types of signal that can reasonably be sent is quite restricted. Let's examine the advantages and disadvantages of four methods of communication: signals using electromagnetic waves, particle beams, gravitational waves and hypothetical tachyon beams.

## Electromagnetic Signals

The most obvious way to send information is via electromagnetic (EM) radiation. Not only does electromagnetic radiation propagate at $c$, the fastest possible speed, it is capable of propagating over interstellar and intergalactic distances. We know that EM signals can operate over such distances because many natural objects indicate their presence in this way over vast reaches of space. After all, astronomy is essentially the science of recording and interpreting EM signals. We use visible light when we look at stars with our eyes or photograph them with optical telescopes; we use radio waves when we study the sky with radio telescopes; increasingly, particularly in satellite experiments, we use the infrared, ultraviolet, X-ray and gamma-ray wavelengths. If we can study natural objects over interstellar distances using the EM radiation they emit, then in principle we can do the same with artificial objects.

For many years the working assumption made by researchers looking for ETCs has been that technological civilizations will build powerful EM transmitters, broadcast a signal, and modulate it in order to convey useful information—perhaps, if we are lucky, they will broadcast their "Encyclopædia Galactica". In Solution 26 I discuss in detail how we might detect purposeful EM signals. Here, I want to argue that it may even be possible to detect EM radiation that leads to the discovery of *inadvertent* markers or beacons of KII civilizations. (Detecting inadvertent markers of a KIII civilization might be even easier.) Even an inadvertent beacon would convey a tremendous amount of information: that life exists on another world, that such life on that world is technologically advanced, the location of the world, and so on.

We've already discussed why a KII civilization might choose to construct a Dyson sphere. Such an object would radiate just as much energy as the central star—the energy has to go somewhere—but it would presumably do so in the infrared. In essence, the sphere would radiate because it is warm, about 200–300 K. So one way to search for an ETC would be to look for bright infrared sources at a wavelength of around 10 microns; the radiation might be the waste heat from astroengineering projects. It's not an easy task, since many stars will exhibit a large infrared excess simply because they are shrouded in dust, but it can be done.

In the early 1990s, a search by Jun Jugaku and Shiro Nishimuro for artificial infrared sources out to a distance of 80 light years found no plausible signatures[165] from Dyson spheres. A few years later, a search at 203 GHz of 17 stars known to produce excess infrared radiation found nothing unusual.[166] In 2009, Richard Carrigan carried out[167] an analysis of the historical IRAS catalogue. (The Infrared Astronomical Satellite, or IRAS, was one of the most important satellite missions of the 1980s and the first space-based observatory to survey the entire sky in the infrared.) Of more than 250,000 IRAS objects, only a few were in any way plausible Dyson sphere candidates; follow-up observations of the 16 least-unlikely candidates using radio telescopes found nothing of interest. Jason Wright and his colleagues[168] are searching databases from more recent and more sensitive satellite observations—those from the Wide-field Infrared Survey Explorer WISE and the Spitzer Space Telescope—to search for waste heat from alien technologies. Their search will enable them to put limits on the activities of any KIII civilizations; for example, they can look for "Fermi bubbles"—patches of a galaxy with high infrared emission, which could be a possible sign of a civilization transforming its galactic neighborhood.

Of course, we can't conclude from the negative results to date that there are no ETCs in the solar neighborhood; civilizations might have chosen not to construct Dyson spheres here for a variety of reasons. Besides, *really* advanced civilizations—as Marvin Minsky pointed out[169]—would consider radiation at any temperature above the cosmic background temperature of 2.7 K to be wasteful. Perhaps an ETC advanced enough to construct a Dyson sphere is advanced enough to squeeze every last drop of useful work out of a star's radiation, leaving waste heat at a few kelvin. Perhaps Dyson spheres are common but we should be looking for them by searching for points in space that possess a small temperature excess over the microwave background?

In 1980, Whitmire and Wright gave another example of how inadvertent beacons can be transmitted[170] by electromagnetic radiation. They asked what would happen if a civilization used fission reactors as an energy source over long periods of time. One of the problems with fission reactors is the need to dispose safely of radioactive waste material. And one proposed disposal method is to launch it into the Sun (though I, for one, wouldn't be too thrilled at the prospect of having tons of radioactive waste perched on top of a chemical rocket). If an ETC used its star as a dumping ground for radioactive waste, then the spectrum of the star could exhibit characteristics that couldn't easily be interpreted as natural. For example, if we saw a stellar spectrum containing large amounts of the elements praseodymium and neodymium, then our interest would be caught. Furthermore, the alteration in the spectrum would not be a brief flicker; the spectral evidence of their nuclear waste disposal policy would be visible for billions of years. (A civilization might even *deliberately* alter its star's spectrum in this way to create a beacon. This possibility was first suggested by Drake. Another method of using one's home star as a beacon was suggested by Philip Morrison: put a large cloud of small particles in orbit around the star in such a way that the cloud cuts off the starlight for a viewer who is in the plane of the cloud's orbit. Move the plane of the cloud and the distant viewer sees the star flash on and off. Variable stars naturally alter in brightness, but if the star flashed in a pattern that represented prime numbers, for example, then the distant viewer could quickly rule out a natural phenomenon.[171] The beauty of this approach from the signaler's point of view is that the signal is likely to be detected by less advanced civilizations during their routine astronomical observations: human astronomers, for example, search for the dimming of starlight caused by transiting planets—it's one of the best ways of finding exoplanets.)

So far, no EM beacons—inadvertent or not—have been identified.

## Particle Signals

Cosmic rays, in the form of electrons, protons and atomic nuclei, can reach Earth over interstellar distances—and cosmic-ray astronomy is a thriving research field. However, charged particles would constitute a poor choice of communication channel because a transmitting civilization couldn't guarantee where the particles would end up: twisting magnetic fields throughout the Galaxy make the paths of these particles quite tortuous. Neutrinos are electrically neutral, so at first glance they seem a better choice for a communication channel. Unfortunately, neutrinos are difficult to study because they react so infrequently with matter; typically, a neutrino will pass through 1000 light years of lead before stopping! Nevertheless, despite the tremendous difficulties involved, astronomers have developed neutrino telescopes. Physicists have even proposed an experiment that would generate trillions of neutrinos every second, in the Fermilab facility in Illinois, and send them to a detector in South Dakota some 1300 km away. The purpose of the experiment is to learn more about neutrino masses, but in principle I guess it could be used to send a signal between Illinois and South Dakota. Could ETCs do something similar but on a much grander scale?

**Neutrino Telescopes** The first neutrino telescope was the brainchild of Ray Davis[172], who developed it in order to study the nuclear fusion reactions that take place in the heart of the Sun. His telescope was in essence a 100,000-gallon vat of perchloroethylene (dry-cleaning fluid) buried almost a mile beneath the ground in a gold mine in South Dakota. It was the strangest telescope anyone had ever constructed (there are stranger telescopes nowadays), but the setup was necessary because neutrinos are so elusive. The dry-cleaning fluid provided enough chlorine atoms to guarantee detectable numbers of neutrinos, while the depth of the mine shielded the vat from other subatomic particles that bombard Earth. His telescope found only one third of the expected number of solar neutrinos, a finding that was a significant result for particle physics: it turns out that neutrinos come in three "flavors"—electron, muon and tau—but the Davis telescope was sensitive to only one type. Nuclear reactions in the Sun create the expected number of neutrinos, but on their way to Earth the neutrino flavors "oscillate".

A more recent and sensitive neutrino telescope is IceCube, which has detectors buried deep in Antarctic ice. In 2013, the IceCube collaboration announced that it had detected 28 high-energy neutrinos that came from some extremely powerful events in deep space. The age of neutrino astronomy is upon us.

In February 1987, the Kamiokande detector in Japan and the IMB detector in America between them stopped 20 neutrinos in a period of a few seconds. Those neutrinos were produced in the famous supernova of that month:

**Fig. 4.10** The IceCube Laboratory in March 2012. The lab is located at the Amundsen–Scott South Pole Station in Antarctica, and hosts the computers that collect raw data. The neutrino detectors themselves, though, are buried deep beneath the ice: the sensors are spread throughout a cubic kilometer of ice and look for flashes of Cherenkov radiation that might indicate the interaction of a high-energy neutrino from space with an atom here on Earth. Although IceCube is based in Antarctica it's actually looking "down" through the bulk of the Earth: it aims to capture neutrinos that arrive from the northern hemisphere. (Credit: Sven Lidstrom; IceCube/NSF)

SN1987A. The supernova SN1987A occurred in the Large Magellanic Cloud, about 170,000 light years away. Demonstrably, then, it's possible for neutrinos to travel interstellar, even intergalactic, distances and for a primitive technological civilization such as ours to detect them. Perhaps ETCs use modulated neutrino beams to communicate[173] with each other? Well, perhaps. Since we are beginning to possess the telescopes that will enable us to look seriously for cosmic neutrinos it can do no harm to keep an eye open for the possibility of artificially generated neutrinos. One wonders, though, whether ETCs would bother with neutrino-based communication when electromagnetic waves do the job with much less fuss. Cost, too, will surely always be a consideration. If it goes ahead the neutrino experiment alluded to above, which sends neutrinos from Fermilab to the mine where Ray Davis did his groundbreaking work, will cost $1.5 billion. That's cheap if you want to learn more about some of the fundamental constituents of the universe; it's hideously expensive if the intention is to send a message.

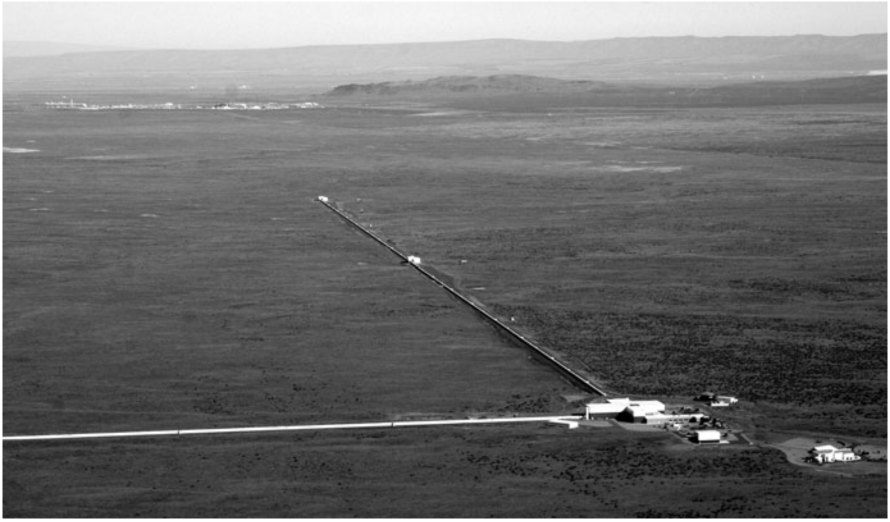


**Fig. 4.11** The LIGO Hanford Observatory, in Washington State, consists of two 4-km arms at right angles, each with laser beams in high vacuum. There is an identical observatory in Louisiana, and the two installations work in tandem. The objective is to detect gravity waves by searching for changes in arm length that are a thousand times smaller than an atomic nucleus. (Credit: LIGO Laboratory)

## Gravitational Signals

Apart from electromagnetism, the only known force that acts over astronomical distances is gravity. It too propagates at the speed of light, so perhaps ETCs might use gravitational waves to signal each other? Gravity, however, is a *much* weaker force than electromagnetism. To build a gravity-wave transmitter you have to be able to take large masses (of the order of a stellar mass) and shake them violently. It's debatable whether even a KII civilization would possess that sort of technology. A KIII civilization might be able to build such a gravity wave transmitter, but why would it bother when EM waves do the job just as well and EM transmitters are so much easier to construct?

Then there's the problem of detecting gravitational waves.[174] It's far more difficult to detect gravitational waves than it is to detect EM waves. It's so difficult, in fact, that terrestrial science has yet to directly detect gravitational waves. Detectors such as LIGO and VIRGO are searching for gravitational waves, but even if they are successful they'll detect gravitational radiation from only the most violent astronomical phenomena. This would be exceptionally interesting scientific data, but we won't find modulated signals in that data.

So, given the difficulties of transmitting and receiving gravitational waves, it seems unlikely that an ETC would choose to use them for communication.

## Tachyon Signals

We can speculate that extremely advanced ETCs will use tachyons—particles that move faster than light—to signal each other. If tachyons exist, and if it's possible to modulate a beam of them to carry signals, then no doubt they'd be an attractive option for interstellar communication. Tachyon-based communication would obviate that irritating delay between asking a question and receiving an answer—a delay that can be hundreds or thousands of years. Unfortunately, as we saw earlier (see page 85), there's absolutely no evidence that tachyons exist, let alone that it's possible to use them to send signals.

Some SF authors have considered a related proposal. One of the oddest features of quantum mechanics is a phenomenon called entanglement. Suppose you have a pair of particles that were created in such a way that you can't describe the quantum state of each particle independently; rather, you can only describe the quantum state for the whole system. For example, you might have created a pair of particles with the property that the overall spin of the system is zero—you know that one of the particles is spinning "up" and one of the particles is spinning "down" but until you measure the spin of an individual particle you are forced to conclude that both particles are in a quantum superposition of spin-up and spin-down. In a sense, until a measurement is made, both particles are spinning up *and* down. The particles are entangled. Suppose you separate the particles by one light year. If you then observe the particle *here* to be spin-down then—instantaneously—the distant particle *there* becomes spin-up. It's as if some influence has travelled one light year in an instant. So can the phenomenon of entanglement be pressed into service as a tachyonic communication channel? Unfortunately not. No information can be transmitted in this way. Besides, whether you believe that making a measurement *here* somehow directly affects a quantum system *there* depends on how you choose to interpret quantum mechanics.

* * *

Perhaps there are many civilizations out there, communicating with each other using gravitational waves, neutrinos and tachyons. Or perhaps they send signals using techniques we have yet to dream of—techniques that break no laws of physics but that are as exotic to us as fiber-optic communication channels would be to a 1939 radio engineer. Since we can't detect such signals it would explain why we haven't heard from ETCs; it would explain the "Great Silence".

On the other hand, even for advanced civilizations, communication by EM waves seems to be a logical choice: the signals are cheap to produce, the information moves as fast as is possible in a relativistic universe, and the messages are easy to receive. And if an ETC *wanted* to make its presence known to other perhaps less developed civilizations (such as ours), then the EM spectrum might be its only option.

For these reasons, although it might seem conceited and it might mean we're missing out on some wonderful conversations, many physicists would argue that we *do* know how to listen for signals from extraterrestrial civilization: we should listen for EM radiation. In fact, given the level of our present technology, we have little option but to try and detect such radiation. But. . . at what frequency should we listen?

## Solution 26 They Are Signaling but We Don't Know at Which Frequency to Listen

*57 channels and nothing on.*

Bruce Springsteen

If extraterrestrial civilizations do indeed use EM radiation to keep in touch with each other, as a means of notifying less advanced civilizations of their presence, or simply as a tool for internal communication, then there are several different types of signal we might search for.

What hope do we have of detecting a signal that wasn't directly intended for us? For example, could we detect leakage radiation from other activities? Well, for several decades EM signals have leaked from Earth due to our television transmissions and our use of military radar; perhaps we could detect the alien equivalent. On the other hand one can reasonably argue that developments in cable and satellite telecommunications systems mean that radiation leakage from Earth will soon cease; and if that happens to us then we might expect it to happen to ETCs. Perhaps the period over which a technological civilization is "radio bright" can be measured in decades, in which case we have essentially no chance of discovering this type of signal. It's possible to imagine EM leakage emanating from future technological developments—such as solar satellites beaming energy back to the home planet in the form of microwaves, perhaps, or navigational beacons for steering through a crowded planetary system—but it would be tough to find them. At first glance a better approach would be to "eavesdrop" and search for inter-civilization communications. However, when one looks at the numbers it becomes apparent that the chances are small of intercepting a communications channel not specifically intended for us. By far the easiest type of signal to detect would be one that an ETC intended for us to receive—either an omnidirectional beacon that *everyone* has a chance of seeing or, even better, a signal deliberately targeted at us.

It's not too arrogant to suppose a nearby ETC would beam signals toward the Sun. Technologically advanced civilizations would surely classify the Sun as a good candidate for possessing life-bearing planets. Furthermore, they could detect the existence of Earth over interstellar distances. We know this is the case because *we* are at the early stages of being able to do this. The success of NASA's *Kepler* mission proves the technology works. For example, in 2013 the mission discovered Kepler-37b—a planet 210 light years away that has a radius not much larger than the Moon. The technology available to human astronomers is advancing steadily, and in a decade or two we'll be able to detect atmospheric biosignatures—oxygen and methane, for example—on distant exoplanets. If we can do it then we must assume that any technologically advanced civilization in our cosmic neighborhood would be aware of Earth's potential for hosting life. If they beam signals to target stars in the hope of making contact then our Sun would be on their list. (Yes, this statement does sound rather too definite. We are trying to second-guess the motives and intentions of putative aliens—an enterprise fraught with risks. But we have to begin somewhere.)

With our present level of technology it makes much more sense to search for targeted communication rather than hope to eavesdrop on others' conversation or expect to find leakage radiation. But at which wavelength will ETCs choose to transmit? In other words: at what frequency should we listen?

**The Electromagnetic Spectrum** The hertz (Hz) corresponds to one cycle of vibration per second; 1 MHz is $10^6$ or 1 million vibrations per second; 1 GHz is $10^9$ or 1 billion vibrations per second. In these units it's easy to appreciate the extreme breadth of the EM spectrum.

Visible light reaches from $7.5 \times 10^{14}$ Hz (deep violet) to $4.3 \times 10^{14}$ Hz (red) and forms only a minuscule part of the spectrum. Ultraviolet, X-rays and gamma-rays have progressively higher frequencies, reaching up to $3 \times 10^{19}$ Hz or higher. Infrared, microwaves and radio waves have progressively lower frequencies, reaching down to $10^8$ Hz.

Our technology employs all these wavelengths for a variety of purposes, ranging from medical applications (X-ray frequencies) to household devices (garage door openers work at 40 MHz, for example, and baby monitors at 49 MHz). There seems to be a frequency for everything. So which frequency is best for interstellar communication?

In the late 1950s, Philip Morrison and his colleague Giuseppe Cocconi were among the first to consider this question.[175] Astronomers had by then developed radio telescopes and were making significant discoveries by looking at the universe through this new window. It was against this background that Morrison investigated the possibility of observing the universe using gamma-rays. As part of this work he showed how gamma-rays, unlike visible starlight,

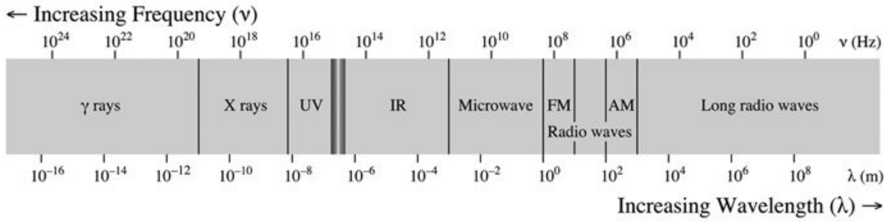



**Fig. 4.12** The wavelengths and frequencies of the electromagnetic spectrum. The horizontal lines appear on a logarithmic scale. It's clear from this diagram that visible light, which is sandwiched between UV and IR, corresponds to only a small fraction of the electromagnetic spectrum. (Credit: Philip Ronan)

could travel across the dusty plane of the Galaxy. He told Cocconi of this result, and his colleague pointed out that particle physicists already generated gamma-ray beams in their synchrotrons; why not send the beam into space and see if an ETC could detect it? It was a fascinating question, and it got Morrison to thinking about the prospects for interstellar communication. He replied that they should consider not just gamma-rays but the whole EM spectrum—from radio waves all the way up to gamma-rays—and choose the most effective band for signaling.

They quickly concluded that visible light would be a poor choice for signaling, since the signals would have to compete with starlight; X-ray and gamma-ray telescopes were not feasible at that time; the radio band seemed the best bet. Furthermore, there were radio dishes already being planned that could take part in the search. If an ETC had dishes of the same size, and used them to transmit directed beams at a sharply tuned frequency, then our radio telescopes could detect their signals from halfway across the Galaxy.

Narrowing down the search to the radio band was a major advance, but it still left many possible frequencies. Radio waves can be anywhere between about 1 MHz and about 300 GHz. This is bad news, for the following reason. If an ETC wants to catch our attention then there are some good reasons for supposing that it will transmit at a precise frequency—it will send a *narrowband signal*;[176] wideband signals are easily mistaken for background noise. (When you twiddle the dial on a radio—the old-fashioned type, not those new-fangled DAB radios with push-buttons—you hear the background hiss of wideband noise between the narrowband signals of the radio stations.) Interstellar masers, which amplify microwaves and act in very much the same way as lasers, generate the narrowest naturally-occurring frequencies: an interstellar maser can emit radiation with a width as small as 300 Hz. Transmissions therefore probably require a bandwidth of much less than 300 Hz in order to be noticed. Suppose, then, that ETCs transmit signals with a bandwidth of 0.1 Hz. (It makes little

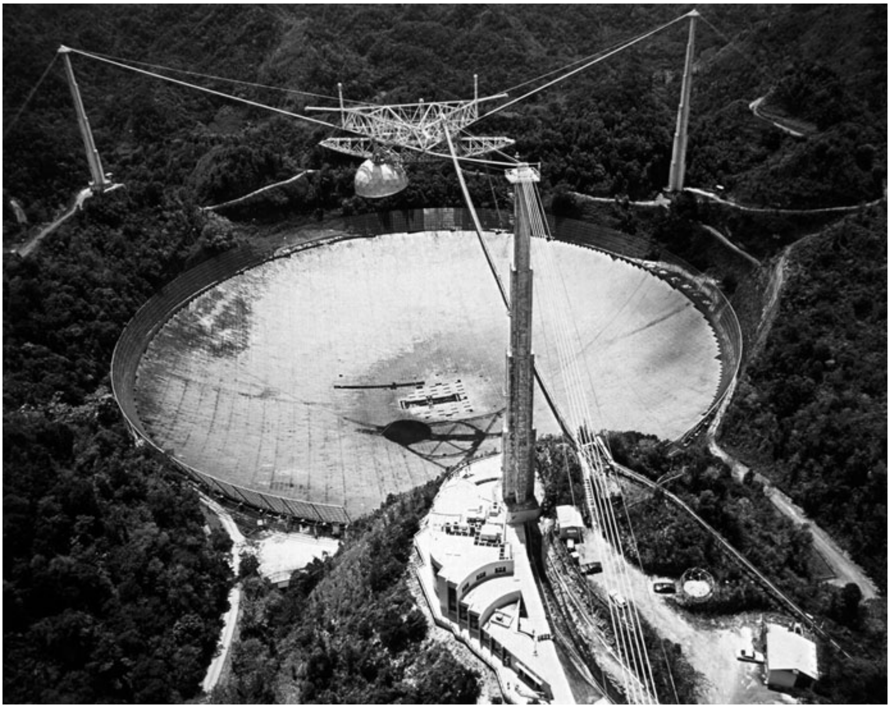


**Fig. 4.13** Since its construction in the early 1960s, in a karst sinkhole in Puerto Rico, the Arecibo Observatory has been home to the world's largest single-dish telescope: the dish is 305 m in diameter, 51 m deep, and covers an area of about 8 hectares. The Chinese Five-hundred-meter Aperture Spherical Telescope will eventually outperform Arecibo, but the Puerto Rican telescope remains a formidable instrument. In principle, it could detect an alien transmission from the other side of the Galaxy. (Credit: H. Schweiker/WIYN and NOAO/AURA/NSF)

sense to transmit over interstellar distances with a bandwidth less than 0.1 Hz, since electrons in interstellar clouds will tend to disperse the signal.) This means that we have a *huge* number of radio frequencies to comb through: there are lots of 0.1 Hz-sized channels in the region between 1 MHz and 300 GHz. Unless we narrow the search even further, or we get lucky, we could be searching for a long time.

Cocconi and Morrison pointed out that the Galaxy is noisy at frequencies less than about 1 GHz. It therefore makes little sense to send a signal at a frequency lower than 1 GHz because background noise would drown it. On the other hand, Earth's atmosphere is noisy at frequencies higher than about 30 GHz. Presumably a technologically advanced ETC would know that beings living underneath the blanket of a water-rich atmosphere would be unlikely to detect a signal at frequencies higher than 30 GHz because of atmospheric interference. In fact, the quietest region is between about 1 GHz and 10 GHz.

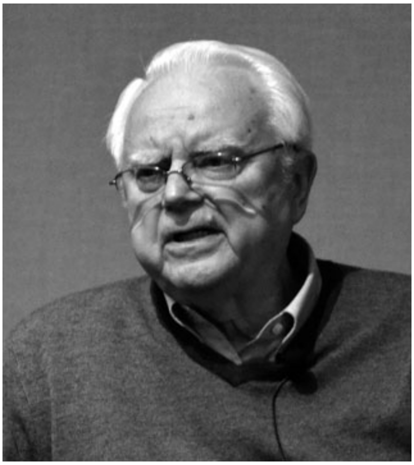


**Fig. 4.14** Frank Drake is a towering figure in the SETI field. In addition to the eponymous Drake equation, he is known for carrying out the first radio search for an ETC. (Credit: Raphael Perrino)

Cocconi and Morrison suggested that it makes most sense to search for radio signals in that region, where an artificial signal would really stand out.

They refined the frequency range even further. Cocconi and Morrison pointed out that clouds of neutral hydrogen—the simplest and most common element in the universe—strongly emit radiation at 1.42 GHz. Every scientifically competent observer in the universe will know about the hydrogen line. It makes sense to look there. There's one further twist: the hydroxyl radical radiates prominently at 1.64 GHz. Hydrogen, H, and hydroxyl, OH, together make up the compound water: HOH—or $H_2O$. Water, as far as we know, is absolutely necessary for the existence of life. Find water, and you have a chance of finding life. And since the region between 1.42 and 1.64 GHz is about the quietest part of the radio spectrum, it seems a logical place for a civilization to broadcast if it wants to attract attention. This band has been dubbed the *waterhole*. It's a beautiful name, conjuring up visions of many different species coming together at a life-giving source of water.

At about the same time that Cocconi and Morrison presented theoretical reasons for why we should listen in the long-wavelength region near the hydrogen line, Frank Drake was *doing* exactly that. Drake had built equipment to study this part of the radio spectrum for mainstream astronomical purposes, but he had an abiding interest in the possibility of extraterrestrial life. He used the radio telescope at Green Bank to listen to two stars—Tau Ceti and Epsilon Eridani—for signals. His Project Ozma was the first time mankind had searched for an ETC. Although the results were negative, Drake's observations—along with the Cocconi–Morrison paper—proved to be a watershed for SETI.

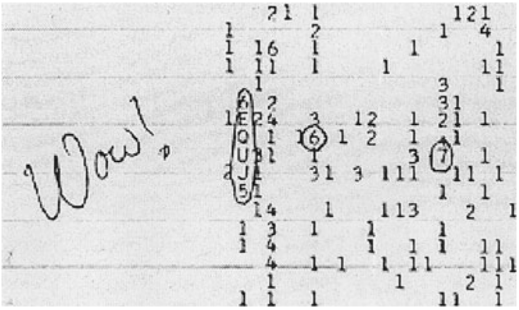



**Fig. 4.15** The famous "Wow" signal. The Ohio State University Big Ear Observatory scanned 50 channels and recorded the observations on a printout sheet. For each channel a list of letters and numbers appeared on the printout. In the Big Ear system, numerals 1 to 9 represented a signal level above background noise. For strong signals, letters were used (with Z being stronger than A). On the night of 15 August 1977, Jerry Ehman spotted the characters "6EQUJ5" on channel 2. This signal started from roughly background level, rose to level U, then decreased back to background level in 37 seconds. This was exactly what an extraterrestrial signal might look like; Ehman circled the characters and wrote "Wow!" in the margin. (Credit: Ohio State University Radio Observatory)

The situation now seems much more complicated than it did four decades ago for Drake, Cocconi and Morrison. The pioneers of radio astronomy had access to only a few spectral lines, so the choice of where to search seemed quite clear. Modern astronomers, however, are aware of tens of thousands of spectral lines emanating from more than 100 types of molecule in interstellar space. One can come up with good arguments for why we should study other frequencies.[177] Important examples include 22.2 GHz, which corresponds to a transition of the water molecule; simple multiples of the hydrogen-line frequency—twice the hydrogen-line frequency, $\pi$ times the hydrogen-line frequency, and so on; and there is a particularly attractive "natural" frequency for intergalactic communication, which I discuss in a later section. Although many authors maintain that the waterhole is the "natural" place to search for signals from within our Galaxy, we might find ourselves being forced to search through the whole window from 1 to 30 GHz.

In more than 50 years of listening, none of the radio searches has found an extraterrestrial signal that is clearly artificial in origin. That's not to say that no signals have been found. Drake himself detected a signal emanating from the general direction of Epsilon Eridani, just a few hours after the commencement of Project Ozma; further investigation, however, showed that the signal was clearly terrestrial in origin. Subsequent radio searches have detected many signals, some of them rather intriguing. The famous "Wow!" signal is typical of the best signals found so far. It was a powerful narrowband spike, with characteristics indicating that it almost certainly came from space, but when Big Ear listened again to that part of the sky the signal had gone. Several attempts to relocate the "Wow!" signal have failed. For example, searches with the Very Large Array enabled astronomers to investigate two hypotheses regarding the signal. First, perhaps it came from a weak yet steady transmission from an extraterrestrial civilization, a signal that momentarily increased in strength due to scintillation (similar to the twinkling of a star). Second, perhaps the signal was a powerful pulse, designed to attract attention to a much weaker continuous signal. Both possibilities seem to have been eliminated. Nothing interesting was found, down to a level that was 1000 times weaker than the original signal.

Another intriguing candidate is GCRT J1745-3009, a radio source that emitted five bursts of low frequency radiation in October 2002. Each burst was equally bright, lasted about ten minutes, and occurred every 77 minutes. A similar burst was observed a year later. Six months after that, astronomers observed a weaker burst. There has been nothing since. Could GCRT J1745-3009 and the "Wow!" signal be examples of extraterrestrial activity, communications not specifically aimed at us but that we happen to have detected? If so, it suggests a new search strategy:[178] build up a catalog of "interesting" radio transients and then apply statistical techniques to construct a probabilistic argument for the existence of extraterrestrial intelligence. However, distinguishing needles from hay is difficult. Although we don't know for certain what type of object GCRT J1745-3009 is, there are plenty of candidates for what it could be: a precessing pulsar, orbiting neutron stars, a radio-emitting white dwarf. . . And although the "Wow!" signal *might* have emanated from a distant civilization, a beam that happened to sweep across Earth's path one August night and then moved on, it seems much more likely that the signal came from some unknown terrestrial source.[179]

Despite the increasing sophistication of radio SETI, sorting through billions of channels in the hope of finding a signal remains a laborious task. Is there really no alternative to the microwave/radio part of the electromagnetic spectrum? It happens that there is.

At about the same time that Cocconi and Morrison were suggesting that we should listen for radio transmissions, other physicists were outlining the working principles of lasers. Early devices were feeble, but just as computing power has increased geometrically, so has the power of lasers. It now seems clear that a technologically advanced ETC could communicate its presence using laser pulses and, as Townes first suggested, they might even prefer this method over radio. Not only would a short pulse of laser light stand out even over interstellar distances, it would plainly be artificial. Furthermore, an ETC could send beacon signals to millions of stars each day. Perhaps we shouldn't be listening for radio signals alone; perhaps we should also be looking for signals in the visible spectrum.

**Projects in the Search for Extraterrestrial Intelligence** Since Project Ozma there have been dozens of SETI projects, most of which have searched the waterhole region. The projects have increased in sophistication over time.[180]

Project META (Million-channel Extra-Terrestrial Array), developed in 1985 by Paul Horowitz,[181] could study a million channels at once in the waterhole region. In 1990, META II started searching the southern sky, monitoring 8 million extremely narrow 0.05-Hz channels near the hydrogen line at 1.42 GHz, and also at twice this frequency, 2.84 GHz. In 1995, Horowitz initiated Project BETA (Billion-channel Extra-Terrestrial Array), which scans the waterhole region at a resolution of 0.5 Hz. From META to BETA in just ten years was significant progress!

Between February 1995 and March 2004, Project Phoenix was the world's most sensitive and comprehensive search for radio signals. It observed 800 stars within 200 light years of Earth, listening for signals between 1.2 GHz and 3 GHz in 1 Hz-wide channels. (At the end of the search the project's leader concluded that "we live in a quiet neighborhood".)

Project SERENDIP (Search for Extraterrestrial Radio Emissions from Nearby Developed Intelligent Populations) piggybacks on radio telescopes[182] being used for other astronomical purposes. The drawback with this approach is that there is no choice over where to listen; it can look for signals only where the telescope happens to be pointing. On the other hand, since it doesn't interfere with the normal functioning of the telescope, the project can be run continuously. The present incarnation of the project, which is SERENDIP V, started in earnest in 2009. It piggybacks on the Arecibo telescope and searches 128 million channels over a 200 MHz bandwidth centered on 1.42 GHz.

The Allen Telescope Array (ATA) is an ambitious project that aims to combine a wide field of view with a large frequency coverage. Rather than having a single large dish, the ATA plans to combine signals from a large number of small dishes. The project, which was made possible by a grant from Paul Allen, the co-founder of Microsoft, has great potential[183] for SETI studies—but its future is unclear. The first phase of the ATA became operational in 2007; it had 42 antennae, which was enough to commence observations. The long-term plan was for the Array to have 350 antennae, but in April 2011 the ATA was put into hibernation because of funding difficulties. Some short-term funding was found, and operations resumed in December of the same year; by finding a source of some further funding, the team have even been able to upgrade the dish receivers. Nevertheless, at the time of writing, the timescale for competing the original plan is far from certain.

One of the scientific goals of the ATA is to act as a stepping stone towards what will be one of the most significant telescopes of the first half of the twenty-first century: the Square Kilometer Array (SKA). As its name implies, the SKA will have an array of dishes whose combined collecting area will be about 1 km. The dish arrays will be based in Australia and South Africa, with the mission headquarters being based in the UK. If everything goes to plan, the SKA will begin full operation in 2024. It will be 50 times more sensitive than previous radio instruments (and capable, for example, of detecting an airport radar from tens of light years away). It will survey the sky thousands of times faster than was possible before. It will provide images of exquisitely high resolution. Although the SKA is an instrument for astronomy, it could also play a role[184] in SETI.

Optical SETI is not as advanced[185] as traditional radio SETI, but this is changing. For many years Stuart Kingsley used his COSETI (Columbus Optical SETI) Observatory to look for narrowband laser signals from a list of target stars; he demonstrated that the equipment required for such a search is relatively simple and within the range of the dedicated amateur astronomer. Professional SETI scientists eventually caught on, however, and are starting to

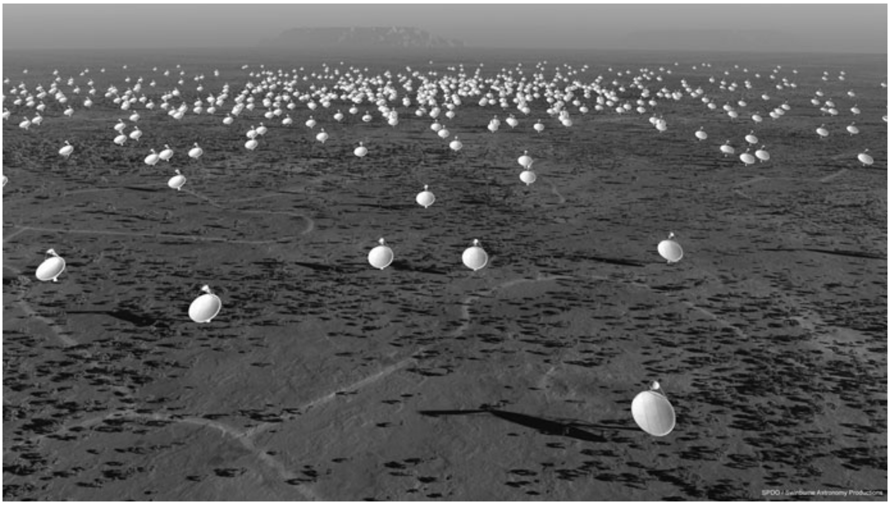


**Fig. 4.16** An artist's impression of how the 5 km diameter central core of the Australian Square Kilometer Array antennae will look. This incredible telescope will combine the signals received from thousands of small antennae spread over a distance of 3000 km in the southern hemisphere. The Square Kilometer Array has the potential to transform astronomy; will it also transform SETI? (Credit: SKA Project Development Office/Swinburne Astronomy Productions)

develop large-scale projects.[186] For example, Project SEVENDIP (Search for Extraterrestrial Visible Emissions from Nearby Developed Intelligent Populations) is an optical SETI initiative that complements the Project SERENDIP radio approach.

Even gamma-rays have been suggested as a communications channel for civilizations that are in contact over intergalactic distances. (John Ball once hypothesized[187] that gamma-ray bursts might be messages sent by ETCs. However, although the detailed origin of these events is still being debated, it is now clear that bursts are a natural phenomenon. We must employ Occam's razor once more: since we can explain bursts as natural phenomena, Ball's hypothesis is simply unnecessary.) The advantage of gamma-rays is that they offer the widest bandwidth in the EM spectrum: if you want to send your "Encyclopædia Galactica" over intergalactic distances then gamma-rays would be the way to go. However, gamma-rays are difficult to detect with ground-based receivers (fortunately for our health, Earth's atmosphere absorbs them) so it's unlikely that gamma-rays will play a direct role in SETI in the foreseeable future. Even if we don't look for gamma-ray encoded messages, however, gamma-ray bursts might still play a role in SETI: they could play the role of "synchronizers".[188] The idea here is that ETCs might decide to transmit signals at the occurrence of some particular event, and gamma-ray bursts—because of their easy detectability—would be a good choice for that synchronizing event.

In more than 50 years of searching—mainly in the radio, but occasionally in the infrared and increasingly in the visible—astronomers have detected no signal. To rephrase Fermi's question: where are the signals? The lack of signals means that we can now start to place limits on the number and type of ETCs in our neighborhood. Some authors claim this null result means we can rule out the presence of KII and KIII civilizations not only in our Galaxy, but beyond even our Local Group of galaxies.[189] This claim is perhaps overstated, since it rests on several assumptions that might not be valid. Nevertheless, taking a conservative viewpoint, we can probably rule out the existence of a KIII civilization anywhere in our Galaxy, a KII civilization in our particular part of the Galaxy and a KI civilization within a 100 light years or so: if they were there, we would surely have heard from them.

Billions of channels and—so far—nothing on.

## Solution 27 They Are Signaling but We Don't Know Where to Look

*We seek him here, we seek him there.*

Baroness Orczy, *The Scarlet Pimpernel*

Even if extraterrestrial civilizations are broadcasting radio signals or sending out laser pulses, and even if we are tuned to the correct channels, where should we point our telescopes? The sky is large and our resources are few. It would be tragic to train our telescopes on Canopus, say, if the civilization on Capella were trying to catch our attention.

We can employ two search strategies. A *targeted search* focuses upon individual nearby stars. It uses instruments of great sensitivity in the hope of detecting signals deliberately beamed toward us or leakage radiation that happens to pass our way. A *wide-sky survey* scans large areas of the celestial sphere and thus encompasses a myriad of stars. The sensitivity of a wide-sky survey is inferior to a targeted search.

The first modern SETI activity—Drake's Project Ozma—was a targeted search, but it targeted just two stars: Tau Ceti and Epsilon Eridani. Since then, astronomers have learned much about "habstars"—stars with the potential to possess habitable planets. Current thinking is that a habstar is likely to have a stable luminosity over long periods of time, it's likely to have a chemical make-up that allows Earth-like planets to form, and it's going to possess a zone in which at least one of those Earth-like planets can possess water in the liquid phase. So if you have a large catalog of stars and you want to prioritize them for SETI purposes, it makes sense to ignore cataclysmic variable stars, for example: their luminosity changes make them unlikely homes for life as we know it, much less technological civilization. There are various other "cuts" you could make to the catalog so that SETI resources would be best employed. Margaret Turnbull and Jill Tarter have taken this approach:[190] they analyzed the 118,218 stars of the *Hipparcos* catalogue—a listing published in 1997, from measurements taken by an ESA satellite devoted to measuring the parallaxes and proper motions of stars—and whittled them down to 17,129 habstars (three-quarters of which are within 140 pc of the Sun). If you're going to do a targeted search, you could do worse than concentrate on these stars.

The phenomenal success of NASA's *Kepler* mission has opened up other possibilities for targeted searches. Whereas the *Hipparcos* mission was specifically designed for astrometrical studies of stars, the *Kepler* mission was specifically designed to search for planets. *Kepler* "objects of interest" (KOI) are those stars that are known to possess planets and that are judged to be most amenable[191] to the presence of Earth-like life. Targeted searches of KOI have already taken place, and more will surely follow.

Some scientists suggest that we can further refine the list of targets if we put ourselves in the aliens' "shoes". If we assume that technologically advanced ETCs won't waste energy by broadcasting omnidirectionally, and will instead choose likely targets at which to send their signals (in the same way that we're discussing likely targets from which to listen), then it follows that we need only concern ourselves with habstars that have had a reasonable chance of having detected Earth. In other words, let's assume that advanced ETCs have their own (no doubt far superior) version of *Kepler*: if they saw Earth transiting the Sun then our system would be an "object of interest" to them and they might well choose to send signals our way. Well, it turns out that because the planets in our Solar System orbit in a plane that's tilted about 60° with respect to the disk of the Galaxy, Earth's existence would be more readily discovered[192] by ETCs in certain directions of the sky; perhaps we should concentrate on listening to stars in those directions?

Yet another suggestion is to look for straight-line alignments[193] of Earth, a habstar and a pulsar. We can assume that an ETC will classify our own Sun as a habstar and they will surely possess their own catalogues of pulsars; they will be able to generate their own list of straight-line planet–habstar–pulsar alignments. The idea here is that an ETC will pick one of the obvious frequencies for communication, and generate a pulsed transmission at a period determined by the pulsar in the alignment.

Could it be, however, that targeted searches are the wrong way to approach SETI? If we restrict our searches based on our understanding of habitability and our best guesses about the motivations of ETCs then we might be missing all sorts of possibilities. Instead of looking hard and long and deep at those planetary systems we believe might harbor life, perhaps we should instead use our telescopes to skim across the sky?

An analysis by Nathan Cohen and Robert Hohlfeld showed why we should play the numbers and look at as many stars as possible.[194] In Nature, we often find that objects with a large value of some property are rare, while objects with a smaller value of that property are common. Thus bright stars of spectral class O are few in number, while dim M-class stars are widespread. Strong radio sources such as quasars are rare, while weak radio sources like stellar coronas are common. Which are we more likely to detect: the rare "bright" objects or the common "dim" objects? It depends on the strength of the rare sources compared to the common sources. For example, quasars are *incredibly* strong radio emitters; it doesn't matter that they are at extreme distances because they far outshine the closer but weaker stellar sources. Thus, radio telescopes in the early 1960s could detect rare, distant quasars more readily than common, nearby sources. In the same way, even if advanced ETCs are incredibly rare, Cohen and Hohlfeld showed that we are more likely to detect their beacons than the weak signals from a host of ETCs not much more advanced than ourselves. The only way to avoid this conclusion is if the stars are teeming with intelligent life: if ETCs were common then targeted searches, such as the KOI search, would be likely to find one. Wide-sky surveys are therefore more likely to produce positive results; at the very least, when we pick targets for in-depth study we should try to ensure that the receiving beam contains galaxies or large clusters of stars behind the target.

So is this the explanation for the great silence? We haven't heard from ETCs because our focus has been too narrow? Well, no. There have been plenty of wide-sky surveys, and more are planned. Astronomers have so far surely not listened long enough, and perhaps they haven't listened at the right frequencies, but it's not correct to say they've ignored wide-sky surveys.

One of the most innovative scientific projects of recent years—an initiative that captured the enthusiasm of the general public and spawned a variety of "citizen science" efforts—is SETI@home. The project, which was initiated by David Gedye, was released to the public in 1999. Participants download a small client program for their home or work computer. The program usually works as a screensaver; in essence, when the user's computer is not engaged in "proper" work, the client program comes to life and begins calculations on a packet of data—known as a work unit—taken by the Arecibo radio telescope. It's important to note that the Arecibo data come from the telescope's normal science work; stars aren't being targeted for SETI purposes, but instead SETI scientists are analyzing whatever data comes in. Once the calculations are complete, the program sends the work unit back to SETI@home, where it's merged with all the other results from around the world, and a new work unit is downloaded. The effect is that volunteers have combined to make

SETI@home one of the world's most powerful computers. So not only can astronomers undertake wide-sky surveys, they have the computing resources to analyze the data in ways that Frank Drake could surely never have imaged when he first pointed a telescope to Tau Ceti in the hope of finding a signal.

I have one tiny feeling of unease with wide-sky surveys, and this harks back to the problem of the frequency at which we should listen. The surveys take in distant galaxies, and most surveys listen at or around the waterhole. But there's a better frequency than the waterhole for intergalactic (as opposed to interstellar) communication: 56.8 GHz.

**A Frequency for Intergalactic Communication** A "natural" frequency for intergalactic communication is represented by

$$f = \frac{k}{h} T_0 \approx 56.8 \text{ GHz}$$

where $T_0$ is the observed temperature of the cosmic background radiation, $k$ is the Boltzmann constant and $h$ is the Planck constant (it thus links the regimes of cosmology and quantum physics). This frequency was originally proposed in 1973 by Drake and Sagan, and independently by Gott in 1982.

The frequency 56.8 GHz is tied to the observed cosmic microwave background, so it's a universal frequency.[195] If an ETC in a distant high-redshift galaxy emitted a signal at a frequency related to the above, then it could be sure the signal might be received at any future time. The signal could potentially reach large numbers of galaxies. (We have another factor to consider here. On Earth it took about 4.5 billion years for a technological civilization to arise. If this is the time it takes other civilizations to arise then it's not worth looking at galaxies with redshifts much larger than 1. The light we now see from these distant galaxies set off when the universe was only about 4.5 billion years old; there would have been insufficient time for a KIII civilization to arise.) Unfortunately, Earth's atmosphere has a wide oxygen absorption band at 60 GHz, which means our radio telescopes can't carry out a search at 56.8 GHz. Observations at this frequency will have to be performed from space. In the meantime, perhaps a KIII civilization in a faraway galaxy is signaling us right now.

## Solution 28 The Signal is Already There in the Data

*I do not search; I find.*

Pablo Picasso

In more than half a century of searching, SETI projects have amassed a huge amount of data. Is it possible that somewhere in all that data there is a thumbprint of an ETC, a signal we haven't yet recognized?

A host of everyday terrestrial signals have the capacity to fool sensitive SETI detectors—military radar, mobile phones and communications satellites all generate potentially confusing radiation. The SETI astronomers are alert to these sources of interference, of course, and can usually identify them for what they are. But there remain a few tantalizing exceptions, detections that probably have a terrestrial source but nevertheless remain unidentified.

For example, between 1972–6, Zuckerman and Palmer examined more than 650 nearby Sunlike stars at a frequency of 1420 MHz and logged ten pulses that could, at a stretch, have been artificial. Between 1985–94 the META project logged several pulses[196] that could possibly have been artificial. We've already discussed the "Wow!" signal. The trouble is, whenever astronomers redirect their telescopes in the direction from whence the radio pulses came, they find nothing. The "signals" never repeat. These pulses might indeed have been the intermittent broadcasts of ETCs, a lighthouse beam that swept across Earth before moving away—or just an as-yet unidentified source of radio interference.

Another problem arises with the interpretation of the data from telescopes. We collect photons from gamma-ray bursts and explain their origin in terms of a cataclysmic fireball; we collect photons from stars with an infrared excess and deduce that the star is shrouded in dust; we find a thermal spectrum and infer that it comes from a blackbody. We could explain all these observations in terms of ETC activity. As we've already seen, Ball suggested that ETCs might communicate by exchanging bursts of gamma-rays; one of the signatures of a Dyson sphere is an infrared excess; the most efficient mode of communication that an ETC could employ would be indistinguishable from blackbody radiation to observers such as us, who are not privy to the system being used.

Ultimately, the difficulty is that we are stuck on a tiny rock, at the bottom of a thick atmosphere, trying to make sense of the universe by interpreting the occasional photons that our telescopes catch. This is a challenge and sometimes scientists might get it all wrong. But if we *can* explain observations in terms of natural phenomena then we need not postulate the existence of ETCs. Occam, again. So when we observe, for example, that the spectra of almost all galaxies show a redshift, it's enough to explain the observation in terms of the expansion of the universe—an explanation fantastic (and beautiful) enough in itself. We don't need to suppose, as did one SF story, that redshifts are the exhaust gases of alien craft fleeing from mankind.

We have to hope that advanced ETCs will make their signals unambiguous and clearly distinguishable from noise. We have to hope that their signals will be strong enough for us to detect. And we have to hope that they repeat their signals often. If they keep their end of the bargain then we have a chance of recording their signals. What a pity it would be, though, if we've already recorded their messages but don't recognize them as such.

## Solution 29 We Haven't Listened Long Enough

*Patience is bitter, but its fruit is sweet.*

Jean-Jacques Rousseau, *Emile*

In 1991, Drake wrote about his hopes for detecting signals from an ETC: "This discovery, which I fully expect to witness before the year 2000, will profoundly change the world."[197] More than two decades later, much has happened in SETI research. But the discovery has not been made. Was Drake simply being impatient? Perhaps the answer to the Fermi paradox is that ETCs are out there, communicating with each other and perhaps even attempting to communicate with us, but that we simply haven't listened long enough for our search to bear fruit.

This is the position most SETI enthusiasts take, and for good reason. Consider, for example, some of the difficulties that radio telescopes face when engaged in the search for extraterrestrial messages. First, the receiving beam area covers only a small patch of sky so there are millions of slightly different directions in which astronomers can point the telescope. Second, for each patch of sky there are billions of frequencies to check. Third, a signal might take the form of a burst rather than a continuous beacon—the telescope has to be on duty or it will miss the message. In short, to detect a radio signal from an ETC, a telescope must be pointing in the right direction at the right time and tuned to the right frequency. There are trillions of possible combinations of these parameters. If ETCs chose to chatter at each other using lasers rather than radio then it's extremely unlikely Earth would just happen to be in the path of any of the beams; billions of civilizations could be out there, talking to each other, and we wouldn't hear them. It seems not unreasonable, then, to say we haven't searched long enough. Perhaps we simply have to be patient.[198]

Some people, however, believe this to be an unsatisfactory resolution of the Fermi paradox. In a sense, the crux of the paradox is that we have been "waiting" for evidence of extraterrestrials for billions of years: they themselves, or their probes, or at least their signals, *should already be here*. Evidence of their existence, whatever form such evidence might take, should have been here long before humankind began to wonder if other species were out there. Spending a few more decades observing, with admittedly much more powerful technology, is missing the point.

Let's consider it another way. How many ETCs presently inhabit the Galaxy? Sagan and Drake suggested there might be $10^6$ ETCs in our Galaxy at or beyond our present level of technological development, so on average there should be an ETC within 300 light years of Earth. A more conservative estimate by Horowitz is that here might be $10^3$ advanced ETCs in our Galaxy so, if they are randomly distributed through space, there will be an ETC within 1000 light years of Earth. If these $10^3$ to $10^6$ civilizations are long-lived, perhaps billions of years old, then they must surely have a Clarke-level of technology—one that, to us, is indistinguishable from magic. Even if they don't want to travel, or find it impossible to travel, surely such civilizations could make it easy for us to see them or hear them. Why don't they? Alternatively, the civilizations might be short-lived. If there are 1000 civilizations now, and if the rate of formation of technological civilizations has been more or less constant over the history of the Galaxy, then about 10 billion civilizations will have lived and died in our Galaxy alone. Is it likely that not one ETC left any visible record of their hopes, their achievements, their existence? (If true, it's an almost unbearably sad thought.)

We return to the question: where are they—either their craft or their probes or their signals? We shouldn't have to *wait* for evidence of their existence—the evidence should already be here.

## Solution 30 They Are Signaling but We Aren't Receiving

*I really do not see the signal.*

Nelson, at the Battle of Copenhagen

Let's suppose that extraterrestrial civilizations are relatively common. Let's further suppose that ETCs are distributed evenly throughout the Galaxy. (Their spatial distribution is unlikely to be uniform because, as we shall see later, some regions of the Galaxy appear to possess habitable conditions and some regions are inimical to life. Nevertheless, it's a reasonable first approximation.) Finally, let's suppose that interstellar travel and colonization is impossible but that, as Drake suggests, ETCs spend some time in a communication phase: they broadcast to the stars for some time and then (for whatever reason) cease. This all seems relatively plausible and a simple analysis suggests that, in this scenario, we should eventually expect to detect a signal. However, Reginald Smith—an amateur scientist[199] with eclectic interests—adds one more assumption to the scenario: Smith assumes that there is a maximum distance within which a signal can be detected. Beyond that horizon the signal becomes so weak as to be undetectable. That additional assumption changes the analysis.

Smith considers a simple model in which an ETC broadcasts isotropically for its entire lifetime $L$. After time $L$ the broadcast stops, but the signals continue through space and are detectable up to a distance $D$ from the original planet. The signals will reach this maximum distance after a time $D/c$. (There are thus two possibilities. If $L > D/c$ then the signals will reach their maximum distance even while the civilization is broadcasting. If $L < D/c$ then the civilization will have stopped broadcasting before the signals reach their maximum distance. This has a bearing on the likelihood of two-way communication being established.) One can calculate the volume of space filled by the signal over its broadcasting period, and thus—for different densities of ETCs given by the Drake equation—the probability that a civilization will be within range. If there's a high probability that an ETC is in the volume of space occupied by the signal then contact is likely; if the probability of finding an ETC in that volume is small then contact is unlikely.

Of course, we don't know the values of the relevant numbers in this model: we can perhaps make estimates for $D$ but we've essentially no idea what might be a reasonable value for $L$. If we make an estimate of $D$ and $L$, however, we can estimate the minimum number ETCs that are required to make contact likely; it's just basic arithmetic. The extreme cases are perhaps what one would expect. If lifetimes or signal horizons are very short then there has to be many civilizations out there for contact to be probable; if lifetimes or signal horizons are very long then we would expect contact even if there were just one or two civilizations in the Galaxy. The intermediate case is the most interesting. If an average ETC remains in the communication phase for one millennium, and if the signal horizon is 1000 light years, then we would need at least a thousand ETCs in our region of the Galaxy in order for contact to be likely. In this scenario there might be 500 technologically advanced civilizations in our neighborhood, and the chances are we would never know.

Could a signal horizon explain the paradox, then? Aliens exist and they are broadcasting—we're just not receiving the signals? It's a thought. To my mind, however, there are too many ways of avoiding the conclusion for it to be the solution we're looking for.

## Solution 31 Everyone is Listening, No One is Transmitting

*Never the least stir made the listeners.*

Walter de la Mare, *The Listeners*

Although it's difficult to detect a signal from an unspecified planetary system among the Galaxy's hundreds of billions of stars, consider how much more difficult it must be to *send* a signal to the stars—at least, to send it with any expectation that someone or something will detect it. And even if a civilization possesses the technology to broadcast a detectable signal, would it *want* to? After all, there might be risks attached to broadcasting the fact of one's existence. Perhaps every civilization worries about the Fermi paradox and concludes that there must be a good reason why everyone else has decided to keep quiet; why be the first to break ranks? Could it be everyone is listening and no one is transmitting?[200]

In a sense, our civilization already transmits signals to the heavens. For several decades, our radio and TV transmitters have been leaking EM radiation into space. As I write, live broadcasts of the fall of the Berlin Wall could be sweeping across the star Vega; the soundtrack to *Saturday Night Fever* is now reaching Arcturus for the first time; cricket lovers in the Hamal system might soon receive word of Bradman's last Test innings. However, it's debatable whether leakage transmissions could be detected, even if ETCs are listening. Our transmitters direct their beams horizontally, to be picked up by individual antennae, so although some of the output is lost to space—a beam of EM radiation sweeps across space as Earth rotates on its axis and as it orbits the Sun—it's down to luck whether any of it intersects with a distant star. Furthermore, the high bandwidth and relatively low power of our transmitters mean even an Arecibo-size telescope would struggle to detect our broadcasts much beyond the orbit of Pluto. So unless ETCs are nearby, extremely lucky and have a level of receiving technology far beyond our own, they are unlikely to detect our inadvertent[201] (or even our advertent, if that's a word) transmissions. Besides, the amount of this leakage radiation is lessening as we increase our use of cable. (The radiation from powerful military radars, and the signals that astronomers bounce off Venus and Mars to map the topography of those planets, has more chance of being detected over interstellar distances. On the other hand, such radiation is highly focused; the beam is unlikely to intersect with an alien receiver.)

What if we *wanted* to be noticed? Rather than trusting to luck and hoping an ETC spots our TV (hoping too, perhaps, that they receive *Cheers* rather than *Charlie's Angels*), we'd need a means of transmitting a powerful narrowband signal. This is "active SETI", the flip side of traditional SETI: instead of pondering how best to *listen* we consider the practicalities of how to *transmit*. Furthermore, by studying the problem of how to transmit a signal over interstellar distances we can learn a lot that will help us to listen for signals. There have been some deliberate transmissions,[202] including a message to mark a book launch, a broadcast of *The Beatles*' song "Across the Universe" to commemorate NASA's 50th anniversary, and an ad for *Doritos* transmitted towards the star 47 Uma. These transmissions were essentially gimmicks, of course, but as we shall see below there have been some serious attempts to send a message to the cosmos.

Suppose we decide to use radio. The first problem is which transmission frequency to use. Well, the logic that makes us *listen* for signals at the waterhole suggests we should *transmit* somewhere in that region, although arguments could be made for several other frequencies. Once a transmission frequency has been decided upon—and let's assume for the moment that we should broadcast in the waterhole—what technology would be required?

Since we don't know in advance where an ETC might reside, the safest option is to transmit isotropically—with the same power in all directions. Unfortunately, isotropic transmission is costly. If we wanted to send a narrowband signal so it could be detected by a small antenna at a distance of 100 light years, say, then the power required by the transmitter would exceed the present total installed electricity-generating capacity of the world. And 100 light years barely extends beyond our immediate neighborhood. The farther away we want the signal to be received, the larger the power requirement of the transmitter. Isotropic transmission is thus an activity that we can't presently undertake. Even if we *could* build such a device, would we commit such a large level of resource to a project that has no guarantee of success?

If we are content to assume that ETCs will listen with an Arecibo-size telescope rather than a small antenna then the power requirements for the transmitter lessen. Indeed, if we knew the precise location of an Arecibo-type telescope on the other side of the Galaxy, then our own Arecibo could send it a signal. The problem is, we don't know in advance where to point the transmitter. An Arecibo-type dish, operating at a frequency in the waterhole region, has an extremely narrow beam. The old needle–haystack dictum doesn't begin to convey the improbability of sending a narrow beam that just happens to align with a large receiver somewhere in the depths of space.

Isotropic transmission guarantees that anyone with an ear can hear you, but it's exceedingly expensive. Beamed transmission is cheap, but it excludes most of your potential audience. These are the two extremes for a radio transmission strategy. We could make various trade-offs and compromises, of course, but interstellar radio transmission isn't easy for us. Could it be that ETCs decide to let others do the hard work of transmission? Perhaps the Galaxy is full of civilizations waiting for others to pay for the phone bill?

The economic argument doesn't entirely convince me. For humanity, at our present stage of development, it's certainly more cost-effective to listen[203] than to transmit. Technologically advanced ETCs, however, would presumably have more resource they could commit to transmission; what's ruinously expensive for us would be small change for a KIII civilization. Besides, they—and us—aren't limited to radio. Even with our present laser technology we can generate a pulse of light that, for a short duration, outshines the Sun. An advanced ETC would presumably have no trouble in generating a pulse that is, briefly, billions of times brighter than its star. Such pulses can be detected with a relatively small optical telescope connected to a charge-coupled device. Furthermore, over distances of a few thousand light years, the interstellar medium has relatively little effect on a visible light signal; unlike radio, optical communication isn't corrupted. In many ways, lasers are more effective than radio dishes for the purposes of interstellar transmission.

The drawback with optical-based communication is that the beam is *extremely* narrow. The transmitting civilization must therefore know the *precise* location of the receiving telescope. Firing randomly into the sky is futile; the laser beam is unlikely ever to be detected. The transmitting civilization must therefore draw up a list of target planetary systems along with precise and accurate values for the positions of those systems. Furthermore, stars aren't at rest. If an ETC sends a signal to where the star is *now*, then by the time the light reaches it the star will have moved on. So the transmitting civilization also needs accurate information about the velocities of the target stars. Gathering information about other planetary systems and the precise location and velocity of stars isn't easy, but neither is it impossible. The *Hipparcos* mission,[204] which observed the heavens between 1989 and 1993, obtained accurate positions and velocities for thousands of stars; the *Kepler* mission, which was launched in 2009, has discovered hundreds of planets; and the *Gaia* mission, which was launched in 2013, will determine the positions and velocities of about a billion stars and detect many more planets. If *we* can implement such missions, then a civilization that's much more advanced than ours should be able to use optical communication over interstellar distances—and radio signals too, if they choose.

Putting economic and technical arguments aside, perhaps we simply *shouldn't* transmit. Many respected thinkers are opposed[205] to active SETI, on the grounds that we don't understand the risks involved in letting advanced and potentially hostile civilizations know about our existence.

As mentioned above, humankind has already sent messages to the sky—not just leakage radiation but also intentional signals. Indeed, as long ago as 1820, the great mathematician Gauss was thinking about ways of signaling[206] our presence to intelligent beings on Mars. Gauss' ideas were impractical, but in 1974 Frank Drake, took the opportunity to use the inaugural ceremony of the refurbished Arecibo telescope to send a message at 2.38 GHz in the direction of M13. (This is a globular cluster containing about 300,000 stars, but unfortunately not of the type we expect to possess Earth-like planets.) The message lasted 3 minutes and was only 1679 bits long, but Drake managed to pack in a lot of information. When the signal reaches M13 in about 24,000 years time astronomers there will, if they can decode it, learn a surprising amount about us. Even if they can't decode it, the very detection of the signal would convey information; it would tell them an intelligent species was here and had advanced to the radio stage—the very fact of the signal carries a message. Several other communications[207] have been sent to the sky, notably from the Evpatoria Observatory in Crimea under the supervision of Alexander Zaitsev.

Both Drake and Zaitsev have been criticized for making these broadcasts without consulting widely. The transmissions represented Earth, yet no national governments were asked their opinion about the content of the signal.[208] Perhaps future large-scale transmissions from Earth will require a planetary government that can speak for us all. Perhaps an advanced ETC only transmits when it has achieved a level of unity such that its signals represent a consensus of their entire world. Is that why we are still waiting to hear from them—they listen not because of technical or economic issues but because of ethical difficulties?[209]

**The Prisoner's Dilemma** Two members of a criminal gang are arrested and imprisoned. Each prisoner is in solitary confinement with no means of speaking to or exchanging messages with the other. The police admit they don't have enough evidence to convict the pair on the principal charge. They plan to get both prisoners sentenced to a year in prison on a lesser charge. Simultaneously, the police offer each prisoner a bargain. Each prisoner is given the opportunity either to betray the other, by testifying that the other committed the crime, or to cooperate with the other by remaining silent. If both prisoners betray each other, each of them serves two years in prison. If A betrays B but B remains silent, A will be set free and B will serve three years in prison (and vice versa). If both prisoners remain silent, both of them will only serve one year in prison (on the lesser charge).

The purely rational, self-interested prisoner should always betray the other. But if both prisoners reason that way they end up with a worse result than if they cooperated.

Could something along these lines really be the resolution to the paradox? That no-one wants to be the first to break the silence? The situation appears to be rather like the famous prisoner's dilemma in game theory: each civilization can choose to passively search (betray) or actively search and broadcast (cooperate). If we really believe the costs of broadcasting are so high then we'll never see a beacon: we might as well close down the SETI program. But there are potential benefits as well as possible dangers to broadcasting, and game theory can be used to analyze this situation. One such game-theory analysis[210] of the problem suggests that the most effective approach for us is to adopt a mixed strategy: listen passively most of the time, but broadcast occasionally. If we adopt such a strategy we can expect other civilizations to do so. And it would only take one civilization to break the ice . . .

## Solution 32 They Have No Desire to Communicate

*Speech is great; but silence is greater.*

Thomas Carlyle, *Essays: Characteristics of Shakespeare*

One can think of any number of reasons why ETCs might want to initiate a conversation—curiosity, pride, loneliness . . . . But perhaps they simply don't feel like talking?

Resolutions of the paradox based on the idea that extraterrestrial civilizations keep themselves to themselves depend on making assumptions about the motives of alien beings. If such beings exist, they'll presumably be the product of eons of evolution in unearthly environments and so possess senses, drives and emotions different from our own. Or they might be artificial intelligences that have taken over from their biological creators. Or they could be of a form quite beyond our imagining. How can we pretend to understand the motives of intelligences so vastly different from ours? Perhaps we *can't* understand alien motives—but it's fun to speculate.

We've already touched on one reason why ETCs might choose to keep quiet: fear. If we broadcast to space we reveal our location and level of technology. If we think the neighbors might be aggressive or, worse still, berserkers then silence might be the best policy. We've no idea whether aliens would think this way, but many humans certainly do. Perhaps caution is a general trait[211] among advanced intelligences.

Others have suggested that the spirit of curiosity pervading humanity (and many other terrestrial species) might be lacking in intelligent extraterrestrials. Perhaps ETCs simply have no interest in exploring the universe or in communicating with other civilizations? One could argue that extraterrestrials lacking curiosity and a desire to learn how the universe works would never develop the technology to communicate over interstellar distances anyway; that any intelligent species we meet *must* have curiosity about the external world. But a glance through the history books shows how some human cultures have been isolationist. Perhaps a similar philosophy is common among ETCs?

A more common argument, usually advanced in a spirit of humility, is that ETCs would be so far beyond us intellectually they'd be indifferent to our existence. I heard one astronomer say that advanced civilizations "wouldn't want to communicate with us because we could teach them nothing; after all, we don't want to communicate with insects". Yet is that true? We're unlikely to be able to teach an advanced ETC anything about a "hard" science such as physics. But actually, physics is relatively easy: the universe is constructed from a small number of basic building blocks that interact in a small number of well defined ways. Advanced ETCs are therefore unlikely to spend much time discussing physics; they'll all have the same physical theories because they all inhabit the same universe. The areas of study that are *really* hard—in the sense of difficult to master—are subjects such as ethics, religion and art. Advanced ETCs wouldn't expect to learn anything interesting about electromagnetism from us, but they might be fascinated in trying to comprehend and understand how we see the universe—that would be a challenge worthy of them. Furthermore, it's not quite correct to say "we don't want to communicate with insects". At the very least we're interested in how insects might communicate between themselves: biologists have gone to great lengths to interpret signals that might be encoded in the dance of the honeybee; pheromone communication by ants has long been studied; the bioluminescence of fireflies, and the way in which these creatures employ light pulses in courtship dialogs, is fascinating. Such investigations are part of a wider study of animal communication and animal cognition. Indeed, the possibility of communication with "lower" species has intrigued humans for thousands of years. Just because *Homo sapiens* might be a "lower" species compared to others out there doesn't mean we are inherently uninteresting. (Besides, even if ETCs *are* indifferent to primitive life-forms such as us, it doesn't necessarily explain why we haven't seen them or their possible interactions with peers.)

Another commonly presented argument is that super-intelligent ETCs refrain from communication with us in order to protect us from developing an inferiority complex; they are waiting until we can provide worthwhile contributions to the conversations taking place in the Galactic Club.[212] As Drake pointed out, however, on an individual basis all of us routinely deal with minds superior to our own. As children we learn from our elder siblings, parents and teachers; as adults we learn from the great authors, scientists and philosophers of the past. It's no big deal: at worst, when we find we'll never write as well as Shakespeare or have insights as profound as Newton, we might be disappointed—but then we shrug and we do the best we can. At best, viewing the accomplishments of others serves to inspire us. Why should it be different for societies?[213]

It's possible to dream up many other reasons why intelligent extraterrestrials are reserved. Perhaps they quickly reach spiritual fulfillment on their home planet and see no need to search for others. Perhaps they believe only ethically advanced species should attempt to spread into space and are waiting for the day when their own species has progressed to such a stage. Perhaps the time delay involved in interstellar communication makes interaction with other species appear less attractive; it would have to be one-way. (But we engage in one-way communication all the time. Two-way communication with

Homer is impossible, but we continue to read him because his works are *interesting*.) Perhaps—and this is a depressing thought, given our lack of progress in spaceflight since the Apollo missions—they just can't be bothered.

The trouble with this and similar resolutions of the Fermi paradox is that they require an unlikely uniformity of motive. If the Galaxy is home to a million civilizations, as the optimists suggest, then perhaps *some* of them have no desire to communicate with others. But to explain the paradox requires *all* civilizations to behave that way. And surely that is unlikely. Indeed, the problem might be even more acute than this. Some authors have argued that, in order to develop the capacity for interstellar communication, a civilization might require a community of billions of minds. Humanity, for example, has over the centuries drawn on the genius of a vast number of minds to develop our present level of technology. If this holds true for other ETCs then there could be trillions of intelligent individuals out there—some of whom, if they belong to a KIII civilization, will have access to unimaginably powerful technology. In this case, these resolutions of the Fermi paradox demand a uniformity of motive not only *between* ETCs but also of individual members or groups *within* an ETC.

## Solution 33 They Develop a Different Mathematics

*The integers were created by God; all else is the work of man.*

Leopold Kronecker

One of the abiding mysteries of science is, as Wigner put it,[214] "the unreasonable effectiveness of mathematics." Why should mathematics describe Nature so well? Whatever the reason, we should be grateful we can comprehend the universe mathematically. It means we can assemble aircraft that remain aloft, build bridges that stay up and construct cars that almost drive themselves. Ultimately, all modern technology is dependent upon mathematics. (People have made aircraft, bridges and cars by trial and error, but I wouldn't want to use them.)

Many mathematicians, perhaps most of them, subscribe at least tacitly to Platonism. The Platonic philosophy holds that mathematics and mathematical laws exist in some sort of ideal form outside the realm of space and time. The work of a pure mathematician is therefore akin to that of a gold prospector; a mathematician searches for nuggets of pre-existing absolute mathematical truth. Mathematics is discovered, not invented.

Some mathematicians, though, take a strong anti-Platonic stance.[215] They claim mathematics is not some sort of idealized essence independent of human consciousness, but is rather the invention of human minds. Mathematics is a social phenomenon, part of human culture. The anti-Platonist contends that mathematical objects are *created* by us, according to the needs of daily life. Mathematics comes from our brains.

It's possible that evolution has hard-wired an "arithmetic module" into our brains. Neuroscientists even have a possible location for this module: the inferior parietal cortex, a comparatively poorly understood area of the brain. This is not to say that arithmetic is the whole of mathematics. Indeed, it's hardly anything compared to the vast edifice built by mathematicians, so perhaps other brain regions play important roles. (Psychologists have recorded the case of a man with a PhD in chemistry who was unable to solve basic problems in arithmetic—$5 \times 2$ was beyond him—yet could manipulate algebraic expressions such as simplifying $(x \times y)/(y \times x)$ into 1. Does this imply that arithmetic and algebra are processed by different regions of the brain?) Nevertheless, it's on the foundations of arithmetic that the worldwide community of mathematicians has constructed such a wonderful cathedral of abstract thought. And if it turns out we do possess an arithmetic processing unit in our heads then we shouldn't be too surprised. After all, our ancestors lived in a world of discrete objects in which the ability to recognize numbers of predators or numbers of prey would have been extremely advantageous. In fact, since the ability to make rapid judgments based upon the perceived numbers of objects is so clearly useful, we might expect animals to possess some sort of "number sense". There is indeed evidence that rats and raccoons, chickens and chimpanzees can make rudimentary numerical judgments.[216] So, although the ability to do integral calculus is not innate, one might argue that the foundations of arithmetic *are* innate. The integers aren't ideal Platonic forms existing independently of human consciousness; rather they are creations of our minds, artifacts of the way the brains of our ancestors interpreted the world around them.

**Counting or Subitizing?** It's unlikely that animals can count in the sense that we understand it. In those experiments claiming to demonstrate counting ability in animals it's difficult to rule out the possibility that animals are using much simpler cognitive processes. For example, when small numbers of objects are involved, the animals might be subitizing. We do the same ourselves: if we are presented with a plate containing 3 biscuits we *know* there are 3 biscuits, not 2 or 4, without having to count them. Subitizing is a perceptual process that works for numbers of objects up to about 6. The process works well for 3 objects, say, because there's only a limited number of ways of arranging them (variations on the patterns $\therefore$ and $\cdots$ pretty much exhaust the possibilities). There are so many different ways of arranging 23 objects, say, that no perceptual clue enables us to readily distinguish a group of 23 objects from 22 or 24 objects. Similarly, many animals can judge relative numerousness. They'll prefer a large quantity of food to a smaller quantity, for example. Again, though, the animals needn't be counting—after all, a pile of 500 birdseeds simply looks bigger than a pile of 300 birdseeds.

If this is correct, then a fascinating question arises: what would the mathematics of an ETC be like? The symbols they used would, of course, be different—but that's a trivial difference. Instead of superficial differences we want to know whether they develop the prime number theorem; the min–max theorem; the four-color theorem. If their evolutionary history were completely different from our own, then perhaps they wouldn't develop the theorems that humans have done. Why should they?[217] If they evolved in an environment in which variables changed continuously rather than discretely, then perhaps they wouldn't invent the concept of an integer. Or perhaps it's possible to develop a mathematical system based upon the concepts of shape and size, rather than number and set as humans have done. Or perhaps the brains of extraterrestrials are so much more powerful than ours they can run numerical simulations in their heads (or whatever serves as their heads). I personally find it difficult to imagine such alien mathematics, but that's almost certainly a deficiency in my imagination; it hardly proves that such different systems can't exist.[218]

None of this is to say that our own mathematics is *wrong*. Surely the relation $e^{\pi i} = -1$ is true and unavoidable anywhere in the universe. At least, I don't see how it could be otherwise. But other intelligences, which have a different evolutionary history, might simply fail to see the relevance of concepts such as e or $\pi$ or i or = or $-1$. Equally, they might possess concepts—important in their own environments—that we have failed to invent.

The point here is that human mathematics enabled us to develop aircraft and bridges and cars. Perhaps this type of mathematics is *required* for the development of technology. For a civilization to build radio transmitters capable of broadcasting over interstellar distances, it has to understand the inverse-square law and a host of other "terrestrial" mathematics. Could a solution to the Fermi paradox be that other civilizations develop other systems of mathematics—systems that are useful for the local conditions in which they find themselves but inapplicable for use in building interstellar communication or propulsion devices?

As a resolution to the paradox this suffers from the same difficulty as several others: even if it applies to *some* civilizations (and many would deny even that possibility), it surely can't apply to *all* civilizations. I can imagine a race of super-intelligent ocean-dwelling creatures developing a mathematical system without the Pythagorean theorem (would they even know about right angles?), but not *every* species will live in the ocean. Some will be land creatures, like us, and it seems reasonable to suppose that at least some of them would develop familiar mathematics.

One final thought. Mathematics, at its heart, is all about patterns. Even if mathematics itself is universal,[219] perhaps different intelligences appreciate and investigate different types of pattern. There could be nothing more interesting for mathematicians than to learn about different mathematical systems. To me, this provides yet one more reason why intelligent beings would choose to try and communicate.

## Solution 34 They Are Calling but We Don't Recognize the Signal

*The true mystery of the world is the visible, not the invisible.*

Oscar Wilde, *The Picture of Dorian Grey*

There's a more subtle argument relating to the previous section. Suppose that an advanced ETC does indeed create "different" mathematics, or—which is perhaps easier to accept and might amount to the same thing—suppose its mathematics was millions of years in advance of ours. If members of that civilization were transmitting to us right now, would we even recognize their transmissions as being artificial?

Much of the present SETI effort concentrates on the waterhole region and on multiples of the hydrogen line frequency (2, 3, $\pi$ times the frequency, and so on). Perhaps ETCs using a different mathematics see nothing special about such frequencies. The "obvious" frequencies for them might be something quite different. But that's a minor point. Let's suppose they broadcast in the waterhole region. Although one can imagine various options[220] for a "lingua cosmica", the hope of communicating with ETCs is usually predicated on finding signals containing simple mathematical patterns and developing from this a shared language. In other words, we hope to receive signals encoded in some math-based language such as Hogben's *Astraglossa*[221] or Freudenthal's [222] LINCOS. Is this hope reasonable?

If advanced ETCs *want* us to find them, then they could easily encode messages we'd recognize as artificial. A signal containing pulses distributed according to some obvious pattern—the first few prime numbers, say—would leave no doubt in our minds about its origin. We have to hope, then, that ETCs *want* to be noticed. But even if we detect a message, could we decode the contents? Consider the Voynich Manuscript.[223] In 1912, Wilfred Voynich, a collector, claimed to have bought a 234-page book from the Jesuit College at the Villa Mondragone, Frascati, in Italy. It presently resides in the Rare Book Room at the Library of Yale University, where its less romantic catalog name is MS 408. The book is about the size of a modern-day paperback, and is bound in a soft, ivory-colored vellum. Many Voynich scholars believe the book was written some time between the 13th century and 1608; radiocarbon

**Fig. 4.17** Folio 78r from the Voynich Manuscript. Note the strange text characters. At first glance they seem to be from a foreign language that you can't quite place; but detailed researches have shown that the characters belong to no known language. Are they characters in some private code? Is the whole thing simply a hoax? No one is sure

dating suggests that the vellum was made from animal that was alive in the early 15$^{th}$ century.[224] And this is pretty much everything we know about the manuscript: it was written in a language or code that no one has yet deciphered. It seems to contain information about herbalism and astrology, among other things, but no one is sure; it could, for example, be a medieval hoax[225] (or a more recent hoax by someone who had access to medieval parchment—quite possibly Voynich himself, who wouldn't have been the first rare book dealer to forge a manuscript).

Whatever information the Voynich Manuscript contains, we know it was written by a human being in the not too distant past. So the author had the same sensory inputs as the rest of us; a cultural background that is recognizable, if not identical to our own; human emotions that drove him (or her) in exactly the same way they drive us. And yet he (or she) wrote a book we can't decipher. If such a situation can occur with a member of our own species, what chance do we have of understanding a message from an ETC?

If aliens exist, they'll surely possess different sense organs, different emotions, different philosophies and, perhaps, even different mathematics. I suspect if astronomers ever detect a message from intelligent extraterrestrials, the dominant emotion mankind would feel—after an initial period of excitement and euphoria—would be frustration.[226] We might struggle for millennia without ever deciphering the meaning of the message. How maddening it would be, particularly in this world of instant access to information, if we could do nothing more than speculate about the contents of a communication from the stars!

Even if we failed to decipher a message, however, the detection of the message itself would give us hugely important information: we would know that we are not alone. So whether we can understand aliens is a quite separate question from whether they exist, and has no real bearing on the Fermi paradox. But there's another question: can we be sure that we would recognize a signal as being artificial? The efforts of SETI scientists are surely doomed if they can't distinguish between an artificial transmission and a natural emission.

One problem with signal recognition is the following: physicists have shown that if a message is sent electromagnetically and has been encoded for optimal efficiency, then an observer who is ignorant of the coding scheme will find the message indistinguishable from blackbody radiation.[227] Now, blackbody radiation is simply the radiation an object emits because it is hot. Astronomers detect blackbody radiation all the time, and of course they apply the simplest explanation to their observations—namely, that they are seeing some natural object that happens to be hot. But they *could* be observing messages that have been encoded for optimal efficiency! If advanced ETCs don't care whether primitive species know about them, and if they encode their communications to each other with optimal efficiency, then we could intercept their messages and remain unaware of their existence.

And the relevance of this to the Fermi paradox? Well, one scenario people have offered is that ETCs long ago accepted the impracticality of interstellar travel, made contact with each other through EM signals and, over the aeons, agreed to communicate with each other with messages encoded for optimal efficiency. They then lost interest in contacting younger civilizations such as our own, so we find the Galaxy filled with blackbody radiation. That *might* have happened, I guess—but it's another example of a "just-so" story that offers no testable prediction.

## Solution 35 Message in a Bottle

*Words fly, writings remain.*

Latin proverb

We know it's possible to transmit information over interstellar distances using electromagnetic radiation. Furthermore, using EM radiation for communication has the advantage that it travels straight and at the fastest possible speed—the speed of light. But, as we've seen, EM broadcasts are not without their problems. An omnidirectional broadcast covers lots of stars but is hugely expensive; a targeted message is cheaper but the size of the potential audience is reduced. Then there's the problem of requiring the audience to be listening at precisely the right time. If an ETC proudly broadcasts one of its greatest screenplays to the universe, but all a listener catches is "Forget it, Jake. It's Chinatown.", then the exercise has been pretty much wasted. Of course if the listener caught the tail-end of a long transmission it could deduce the presence of a transmitting civilization, and that in itself would be hugely important, but the same result could be obtained much more cheaply and reliably through a "We are here" beacon. If you want to transmit large amounts of information, to share with the community of intelligences your cultural highlights, scientific knowledge and accumulated wisdom, is radiation the best way to do it?

Questions involving the cheapest, most accurate, most effective way of transmitting information are perhaps best addressed by communications theorists—after all, these are the people who developed the theories that enable the internet and wifi to function effectively. Well, in 2004 Christopher Rose (a professor of electrical engineering at Rutgers University) and Gregory Wright (an astrophysicist) took a communications-theory approach to the question of interstellar communication. In particular, they dropped the requirement that information had to be sent at the fastest possible speed and then investigated how much energy would be required to send a message. Their result was startlingly clear but counter-intuitive[228] (at least, it was counter-intuitive to me): from an energy perspective it makes *much* more sense to write down a message on some material and hurl it into space than it does to broadcast the message. Sending a physical message has the added advantage that if the message is intercepted and decoded then the entirety of the information gets through, without the need for repetition: you can guarantee the recipient has a chance of watching the whole of *Chinatown* rather than risk them seeing just the final few seconds.

Rose and Wright thus make a compelling case that ETCs are more likely to send a message in a bottle than to broadcast radio. The starting point of their

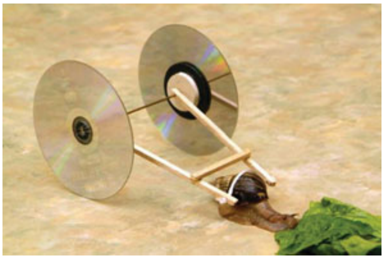


**Fig. 4.18** Ben-Bassat et al. (2005) showed how a Giant African Snail acting as a data transfer agent could exceed all known "last mile" communications technologies in terms of bit-per-second performance. Hitch a couple of information-laden DVDs to the shell, provide motivation in the form of a lettuce leaf, and "hey presto": blistering data transfer speeds. (Credit: Herbert Bishko)

argument is the following everyday insight: if you need to transfer extremely large amounts of data from one side of town to another then a reliable way of doing so is to fill a truck with blu-ray discs and drive to your destination. Furthermore, a simple physical exchange often has quicker rates of data transfer than radiation. Consider this example: the theoretical maximum information rate over optical fiber is about 100 terabits per second, but you can easily exceed that rate simply by pushing a box full of 5 TB hard disks across your desk.

We tend not to use "physical" techniques in modern communication networks; we usually want information to be transferred quickly and for most purposes, in everyday life, electromagnetic signaling is essentially instantaneous. But when we send a radio message to the stars those electromagnetic waves will be traveling for hundreds or thousands of years; in this case urgency seems to be much less of a factor, and we might reasonably tolerate a delay. Rose and Wright applied this thinking to the case of interstellar communication, and in that context asked: when is it better to write, and when is it better to radiate?

A key point in their argument is the observation that we are storing ever-larger amounts of data in ever-smaller volumes of material. When I was young, my music collection consisted of shelves of black plastic; when I moved to CDs, the physical volume taken up by my collection shrank even though the amount of music I owned had increased; my wife and I eventually combined our collections and now, on a flash drive that I can pop into my jeans pocket, I've access to more music than I'm ever likely to listen to (or frankly, since our tastes differ, than I would want to listen to). There seems to be no reason why this trend can't continue for many years to come, and eventually it should be possible to store all of the world's written and electronic libraries—say, $10^{20}$ bits of information—in a grain of material that weighs no more than a gramme. How much energy does it take to inscribe this information in a substrate of mass 1 g and then send it off into space at, say, one-thousandth of the speed of light? How much energy does it take to broadcast this number of bits? Rose and Wright worked out the numbers and made the comparison. They showed that there's always a break-even distance beyond which it's better to write. The break-even distance depends upon several factors, but on an astronomical scale it's never particularly large. Here then is their general conclusion: in terms of energy per bit, it's *overwhelmingly* more efficient to write than it is to radiate. Depending upon the details, such as the distance over which the message travels and the speed at which it does so, the difference in efficiency can be a factor as large as $10^{24}$.

One could make the reasonable objection that any information inscribed on a 1 g speck of material won't survive an interstellar journey: cosmic rays and other insults will degrade the message. Furthermore, over the millennia during which the message is in transit a destination star will drift in position—so there'll be a need for some sort of propulsion system to nudge the message back on track. And there'll be a need for a breaking system to be deployed once the "bottle" has reached it's destination. Fine. You could provide the 1 g of inscribed material with 10 tonnes of fuel and shielding and it would *still* be much more favorable than broadcasting the message. You could send out whole fleets of these information-rich grains and it would still make more sense than broadcasting the information, at least from the perspectives of energy use and message persistence.

Of course, we have only a hazy understanding of how economics works here on Earth so we can have absolutely no idea of how economics will work for an ETC. Perhaps energy-use-per-bit is not an important consideration for technologically advanced civilizations, and they can afford to take a money-no-object approach to the problem of interstellar communication. Perhaps they reason that there's no point sending out such small packages into the immensity of the universe since they are unlikely to be found or recognized as artificial—why go to all that bother if the bottles are never going to be opened? Perhaps. But it's difficult to see past their numbers. Rose and Wright published their calculations in a letter to *Nature* and in doing so presented a compelling alternative to an argument that appeared more than four decades earlier, also in letter to *Nature*—the Cocconi and Morrison paper that kickstarted the radio search for extraterrestrial intelligence.

So here is an answer to the Fermi paradox: we've been looking for a broadcast when we should have been looking for a message in a bottle. (We might argue, however, that if ETCs would find it so easy to send a physical message then why haven't we already seen one? Since it would be pointless to hurl a small bottle into space by itself, they would surely attaching a clear, noticeable, persistent beacon to the bottle. Where are the beacons?)

The Rose–Wright argument raises several interesting questions. For example, assuming that a message has reached the Solar System and some sort of beacon was indeed attached to the message, where precisely should we search? (This leads to a discussion similar to that given in Solution 5.) Since the RNA molecule can store a vast amount of information in a small mass, perhaps life itself is the message? (This brings us back to Crick's concept of directed panspermia, as discussed in Solution 6.) Above all, perhaps, should we change the focus of SETI away from radio and optical telescopes and onto direct searches for inscribed material? Even if the answer to that question were "yes", however, it would be difficult to sell it to the relevant stakeholders. Traditional SETI can piggyback on mainstream astronomical research: there'd be little cost attached to a search for transmissions from Vega, say, if radio telescopes were already pointing at that star. How would one get funding to search for an object of unknown form, that possesses unknown properties and is in an unknown location (Earth–Moon Lagrangian point? Asteroid Belt? Oort Cloud?) ... no agency would approve such a mission. So, like the drunk who looks for his lost keys at night under a lamppost, not because that's where he lost them but because that's where he can see, we might be condemned to search for electromagnetic broadcasts because we can.

## Solution 36 Oops ... Apocalypse!

> ... *we make guilty of our disasters the sun, the moon, and the stars;*
> *as if we were villains on necessity, fools by heavenly compulsion.*
>
> William Shakespeare, *King Lear*, Act I, Scene 2

One obvious, if gloomy, resolution of the Fermi paradox occurs if $L$—the factor in the Drake equation that describes the lifetime of the communicating phase of an ETC—is small. As we shall see later, there are a variety of ways in which Nature might kill off life. In the next three Solutions, however, I'd like to explore the idea that intelligent species might be the authors of their own doom. Let's look here at the possibility that curiosity can kill civilizations as well as cats.

### Particle Physics—A Dangerous Discipline?

Over the past century or so, physicists have been probing the fundamental nature of matter. They are interested in learning about the basic building blocks of the universe and the ways in which they can interact. The way they go about this is to smash particles together, at high energies, and then see what happens. It's a crude way of studying the physical world, but it's remarkably effective. However, some people believe that the high energies involved in such experiments could initiate some sort of global disaster. If particle physics experiments can indeed produce doomsday, and if an intelligent species' natural curiosity about the universe leads them inexorably to build such experiments, perhaps we have a solution to the Fermi paradox?

The concern that developments made by physicists might turn out to be catastrophic isn't new. In 1942, Teller wondered whether the high temperatures in a nuclear explosion might trigger a self-sustaining fire in Earth's atmosphere. Calculations by other physicists, including Fermi, put minds to rest: a nuclear fireball cools too quickly to set the atmosphere on fire. A more recent scare came in 1995 from Paul Dixon, a psychologist with only a hazy grasp of physics, who took to picketing the Tevatron particle accelerator at Fermilab with a home-made sign warning that Fermilab would become "home of the next supernova".[229] At the time the Tevatron was the highest energy particle collider in the world, and since then only CERN's Large Hadron Collider (LHC) has surpassed it. As the Tevatron increased the energy of its particle collisions, so Dixon's worry increased. He became convinced that collisions at the Tevatron might trigger the collapse of the quantum vacuum state.[230]

A vacuum is simply a state of least energy. According to some current cosmological theories, the early universe might briefly have become trapped in a metastable state: a "false" vacuum. The universe eventually underwent a phase transition into the present "true" vacuum, unleashing in the process a colossal amount of energy—it's similar to what happens when steam undergoes a phase transition to form liquid water. But what if our *present* vacuum is not the "true" vacuum? Rees and Hut published a paper in 1983 suggesting this could be the case.[231] If a more stable vacuum exists, then it's possible for a "jolt" to cause our universe to tunnel to the new vacuum—and the point at which the jolt occurs would see a destructive wave of energy spread outward at the speed of light. The very laws of physics would change in the wake of the wave of true vacuum.

Dixon needn't have worried unduly about this particular accelerator-induced apocalypse. As Rees and Hut themselves pointed out in their original paper, Nature has been using cosmic rays to carry out particle-physics experiments for billions of years—and at energies much higher than anything physicists can achieve.[232] If high-energy collisions made it possible for the universe to tunnel to the "true" vacuum—well, cosmic rays would have caused the tunneling to occur long ago. In case you're still worried I should point out that budgetary cuts, and competition from the LHC, caused the Tevatron to close in 2012; we've already dodged that particular bullet.

A related scare made the news in 1999. Various newspapers and magazines reported that experiments at a new facility, the Relativistic Heavy Ion Collider (RHIC) on Long Island, might trigger a catastrophe. Physicists built the RHIC in order to accelerate gold nuclei and other particles to high energies and then smash them together; conditions at the point of collision could replicate the conditions that were present in the universe just a microsecond after the Big Bang. These experiments, it was suggested, might destroy Earth. This particular flurry of concern began when someone calculated that the energies involved in the RHIC experiments would be enough to create a tiny black hole. The fear was that the black hole would tunnel down from Long Island to Earth's center and proceed to devour our planet. Fortunately, as more sensible calculations quickly showed, there's essentially no chance of this happening. To create the smallest black hole that can exist requires energies about 10 million billion times greater than the RHIC can generate. Even if the RHIC were able to produce a black hole it would be a puny object with only a fleeting existence. Such a tiny black hole would struggle to consume a proton, let alone Earth.

**Small Black Holes** The smallest possible black hole is about $10^{-35}$ m across—the so-called Planck length. Smaller structures get wiped out by quantum fluctuations. The creation of even the smallest black hole would require energies of around $10^{19}$ GeV, which is billions of times larger than RHIC energies. And even if it *could* create such an object, the black hole would evaporate on a timescale of $10^{-42}$ s. There are certainly more pressing things to worry about.

We can sleep soundly, safe in the knowledge that the RHIC won't produce a black hole. (The RHIC has been running since 2000 so even if we don't believe the theoreticians we can be pretty confident that any black-hole-related disasters would have occurred by now.) I think we can rest assured, too, that it won't destroy Earth through the production of *strangelets*—chunks of matter containing so-called strange quarks in addition to the usual arrangement of quarks.[233] So far no one has seen strangelets, but physicists wondered whether experiments at the RHIC might produce them. If strangelets were produced, then there's a risk they might react with nuclei of ordinary matter and convert them into strange matter—a chain reaction could then transmute the entire planet into strange matter, and Earth would end up as a dense sphere about 100 m across. However, having raised the possibility of catastrophe, physicists were quick to reassure everyone. Calculations show that strangelets are almost certainly unstable; even if they are stable, the RHIC wasn't operating at the energy where they were most likely to be created; and even if they were created at the RHIC, their positive charge would cause them to be screened from interactions by a surrounding electron cloud.[234] Once they've been raised, however, concerns tend not to die away. As I began writing this section I

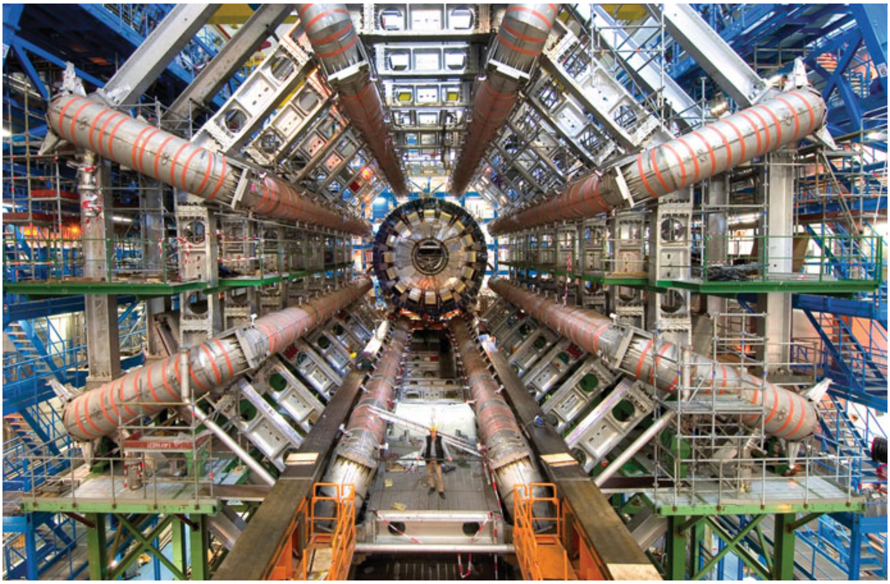


**Fig. 4.19** The LHC is quite possibly the most complex and impressive machine ever built. The ATLAS detector shown here is one of several detectors attached to the LHC; you can get a sense of its scale when you see the man standing in front of it. A 27-kilometer-long tunnel contains a ring of superconducting magnets that accelerate charged particles to incredible energies. It's certainly an amazing machine, but it won't destroy the universe. It won't even destroy Earth. (Credit: CERN)

came across a piece by two lawyers[235] suggesting that an upgrade to the RHIC was dangerous because it could now collide gold nuclei at *lower* energies than before—energies at which strangelet production is more likely. At the time of writing the RHIC has been doing groundbreaking physics in perfect safety for 14 years, but it seems some people will always consider it to be dangerous.

The LHC collides particles at energies in excess of the Tevatron, the RHIC or any other collider experiment ever built. It's perhaps not surprising, then, that just before it went live in 2008 there were lawsuits filed in various courts, protests at the European Commission and death threats aimed at members of the LHC team. All the worries expressed about previous collider experiments were trotted out before the LHC began operations, along with another possibility: that particle collisions might generate monopoles—hypothetical particles that are, in essence, isolated magnetic poles. Physicists at CERN patiently answered the worries,[236] but to my mind that was unnecessary. As Rees and Hut remarked when discussing the possibility that the Tevatron might collapse the vacuum, the LHC isn't doing anything that Nature isn't already doing daily and on a much larger scale. High-energy particles collide with nuclei in Earth's atmosphere all the time. Fortunately the lawsuits and the fearmongering got nowhere, and in 2012 the LHC made one of the great achievements of 21st century science when it discovered the Higgs boson.

The concept of an accelerator accident causing the destruction of a world via black hole or strangelet production (or the destruction of the entire universe, in the case of a vacuum collapse) is really a non-starter. The physics of these events isn't known perfectly—that's precisely why physicists carry out the research, after all—but they are well enough known for us to realize that the doom-merchants have it wrong in this case. We must look elsewhere for a resolution of the paradox.

## Macro-Engineering Goes Awry

As we shall discuss on page 174, and as I'm sure everyone already is aware, most climate scientists believe human activity is warming the planet. Since climate change can have potentially disastrous results (indeed, I present it as another solution to the paradox) there are serious proposals for a geoengineering approach to control the warming. One method would be to alter Earth's albedo and reflect more of the Sun's light; this could be done by using reflectors out in space or by releasing stratospheric aerosols. A different approach would be to reduce atmospheric carbon; one way of doing this would be to fertilize the ocean so that surface algae increase their uptake of carbon and, upon their death, take that carbon to the ocean floor. The problem with these projects is that, by definition, they need to act on a global scale. It's fair to say that we don't fully understand all the side-effects that such macro-engineering projects would entail. Could such projects put our civilization at risk? (Perhaps they could, but the situation might become so bad that we'll be forced to take the risk.)

There might be other projects that entail existential risk. In 2003, for example, the planetary scientist David Stevenson published a (tongue-in-cheek) proposal to investigate Earth's core.[237] The idea was to use nuclear weapons to open a crack in Earth's crust and then fill the crack with molten iron containing a probe. The iron would fall under gravity and eventually reach Earth's core, carrying the probe with it. In case anyone was dreaming of actually doing this, Ćirković and Cathcart pointed out that it would be a rather dangerous activity:[238] large deposits of carbon dioxide could be released, causing a global warming effect much greater than mankind is producing. Earth might end up like Venus.

Ćirković and Cathcart weren't suggesting that macro-engineering disasters are *the* solution to the Fermi paradox, but they did offer it as a partial solution. Perhaps engineering on a large scale poses existential risks?

## The Gray Goo Problem

The emerging field of nanotechnology appears to be the natural outcome of converging advances in many different subject areas.[239] The term refers to engineering that takes place at the nanoscale, a scale where the dimensions of objects are typically measured in nanometers (billionths of a meter). Since molecules are of this size, it also goes by the name of molecular engineering. Future nanotechnologists will have the ability to assemble custom-made molecules into large, complex systems; their capacity to create materials will be almost magical. Since this capacity appears to be so wonderful, and yet is presently beyond our abilities, several commentators are skeptical of nanotechnology. So it's worth emphasizing that there seems to be no *fundamental* reason why we won't be able to develop the technology. Nature herself is a "nangeneer": enzymes, for example, are nanotechnological devices that employ biochemical techniques to carry out their tasks. If Nature can do it, so can we. (It's also worth pointing out that the success or failure of nanotechnology will determine whether we ever develop Bracewell–von Neumann probes.)

One element of any future nanotechnology is likely to be the *nanorobot*—nanobot, for short. We should welcome the coming of the nanobots because they have the potential to improve health care:[240] they'll diagnose medical problems at an early stage, monitor the body's processes and target the delivery of drugs. They'll have applications in other fields too, including energy generation, pollution control and water treatment. It's an exciting technology. At present, nanobots are exceedingly primitive, but undoubtedly they'll improve. Looking a couple of decades into the future theoretical studies suggest we could construct nanobots from several types of material, with carbon-rich diamondoid materials perhaps being a good choice. Studies also suggest that one of the most useful types of nanobot will be a self-replicating machine.[241]

Alarm bells start to ring whenever self-replication is mentioned. The danger inherent in producing a self-replicating nanobot in the laboratory is clear upon answering the following question. What happens when such a nanobot escapes into the outside world? In order to replicate, a nanobot made of a carbon-rich diamondoid material would need a source of carbon. And the best source of carbon would be the Earth's surface biosphere: plants, animals, humans—living things in general. The swarms of nanobots (for soon there would be many copies of the original) would dismantle the molecules in living material and use the carbon to produce more copies of themselves. The surface biosphere would be converted from the rich, varied environment we see today into a sea of ravenous nanobots plus waste sludge. This is the gray goo problem.[242]

Exponential growth, as I've emphasized several times before, is an extremely powerful phenomenon. Freitas has shown that, under ideal conditions, a population of nanobots growing exponentially could convert Earth's entire surface biosphere in less than three hours![243] We can add this, then, to a depressing list of ways in which the lifetime of the communicating phase of an ETC might be shortened: a laboratory accident, involving the escape of a nanobot, turns biosphere to sludge.

This solution to the paradox, which has been seriously proposed, suffers the same problem as many other solutions: even if it can occur it's not convincing as a"universal" solution. Not every ETC will succumb to the gray goo.

The young boy in Woody Allen's *Annie Hall* gets depressed at the thought that the universe is going to die, since that'll be the end of everything. I'm becoming depressed writing this section, so to cheer up myself—and any young Woodys that might be reading—I think we need to ask whether the gray goo problem is even remotely likely to arise. As Asimov was fond of pointing out, when man invented the sword he also invented the hand guard so that one's fingers didn't slither down the blade when one thrust at an opponent. The engineers who develop nanotechnology are certain to develop sophisticated safeguards. Even if self-replicating nanobots were to escape or if they were released for malicious reasons, then steps could be taken to destroy them before catastrophe resulted. An exponentially growing population of nanobots would immediately be detectable by the waste heat it generated; defense measures could be deployed at once. A more realistic scenario, in which a population of nanobots grew slowly to evade detection, would take years to convert Earth's biomass into nanomass; that would provide plenty of time to deploy safeguards. Gray goo might be not such a difficult problem to overcome. It's simply one more risk that an advanced technological species has to live with.

## Solution 37 Ouch . . . Apocalypse!

*War does not determine who is right—only who is left.*

Bertrand Russell (attributed)

To more than a few scientists working during the Cold War it seemed quite certain that ETCs would discover the interesting properties of element 92 (known to us as uranium) and therefore learn how to construct nuclear weapons. For several scientists, then, the reason for a short lifetime (in other words, a small

**Fig. 4.20** The Castle Romeo test, a thermonuclear explosion on Bikini Atoll in 1954, yielded 11-mTon. The power of such bombs soon became even larger. (Credit: US Dept of Energy)

value for *L* in the Drake equation) was obvious: advanced civilizations inevitably annihilate themselves in a nuclear holocaust, as the human race was apparently on the verge of demonstrating.[244]

It hardly seems worth mentioning that, depending upon the severity of a nuclear war, the extinction of an intelligent species might follow. (One hesitates to use the word "intelligent" in this context, but the meaning is clear.) The world's arsenals still contain many thousands of nuclear weapons, and if they were ever used in large numbers then they'd certainly destroy *Homo sapiens*. Even a limited nuclear war might be ruinous for our species.[245]

Nevertheless, as many SF writers have demonstrated, it's possible to imagine scenarios in which members of a warring species survive a limited war and, over a period of thousands of years, recreate their civilization. One of the earliest post-apocalyptic novels, and certainly one of the best, is Miller's *A Canticle for Liebowitz*. Miller describes how a flicker of knowledge is preserved[246] by monks after a nuclear war has decimated the population. In *Canticle*, humankind eventually rediscovers the power of science and, a few millennia after the first nuclear holocaust, has "advanced" to the stage where the Bomb can be dropped once

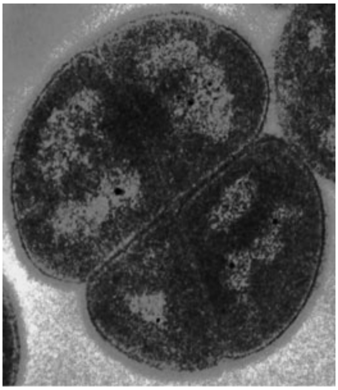


**Fig. 4.21** A transmission electron micrograph of the organism *Deinococcus radiodurans* growing on a nutrient agar plate. This bacterium can survive extremes of radiation and desiccation. (Credit: Michael Daly, Uniformed Services University, Bethesda)

again. Is the urge to war so deeply ingrained that a civilization learns nothing? Are civilizations somehow *compelled* to drop bombs as soon as they can? Unless that's the case, limited nuclear war can't provide an explanation of the paradox.

**Conan the Bacterium** Even a total, all-out, no-holds-barred nuclear war wouldn't destroy *all* life on a planet. Consider the organism *Deinococcus radiodurans*. Scientists first isolated it in 1956 from a can of ground beef; the beef had been radiation-sterilized, but the meat still spoiled. It turns out that *D. radiodurans* can survive an exposure to gamma-radiation of 1.5 million rads. For comparison, a dose of 1000 rads is usually enough to kill a man. Exposure to intense radiation blasts apart its DNA—but within a few hours the organism reforms its entire genome with seemingly no deleterious effects. This organism can withstand other extreme conditions, such as prolonged desiccation, which is why it's sometimes called "Conan the Bacterium". A nuclear war wouldn't unduly inconvenience Conan the Bacterium.

Not just bacteria would survive; various other organisms could survive a nuclear war. If intelligence is an inevitable outcome of evolution (this is contentious, as we shall see later, but is presumably the viewpoint of those who argue there are a million ETCs in the Galaxy) then the wait for intelligence to emerge after a nuclear holocaust wouldn't be endless: a few hundreds of millions of years, perhaps. This is an unimaginably vast reach of time on a human scale, but, again, it's not particularly significant when compared to the age of the Galaxy.

For seven decades our various governments have managed to negotiate the threats posed by nuclear weapons. We can only hope that this situation will continue and that ETCs are equally successful in averting a nuclear holocaust. Those civilizations that avoid the Scylla of nuclear war, however, must still navigate the Charybdis of biological and chemical warfare. Chemical weapons can be used to destabilize ecosystems while genetically engineered biological weapons can destroy food supplies or decimate populations directly; it's not just H-bombs that have the capacity to destroy civilization. Even more worrying is the possibility that biological weapons could be deployed by groups or even individuals. Could it be that some crazed individual, or simply someone with a grudge, might bring the world to an end? The mathematician Joshua Cooper offers bioterrorism[247] as a possible cause of the great silence.

Cooper argues that we can reasonably assume two things about any civilization that has reached a point where it is engaging in space travel. First, the civilization will consist of many individuals. (Why should we expect alien numbers to be legion? Well, Cooper argues that it's hugely expensive to escape the gravitational field of a planet large enough to possess an atmosphere. In humanity's case it wasn't until billions of individuals were available that sufficient resources could be brought to bear on the problem, and Cooper argues that the same will be true of ETCs. Technological and scientific developments might eventually mean these individuals consolidate into a smaller number of entities, but during the dawn of their space-faring epoch there must have been thousands of individuals working on the project supported by the collective funding of billions.) Second, their scientists will have mastered the chemistry of life—whatever version of life they happen to possess. (Why should we assume that they'll understand and have mastered their biochemistry? Well, Cooper again argues from a comparison with the development of human civilization. The same computational and technological abilities required for successful space travel are necessary for investigating the physical underpinnings of life. In our case, the development of space technology and biotechnology happened at essentially the same time; Cooper argues that, when one looks at these developments on a cosmic timescale, alien civilizations will learn to master their biology and their space environment at the same instant.) If one accepts these two points, a disturbing suggestion follows.

In recent decades, biochemistry has followed the same trajectory as computing: each year available power increases and cost drops. Watson and Crick published the structure of DNA in 1953; fifty years after their discovery, a typical student of biology would have been expected to sequence DNA in the undergraduate lab; two decades from now, the typical undergraduate will probably be expected to create an artificial organism from scratch. The Human Genome Project, which was founded formally in 1990, made a rough draft of the genome available in 2000 at a cost of a couple of billion pounds; when I published the first edition of this book, the cost of sequencing a human-sized genome had fallen to about sixty million pounds; the equivalent cost today is about four thousand pounds and soon cost simply won't be an issue. Progress in genomic sequencing is following a path that makes Moore's law seem tardy. It seems certain that within a few decades the billions of individuals here on Earth will have the capacity, if they so wish, to create artificial life. And any population of several billion will contain individuals who are insane, hateful or vengeful: we have plenty of such people among us right now. The difference is that in a few years those people will be able to create pathogens that target those possessing the "wrong" number of X chromosomes, "too high" a production of melanin, or otherwise "undesirable" genetic traits. The equal-opportunities misanthrope could unleash an engineered bioweapon to kill us all. So Cooper offers this as a possible resolution of the paradox: any spacefaring civilization will possess a knowledge of how to destroy its own type of life, and it's likely that one individual from the billions that make up the civilization will—for whatever reason—apply that knowledge.

Personally, I find Cooper's assumptions too grounded in anthropocentrism. Science fiction writers have imagined worlds in which alien civilizations *don't* consist of myriad individuals and in which science develops in a quite different way to the historical development here on Earth. Those writers might well be wrong, but in a field such as this their ideas surely have as much claim to legitimacy as Cooper's. I don't see that this is a plausible resolution to the Fermi paradox. However, Cooper's argument does contain a clear warning: unless we start to think about the threat now, and the countermeasures that we might be able to put in place, our own future is far from assured. At present, the fanatics and the lunatics of this world can engage only in localized murder; that situation might change. The possibility of devastation through nuclear annihilation might forever be beyond them; certainly the technology required to produce H-bombs will remain at the state level for many decades to come. The possibility of devastation through bioterror is much more likely.

## Solution 38 Heat Wave

*If you can't stand the heat, get out of the kitchen.*

Harry S. Truman

A necessary ingredient for the emergence of a technological civilization is—presumably—a planet that possesses a temperate climate for extended periods. Single-celled organisms are resilient, but it's difficult to see how complex multicellular life could thrive on a frigid planet, where water is locked up in the solid form. On the other hand, complex life would be scalded on a hot planet, where water is in the gaseous form; indeed, temperatures don't have to reach anywhere close to boiling point for complex life to suffer. What's required is that Goldilocks "just right" object, a planet on which water is free to flow and work its magic. Earth is clearly a Goldilocks planet in this respect, but it's not immediately obvious *why* Earth possesses the surface temperature it does. Clearly, Earth receives energy from the Sun and that warms our planet—but then why isn't the Moon at the same temperature as Earth? After all, both Earth and Moon are the same distance from the Sun. (The temperature on the surface of the Moon varies significantly, depending on whether it's night or day. When the Sun is overhead, the Moon's surface temperature can exceed 100 °C; once the Sun has set, though, the temperature can drop below −150 °C. This just emphasizes the difference between Earth and its satellite.)

We have the atmosphere to thank for Earth's temperate nature. Earth receives energy from the Sun at a variety of electromagnetic wavelengths—ultraviolet, visible and near-infrared. Almost all of this energy passes straight through the atmosphere and about half of it is absorbed at Earth's surface, which subsequently becomes warm. Any warm surface will radiate simply because it's warm, and the peak wavelength of the radiation depends upon the surface's temperature. In the case of Earth, most of the thermal radiation it emits is in the far infrared region. Here's the wonderful thing: the chemical make-up of Earth's atmosphere is such that it's almost transparent to the incoming short-wavelength ultraviolet, visible and near-infrared radiation but it's almost opaque to the outgoing longer-wavelength far-infrared radiation. The radiation emitted by Earth's surface is absorbed by the atmosphere, which then re-radiates it—and the radiation that's emitted downwards is absorbed by Earth's surface. Our atmosphere thus keeps us warm. Not only that, the atmosphere has a moderating effect; winds carry heat from the equator to the poles and from the day side of the planet to the night side. Without an atmosphere, there would surely be no life on Earth.

This atmospheric trapping of solar radiation is known as the greenhouse effect and it was first quantified by Svante Arrhenius (he of panspermia fame) as long ago as 1896. The underlying idea dates back seventy years before that. So it's not a new suggestion that the so-called atmospheric greenhouse gases—primarily water vapor, carbon dioxide, methane and ozone—play a crucial role in determining Earth's climate. And given the fundamental importance of climate for life, you'd think it would be foolish in the extreme for a civilization to mess with atmospheric greenhouse gases. But that's precisely what humanity is doing.

Since about 1850, global energy use has rocketed. Those of us living in the developed nations have access to a myriad of technologies that make our lives more comfortable than those of our Victorian ancestors: we have access to cars, air travel, the Internet, powerful illumination, central heating, mobile phones, exotic foodstuffs, clean water on tap . . . but all those taken-for-granted conveniences of modern life require energy. Lots and lots of energy. Since the Industrial Revolution, humanity's insatiable demand for energy has been met mainly by extracting fossil fuels—coal, petroleum, natural gas—and burning them. Had humanity not discovered vast reserves of these energy-dense materials then our present civilization would likely be quite different: scientific and technological innovation would no doubt have continued, but progress would surely have been much slower and our choices would have been constrained. Our current level of civilization, which permits us to at least contemplate the exploration of space, requires lots of cheap energy—and for decades to come it's likely that cheap energy will be provided by the combustion of fossil fuels.

There are at least two aspects of this situation that are relevant to the Fermi paradox (if we make the inevitable anthropocentric assumption that all ETCs will need to go through a phase of meeting energy needs by burning fossil fuels). First, fossil fuels constitute a finite resource. An inexorable increase in energy demand will eventually exhaust fuel reserves. If our access to fossil fuels were to end abruptly, right now, the consequences would be unthinkable. Our civilization would collapse. One suggested resolution of the Fermi paradox, then, is that the inevitable depletion of fossil fuels means that civilizations never make it into deep space. They collapse before they can colonize a world that contains more energy resources. Personally, in this case I'm an optimist. We have a few decades yet before the crunch comes and I'm sure that, before then, politicians will wake up to the danger: they'll pour resources into the problems of energy generation and some other fuel will allow us to maintain our lavish standards of living. Second, and more insidiously, when we burn fossil fuels we release greenhouse gases. The large-scale burning of fossil fuels can change the amount of greenhouse gases in the atmosphere and that in turn can change the climate.

Fossil fuels were formed through the decomposition of buried dead organisms. Oil and natural gas comes from organisms that lived in rivers or oceans and were buried under layers of silt; over millions of years this organic material was "cooked" under pressure to create the deposits we tap into today. Coal formed in a similar way, except that the original material was trees, ferns and plants. Since fossil fuels came from organic material, they contain carbon—anthracite coal, for example, is almost pure carbon—and so the burning of these fuels releases this element. The carbon that's released combines readily with oxygen to form the greenhouse gas carbon dioxide. Since the start of the Industrial Revolution, a century and a half ago, humankind has been releasing carbon that took tens of millions of years to store. Not surprisingly, the levels of atmospheric carbon dioxide have been rising steadily.

The best data on levels of atmospheric carbon dioxide come from measurements taken two miles above sea level at the Mauna Loa Observatory in Hawaii. Charles Keeling began measuring[248] carbon dioxide levels there in 1958, and observations continue. Keeling's observations were exquisite: the

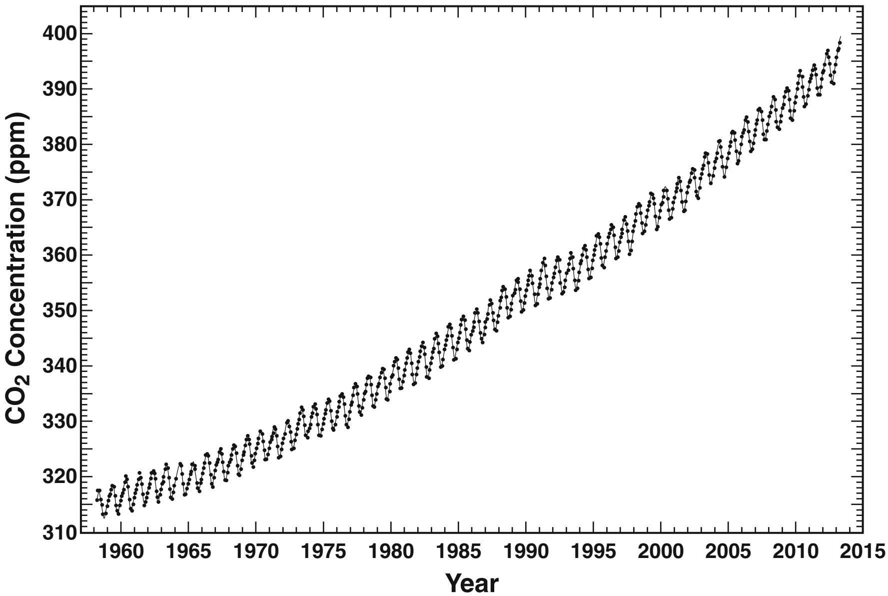



**Fig. 4.22** The monthly average concentration of carbon dioxide as measured at the Mauna Loa Observatory in Hawaii. The data come from a long-running program by the Scripps Institution of Oceanography. On 9 May 2013, the concentration reached 400 parts per million; a comparison of carbon dioxide trapped in air bubbles taken from Antarctic ice cores suggest that atmospheric carbon dioxide is now at its highest level for at least 800,000 years. (Credit: NOAA)

Keeling curve, shown in figure 4.22, is one of the most beautiful in all of science. At least, it would be beautiful if it weren't so chilling. The Keeling curve shows Earth "breathing": starting in spring, the growth of plants and trees in the large land masses of the northern hemisphere causes the level of carbon dioxide to drop; as plant growth stops later in the year, the level of carbon dioxide increases. On top of this seasonal variation, however, we see a yearly increase in the total amount of carbon dioxide in the atmosphere. Various lines of evidence demonstrate that this increase comes from the burning of fossil fuels: our demand for cheap energy means we add about 11 billion tonnes of carbon dioxide to the atmosphere each year.

One might reasonably expect that pumping billions of tons of a greenhouse gas into the atmosphere would cause Earth to warm. And there is indeed clear evidence for global warming: the average surface temperature[249] has increased by about 0.85 °C since 1880. Global warming can in turn cause climate patterns to change. (A thorough discussion of climate change would involves more factors than just average temperature, but in this context it's appropriate to focus on global warming.) Although a vocal minority of commentators deny the existence of any link between human activity and global warming, the scientific community is speaking with compelling clarity: human activities over the past century or two have released large amounts of greenhouse gases into the atmosphere and this has caused the Earth to warm. There are two outstanding questions. How much will temperatures rise in the coming decades? And how will rising global temperatures affect humanity?

The worst-case scenario for global warming would be a runaway greenhouse effect. A runaway effect can occur when there's positive feedback in a system. In this case, the fear is that increasing temperatures cause more water vapor to be released into the atmosphere which, because water vapor is a greenhouse gas, causes an increase in global temperature that in turn causes more water vapor to be released which . . . the end result is that the oceans boil away. The temperature only stabilizes when the surface temperature reaches about 1400 K and Earth starts radiating in the near infrared, at wavelengths for which water vapor isn't a greenhouse gas. A runaway greenhouse would, of course, spell the end of complex life on Earth. Fortunately, the latest research suggests[250] that the burning of fossil fuels almost certainly *won't* trigger a runaway effect. One bullet dodged. (A runaway greenhouse *is* Earth's probable long-term fate—the Sun is getting hotter as it ages and eventually it will trigger some sort of runaway process—but we have a billion years or so before we need worry.)

Although an anthropogenic runaway greenhouse effect is unlikely, it seems inevitable that over the next century we'll experience a human-induced increase in average global temperature. Would that be such a bad thing? After all, one can argue that civilization wouldn't have arisen had the last Ice Age continued. If you need to feed billions of mouths, warm is certainly good; perhaps warmer is better? Well, if the temperature rise turns out to be at the lower end of the predictions it mightn't be so bad. There are likely to be winners and losers. Some low-lying countries will disappear, and it turns out that the poorest countries—those most lacking the resources to cope with the effects of climate change—are most likely to be hit hardest. Overall, though, if the temperature rise is limited and it takes place gradually then humans will cope. If the temperature rise turns out to be at the upper end of the predictions, however, then there can only be losers. It's difficult to imagine our civilization continuing in a world six degrees warmer than it is now.

It seems we are stuck between the proverbial rock and a hard place. We can't turn off the taps because our civilization would collapse without the cheap energy provided by fossil fuels. But if we continue to burn carbon we risk a level of climate change that will cause our civilization to collapse.

So—is this a resolution to the Fermi paradox? That for civilization to get going requires the cheap energy provided by burning fossil fuels, but that the very act of burning such fuels causes the end of civilization? Well, disregarding the objection that this argument is anthropocentric, we can hope that humankind will find a way to navigate between Scylla and Charybdis. Soon, perhaps, people in the developed nations will understand that it's cheaper to develop alternative energy supplies than it is to rebuild after climate-change-induced floods and fires and typhoons. In the worst case, we might have to engage in some form of geoengineering to keep ourselves cool. However it's done, we at least have a chance of mitigating the effects of climate change. And if we can do it, so can others.

## Solution 39 Apocalypse When?

*No man has learned anything rightly until he knows that every day is Doomsday.*
Ralph Waldo Emerson, *Works and Days*

Humankind could destroy itself in a variety of ways. In addition to the calamities discussed in previous Solutions one could add genetic deterioration, overstabilization, epidemics and a dozen other problems. And this is without mentioning the many external factors that threaten us, such as meteor impact, solar variability and gamma-ray bursts. It barely seems worth getting out of bed in the morning. Surely, though, an intelligent species such as *Homo sapiens* will learn how to navigate these problems? Remarkably, a line of reasoning called the delta *t* argument suggests not.

In 1969, while he was still a student, Richard Gott visited the Berlin Wall. He was on vacation in Europe at the time, and his visit to the Wall was one of several excursions; he'd seen the 4000-year-old Stonehenge, for example, and was suitably impressed. As he looked at the Wall, he wondered whether this product of the Cold War would stand as long as Stonehenge. A politician skilled in the nuances of Cold War diplomacy and knowledgeable about the relative economic and military strength of the opposing sides might have made an informed estimate (which, judging by the track record of politicians, would have been wrong). Gott had no such special expertise, but he reasoned in the following way.[251]

First, he was there at a random moment of the Wall's existence. He wasn't there to see the construction of the Wall (which happened in 1961), nor was he there to see the demolition of the Wall (which we now know happened in 1989); he was simply there on vacation. Therefore, he continued, there was a 50:50 chance he was looking at the Wall during the middle two quarters of its lifespan. If he was there at the *beginning* of this interval, then the Wall must have existed for 1/4 of its lifespan, and 3/4 of its lifespan remained. In other

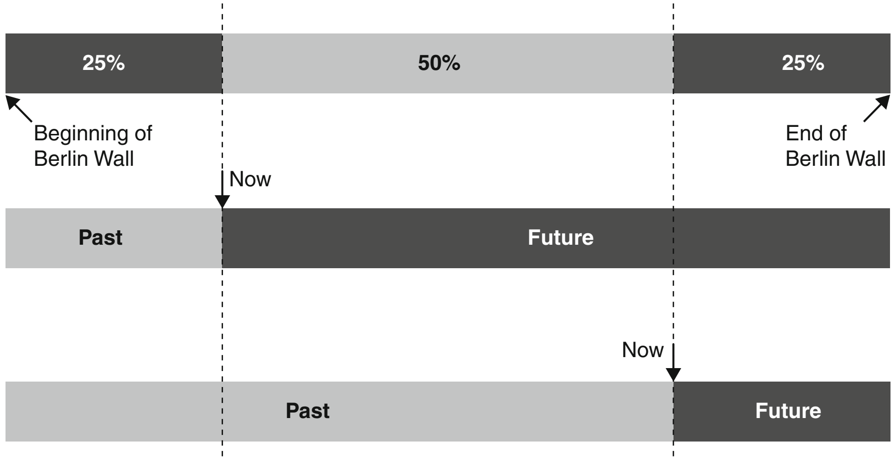



**Fig. 4.23** An illustration of Gott's prediction that the Berlin Wall would last for between another 2 years 8 months to 24 years after he first saw it in 1969

words, the Wall would last 3 times as long as it already had existed. If he was there at the *end* of this interval, then the Wall must have existed for 3/4 of its lifespan, and only 1/4 was left. In other words, the Wall would last only 1/3 as long as it already had existed. The Wall was 8 years old when Gott saw it. He therefore predicted, in the summer of 1969, that there was a 50% chance of the Wall lasting a further 22/3 to 24 years (8 × 1/3 years to 8 × 3 years). As anyone who saw the dramatic television pictures will remember, the Wall came down 20 years after his visit—within the range of his prediction.

Gott says the argument he used to estimate the lifetime of the Berlin Wall can be applied to almost anything. If there's nothing special about your observation of a thing, then, in the absence of relevant knowledge, that thing has a 50% chance of lasting between 1/3 to 3 times its present age.

In physics, standard practice is to talk about predictions that have a 95% chance of being correct rather than a 50% chance. Gott's argument remains the same, but there's a slight change in the numbers: if there's nothing special about your observation of an entity, then that entity has a 95% chance of lasting between 1/39 to 39 times its present age. It's important when applying Gott's rule to remember that the observation must not have any particular significance. Imagine you've been invited to a wedding and, at the reception, you start chatting to a couple you've never seen before. If they tell you they've been happily married for ten months, then you can inform them their marriage has a 95% chance of lasting between just over a week to 32 1/2 years. On the other hand, you can predict nothing about how long bride and groom will be

**Fig. 4.24** A hole in the Wall. There's a remarkable argument that links the lifespan of the Berlin Wall to the lifespan of our species! (Credit: Frederik Ramm)

together: you are at the wedding precisely in order to observe the beginning of the marriage. (The flaw in applying the rule to funerals should be obvious.)

Using the delta *t* argument to estimate the longevity of concrete walls and human relationships is amusing, but we can use it to estimate something more serious: the future longevity of *Homo sapiens*. Our species is about 175,000 years old. Applying Gott's rule, we find there's a 95% chance that the future lifetime of our species is between about 4500 years and 6.8 million years. That would make the total longevity of our species somewhere between about 0.18 and 7 million years. (Compare this with the average longevity for mammalian species, which is about 2 million years. Our closest relatives, *Homo neanderthalensis*, survived for maybe 200,000 years; *Homo erectus*, another Hominid species and possibly one of our direct ancestors, lasted for 1.4 million years. So Gott's estimate is certainly in the right ballpark for species longevity.) The argument says nothing about *how* we are going to meet our end; it could be by one or more of the methods discussed elsewhere, or by something quite different. The argument simply says it is highly likely our species will perish some time between 4500 years and 6.8 million years from now.

If this is the first time you've met Gott's argument, then you might well think (as I confess I did) that it's nonsense. However, try to pinpoint exactly where the logic is faulty—that's far from easy. The "obvious" objections to the argument have been robustly refuted. Before examining possible objections to Gott's line of reasoning, and looking at the implications of the delta *t* argument for the Fermi paradox, it's worth considering a slightly different version of the same idea.

Imagine you are a contestant on a new TV game show. The rules of the game are simple. Two identical urns are put in front of you and the host tells you one urn contains 10 balls and the other contains 10 million balls. (The balls are small.) The balls in each urn are numbered sequentially (1, 2, 3, . . ., 10 in one urn; 1, 2, 3, . . ., 10,000,000 in the other). You take a ball at random from the right urn and find the ball is number 7, say. The point of the game is for you to bet whether the right urn contains 10 balls or 10 million. The odds are not 50:50. Clearly, it's far more likely that a single-digit ball comes from the urn with 10 balls than from the urn with 10 million. Surely, you'd bet accordingly.

Now, instead of two urns consider two possible sets of the human race, and instead of numbered balls consider individual human beings numbered according to their date of birth (so Adam is 1, Eve is 2, Cain is 3, and so on). If one of these sets corresponds to the real human race, then my personal number will be about 70 billion—as will be any of the readers of this book, since of the order of 70 billion people have lived since the beginning of our species. Now use the same argument as we did with the urns: it's much more likely you'll have a rank of 70 billion if the total number of humans who will ever live is, say, 100 billion than it is if the total number is 100 trillion. If you were forced to bet, you'd have to say the likelihood is only a few more tens of billions of people will live. (A few tens of billions of people sounds a lot, but at the present rate we add a billion people to Earth's population every decade.)

The delta *t* argument is an extension of the Copernican principle. The traditional Copernican principle says we aren't located at a special point in space; Gott argues we aren't located at a special point in time. An intelligent observer, such as you, Gentle Reader, should consider yourself to be picked at random from the set of all intelligent observers (past, present and future), any one of whom you could have been. If you believe mankind will survive into the indefinite future, colonize the Galaxy, and produce 100 trillion human beings, you have to ask yourself: why is it that I'm lucky enough to be among the first 0.07% of people who will ever live?

Gott uses the same type of probabilistic argument to deduce a variety of features of galactic intelligence, some of which are directly relevant to the Fermi paradox. They all depend upon the idea that you are a random intelligent observer—with no special location in either space or time. First, the colonization of the Galaxy can't have occurred on a large scale by ETCs (because if it had, you—yes, *you*—would probably be a member of one of those civilizations). Second, applying the delta *t* argument to the past longevity of radio technology on Earth and combining this with the Drake equation, Gott finds at the 95% confidence level that the number of radio-transmitting civilizations is less than 121—and possibly much less than this, depending upon the parameters fed into the Drake equation. Third, if there's a large spread in the populations of ETCs, then you probably come from an ETC having a population larger than the median. Thus, ETCs with populations much larger than our own must be rare—rare enough that their individuals don't dominate the total number of beings, otherwise you'd be one of them. From which we deduce there is probably not a KII civilization to be found in the Galaxy, nor a KIII civilization anywhere in the observable universe.

As I indicated earlier, there seems to be something not quite right with the argument; it *feels* wrong—but where exactly is it wrong? There are philosophical opinions both for and against Gott's argument, and perhaps the safest course of action is to let the philosophers slug it out. Personally, I'm uneasy with the assumption that intelligent species necessarily have a finite lifespan; recent observations indicate the universe will expand forever and so it's *possible* for humankind to survive forever (in which case a straightforward application of a doomsday argument becomes problematic). What is the definition of "humankind" in this case anyway? When, exactly, does Gott believe humankind "started"? And if our species evolves into something else, does that count as the *end* of humankind? Nevertheless, despite the feelings of unease one might have, the doomsday argument still stands.

Various aspects of the doomsday argument are addressed in an ingenious manner[252] by Willard Wells in his book *Apocalypse When?* Wells takes the argument in a mind-bending direction and, in addition to quantifying the existential risks facing us, he presents another possible answer to Fermi's question. He notes that evolution has caused humans to identify and deal quickly with short-term hazards; we have no such instinct for recognizing or appreciating long-term threats. If this feature is typical of intelligent species then perhaps doomsday inevitably arrives for them through their failure to anticipate long-term consequences.

## Solution 40 Cloudy Skies Are Common

*The long night had come again.*

Isaac Asimov, *Nightfall*

Whenever polls are taken of such things, Asimov's story *Nightfall* is routinely voted as the greatest piece of SF below novel length. *Nightfall* tells the tale of scientists on Lagash, a planet in a system of six stars.[253] In reality, the

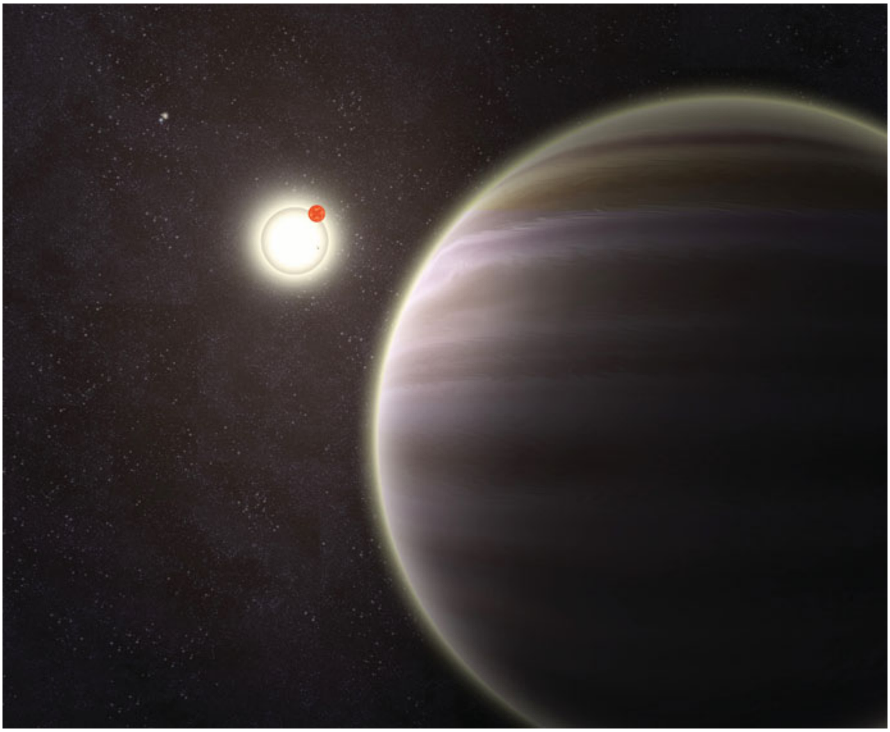


**Fig. 4.25** An artist's conception of planet Kepler-64b, also known as PH1. It was discovered in October 2012, and is the first example of a planet in a four-star system. The planet orbits a double star; a more distant pair of stars orbits the system. PH1 is a gas giant, but if you could stand on it you'd experience a double sunset and see two bright stars in the night sky. The situation on Lagash, with its six suns, would be even stranger. (Credit: Haven Giguere/Yale)

chaotic orbit of Lagash surely wouldn't permit the existence of advanced life-forms. For the sake of the story, however, Asimov postulates that intelligent, technologically advanced creatures have developed on the planet. The story takes place soon after physicists on Lagash have discovered the law of universal gravitation, which allows them to predict the position of any of their planet's six suns. Their newfound knowledge also enables them to deduce the existence of a moon orbiting Lagash.

The presence of Lagash's moon has to be *deduced* because the moon isn't visible: the presence of six suns means darkness never falls on Lagash. The planet never has night. *Nightfall* describes what happens on Lagash when a rare alignment of the moon and the six stars produces an eclipse, and the beings of Lagash for the first time see the night sky. It's a wonderful story.[254]

The astronomers on Lagash would find it difficult to develop what we call astronomy. Since light from their six suns drowns out light from any other astronomical body, they couldn't learn of the existence of planets or stars. Without a clear view of the skies, how could Lagash astronomers possibly develop an understanding of the physical universe or of their place in it?

Although the situation in *Nightfall* is unlikely, one can think of many cases where the physical environment in which an intelligent species finds itself would prevent it from exploring the cosmos. What if, as one philosopher asked, cloudy skies are common? Or what if intelligent species are more likely to evolve in the sea rather than on land? No matter how intelligent the species, no matter how advanced its technology or how high its civilization, if it has no inkling of the existence of other worlds, no reason to suspect that other beings might be out there, the notion of attempting contact would simply not arise. Interstellar communication wouldn't take place. Perhaps there are thousands of ETCs out there in our Galaxy—but they are behind cloud cover, or stuck near the galactic center where the sky is eternally bright, or in any of a hundred environments that would render astronomy difficult. Does this explain the paradox?

This idea has made for some of the greatest SF stories, but it's difficult to accept it as an explanation of the Fermi paradox. As we shall see later, the Galaxy probably contains a trillion or more planets. It's inconceivable that Earth is the only planetary environment with a clear view of the skies.

## Solution 41 As Good as it Gets

*The only real progress lies in learning to be wrong all alone.*

Albert Camus

Implicit in Fermi's question is the notion of scientific and technological progress. When Fermi discussed flying saucers with York, Konopinski and Teller they all gave serious consideration to the possibility of faster-than-light travel; but if superluminal travel is possible, it requires a knowledge of physics that's far beyond our own. When researchers model galactic colonization using cellular automata or Monte Carlo methods or some other computational technique, they assume that the vast reaches of interstellar space can indeed be traversed; but at present humans certainly don't possess the necessary technology to colonize the Galaxy. When we puzzle over the lack of evidence for the byproducts of alien technology we assume that it is in fact possible to build Dyson spheres or Shkadov thrusters or antimatter rockets; but although we can imagine such technologies we certainly couldn't develop them. If a civilization is a million years older than us then it will possess an understanding of science and technology that appears almost magical—that's the assumption we tend to make. But what if Earth science of, say, the year 2020 is as good as it gets? What if our current level of scientific understanding is as deep as it's possible to obtain?

Consider the world of the very small. Over a period of several decades, physicists have developed the so-called standard model of particle physics. The model tells us that all matter consists of a limited number of particles (three pairs of quarks; three pairs of leptons) interacting through a limited number of forces (electromagnetic, weak and strong). In 2012, physicists working at the Large Hadron Collider discovered the final element of the standard model,[255] the Higgs boson. The standard model is *stunningly* successful; it's consistent with the results of every subatomic experiment ever made. But the model isn't complete. It doesn't incorporate gravity; it applies to only about 4% of the mass–energy content of the universe, since it doesn't include dark matter or dark energy; and it contains 19 parameters, the values of which are unexplained and must be inserted "by hand". Physicists desperately want to find evidence of physics beyond the standard model, but so far no cracks have appeared: the standard model remains solid even though we know, somewhere, it must break down.

Or consider the world of the very large. Over a period of several decades, cosmologists have developed the so-called standard model of cosmology. Take general relativity, which is perhaps the most beautiful theory in all of physics, add just six parameters (these are numbers such as the densities of "normal" matter, dark matter and dark energy) and one obtains a model that's consistent with all cosmological observations ever made. It seems that we live in a universe that underwent a period of cosmic inflation and then expanded in a Big Bang some 13.8 billion years ago; the gravitational pull of matter and dark matter slowed the expansion at first, but eventually the effects of dark energy started to cause the expansion to accelerate. The standard model of cosmology is a *stunning* achievement: it's quite amazing that we can make precision measurements of the large-scale structure of the universe and develop a model to explain them. But the model isn't complete. We have little idea about the nature of dark matter; dark energy is a complete mystery; and the underpinning theory, general relativity, can't be the whole story be causes it isn't a quantum theory (and if we can be confident in anything it's that the universe is fundamentally quantum in nature). Physicists desperately want to reconcile gravity with quantum theory, and they dearly want to understand the origin of dark energy, but there's no consensus on how to do it: the standard model works, but we don't know why.

Or consider the living world. Crick and Watson proposed a double-helix structure for DNA in 1953, and since then biologists have made huge strides in understanding the biochemical underpinnings of life on Earth. In recent years, advances in genetic technology have probably outpaced even those in computing technology. And yet we still don't know how life came into being in the first place. As for understanding a phenomenon such as consciousness in biophysical terms . . . well, how do we even begin?

As the 19$^{th}$ century turned into the 20$^{th}$, some eminent scientists believed that physics was essentially complete; all that remained was to measure known quantities with ever-increasing accuracy. There were just two clouds on the horizon back then: the continuing inability of experimentalists to detect luminiferous ether and the inability of theorists to explain the ultraviolet catastrophe of blackbody radiation. By the time those clouds had cleared, we had completely new physics: special relativity (which explained the problem of the ether) and quantum theory (which explained the blackbody problem). Perhaps we are now experiencing the opposite situation to the scientists of over a century ago: no-one believes that physics is over—there are so many reasonable questions to which we don't have answers—but perhaps the theories now in our possession are good enough to explain any observation we can make. Perhaps we are in the exasperating position of knowing that our theories are incomplete—that they are wrong, even—but being unable to perform the experiments that would point to better theories. Perhaps the physical make-up of our universe is such that this is the fate of any intelligent species.

If that's the case then it would explain why we haven't heard from ETCs: they are all stuck at more or less the same level of scientific understanding—and thus technological competence—as we are. They know about quarks and cosmic inflation and dark energy, but they don't know how it all fits together any more than we do. They wonder whether anybody is out there, but they have the same limited lack the capacity to broadcast their existence to the wider universe as we do. It's a frustrating thought.

Personally, I can't bring myself to believe that scientific progress is going to halt any time soon. The hunt for dark matter will surely soon turn up new physics; while writing this section I've learned that work has been completed on the upgrade of the Large Hadron Collider, which means that particle physicists will soon be able to probe the subatomic world at distance scales never before available; in forthcoming years, astronomers and cosmologists will be studying dark energy and high-energy cosmic rays and the gravitational universe with telescopes of quite astounding capability.[256] And as for biology, the rate of increase in knowledge shows no signs of slowing. Something will turn up. If we ever meet an extraterrestrial civilization then I'm certain its science would be far in advance of ours; to imagine they know no more than we do now is surely untenable.

## Solution 42 They Are Distance Learners

*Give me just enough information so that I can lie convincingly.*

Stephen King

The Fermi question—"Where *is* everybody?"—resonates because it contrasts two simple observations. First, the Galaxy is old enough for intelligent life to have arisen long ago. Second, claims of flying saucers notwithstanding, we see no signs of the artefacts of intelligent life—no alien spacecraft, no self-replicating probes, no astroengineering projects. When we unpick that second observation, however, we find inside it various assumptions. For example, we tacitly assume that the only way to explore the Galaxy is through a "brute force" application of fleets of spaceships or self-replicating probes. Perhaps ETCs have more subtle ways to gather information?

Mike Lampton, a scientist[257] at the Space Sciences Laboratory in the University of California, Berkeley, points out that when we model galactic exploration via spacecraft we betray an "Earth-2000" mindset: physical exploration might make sense to Earthbound physicists in the year 2000 but would it make sense to the physicists of an ETC? Won't alien physicists have other options, particularly since they'll have been studying their subject for millions of years and will thus understand much more about Nature than our physicists? (Well, that's the hope—but see the discussion in the previous Solution. Note that when writing "2000" here it should be understood that we can give or take a few decades. Galactic colonization by probe was being discussed in the 1960s and it's being discussed as I write this book. A few years either way doesn't matter; it's all "Earth-2000" physics.)

It's quite widely accepted by physicists that Earth-2000 physics is incomplete. As described earlier, the standard model of particle physics and the standard model of cosmology might stand as tributes to the human intellect but they rest on incompatible foundations and they fail to describe 96% of the mass–energy content of the universe. Moving beyond these models poses a deep problem for Earth-2000 physicists, but they mightn't be an issue for Trantor-20000 physicists. And solving these fundamental mysteries of physics might well open up new ways to explore the universe.

Lampton points out that information is becoming increasingly important in our society. We are mining more data and mining less coal. As technology advances, there'll be less need to move goods or people. (Fancy a juicy jabuticaba? No need to ship them in from South America. Just print one at home.) The same applies to space travel. It's already safer, cheaper and more practical to explore Mars using telepresence than it is to attempt to send manned craft. We can even do science at a distance. For example, if a rover found alien microbial life hiding under the sands of Mars then we wouldn't need to send astronauts[258] to investigate: genome sequencers inside the rover could transmit genetic information back to Earth and we could reconstruct the life form in labs using biological printers. If we can contemplate learning about the universe at a distance using Earth-2000 physics, surely any ancient information-driven society could do the same, only much more efficiently.

Lampton's take on the Fermi paradox, therefore, is that all technologically based societies eventually make a transition: pre-transition societies are motivated by colonization, conquest and trade; post-transition societies are driven by information. A post-transition society doesn't have to "be there" if it has a complete but remote knowledge of "there". When an individual from a post-transition society really wants to visit "there", it need simply construct a local simulation. If such a societal transition occurs on a short timescale compared to the colonization timescale—and if we extrapolate current trends here on Earth that would appear to be likely—then the paradox disappears.

I'm not completely convinced. I'm not entirely sure that developments in information technology have been more important than developments in transportation technology. I was an early adopter of the Web, I own numerous internet-connected devices, and I find it difficult to remember a time when I wasn't permanently online—but the truth is that I managed fine without the internet. I was just as productive then as I am now. (Isaac Asimov produced over 500 books, many of them science popularizations. He wrote them using a typewriter. Would he have been more productive if he'd had access to Google? Charles Dickens didn't even have a typewriter. Would he have written more novels if he'd had access to our cut-and-paste and spellcheck technologies? I doubt it.) We've had access to speed-of-light communications technology since about the time Queen Victoria ascended to the throne: Cooke and Wheatstone patented an electrical telegraph system in May 1837. Haven't advances since then simply been ways of sending richer types of postcard? Advances in transportation, though, really have made a difference; they've given me a way of life that my ancestors could never have dreamed of. Of course, I might be displaying symptoms of late-onset Luddism here; people younger and more intelligent than me certainly seem to regard advances in information technology, and the blurring between the real and the virtual, as both inevitable and profoundly transformative—so perhaps this approach does begin to address the paradox.

Note that the transition to which Lampton refers is essentially instantaneous on astronomical timescales, but it isn't a change of *kind*: a post-transition society might have different motivations to its pre-transition forebear, but its essence wouldn't change. The authors of the following set of Solutions argue that technological societies inevitably undergo a quite different type of transition.

## Solution 43 They Are Somewhere but the Universe is Stranger Than We Imagine

*listen: there's a hell*
*of a good universe next door; let's go.*

e. e. cummings, *pity this busy monster, manunkind*

Physical science contains theories that are remarkable in their range of applicability. The standard model of particle physics explains phenomena that occur at subatomic scales while the standard model of cosmology describes the universe on the largest scales. Our theories explain events that happened a tiny fraction of a second after the Big Bang and they predict what will be the ultimate fate of the universe. And our theories aren't too bad at making sense of the middle range of phenomena, those that occur on everyday scales; our technology is proof of that.

Some people—in my experience those who tend to accept the UFO explanation of the Fermi paradox—argue that physical scientists are filled with hubris for daring to claim such success. Science, being the product of the human brain, can't possibly capture the subtleties and mysteries of the universe. These people try to explain the paradox by supposing the universe isn't what we think it is. Similar-sounding suggestions made by scientists and SF writers are more interesting.

For example, perhaps intelligent species evolve to a non-physical state that transcends the limitations of spacetime. Clarke's novel *Childhood's End* describes humanity's transition from our present rather immature state to a merger with the galactic "overmind" (some sort of spiritual union, the precise nature of which is never made clear). According to this suggestion, we don't hear from ETCs because they've evolved beyond our secular existence.

Another suggestion: intelligent species eventually evolve telepathic abilities and can communicate directly, from mind to mind, even over interstellar distances. Not for them the difficulties of radio communication. Perhaps they even *travel* using the power of the mind—as with the jaunt in Bester's novel *The Stars My Destination*. If this were true, ETCs might not bother trying to communicate with those of us who live a psi-challenged existence.

Yet another suggestion, just as outrageous but based on more conventional ideas, is that ETCs are busy exploring parallel universes. The many-worlds interpretation of quantum mechanics suggests that every time we make a measurement on a quantum system possessing two possible states, the universe splits into universe A and universe B.[259] An observer in universe A measures one outcome of an experiment, an observer in universe B measures the other possible outcome. The result is a never-ending branching of universes. In the totality of universes, all possibilities are realized. If the many-worlds interpretation is correct (a big "if"—there are several competing interpretations of quantum mechanics) and if it's possible to move between universes (an absolutely huge "if"—there's absolutely no indication that such travel could occur) then perhaps ETCs are elsewhere. Why stick around a dull place such as this universe when you can explore *really* interesting places?[260]

One last suggestion, based on recent advances in the arcane subject of string theory, takes a slightly different approach to the notion of parallel universes. In string theory, branes are physical objects that exist in higher dimensions. A point particle can be thought of as a brane of zero dimension; a string is a one-dimensional brane; and if a brane has dimension $p$ then it's a $p$-brane. Theoretical physicists investigating cosmological models based on the brane idea have proposed that our four-dimensional universe might be restricted to a brane inside a higher-dimensional space. In brane cosmology, one or more of those extra dimensions can be large. We don't see these large extra dimensions because photons, which are the particles that allow us to see something, are restricted to the brane; in fact *all* particles, including those that constitute our bodies, are restricted to the brane. The gravitational force, however, can "leak" into the large extra dimensions; indeed, in braneworld cosmology it's this "leakage" that explains the weakness of gravity. If brane cosmology turns out to be a true description of the universe then it's possible that there are other branes—other worlds—sitting literally parallel to our own, stacked together like slices of bread in a loaf. These universes could be just a millimeter away from each other in the higher dimensions, but matter and radiation would be stuck to each of the branes; only through gravity can branes affect one another. Regarding the Fermi paradox the suggestion is, of course, that advanced civilizations learn how to move through the "bulk";[261] they find it energetically more favorable to move that millimeter through higher-dimensional space than to move light years through our four-dimensional spacetime. Needless to say, there's no experimental evidence for the existence of branes existing in a higher-dimensional bulk; even if these large extra dimensions exist there's no reason to suppose they are navigable.

While it's certainly true that science has not told us *everything*—indeed, what remains to be discovered seems to grow exponentially—it's wrong to say science has told us *nothing*. Over the past 400 years science—a process involving hundreds of thousands of people working individually and cooperatively—has yielded reliable knowledge about the universe. Any new theories not only have to explain new observations and experimental findings, but also the accumulated set of observations and findings—which makes it extremely tough to develop new theories. No one has succeeded in developing useful theories of phenomena such as transcendent spiritual unions, interstellar telepathic communication, inter-universe travel—or any of the other imaginative suggestions that have been made. In fact, since at present we can understand the universe without invoking the existence of such phenomena, we don't *need* to develop new theories to explain them. That doesn't mean such phenomena are impossible; but we require evidence before we need to study them in earnest.

So while these suggestions all make for good stories, it's difficult to take them seriously as resolutions of the Fermi paradox.

## Solution 44 Intelligence Isn't Permanent

*No permanence is ours; we are a wave*
*That flows to fit whatever form it finds.*

Hermann Hesse, *The Glass Bead Game*

In 2002 Karl Schroeder, a Canadian SF writer, published a novel called *Permanence* which contains dozens of interesting scientific and philosophical speculations[262]—including, as Milan Ćirković later emphasized, a possible solution to the Fermi paradox. Ćirković called it an "adaptationist" solution.[263]

In biology, adaptation is the evolutionary process whereby a population of organisms becomes better suited to the habitat and environment in which it lives; an adaptive trait is a common feature that results from that process, some physiological, behavioral or life-cycle aspect that improves the chance of organisms surviving and reproducing. There are any number of examples of adaptation: bats use echolocation to catch insects; katydids mimic leaves in order to evade predators; the claws of a cheetah help it catch prey.

**Adaptation Doesn't Explain Everything** It's tempting to find "just-so" stories that explain everything in terms of adaptive traits, but not everything is the result of adaptation. Some structures, for example, are vestigial: fish that live in completely dark caves have sightless eyes—the function was lost when their sighted ancestors began to inhabit environments in which there was no pressure to maintain eyesight. In other words, in the blackness of those caves fish with good eyesight no longer outcompeted fish with poor eyesight. The sightless eyes of the present day are a result of evolutionary history, not adaptation. Some phenomena are byproducts: the redness of blood, for example, is due to the particular properties of haemoglobin rather than adaption. And some features might be the result of exaptation: perhaps bird feathers arose for insulation and were only later co-opted for flight, in which case feathers would be an exaptive trait for flight (but adaptive for insulation).

In *Permanence*, Schroeder discusses the idea that intelligence and consciousness have no more significance than adaptive traits such as echolocation in bats or leaf mimicking in katydids. And just as sighted fish can lose eyesight if the selective advantage of possessing eyesight disappears, so too might intelligence and consciousness wither as the environment in which it finds itself changes.

**Fig. 4.26** Two male red crossbills photographed in Oregon. These types of finch possess an unusual trait: their bills cross (hence the English name of the bird). These birds feed on the seeds of conifer cones, and the shape of their bills—crossed at the tips—helps them extract the seeds. The shape of their bills is an adaptive trait. Might human intelligence and consciousness be an adaptive trait, with as much or as little significance as the crossbill's distinctive mandibles? Is human intelligence important only to the extent that it provides an evolutionary advantage in a fluctuating environment? (Credit: Elaine R. Wilson)

For Schroeder, intelligence isn't a prerequisite for toolmaking or civilization. He has one of his protagonists state that: "consciousness appears to be a phase. No species we have studied has retained what we could call self-awareness for its entire history. Certainly none has evolved into some state above consciousness." And later: "Originally, we must have had to put a lot of thought into throwing things like rocks or spears. We eventually evolved to be able to throw without thinking—and that is a sign of things to come. Some day, we'll become . . . able to maintain a technological infrastructure without needing to think about it. Without need to think, at all . . . ."

In *Permanence*, then, intelligence is impermanent. The communicating lifetime, *L*, of an ETC is limited not through apocalypse but through selection pressures. It isn't that intelligent species destroy themselves before communicating; it's that they better adapt to their environments and, in doing so, lose the capacity to communicate over interstellar distances. We don't see Galaxy-spanning civilizations because technologically advanced societies inevitably revert to direct biological adaptation. Life continues; it just doesn't possess the intelligence required to reach out over interstellar distances.

Is this a plausible solution to the Fermi paradox? I'm not won over. Although I'm sympathetic to Schroeder's view about the *significance* of intelligence, I'm far from convinced about the inevitability of the result he foresees. One can raise a variety of objections, some more hard-nosed than others. Since we're in the realms of speculation anyway, here's an objection based on speculation: at a certain point in the development of a civilization it's possible that biology becomes essentially irrelevant. This scenario, which I discuss in the next Solution, seems at least as plausible as the idea Schroeder discusses in *Permanence*.

## Solution 45 We Live in a Postbiological Universe

*All things must change,*
*To something new, to something strange.*

Henry Wadsworth Longfellow, *Kéramos*

Steven Dick, a noted historian of science,[264] has long argued that we need to employ a "Stapledonian" mindset if we want to think about the nature of intelligence in the universe. Olaf Stapledon was a British philosopher[265] who, in several science fiction novels, considered human evolution over astronomical timescales. His novel *Last and First Men*, published in 1930, outlined a "future history" and described 18 distinct species of human over the next two billion years (our own species are the First Men). His novel *Star Maker*, published in 1937, was even grander in scale: it described the entire history of life in the universe—and it contained, among many mind-bending concepts, the first description of a Dyson sphere; Freeman Dyson himself has suggested that such a structure should be called a "Stapledon sphere". Dick's point is that in contemplating intelligence in the universe we need to consider not only astronomical timeframes but the biological and cultural evolution of intelligent species that might take place over such timescales. When thinking about ETCs we need to take the sort of approach that Stapledon took in his novels.

I mentioned the relevant timescales in chapter 1, but it's worth repeating them here. We now know that the universe is 13.798 billion years old, give or take 37 million years.[266] Astronomers still don't have a full understanding of the formation of the first stars, but it seems reasonable to assume that the first Sun-like stars, and thus perhaps the first rocky planets, formed within a billion years after the Big Bang—in other words, about 12.8 billion years ago. If we take Earth as a guide, and assume intelligent life appears 4.5 billion years after a planet forms, then we could argue that the oldest civilizations appeared in the universe 8.3 billion years ago. The oldest stars in our Milky Way galaxy formed roughly 10–11 billion years ago so a similar argument suggests that the oldest neighboring civilizations could quite easily be 5 billion years old. Ray Norris, using an argument based upon stellar evolution,[267] reasoned that the median age of an ETC would be 1.7 billion years. Different astronomers, using different arguments, have argued for a median age somewhere between the Norris figure of 1.7 billion years and the 8.3 billion years mentioned above. The specific age is hardly the point: whether you believe that extraterrestrial civilizations might be 1.7 billion years older than us, 8.3 billion years older or somewhere in between, the take-home message is the following. If ETCs survive natural and home-made catastrophes they are likely to be much, *much* older than us: the genus *Homo* is estimated to be about 2.3 million years old.

If ETCs persist over long reaches of time then we need to explore their likely evolution. As the famous physicist Niels Bohr once remarked, it's hard to make predictions, especially about the future. It's probably impossible to predict how a billion years of biological evolution will pan out. However, one might argue that biological evolution becomes increasingly irrelevant once a civilization reaches a certain level of technical sophistication: *cultural* evolution vastly outpaces biological evolution. The rapidity of cultural evolution means that civilizations can change dramatically on timescales as short as hundreds of years. If the possession of advanced intelligence implies the possession of culture then it follows, Dick argues, that any discussion of extraterrestrial civilizations must necessarily take account of cultural evolution.

How might cultural evolution proceed? Well, even in the case of human civilization the answer is that we just don't know. In the humans-only Galaxy of Asimov's *Foundation* the development of society could be predicted and even shaped through the theory of psychohistory; we lack such a theory. As for the cultural evolution of extraterrestrial civilizations . . . who can tell? In the absence of a theory of universal cultural evolution perhaps the best one can do is to extrapolate the trends we believe we see here on Earth. Dick highlights the following fields as being most relevant in this regard: artificial intelligence, biotechnology, genetic engineering, nanotechnology and space travel. Of these, he regards artificial intelligence as being most important because the other fields can be seen as serving intelligence: through biotechnology and nanotechnology we might construct efficient artificial intelligence; the development of space travel will spread intelligence; genetic engineering might provide a route to increased biological intelligence. Here on Earth these trends are gathering in pace. Dick therefore posits what he calls the Intelligence Principle: "the maintenance, improvement and perpetuation of knowledge and intelligence is the central driving force of cultural evolution, and that to the extent intelligence can be improved, it will be improved".

The Intelligence Principle implies that given enough time—and ETCs will have had enough time—biologically based intelligence will create artificial intelligence. The products of biological evolution will be replaced by, or will merge with, their machine progeny. Stapledonian thinking suggests that we might live in a postbiological universe.

Such an outlook has several implications for both SETI and the Fermi paradox. One implication is that we might be looking in the wrong place: postbiologicals, freed from the shackles of a corporeal existence, need not remain planet-bound. The SETI focus on Earth-like planets orbiting Sun-like stars might be misplaced. A second implication is that postbiologicals might be more interested in receiving signals from biologicals than in trying to communicate with them. A third is that the huge differences between postbiologicals and biologicals—differences in age, abilities, physicality and much else—might lead to a qualitative difference between our minds and theirs: communication might be impossible.

The idea of a postbiological universe is not without its problems. For example, presumably the postbiologicals would also undergo cultural evolution—where would that lead? And the Intelligence Principle itself, on which Dick hangs his argument, hardly has the status of a physical law. The Principle seems persuasive because in our culture increased knowledge and intelligence confers a competitive advantage, but the Principle might have only a local application; perhaps the cultural evolution of extraterrestrial civilizations is driven by hate or by the drive to conquer or by some emotion we lack the words to describe. (George R. R. Martin once wrote[268] a wonderful, haunting story called *A Song For Lya*, in which the prime motivation of one extraterrestrial culture was love. Do read it.) Nevertheless, the notion that we might inhabit a postbiological universe exerts a strong pull. It's possible to take forward Dick's argument in a variety of ways, each of which gives a subtly different take on the Fermi paradox. The next few solutions discuss different aspects of a postbiological universe.

## Solution 46 They Are Hanging Out Around Black Holes

*Many a trip continues long after movement in time and space have ceased.*
John Steinbeck, *Travels with Charley*

When discussing Gillett's take on the Fermi paradox (see page 1) we looked at the Kardashev scale for classifying ETCs. The scale is based on energy consumption. To recap: a KI civilization is capable of harnessing the energy of an Earth-like planet; a KII civilization is capable of harnessing the energy of a star; a KIII civilization is capable of harnessing the energy of an entire galaxy. If the development of our own technological civilization is anything to go by (and, as always, let's take it as read that we've no idea whether this is the case in general) then the Kardashev scale would seem to be a reasonable measure of the state of advancement of an ETC. Humankind consumes ever more energy because there are ever more of us wanting to do ever more things. We don't know what humans might want to do in the future, but whatever it is will require energy. Similarly, whatever wonderful technologies an ETC possesses, we can be fairly sure large quantities of energy will be involved—and the more advanced and widespread the technology, the more energy will be needed.

The Cambridge cosmologist John Barrow has introduced a scale of inward manipulation,[269] which one can argue would be just as applicable to ETCs as the Kardashev energy scale. A civilization at the BI level of advancement can manipulate objects at its own size, or about 1 m (assuming that intelligent beings exist, as we do, at this size). A BII civilization can work with objects at the $10^{-7}$ m scale, which would allow it to manipulate genes. A BIII civilization can work with objects at the $10^{-9}$ m scale, which would allow it to manipulate molecules. Barrow argues that human civilization is now at the BIV level, since various technological advances are allowing us to manipulate individual atoms at the $10^{-11}$ m scale. But, as Richard Feynman once remarked, "there's plenty of room at the bottom";[270] in other words, there's more to explore at small scales than at large scales. Indeed, there are 35 orders of magnitude between the human scale of 1 m and the smallest possible scale defined by quantum physics—the Planck scale; there are "only" 26 orders of magnitude between the human scale and the size of the observable universe. Perhaps, then, as ETCs advance in sophistication they choose to investigate the microworld instead of, or at least in addition to, the macroworld. A better classification of ETCs might be their ability to manipulate smaller and smaller length scales. On the Barrow scale a BV civilization could manipulate atomic nuclei (and thus work at the $10^{-15}$ m distance range); a BVI civilization could manipulate elementary particles ($10^{-18}$ m); and a BΩ civilization could manipulate the structure of spacetime itself ($10^{-35}$ m).

In his PhD thesis, presented in 2013, the Belgian philosopher Clément Vidal argues[271] that ETC development can best be discussed using a two-dimensional metric that combines the Kardashev and Barrow scales. In particular, he argues that SETI researchers should consider what it might mean if civilizations were at the KII–BΩ level—civilizations that could harness the energy of a star *and* manipulate spacetime.

If a civilization has the capacity to manipulate spacetime then its technology will be able to work with black holes—regions of spacetime from which nothing can escape.[272] Black holes are relatively common in the universe: the fate of a high-mass star is to end as a black hole, and astronomers believe that a supermassive black hole lurks at the center of every galaxy. Vidal argues that black holes are "attractors for intelligence": KII–BΩ civilizations will be drawn to using these extreme objects. With our current level of understanding it's impossible to state what such an advanced civilization would choose to do with their black holes, but it's interesting to speculate. Advanced civilizations might, for example, use black holes to store or extract energy[273]—various mechanisms have been proposed for extracting energy from black holes, and they often have excellent efficiency. Or they might use them for science: a black hole gravitational lens could form the basis of a hugely powerful telescope. Some scientists have postulated that the spacetime around a rotating black hole might allow the construction of a hypercomputer—a device that allows the solution of problems that traditional computers can't solve; surely any ETC would be interested in *that* possibility. There might be technological reasons for investigating black holes: perhaps they facilitate travel through spacetime. And then there's the speculation that truly advanced civilizations might use time dilation effects around a black hole in order to survive into the indefinite future (whereas the less advanced civilizations survive only as long as their host star—humankind, for example, has at most only a couple of billion years before the Sun devours all life on Earth). Surely these are reasons enough to consider black holes as an attractor for intelligence? If not, there are plenty of other reasons one can dream up. If nothing else, black holes are the last word in waste-disposal units.

A KII–BΩ civilization could use the energy of a star to drive its black hole technology. Would it be possible to detect such activity? Well in principle, yes, it would. We know such detection is possible because astronomers have observed dozens of objects known as X-ray binaries (XRB). An XRB is a system in which a donor object (usually just a normal star) loses matter to a compact accretor (such as a black hole). Material is sucked away from the star and forms an accretion disk around the dense object. This infalling matter releases gravitational potential energy in a process that's much more efficient than the hydrogen fusion that powers stars: X-rays and high-energy particles pour out of the system. So ... could X-ray binaries be a manifestation of advanced civilizations using black holes for hypercomputation, spacetime traveling or waste disposal? Could XRBs be ETCs?

If one accepts that civilizations will inevitably evolve towards a KII–BΩ state, and if that evolution occurs quickly on an astronomical timescale, then that might explain why we haven't seen those civilizations. We've been listening for low-energy communications instead of looking for signs of high-energy activity in, for example, XRBs. However, Vidal himself points out the inherent weakness in trying to account for X-ray binaries as a manifestation of advanced technology: it's far more reasonable to explain them in terms of natural phenomena. X-ray binaries are just binary star systems in which one of the stars

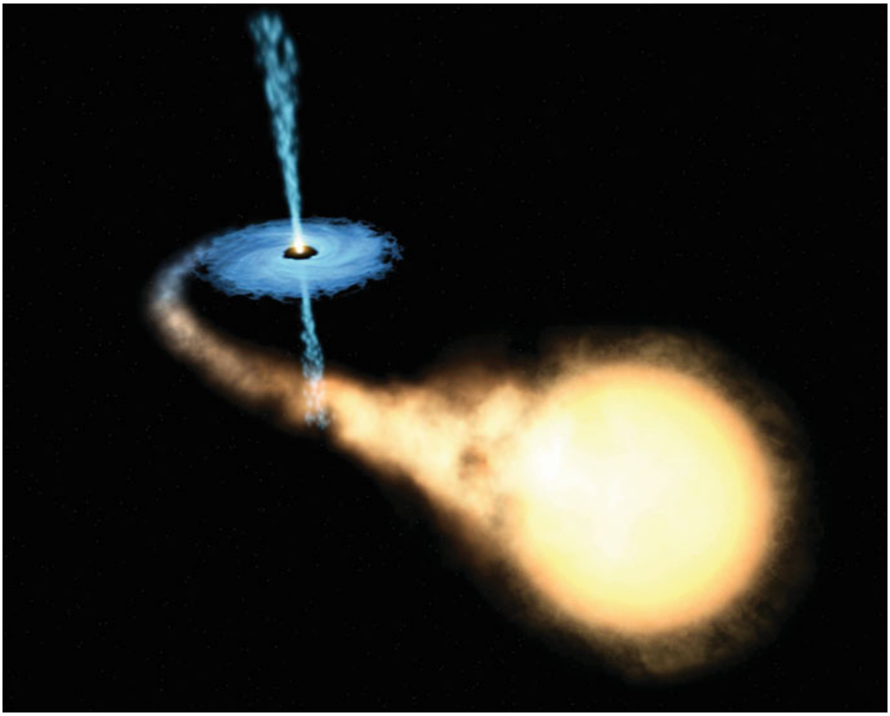


**Fig. 4.27** Artist's impression of the microquasar GRO J1655-40 with its relativistic jets shooting out from the poles. A microquasar is the little brother of a quasar; in both cases there's an accretion disk surrounding a black hole, but the mass of the central black hole is different. Quasars possess supermassive black holes (measured in millions of solar masses) whereas microquasars possess black holes that have the roughly the same mass as a typical star. (Credit: NASA/STScI)

has reached the end of its life and becomes, as happens to old stars, a compact object—either a white dwarf, a neutron star or a black hole. In some systems astronomers have yet to clarify the specific details of the physical processes involved, but there seems little doubt that the phenomena we observe are simply the result of gravity doing its thing. There's as much need to invoke ETCs to explain the output of XRBs as there is to invoke Santa Claus to explain the distribution of presents on Christmas Day. Note, however, that observations could change this conclusion (regarding XRBs, that is, not Santa Claus). Calculations suggest that if a stellar remnant has a mass less than 1.44 times that of the Sun then its fate is to become a white dwarf; if the remnant has a mass between about 1.5–3.0 times that of the Sun then it will end its life as a neutron star. Only if the remnant mass is greater than about three solar masses does a black hole form. So if it could be established that the black hole in an X-ray binary system had a mass equal to that of the Sun, say, then we'd certainly want to investigate the system more closely for signs of technological activity.

I believe the value of Vidal's two-dimensional classification of civilizations, and his focus on KII–BΩ civilizations, isn't so much the suggestion that X-ray binary sources might have a technological origin but that SETI might profitably expand its search space. To date, the focus of SETI has been on detecting low-energy radiation that could be used for communication; we could also investigate high-energy radiation for traces that might be the byproduct of technology at the KII–BΩ level. And consider that even if the black hole in an X-ray binary *formed* by natural means, perhaps a nearby KII–BΩ civilization would choose to make use of it for all the reasons mentioned above. As Vidal points out, a waterfall is a natural phenomenon—but it's a phenomenon where you can often find evidence of technology, because we construct hydroelectric power stations nearby that take advantage of what nature has provided. So we could, for example, look for evidence for the regulation of energy flow within XRBs. Astrophysicists are already investigating the high-energy universe in order to understand more about violent phenomena such as supernovae, microquasars and active galactic nuclei. SETI scientists could quite easily piggyback on these existing and ongoing observations: it would be cheap, and who knows what we might find if we look?

Nevertheless, as things stand at present the high-energy universe and the low-energy universe are alike in that we have no need to invoke extraterrestrial intelligence to explain what we observe. All we need are time and the laws of physics acting on inanimate matter and energy.

## Solution 47 They Hit the Singularity

*Things do not change; we change.*

Henry David Thoreau, *Walden*

Back in 1965,[274] Gordon Moore—the co-founder of the Intel Corporation—remarked how the number of transistors per square inch that one could fit on an integrated circuit seemed to double every 18 months. This remark became known as Moore's law, though it's more an observation than a law of Nature. In its current incarnation, Moore's law states that data density doubles every 18 months. The law has held true in the five decades since its formulation, and certain other computing hardware performance measures have kept pace too. The result: cheap, fast computing power is readily available and thus our world has changed. If the law continues to hold over the next couple of decades, and there seems to be no reason why it shouldn't, then we'll continue to see ever faster and more powerful machines—better tablets and smartphones and wearable tech.

Vernor Vinge, extrapolating these phenomenal improvements in computer hardware and related technologies, argues that mankind will likely produce super-human intelligence some time before 2030.[275] He considers four slightly different ways in which science could achieve this breakthrough. We might develop powerful computers that "wake up"; networks such as the Internet might "wake up"; human–computer interfaces might develop in such a way that users become super-humanly intelligent; and biologists might develop ways of improving the human intellect. Such a super-intelligent entity, however it comes into being, might be humankind's last invention because the entity itself could design even better and more intelligent offspring. The doubling time of 18 months in Moore's law would steadily decrease, causing an "intelligence explosion". A quicker-than-exponential runaway event might end the human era in a matter of a few hours. Vinge calls such an event[276] the *Singularity*.

The term Singularity is unfortunate, in that mathematicians and physicists already use it in a specific sense: a singularity occurs when some quantity becomes infinite. At Vinge's Singularity, however, no quantity need become infinite. The name nevertheless captures the essence of what would be a critical point in history: things would change *very* rapidly at the Singularity, and—as with the singularity in a black hole—it becomes hard to predict what happens after we hit it. The super-intelligent computers (or the super-intelligent humans or human–computer beings) turn into . . . what? It's difficult, perhaps impossible, to imagine the capabilities and motives and desires of entities that are the product of this transcendental event.[277]

Vinge argues that *if* the Singularity is possible, then it *will* happen. It has something of the character of a universal law: it will occur whenever intelligent computers learn how to produce even more intelligent computers. If etcs develop computers—and since we routinely assume they'll develop radio telescopes, we should assume they'll develop computers—then the Singularity will happen to them, too. This, then, is Vinge's explanation of the Fermi paradox: alien civilizations hit the Singularity and become super-intelligent, transcendent, unknowable beings.

Vinge's speculations about the Singularity are fascinating. And as an explanation of the Fermi paradox, the suggestion improves on explanations requiring a uniformity of motive or circumstance. Not *every* ETC will blow itself up, or choose not to engage in spaceflight, or whatever. But we can argue reasonably that every technological civilization will develop computing; and if computing inevitably leads to a Singularity, then presumably all ETCs will inevitably vanish in a Singularity. The ETCs are there, but in a form fundamentally incomprehensible to non-super-intelligent mortals like us. Nevertheless, as an explanation of the paradox, I think it has problems.

First, even if high intelligence *can* exist on a non-biological substrate,[278] the Singularity might never happen. There are several reasons—economic, political, social—why a Singularity might be averted. There are also technological reasons why the Singularity might not occur. For example, for the attainment of the Singularity, advances in software will be at least as important as hardware advances. Without *much* more sophisticated software than we currently possess, the Singularity will just not happen. Now, while it's true that various hardware measures seem to obey Moore's law, improvements in software are much less spectacular. For example, the word processor I use is the latest version of the program. It certainly has more features than the version I was using when I wrote the first edition of this book, but I never *use* those features and, indeed, the program is becoming steadily less useful to me. I persevere with it because everyone else uses it and I need to exchange documents; gradually, alternatives to this workflow are appearing and soon I might be able to ditch the program completely. The program I'm using to typeset this book,[279] which is called TEX, is a wonderful piece of software whose creator froze development on the program several years ago. While there is some progress in the worldwide TEX community toward an even better typesetting program, progress is much slower than would be the case if Moore's law were in operation. Of course, the kind of software required to create the "intelligence explosion" has nothing to do with word processors or typesetting programs. But the point is the same: advances in software and in software methodologies come at a much slower rate. We simply might not be smart enough to generate the software that will lead to a Singularity. Perhaps we'll see a future in which incredibly powerful machines do amazing things—but without self-awareness. Isn't this scenario at least as plausible as a future that contains a Singularity?

Even if a Singularity is inevitable, I fail to see how it explains the Fermi paradox. We can ask, as Fermi might: where are the super-intelligences? The motives and goals of a super-intelligent post-Singularity creature might be unknowable to us—but so, presumably, would the motives and goals of any "traditional" KIII civilizations that might exist. Yet we are happy to think about how to detect such KIII civilizations. In fact, we might have more chance of understanding the post-Singularity beings on Earth than we would of understanding extraterrestrials, because in some sense those entities would *be* us. We would, in some sense, have created them and possibly imprinted upon them certain values. Even if we are unable to understand or communicate with super-intelligent entities, it doesn't follow that those entities must disengage with the rest of the physical universe. A super-intelligence must, like us, obey the laws of physics; and presumably it would make rational economic decisions. So the same logic that suggests an advanced technological civilization would quickly colonize the Galaxy leads us to conclude that a super-intelligence would also colonize the Galaxy—except it would do so more quickly and efficiently than biological life-forms.

Even if they choose not to colonize, even if post-Singularity entities transcend our understanding of reality—perhaps they go off into other dimensions (as described on page 190) or spend their time creating the child universes that Harrison proposed (page 74), or engage in any activity bar exploration of our universe—there would be non-augmented, normal-intelligence beings left behind. In the case of humans, perhaps many of us would choose not to take part in the Singularity. But it doesn't follow that humans would become extinct. Unless the super-intelligences felt they had to destroy us (why would they bother?), we could go on living as we've always done. We might bear the same relation to the super-intelligent beings as bacteria do to us—but so what? Two billion years ago bacteria were the dominant form of life on Earth, and by many measures (species longevity, total biomass, ability to withstand global catastrophe, and so on) they still are. The existence of humans doesn't affect bacteria. In the same way, the existence of super-intelligent beings need not necessarily affect humanity; they could do their weird stuff, and we could continue doing the things we want to do—such as attempting to contact like-minded beings in the Galaxy.

To my mind, the existence of a Singularity doesn't explain the Fermi paradox. It exacerbates it.

## Solution 48 The Transcension Hypothesis

*What's past is prologue.*

William Shakespeare, *The Tempest*, Act II, Scene 1

In a paper published in 2012,[280] in the journal *Acta Astronautica*, John Smart argues that advanced civilizations do indeed hit a technological singularity but that it's possible to predict where that singularity will take them. Smart agrees with Vidal (see Solution 46) that black holes are attractors for intelligence: we don't see advanced civilizations because they disappear into black holes. This is the transcension hypothesis.

Smart's argument takes in a wide range of elements, but the gist of it is as follows. First, let's consider transcension itself. Smart draws attention not only to the acceleration in our computational capability, which we've already considered in different contexts, but also in the efficiencies and densities of the physical inputs to computation. Those inputs are space, time, energy and matter—STEM for short—and Smart defines "STEM" compression to be the phenomenon of increasing spatial, temporal, energetic and material density and efficiency per computation over time. Consider the spatial aspect: over humankind's history we've changed from being nomadic hunter–gatherers to urbanites. This change is relatively recent: according to the World Health Organization,[281] it was only in 2010 that the majority of humans lived in an urban area—just a century earlier only 20% of people lived in cities but the prediction is that, by 2050, 70% of us will be city dwellers. Cities appear to generate greater per capita wealth and higher levels of innovation than rural areas, observations that presumably lure people into cities, and yet one could argue that we are already moving beyond cities: in terms of information generation and computation per resource, companies and corporations—which are even more spatially dense—outperform cities. Smart argues that this increasing spatial density will increase further, and that we'll see similar trends regarding time, matter and energy. Human civilization will evolve in the direction of inner space rather than outer space. Civilization will become denser, quicker, more energy efficient—and change its physical substrate where necessary in order to maintain the process of STEM compression. There's no need to consider the Kardashev scale when discussing intelligent civilizations: STEM compression will mean that advanced civilizations will develop increasingly localized, increasingly dense, increasingly efficient structures and energy flows. And the limit of STEM compression? That's the Planck scale, a limit set by the way the universe is put together. Civilizations that hit the Singularity inevitably disappear behind an event horizon.

In Solution 46 we considered Vidal's arguments for supposing that black holes beckon advanced civilizations. Smart presents further reasons for supposing that black holes are natural attractors for intelligence. In particular, the phenomenon of gravitational time dilation leads to some interesting speculations. For example, the closer one gets to a black hole event horizon, the slower time appears to run (as measured by a distant observer; time runs as normal for the observer who approaches the horizon). The flip side of this phenomenon is that for the observer close to the event horizon, time in the external universe appears to run more quickly. If an observer can hover close to the event horizon of a black hole then he/she/it is able to watch billions of years of universal dynamics unfold in an instant. Smart argues that any civilization that has maximized its local STEM resources, and finds the local universe to be an increasingly uninteresting and unsurprising place, will want time in the rest of the universe to run as quickly as possible: that way any interesting tidbits, items of useful non-local information, reaches it in the shortest amount of local time. (The civilization might want to create a shell of matter around itself to form a focal sphere: gravitational lensing will allow the civilization to observe the distant universe, as discussed on page 56.) Over extremely long time scales, the black holes within a galaxy will collide and merge; another consequence of gravitational time dilation near the event horizon is that this merger process takes only a short time from the point of view of the advanced civilization inhabiting a black hole. This mechanism allows advanced civilizations to meet up, eventually. (So if Smart is right, and humankind reaches the Singularity when he thinks it will, then we'll only have to wait a few hundred years to meet ETCs. Of course, hundreds of billions of years will have passed in the external universe.)

Various questions spring to mind with this scenario, and Smart has responses lined up for them. An obvious question, for example, is: why don't we see or hear civilizations in the run-up to transcension? If such civilizations are numerous—and Smart, in a related publication, estimates that there could be as many as 2.25 billion advanced technical civilizations in our Galaxy; an astonishingly high number—then why don't see *some* evidence for astroengineering, why don't we detect *some* radio beacon from their pre-transcension phase? Well, Smart argues that human civilization might reach its transcension event 600 years from now; a few centuries is a mere eyeblink on a cosmic timescale. It's highly unlikely that we would be close enough to any pre-transcension civilization to pick up any attempt at communication. (Even if we discovered a neighboring civilization that was 100 light years away and at exactly the same technological level as us—a situation that's highly unlikely to occur—then the extent of our conversation would be three two-way exchanges of information before one or other party transcended. We couldn't even share a knock-knock joke.) Smart goes further and argues that a civilization on the path to transcension would intentionally refrain from broadcasting: the transmission of information might alter the transcension path of other civilizations and reduce the variety of information available when civilizations merge. Indeed, Smart argues that a civilization, once it recognizes that its destiny lies in a black hole, will develop a "prime directive"—its morals will prevent it from broadcasting. In this sense, the transcension hypothesis is a variant of the zoo hypothesis (see page 61).

The transcension hypothesis has the benefit of offering some specific and potentially falsifiable predictions. First, although radio SETI probably will not succeed if the transcension hypothesis holds true, the outlook for optical SETI is much better. In particular, optical methods have the potential to analyze exoplanet atmospheres. According to Smart, as a civilization undergoes STEM compression the tell-tale life signatures on a planet will disappear. The transcension hypothesis thus predicts an absence, or at least a lower frequency, of life-signature exoplanets in the inner ring of the galactic habitable zone. Second, Smart predicts the existence of a well-defined and ever-growing edge to the transcension zone, an area in which civilizations that are the right age "flip" their states to become STEM-dense. Third, Earth must be close to the edge of the transcension zone since we appear to be close to our own transcension event.

Perhaps the predictions outlined above will be verified. Personally, I'm not convinced. Many of the concepts invoked in the transcension hypothesis are speculative (Smart's paper is only ten pages long, not counting the abstract or reference sections, and yet it contains 66 "if"s, 6 "assume"s and 3 "presumably"s; the rate of 7.5 hypotheticals per page is worthy of a Rudyard Kipling poem). On the other hand, despite the speculative nature of the concepts, the hypothesis suggests an inevitability to the transcension process. The argument in favor of this inevitability relies on another supposition: that our universe is a system presently engaged in a life cycle. Smart takes as his inspiration here the relatively new ideas of evolutionary developmental biology[282] ("evo-devo", as it's often known). Development is the process by which organisms grow and mature. In organisms that reproduce sexually, a zygote becomes an embryo that eventually gives rise to an individual possessing the same body plan as the original organism, and that individual will age and eventually die. Unlike evolution, which is a random and haphazard process, development is all very directed and constrained: a fly embryo will give rise to a fly, a human embryo will give rise to a human. Evo-devo compares the developmental processes of different organisms in order to understand, among other things, how developmental processes evolved. (One of the amazing advances of evo-devo is that biologists are beginning to understand not only how, for example, a human limb develops—but how a slight alteration to the process would cause the formation of a wing or a flipper. To a first approximation, animals share the same genes. And yet that single genetic toolkit allows embryos to develop the vast range of organizational types we see in animal life on Earth.) Smart argues that our universe displays some aspects of a life cycle (it was "born" in a Big Bang, it has grown and reached maturity, it might replicate via processes we've discussed earlier, and eventually it will suffer some sort of death). And if the universe is engaged in a life cycle then then we need to ask which of its features are evolutionary (and hence unpredictable) and which are developmental (and hence predictable). Smart argues that the ever-increasing exploration of inner space by advanced technological civilizations is guided by a universal process of evolutionary development. Transcension is inevitable.

The "inevitability" of transcension rests on too many hypotheticals for me to accept it. And for the transcension hypothesis to work, not only must all civilizations in the Milky Way march in lockstep to their black hole destiny, so must all civilizations in neighboring galaxies. Indeed, the transcension hypothesis requires that all individual elements in all civilizations in all neighboring galaxies develop in the same way. Personally, I find that unlikely. When I look around me I see contingency rather than convergence.

## Solution 49 The Migration Hypothesis

*Nothing burns like the cold.*

George R. R. Martin, *A Game of Thrones*

In recent years the Serbian astronomer Milan Ćirković has thought more deeply than most about the Fermi paradox. It's interesting, then, that Ćirković can take the same starting points as an author such as Clément Vidal and yet arrive at a quite different conclusion regarding the development of advanced technological civilizations—and a slightly different resolution to the paradox.

In a paper co-written with the futurologist Robert Bradbury,[283] Ćirković argued that intelligent life will arise at various points in the Galaxy and, if such life survives all the natural and self-initiated disasters that fate throws at it, will inevitably pursue a trajectory that leads to postbiological evolution. Ćirković and Bradbury agree with Vidal, Smart, Dick and others that the emergence of artificial intelligence and the ability to manipulate matter on the nanoscale will lead to spatially compact civilizations. They disagree, however, about the likely physical location of these civilizations.

If one accepts that proposition that technologically advanced beings will be motivated by the processing of information—which is essentially a variant of Dick's Intelligence Principle; whether these beings "have" computers or "are" computers doesn't really matter—then one can ask where such processing will occur most efficiently. Ćirković and Bradbury point out that heat is the enemy of computation. While many of the challenges facing present-day computers will eventually be overcome by employing different designs or fancier technology, the issue of heat dissipation arises directly from the laws of thermodynamics. The problem of heat dissipation will restrict the computational processes of even the most advanced technological civilizations—assuming they are bound by the laws of physics—and, since information processing is assumed to be the guiding motivation of such civilizations, Ćirković and Bradbury argue that this restriction will dominate their policies. (Precisely what sort of computation these civilizations would perform is unknown, but the assumption is that they'll prioritize the ability to process information over the physical colonization of the Galaxy. )

The maximum number of bits one can process using a given amount of energy is inversely proportional to the temperature of the processor. It follows that as one decreases the temperature of the heat reservoir in contact with the processor, computation becomes more efficient. The limiting temperature is that of the universe itself, the temperature of the cosmic microwave background: 2.7 K. (It's possible to cool a processor below this temperature, but the efficiency gains are offset by the energy required to do the cooling.) Radiation from stars causes the inner reaches of a galaxy to be rather hotter than the microwave background temperature; the limiting temperature is approached asymptotically as one moves away from the center. From a thermodynamic point of view, therefore, the best places to carry out computation are in the cold outer regions of a galaxy. Interestingly, these are also the places where various astrophysical phenomena that are inimical to life—high-energy events such as supernovae—are less likely to occur. All this leads Ćirković and Bradbury to posit the migration hypothesis as a solution to the Fermi paradox: in order to improve the efficiency of their computing, ETCs will migrate outwards from their original location to the cold outer reaches of the Galaxy. They'll move from a "galactic habitable zone" to a "galactic technological zone"—and the galactic rim will be home to a collection of individual, highly advanced "city states". The reason we don't see advanced civilizations in our neighborhood is because they, or their computers, find it intolerably hot here. As for why we don't hear from them—well, Ćirković and Bradbury agree with other writers that postbiological civilizations would have little interest in trying to communicate with creatures such as us who are so far below their intellectual level. Indeed, as Smart pointed out in a slightly different context, by leaving other civilizations free to explore their own path to a postbiological future an ETC will maximize the amount of interesting information it might learn from them when communication finally becomes worthwhile.

One's initial reaction to the migration hypothesis is likely to be that the huge expense of moving from galactic center to outer rim is likely to dwarf any savings made from computing efficiency. Remember, however, that these civilizations are likely to be high on the Barrow scale—they're going to be small and compact. Interstellar travel need not be hugely difficult for them and, if they are indeed motivated primarily by the wish to perform computation as efficiently as possible, the savings made by relocating to a colder environment can relatively quickly outweigh the transport costs.

So starting from the proposition that civilizations will inevitably follow an evolutionary trajectory that leads to a postbiological future dominated by the desire to perform computation we can conclude that: ETCs will want to get close to black holes with their associated high-energy environments (Vidal's conclusion) . . . or get as far away from them as possible (Ćirković and Bradbury).

## Solution 50 Infinitely Many Civilizations Exist but Only One Within Our Particle Horizon: Us

*We all live under the same sky, but we don't all have the same horizon.*
Konrad Adenauer

Michael Hart has an interesting way of considering the paradox he has done so much to promote.[284] To fully appreciate his argument we have to understand the notion of a particle horizon.

A particle horizon is easiest to explain in a static universe. (The universe is dynamic, not static: it began in the Big Bang, has been expanding ever since, and recent findings suggest it will expand for eternity. Taking into account the expansion of the universe makes a discussion of particle horizons rather subtle. Fortunately, nothing is lost if we discuss the idea in terms of a static universe.) Imagine, then, a universe that is infinite in extent and throughout which galaxies are uniformly distributed. Furthermore, this model universe came into existence about 14 billion years in the past; perhaps the galaxies already existed, and some supreme intelligence "threw the switch" and turned on all the stars at precisely the same instant. What would such a universe look like to an observer on an Earth-like planet, some 14 billion years after this creation event? Would the night sky be blindingly bright, the result of light reaching the planet from the infinite number of galaxies? Those unfamiliar with Olbers' paradox might be surprised to learn that this infinite static universe would look similar to the one we inhabit. The point to remember is that nothing can travel faster than light. So no influence—no light, no gravitational waves, *nothing*—could have reached the observer from regions more distant than 14 billion light years. This distance—the distance to the particle horizon—is the effective size of the observable universe. Nothing from beyond the horizon has had time to reach the observer.

Hart makes the following argument. First, suppose our universe is infinite. The size of the *observable* universe, however, is given by the distance to the particle horizon and that's finite because the universe began some 14 billion years ago. Second, suppose abiogenesis—the development of life from non-living material—is an exceedingly rare occurrence. (We'll discuss the problem of abiogenesis in more detail in the next chapter, but at this point it's sufficient to note that Hart argues the probability of generating the characteristic molecules of life through the random shuffling of simpler molecules is exceptionally small.) It follows that in an infinite universe there'll necessarily be an infinite number of planets with life, but within any given particle horizon there might only be one planet with life. According to this argument, there's a sense in which there's nothing particularly special about Earth: in an infinite universe there'll be an infinite number of other Earths teeming with life. But within *our* particle horizon—within *our* observable universe—only Earth spontaneously gave rise to life.

As Hart points out, his idea can be falsified in a variety of ways. For example, extraterrestrials could visit Earth, or SETI might succeed and detect signals, or astrobiologists might show that life arose spontaneously on Mars and independently of Earth. Any of these developments would disprove the notion of abiogenesis being a rare, once-in-a-universe event. In the absence of these developments, though, Hart argues that the Fermi paradox leads to a chilling conclusion: we are the only civilization within our particle horizon. Although the universe contains an infinite number of advanced civilizations, for all practical purposes we are alone.

The celebrated physicist Alan Guth has presented[285] a rather different cosmological argument to show we are alone. The argument is based on one of the key underpinning concepts[286] in cosmology: *inflation*. Guth and others came up with the concept of inflation during the 1980s in order to explain several observed features of the universe that are puzzles within the traditional Big Bang picture. The basic idea is that the universe began as a sort of vacuum fluctuation, a small patch of spacetime that underwent a brief period of exponential expansion—inflation—that took it almost instantaneously from a subnuclear-sized object to an object the size of an apple. Once inflation stopped, the "traditional" Big Bang expansionary phase took over. Inflation explains how the universe got to be so big, so smooth, so flat. In addition to explaining these observations (and various other properties of the universe), inflation strongly suggests that our universe is part of a multiverse—there are an infinite number of "local universes" or "bubble universes", of which ours is just one. In the particular bubble universe we inhabit the inflationary expansion ceased after a tiny fraction of a second; in other regions of this vast landscape the expansion continues, spawning bubble universes as it does so. In other words, once inflation starts it never stops; it's eternal.

There are many different specific models of inflation, but it's difficult to avoid the general conclusion that eternal inflation creates vast numbers of universes. Guth considers one model in which there are good reasons to suppose that every second the number of bubble universes is multiplied by a factor of $e^{10^{37}}$—a number that makes the googol look vanishingly small. This is an insanely large rate of universe production: you start with one universe, a second later there are $e^{10^{37}}$ universes, and a second after that you have to multiply by the same factor. It boggles the mind, but it's the sort of picture one has to contemplate when discussing cosmological inflation. And in this picture, young universes vastly outnumber old universes. Assuming this scenario holds true, Guth poses the question: is there another civilization in the visible universe (that is, the bubble universe we inhabit) that's as advanced as ours?

Suppose it takes a definite minimum time $t_{civ}$ to develop an advanced civilization. (It doesn't really matter how we define "advanced" here; equally, although a sharp minimum development time is unlikely to be realistic, we don't have to define a more convincing measure. The numbers involved outweigh these considerations.) Since we exist, the age $t_0$ of our bubble universe must satisfy the constraint $t_0 \geq t_{civ}$. Now suppose an ETC exists somewhere in our bubble universe, and that it's one second more advanced than us. Our bubble universe would then also have to satisfy the constraint $t_0 \geq t_{civ} + 1$ s. However, in the scenario we are considering, there are $e^{10^{37}}$ more bubble universes that satisfy the first constraint than satisfy the second constraint. Since we know we live in a bubble universe that satisfies $t_0 \geq t_{civ}$ we are overwhelmingly unlikely to find that our bubble universe also satisfies $t_0 \geq t_{civ} + 1$ s. The conclusion is that we are alone in our own particular part of the multiverse.

Guth wryly points out that although this argument might explain the Fermi paradox, a more plausible interpretation is that we don't fully understand how to formulate probabilities when discussing the infinity of bubble universes that arise in eternal inflation.

* * *

The cosmology-based arguments of Hart and Guth suggest that there might be infinitely many ETCs in the wider universe, but none with which we can communicate. Effectively, we are alone. The idea that we are alone—the third class of solution to the Fermi paradox—is the subject of the next chapter.

# 5
# They Don't Exist

The final class of solutions to the Fermi paradox are based around the notion that for some reason "they"—extraterrestrial civilizations with whom we might hope to communicate—don't exist.

Within this class of solutions, one can discern different approaches to Fermi's question. Ultimately, though, these solutions depend upon making one or more of the terms in the Drake equation tiny. If a single term is close to zero, or else if several of the terms are small, the effect is the same: when all the terms are multiplied together, the result is $N = 0$. There are no others. The only technologically advanced civilization in the Galaxy, and perhaps the whole universe, is our own.

A couple of the terms in the Drake equation refer to suitable environments. Could it be that truly Earth-like planets are rare? Peter Ward and Don Brownlee, scientists at the University of Washington, wrote a stimulating and thought-provoking book[287] called *Rare Earth*. They presented a coherent argument about why complex life might be an unusual phenomenon. (Strangely, they make no mention of the Fermi paradox.) In this chapter I'll discuss several of the ideas made in *Rare Earth*. Since each of these ideas has been proposed individually as a resolution to the Fermi paradox, I discuss them individually. However, I could equally have grouped them as a single "Rare Earth" solution to the paradox.

Or could it be that technologically advanced ETCs don't exist because life itself is a rare phenomenon? Perhaps the emergence of life from non-living material is an almost miraculous fluke. Or perhaps single-celled organisms are common but the evolution of *complex* life-forms is unlikely to occur. I'll discuss several solutions established on these ideas, but it's worth bearing in mind that the discussions will contain a major limitation: I'll assume throughout that naturally occurring life is carbon-based and requires water as a solvent. Some scientists have argued that other chemicals, notably silicon, could be used instead of carbon; some have even argued that other solvents, perhaps methane, could be used in place of water. Personally—and this may be a failure of imagination on my part—I find it difficult to conceive of a biochemistry that doesn't feature water or carbon. Water in particular I am sure is necessary



S. Webb, *If the Universe Is Teeming with Aliens . . . WHERE IS EVERYBODY?*, Science and Fiction, DOI 10.1007/978-3-319-13236-5_5

for life: find water and you have a chance of finding life. If you believe life can take quite different forms—perhaps as persistent patterns in plasma clouds, or as information-carrying whorls in viscous fluids, or whatever—then these discussions will seem narrow-minded.[288]

We might later discover that some of the solutions I discuss here stemmed from a lack of scientific imagination. But we're in the difficult position of trying to generalize from a single instance—as far as we know, Earth is the only planet with life. It's dangerous to draw conclusions from a sample size of one, but in this case what else can we do? Inevitably we'll be influenced—perhaps biased is a better word—by those factors that seem necessary for our continued existence. We are bound by the Weak Anthropic Principle (WAP), which states that what we can observe must be restricted by the conditions necessary for our presence as observers. Since it's impossible to avoid the WAP in a discussion of the Fermi paradox, it makes sense to begin this part of the book with a Solution based upon anthropic reasoning. Anthropic arguments are rather abstract; most of the later Solutions will be based on more concrete proposals.

## Solution 51 The Universe is Here for Us

*Man is the measure of all things.*

Protagoras

A remarkable argument, which predates Hart's seminal analysis of the Fermi paradox, suggests mankind is probably alone. The argument relies on there being a number of "difficult steps" on the road to the development of a technologically advanced civilization. Examples of potentially "difficult steps" might include the genesis of life, the evolution of multicellular animals and the development of symbolic language. I'll discuss these particular steps in more detail later, but for the present argument the details are unimportant. The argument simply requires there to be some number *n* of critical yet unlikely steps on the road to intelligence, with each step only possible after earlier earlier steps in the sequence have taken place. The eminent evolutionary biologist Ernst Mayr once listed over a dozen such steps;[289] other scientists have suggested the number might be even greater, particularly if certain physical and astronomical coincidences are added to the list. The status of these various steps can, of course, be contested. Some of the evolutionary steps we call "difficult" might not be hurdles at all. We regard a particular evolutionary step as being difficult if it occurred only once in Earth's history, but some steps probably *could* be taken only once—the competition they stimulated would have made a second occurrence redundant. On the other hand, some steps probably were genuinely improbable. For example, if a particular critical step required several otherwise worthless mutations to take place at the same time, then it would make sense to regard the step as a fluke.

Now consider a remarkable coincidence, which lies at the heart of the argument presented below.

On the one hand, the lifetime of our Sun is about 10 billion years. The period over which it can sustain life-bearing planets is almost certainly less than this—some astronomers believe[290] the future evolution of the Sun will cause Earth to become uninhabitable in another billion years or so, and the total "useful" lifetime of the Sun might be as little as 6 billion years. The Earth's biosphere is in its old age. On the other hand, *Homo sapiens* arrived on the scene when the Sun was about 4.5 billion years old. These two timescales—the lifetime of the Sun and the time for the emergence of intelligent life to appear around the Sun—are certainly within a factor of 2 of each other, and could quite easily be within a factor of 1.3 of each other. The near equality of these timescales is remarkable. The two timescales are determined by factors that, either individually or in combination, would seem to have nothing to do with one another. The Sun's lifetime is determined by a combination of gravitational and nuclear factors, while a combination of chemical, biological and evolutionary factors determines the time of emergence of intelligent life. We live in a universe in which timescales span a vast range: many subatomic processes occur on timescales as short as $10^{-10}$ seconds, while many astronomical processes occur on timescales as long as $10^{15}$ seconds. The likelihood that two completely independent timescales have almost the same value is remote. How can we explain this observation without resorting to coincidence?

One solution would be if the evolutionary timescale is much *smaller* than 4.5 billion years. Suppose the typical time for the evolution of intelligent life on an Earth-like planet is just 1 million years. The coincidence of timescales would lessen—but at the expense of making the probability of mankind's recent emergence vanishingly small. After all, if we *could* have emerged just 1 million years after the Earth cooled, then why don't we observe our planet to be 1 million years old? At the very least, why don't we observe it to be 2 million years old, or 3, or 4? Why did it take 4.5 *billion* years for us to appear? This is not a good solution.

The other solution requires the evolutionary timescale to be much *longer* than 4.5 billion years. This accords with Mayr's suggestion of a number of difficult steps in the development of intelligence—"difficult" in this sense meaning that, on a given viable planet, the typical time for a step to occur is long (perhaps longer than the present age of the universe). If several difficult steps must be taken, then we'd not expect to be here at all!

Most people, upon hearing this second solution, tend to dismiss it on the same grounds as the first solution: the probability of mankind emerging recently is small. But the two situations aren't equivalent.

Consider the ensemble of all possible universes. (Whether you consider these universes as somehow "real" or as some sort of mathematical idealization is up to you.) In some universes, unlikely things will occur; a chain of improbable events will happen. In some universes, due to the blind workings of chance, the set of difficult steps leading to intelligence will happen. *And it's precisely such a universe that an intelligent species will observe—with themselves in it.* In other words, we can ignore the possible universes in which we don't exist—since by definition they don't exist for us. We *must* observe in those universes in which the difficult steps have occurred and led to us. Now we can pose the following question. Of all the universes that exist for us, when are we most likely to emerge, given that we can only emerge some time in the 10-billion-year total lifetime of the Sun? (Or, if it happens to be the case, the 6- to 7-billion-year *useful* lifetime of the Sun?) A simple calculation shows that if there are 12 difficult steps, then the most likely time of emergence is after 94% of the star's available lifetime has passed.

Our observations seem to be consistent with the results of this simple calculation. If the Sun were able to sustain life on Earth for 10 billion years, then mankind emerged after roughly 50% of the time available had elapsed. If the Sun can sustain life for only another billion years or so, then mankind emerged after roughly 83% of the time available. This is impressively close to the expected time of arrival.

**The Most Probable Time of Emergence of An Extraterrestrial Civilization**

Suppose there are $n$ difficult steps on the road to the development of a civilization capable of interstellar communication. And suppose these steps must take place over the lifetime $L$ (in years) of a star. A straightforward calculation shows that the most probable time of emergence of a communicating civilization is given by the expression $L/(2^{1/n})$. If there are a dozen difficult steps, so $n = 12$, then the most probable time of emergence is $0.94L$. The calculation doesn't determine exactly when an intelligent species will emerge. It simply states that the median time of emergence, if there are 12 difficult steps to negotiate, is 94% of the star's lifetime.

Finally, we come to the key point. Merely because we've selected a universe in which *we* exist (and how could we select any other type of universe?), we can't infer that *other* intelligent species exist. We *have* to be here because we observe ourselves to be here; but the existence of aliens must contend with probabilities, and the odds aren't good. Another calculation makes this clear. If there are a dozen difficult steps to negotiate on the road to high intelligence, then even under generous assumptions there's only one chance in a million billion of there being another intelligent species in our whole universe. No wonder we don't observe them!

**The Number of Intelligent Species in Our Universe** Suppose there are $n$ difficult steps on the road to intelligence and each step typically requires $d$ years to occur. Furthermore, suppose there are $p$ viable planets, each of which could have supported life for $t$ years. The number of intelligent species out there is given by the expression $p \times [t/(n \times d)]^n$. Let's be generous and suppose every star in every galaxy possesses a viable planet; so $p \approx 10^{22}$. Let's be even more generous and suppose every planet has been viable for about the age of the universe, so $t \approx 10^{10}$ years. However, $d$ must be long: that is, after all, what makes the step difficult. So let's suppose $d \approx 10^{12}$ years—100 times the age of the universe. Finally, let us suppose as before there are a dozen difficult steps, so $n = 12$. If we plug these numbers into the expression above, we find the number of intelligent species out there is $10^{-15}$.

This type of argument for the non-existence of ETCs was first presented by Brandon Carter,[291] who called it an an anthropic argument. (We've encountered anthropic ideas before in this book: the doomsday argument of Gott and Hart's suggestion regarding the improbability of life's genesis have anthropic overtones. We'll meet other examples.) Carter originally concluded that $n = 1$ or 2, in other words that there were only at most a couple of difficult steps. A more recent analysis[292] by Andrew Watson suggested $n = 4$. Carter's use of the term "anthropic" was perhaps unfortunate, since it implies mankind is somehow necessary. All that's needed for the argument to work is that intelligent observers—*any* intelligent observers—self-select their universe. It's just that in this universe it is *we* who make the observations.

The status of anthropic reasoning in science is contentious. Some view it as an abdication of a scientist's responsibility to provide explanations. For example, Smolin's idea of natural selection acting on whole universes (see page 72) is an attempt to move away from anthropic reasoning. Nevertheless, many respectable scientists have employed anthropic ideas in an attempt to explain several features of the universe that seem to be "just right" for the evolution of life. For example, if certain physical constants possessed slightly different values then we wouldn't be here: stars wouldn't shine, or heavy elements couldn't form, or the universe would have collapsed in on itself in a fraction of a second, and so on. The fact of our existence can perhaps, in some way, make sense of these observations (but I think one could equally argue that these "explanations" are essentially trivial). At the very least, an awareness of anthropic reasoning can help us guard against a severe case of observational bias.[293] For example, you'll often hear astrobiologists claim that once life gets going it's extreme resilient—and they support their claim by enumerating the many and varied shocks that the universe has thrown life's way, from asteroid impact to catastrophic climate change. Life on Earth has survived all those shocks, so it certainly *seems* to be robust. But how could we possibly observe it to be otherwise? Any intelligent observer *has* to look back on its evolutionary history and see events that failed to wipe out life; if life *had* been wiped out there'd be no intelligent observers to look back and bemoan the fact. We can deduce little about life's resilience from our single observation of past life on Earth. Indeed, in writing this paragraph I'm aware how I've made it seem that Carter's "difficult steps" argument emphasizes intelligence over other attributes—but the choice of intelligence as the focus is arbitrary and is made purely because this attribute is important to humankind. Carter's model is actually quite general, and could be applied to any series of "difficult steps"—to the possession of peacock-like display feathers, for example, if we thought such feathers were an organism's most important attribute. If the number of difficult steps to reach peacock feathers is the same as the number of difficult steps to reach intelligence then the most probable time of appearance of feathers and intelligence would be the same. Peacocks, however, don't ponder this question.

The literature contains several types of anthropic reasoning, corresponding to several anthropic principles each with different shades of meaning. According to Carter, the weak anthropic principle (WAP) is that "what we can expect to observe must be restricted by the conditions necessary for our presence as observers." The WAP seems almost tautologous. The strong anthropic principle (SAP), on the other hand, is more contentious: "the universe (and hence the fundamental parameters on which it depends) must be such as to admit the creation of observers within it at some stage." Barrow and Tipler, in a classic book, also discuss the final anthropic principle (FAP), which they define as "intelligent information-processing must come into existence in the universe and, once it comes into existence, it will never die out".[294] The mathematician Martin Gardner, in his inimitable way, called this latter version the completely ridiculous anthropic principle (CRAP).

It's interesting to note that Tipler expanded upon the notion[295] of the final anthropic principle in a book entitled *The Physics of Immortality*. He considered the far future of the universe, and was lead to a concept not unlike Teilhard de Chardin's Omega Point. Tipler's analysis showed that, if the universe were to collapse in a Big Crunch, then a future intelligence would find it possible to perform an infinite number of computations. Every being who ever lived could be "resurrected" as a computational simulation. According to Tipler's interpretation of the FAP, the universe *must* be such that it allows this infinite amount of information processing. Now, although Tipler's ideas were attacked as being altogether too speculative (and too overtly religious), his hypothesis at least possessed the virtue of being falsifiable. He made a definite, testable prediction: the universe is closed and will collapse on itself. A few years after his book was published, however, cosmologists discovered that the universe is expanding *more rapidly* as it ages; it might end in a Big Rip but it certainly won't end in a Big Crunch. Tipler, it appears, was wrong; his interpretation of the FAP seems disproven. Perhaps one day soon we'll discover signals from extraterrestrials, or even receive a visit from them. Such a discovery would cast doubt on the WAP and SAP. I leave the reader to decide whether such a discovery is probable.

## Solution 52 The Canonical Artefact

*People's interest is in the product, not in its authorship.*

Jonathan Ive

For the past few decades, physicists have been searching for a "theory of everything"[296]—a unification of the fundamental forces, and the fundamental particles on which those forces act, expressed in a mathematical form. After a vast amount of research into this topic we still have no clear idea of what such a unification would look like, but let's suppose that some major breakthrough has allowed physicists to write down the equations of a final theory—and thus fundamental physics is complete. The theory of everything should be able to answer questions such as: why does the universe contain about $10^{80}$ nucleons? Why is the universe so long-lived ($4 \times 10^{17}$ seconds and counting)? Here's another question that a theory of everything might be able to answer: what is the probability that, in a universe governed by this theory of everything, a life-form possessing advanced intelligence will evolve?

This question was addressed by Gerard Foschini,[297] a scientist who spent his career at Bell Labs—an institution whose research has so far led to the award of seven Nobel prizes for physics. In considering the question of the evolution of advanced intelligence, Foschini assumes the context of an extant theory of everything (the precise form of which has no bearing on his argument) and a set of initial conditions that are identical to those in place when our universe was one second old. In other words, he supposes that up until 1 s after the Big Bang all possible universes evolve in the same way. This means all universes have the same number of nucleons (roughly $10^{80}$) from which to build atoms, the same density so the universe becomes large and long-lasting, and roughly the same large-scale structure. After those given initial conditions, however, the universe can develop in any way consistent with the theory of everything. To repeat: within that development, how likely is it that advanced intelligence will evolve?

Even the most steadfast determinist would surely agree that Foschini's assumed context—a theory of everything plus some initial conditions—has absolutely nothing to say about the likelihood of the universe evolving to contain an author who pens, for example, *Hamlet*. We can't use *Hamlet*, or anything else that's specific to our own cultural and evolutionary development, as a determinant of advanced intelligence: the particular history that led to Shakespeare writing *Hamlet* is so inordinately complex that a theory of everything can't possibly be expected to predict it. However, Foschini argues that there's an object, or rather a class of objects, that *would* act as a flag for the presence of advanced intelligent life-forms. Indeed, Foschini argues that any and all advanced life-forms would inevitably want to develop this object—the canonical artefact—not only because they can but also because it would be more interesting for them to affirm the artefact's existence than its non-existence. The notion of an advanced, intelligent life-form therefore becomes synonymous with a life-form that constructs the canonical artefact and, because it's inevitable that intelligence will construct the canonical artefact, the relevant question becomes: what is the probability that, in a universe governed by the assumed theory of everything and the given initial conditions, the canonical artefact will come into being? This is a question that we can meaningfully discuss. (Note that humans haven't yet constructed the canonical artefact—but we could, and one day we might.)

So—what is the canonical artefact? Well, let's begin by stating what it *isn't*. It can't be a work of literature or music or art, for reasons mentioned above. Similarly, it can't be a technological marvel such as a steam engine (the beings on planet Xymphzyk might be clever but lack the materials to construct a working steam engine) or a codification of some advanced ethical principles (our friends on Xymphzyk might develop ethics that are entirely unrecognizable and, in any case, might feel no need to enshrine them). Instead, Foschini argues that the canonical artefact must be *minimal*—so that wildly different histories after the first second of the universe could still contain the object—and yet be so highly *distinguished* that there's essentially no chance of the object appearing through natural physical processes. Such an artefact could be manufactured by creating a simple object made out of atoms (from the initial conditions, we know that atoms will exist) and whose construction depends on some set $N$ of the positive integers, some set of numbers that has a specific and deep meaning within pure mathematics. Furthermore, the canonical artefact must have a *distinct presence*; in other words, it must endure for some minimum period of time and the atoms of which it is made must be distinct from the surrounding material. This requirement helps us identify the artefact without ambiguity. How big must the artefact be and how long must it endure? Well, if $n$ bits of information are required to express all the numbers of $N$ then a convenient choice for $\eta$, the minimum number of atoms in the artefact, is $\eta = n$. A convenient choice for $\tau$, the minimum lifetime of the artefact, is the time $\tau_y$ it takes a ground state electron to orbit the nucleus of a hydrogen atom—so $\tau = \tau_y \approx 10^{-16}$ s. Foschini then goes on to give one possible example of the canonical artefact.

Foschini takes $N$ to be the ordered list of the orders of the 26 sporadic simple groups (see the box for a brief explanation of what this means). In other words, $N$ is a particular sequence of 26 positive integers that relate to a deep area of abstract mathematics. It's the sort of thing that an advanced, intelligent life-form would know about and understand. The first number in the list is 7920, the second is 95040; the 26th contains 54 digits, so I won't bother to write it out. It takes about 1245 bits of information to express these integers, so following on from the discussion above we can say that the canonical artefact must possess a minimum of 1245 atoms. If we were to demand that these integers be expressed in base 10 then we'd be guilty of provincialism; that humankind routinely employs base 10 for its calculations is thanks to a quirk of evolutionary history that delivered us with ten fingers and thumbs. Foschini argues that a better choice would be the following: for each of the 26 integers in the list calculate the smallest number that is relatively prime to those integers, then express each of the integers in the appropriate base. (Two integers are "relatively prime" if their only common positive factor is 1. For example, the integers 4 and 5 are relatively prime since they are commonly divisible by 1 and nothing else; the integers 4 and 6 are *not* relatively prime since they are both divisible by 2.) For example, 7920 is the first number in the list and 7 is the smallest number relatively prime to 7920. So we express 7920 in base 7, which gives us the first integer for the canonical artefact: 32043. The other 25 numbers in the list are handled similarly.

**The Sporadic Simple Groups** A group in mathematics has a very particular meaning. A group is a set of elements and an operation that can act on any two of those elements; in doing so, four conditions met be satisfied. First there is *closure*—the result of the operation must be an element that is a member of the group; the operation of adding two integers, for example, always generates an integer. Second there is *associativity*—an example of this would be $(1 + 2) + 3$ being the same as $1 + (2 + 3)$; for associativity the order in which the operation is applied doesn't matter. Third there is an *identity*—a unique element such that when the operator acts on it and some other element, that other element is unchanged; with integers under addition the identity is zero (for example, $1 + 0 = 0 + 1 = 1$). Fourth there is *invertibility*—for every element in the group there is another element in the group such that once the operation is applied the identity element results; with the addition of integers, for example, every positive integer has a corresponding negative integer that gives the identity (for example, $1 + (-1) = (-1) + 1 = 0$). Thus the set of integers forms a group under addition. The set of integers doesn't form a group under division, however, because it fails the invertibility test.

The order of a group is simply the number of elements in its set. The order can be finite, if there are a countable number of elements, or it can be infinite.

One of the milestones of mathematics has been the complete classification of entities called finite simple groups. These groups all follow a simple pattern—except for 26 so-called sporadic groups. The smallest sporadic group is called $M_{11}$ and it is of order 7920. The largest sporadic group is called the Monster group and its order is approximately $8 \times 10^{53}$. These groups address several deep problems in mathematics.

Finally we're in a position to construct the canonical artefact, and we're free to use whatever method we prefer. Different life-forms will have different construction preferences—the ocean-living, ameliac creatures of Xymphzyk will use a quite different method to the desert-dwelling, multi-limbed creatures of planet Kyzhpmyx—but that doesn't matter; the main requirement the life-form must possess (in addition to an understanding of the mathematics involved) is sufficient manipulative ability to construct something that belongs to the class of canonical artefacts. Foschini gives the following as one possibility. Imagine threading beads on a necklace, with the beads being identical except for mass: they come with a mass of 1 unit (which represents the number 1), a mass of 2 units (representing the number 2) and so on up to a mass of *m* units (representing the base *m*; this can be used if we need to represent the number 0). The material expression of the number 32043 (in other words, the base 7 version of the first number in the list *N*) is just the appropriate five beads sandwiched between some sort of separator—perhaps a bead that differs in shape or substance or size. We proceed in the same way for the remaining 25 integers on the list, adding the appropriate beads to the necklace and differentiating them by using the agreed separator. At the end we have something that is a canonical artefact. To reiterate: this method of construction isn't the only option. We could use tokens instead of mass-graded beads, for example, as long as the tokens carry their meaning without a reliance on historical information. Three discs would be an adequate representation of the number 3 in base 7; but a disc with the symbol "3" inscribed upon it wouldn't suffice—the symbol makes sense only for those who share our particular history.

We have the canonical artefact—an object we can hold in our hands. So what? Well, we can calculate the probability of the event that the universe constructs the canonical artefact. First, let's estimate how much "room" is available in the universe to construct the canonical artefact. Let's be generous and say the universe is 20 billion years old—about $10^{18}$ s. The second, however, isn't a good unit to use in this context; a more appropriate unit would be the "atomic year", $\tau_y$, which we've said is $10^{-16}$ s. In these units the universe is about $10^{34}$ atomic years old. There are about $10^{80}$ nucleons in the universe so the maximum "room" in which the canonical artefact can be constructed is $10^{114}$ nucleon–atomic years.

Now suppose that the posited theory of everything, combined with the initial conditions, is indifferent as to whether the canonical artefact arises. Let's make the construction of the artefact be as simple as possible, by supposing that the universe is full of mass-graded beads—all that needs to be done is to put the beads in the appropriate order and have that order last for a minimum of one atomic year. Given that we need just over $10^3$ bits of information to represent the 26 numbers of $N$, the universe can contain a maximum of about $10^{114}/10^3 = 10^{111}$ of these beads. However, the universe can contain lots of sequences of 26 elements and we've stated that our $N$ isn't to be favored over the other possible sequences. There are about $2^{1245} \approx 10^{375}$ sequence choices; our $N$ is just one of them. So the probability that the universe constructs the canonical artefact is $10^{111}/10^{375} = 10^{-264}$.

A probability of 1 in $10^{-264}$ is as good as zero. One can modify the argument so as to alter the number $10^{-264}$, but even those modifications that increase it can't make a dent in the conclusion—and some reasonable modifications make the construction of the canonical artefact even *less* likely. If the theory of everything, plus the initial conditions, is indifferent to the construction of the canonical artefact then it's a safe bet that we are alone in our actualization of the universe.

One could argue instead that the theory of everything, combined with the initial conditions, somehow strongly favors the emergence of life-forms that have the ability and inclination to construct the canonical artefact. But for this to hold true requires a 264 order of magnitude effect that's quite unknown to physics.

Or one could argue that a theory of everything, if it exists, can't explain—even in principle—the material expression of thought as expressed by the canonical artefact.

Foschini's argument is unusual, but it's difficult to avoid one of those three conclusions. If the first conclusion holds, we're alone.

## Solution 53 Life Can Have Emerged Only Recently

*To everything there is a season, and a time to every purpose under heaven.*

Ecclesiastes 3:1

The astronomer Mario Livio takes issue with[298] the notion, discussed in Solution 51, that the timescale for the evolution of intelligent life is completely independent of the main sequence lifetime of a star. If the two timescales were related in a particular way—if the evolutionary timescale increases as a star's lifetime increases—then we would *expect* to observe the two timescales to be roughly equal. Carter's gloomy conclusion about the non-existence of ETCs would then not follow. But how can the lifetime of a star influence the timescale of biological evolution?

Livio considers a simple model of how a planetary atmosphere such as Earth's develops to the stage where it can support life. It is *not* a serious model of atmospheric development; rather, it attempts to demonstrate a possible link between stellar lifetimes and biologically relevant timescales.

In his model, Livio identifies two key phases in the development of a life-supporting atmosphere. The first involves the release of oxygen from the photodissociation of water vapor. On Earth, this phase lasted about 2.4 billion years and resulted in an atmosphere with oxygen levels at about 0.1% of present values. The duration of this phase depends upon the intensity of radiation emitted by the star in the wavelength region of 100–200 nm, because only this radiation leads to the dissociation of water vapor.

The second phase involves an increase in oxygen and ozone levels to about 10% of their present values. On Earth, this phase lasted about 1.6 billion years. Once oxygen and ozone levels were high enough, Earth's surface was shielded against ultraviolet (UV) radiation in the wavelength region of 200–300 nm. This shield was important because it protected two key ingredients of cellular life: nucleic acids and proteins. Nucleic acids are strong absorbers of radiation in the wavelength region of 260–270 nm while proteins absorb radiation strongly in the wavelength region of 270–290 nm; radiation in the region of 200–300 nm is therefore lethal to cell activity. A vital condition for the development of land-based life is that an atmosphere develops a protective layer for these wavelengths. And of the likely candidates of a planet's atmosphere, only ozone absorbs efficiently in the wavelength region of 200–300 nm: *a planet needs an ozone layer*. Livio argues that, as on Earth, the timescale for developing an ozone shield against UV radiation is roughly equivalent to the timescale for the development of life.

Different types of star emit different amounts of energy in the UV region. High-mass stars are hotter than low-mass stars and thus emit more UV radiation, but they have shorter lifespans. So for a given planetary size and orbit, the timescale for developing an ozone layer depends upon the type of radiation emitted by the star and thus on the star's lifetime. Following a detailed calculation, Livio argues that the time needed for intelligent life to emerge increases almost as the square of the stellar lifetime. If such a relation holds, then we are *likely* to observe intelligent species to emerge on a timescale comparable to the main-sequence lifetime of a star.

The purpose of Livio's model, to repeat, is simply to show whether a relationship *possibly* exists between stellar lifetimes and the timescale for biological evolution. Even with this proviso, one can still disagree with parts of Livio's argument. For example, his model involves a *necessary* condition for land life to evolve (namely, the development of an ozone layer) but this is not a *sufficient* condition. There are many other steps on the road to the evolution of intelligent life, so even if there is a link between stellar lifetimes and the timescale

**Fig. 5.1** The planetary nebula NGC 7027 lies about 3000 light years away. It's a particularly young object, which only started expanding about 600 years ago. Planetary nebulae such as this produce much of the carbon we observe in the universe. (Credit: NASA)

for biological evolution the relationship might play only a minor role. Nevertheless, encouraged by the discovery of a link between these timescales and the possibility therefore that the existence of ETCs isn't ruled out, Livio is permitted to pose the following question: in the history of the universe, when is a likely time for ETCs to emerge?

If life on Earth is typical of life elsewhere, then most life-forms will be carbon-based. Livio therefore suggests that the emergence of ETCs will coincide with the peak in the cosmic production of carbon. And this is something we can calculate.

The main producers of cosmic carbon are planetary nebulae, which occur at the end of the red-giant phase of average-mass stars. Planetary nebulae shed their outer layers into the interstellar medium, and the material is recycled to form later generations of stars and planets. Since astronomers believe they know the historical rate of star formation (it was higher in the past than it is now, with a peak that occurred billions of years ago)[299] and the relevant details of stellar evolution, they can calculate the rate at which planetary nebulae formed in the past—and thus the rate of cosmic carbon production. According to Livio's calculations, the rate of planetary nebula formation peaked a little less than 7 billion years ago. From this, he argues we might expect carbon-based life to have started when the universe was about 6 billion years old. Since the time required for advanced ETCs to evolve is a significant fraction of a stellar lifetime, we'd expect ETCs to develop only when the universe was about 10 billion years old. If this is the case, then ETCs can't be more than about 3 billion years older than us.

Livio's conclusion has been proposed by some authors as a resolution of the Fermi paradox. These authors suggest that life could have emerged only recently. There are presently no ETCs capable of interstellar travel or communication because, like us, they've had insufficient time to develop. Perhaps the Galaxy will one day be aswarm with interstellar commerce and travel and gossip. For now, though, all is silence.

More recent measurements of stellar formation rates imply that the 3 billion year limit could be a significant underestimate. But even if Livio's conclusion is correct, and there are no ETCs more than 3 billion years older than us, I fail to see how it resolves the paradox. An ETC that's had 3 billion years to develop its technology has had *plenty* of time to colonize the Galaxy or at least announce its presence to the universe. (In the Universal Year, ETCs could have reached our present level of technology at about 1st October.) Unless it can be shown that intelligence is only coming into existence *now*, and that life on Earth is among the most "advanced" in the Galaxy, the arguments don't really address the main thrust of the paradox.

## Solution 54 Planetary Systems Are Rare

*A time will come when men will stretch out their eyes.*
*They should see planets like our Earth.*

Christopher Wren, *Inaugural Lecture, Gresham College*

The arguments given so far in this chapter have all been rather abstract. One can think of more tangible reasons why ETCs might not exist. For example, perhaps there's no place for them to develop.

A common assumption is that complex life requires a planet—preferably Earth-like—on which to originate and evolve. Even if technologically advanced species eventually move away from planet dwelling, the evolutionary ancestors of those species must presumably have started out as planet dwellers. (Some SF writers have explored the possibility of life originating in more exotic locales,[300] including the surface of a neutron star and a ring of gas around a neutron star. Although these fictional accounts are often surprisingly plausible, it remains far easier to imagine such possibilities than it is to demonstrate convincingly and in detail how complex life could originate and evolve anywhere other than on a planet.) When Sagan arrived at his figure of 1 million ETCs in the Galaxy, he assumed there might be as many as 10 planets per star. But perhaps planetary systems are rare, and the $f_p$ term in the Drake equation is small? If $f_p$ were small enough, this alone could explain the Fermi paradox.

It's not *too* long ago that this suggestion was, if not likely, at least conceivable. Now, I present it purely as a historical curiosity. The phenomenal strides made by observational astronomers in the past two decades mean we know for certain that planetary systems aren't rare: at the time of writing there are 1779 confirmed exoplanets, but by the time you read this there'll be many more. It's likely that most stars have planets orbiting them.

So if planetary systems are common, why did some astronomers argue until relatively recently that a scarcity of planets might explain the Fermi paradox? Well, even when I was a student—and I'm not *so* advanced in years—astronomy textbooks could still present two competing scenarios for planetary formation.[301] In the first, a planetary system such as ours was pictured as forming in a catastrophic event. In the second, planetary systems were thought to condense out of nebulae.

The nebular hypothesis feels a rather more "natural" explanation than the catastrophe hypothesis, but it seems to possess a fatal flaw. If the Sun, for example, formed from the collapse of a rotating cloud of dust and gas, then calculations show that it should now rotate extremely quickly: the Sun should contain most of the angular momentum of the Solar System. The Sun, however, rotates rather sedately—its equatorial regions rotate once in about 24 days, while its polar regions rotate once in about 30 days. Most of the angular momentum of the Solar System resides in the planets. This observation led many astronomers to prefer models of planetary formation based on catastrophic events. The most popular model had a star almost colliding with the Sun; tidal effects pulled a gaseous filament from the Sun, and the filament later broke up and condensed to form the planets.[302]

If planets really did form in stellar collisions, then the outlook for finding ETCs would be bleak. The density of stars in space is quite low, so collisions would be infrequent; one early estimate put the number of planetary systems formed in this way at just ten per galaxy! In a lecture in 1923, the celebrated mathematician James Jeans said: "Astronomy does not know whether or not life is important in the scheme of things, but she begins to whisper that life must necessarily be somewhat rare." Jeans clearly thought he knew the resolution of the paradox, and the paradox had not yet been formulated.

However, the nebular hypothesis never went away. Theories of planetary formation based upon collisions also possessed problems. The collision theory

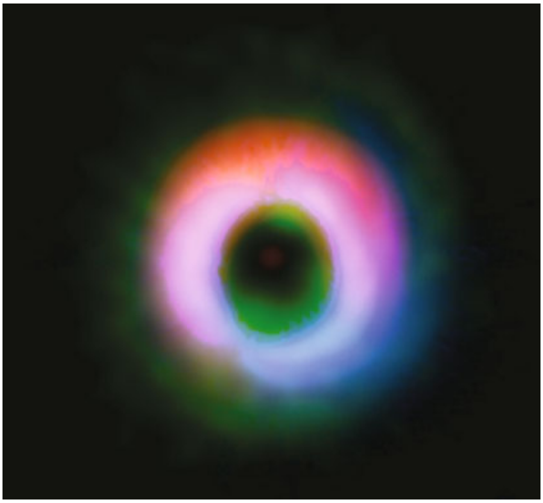


**Fig. 5.2** In 2014, astronomers using the Atacama Large Millimeter/submillimeter Array (ALMA)—one of the most impressive telescopes on the planet—discovered this dusty protoplanetary disk around a youthful star called HD142527. The star is about 457 light years distant from Earth. The high density of dust (shown red) in the northern side of the disk suggests that planets are being formed there right now. Perhaps, in a couple of billion years, the planets now forming will be home to life. (Credit: ALMA/ESO/NAOJ/NRAO/Fukagawa et al.)

failed to explain many observed properties of our Solar System. Furthermore, the major difficulty with the nebular hypothesis—namely, explaining how the bulk of the angular momentum of the Solar System resides in the planets—was eventually resolved. It happens that the young Sun *did* rotate at high speed, but the rotation generated a strong magnetic field. Magnetic lines of force stuck out into the solar nebula, like spokes from a hub, and dragged the gas around with it. This "magnetic braking" effect slowed the Sun, and transferred angular momentum to the gaseous disk. Astronomers observe direct evidence for this: young stars rotate up to 100 times more quickly than the Sun, whereas old stars rotate more sedately. We can now be certain that the planets in our Solar System formed when small planetesimals condensed out of a disk-shaped cloud of dust and gas; in gentle collisions, these planetesimals stuck together and gradually formed the planets we see today. The same process occurred and is occurring around other stars. Planets are common, as Sagan believed.

Astronomers have taken images of protoplanetary disks (see for example Figure 5.2). They've even taken images of planets around distant stars (see for example Figure 5.3, which is a staggering technical achievement: planets shine only by reflecting their star's light, so taking an image of an exoplanet is rather like trying to observe the light of a firefly next to a bomb. The discovery of

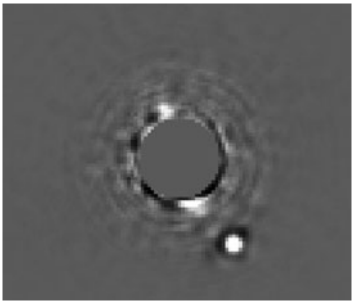


**Fig. 5.3** Astronomers can directly image some exoplanets. In 2014, the Gemini Planet Imager saw first light: this is an infrared image of Beta Pictoris b. Radiation from the central star, Beta Pictoris, is blocked by a mask so that it doesn't drown the reflected light from the planet. Beta Pictoris b is about 63.4 light years distant from Earth, is rather larger than Jupiter, and formed only recently: it's about 10 million years old. (Credit: Processing by Christian Marois, NRC Canada)

exoplanets on a large scale, however, has been achieved not by direct imaging but by observing the effects that planets have on their host star. For example, the gravitational attraction of a large planet on a star causes the path of the star to "wobble" as the planet completes its orbit; if the orbital plane is edge on to our line of sight then astronomers can detect the regular to-and-fro motion of the star through Doppler shifting of its spectral lines. And if a planet happens to transit the star—in other words, if it moves in front of the star's disc as seen from Earth—then there's a tiny, but nevertheless measurable, dip in brightness. These techniques for exoplanetary discovery[303] have been tremendously successful; the NASA *Kepler* mission, in particular, has been particularly fruitful.

Clearly, then, one can't explain the Fermi paradox by arguing that planetary systems are rare. A few decades ago the argument was plausible; advances in astronomy have shown that it's wrong. It's now clear that there are hundreds of billions of planets in the Galaxy. There are many potential homes for life.

## Solution 55 Rocky Planets Are Rare

*Here's metal more attractive.*

William Shakespeare, *Hamlet*, Act III, Scene 2

Thanks to the *Kepler* space mission and various ground-based exoplanet-hunting initiatives, we now know that planetary systems are common. The majority of confirmed exoplanets are much larger than our planet, with about three quarters of them having a radius at least twice that of Earth, but this is hardly surprising because the two commonest techniques for finding exoplanets—the radial velocity (or "wobble") method and the transit (or "brightness-dip") method—are more sensitive to the presence of a big planet such as Jupiter than a small planet such as Earth. Small representatives from any class of physical object tend to outnumber large members of the class, though, so it seems certain there are many Earth-size planets—it's just that we aren't so good at detecting them. Nevertheless, even if the Galaxy contains a host of Earth-size planets, are those planets necessarily going to be terrestrial in nature? This is an important question because in order to develop the sort of technology that enables interstellar travel, or at least interstellar communication, a civilization presumably requires access to workable lodes of metallic ore. (Several SF authors have critically examined this assumption, and based thought-provoking stories set on planets that lack the ores to which we have access, but it's difficult to see how an ETC could fashion a radio telescope from rock, water and organic material.) Could it be that Earth is special, in that its rocks contain lots of metal?

In order to answer that, we need to think about how our Solar System came into being and consider whether its birth might have been in some way peculiar.

As far as we know, the only surviving witnesses to the birth of the Solar System are a group of metal-rich meteorites called chondrites. Inside certain types of chondrite one can find calcium–aluminum-rich inclusions (CAI). These are pockets of minerals ranging in size from less than a millimeter up to a centimeter. One can also find chondrules. These are small spherical inclusions, typically 1 to 2 mm in diameter, and are composed mainly of the silicate minerals olivine and pyroxene. (The name "chondrule", and thus "chondrite", comes from their visual appearance: the Greek word *chondros* means "grain" or "seed".) Using the known decay rates of various radioisotopes, planetologists can calculate when CAIs and chondrules formed. The best estimates imply that CAIs and the oldest chondrules formed about 4.567 billion years ago—slightly before the formation of Earth itself.[304]

Chondrites occasionally fall to Earth, and when they do they are studied intensively. Indeed, chondrites have been studied for centuries, and much is now known about their chemical and physical makeup. At least one mystery remains, however: the precise nature of the chondrules.[305]

What seems clear is that chondrules must have been flash-heated to temperatures of 1000 K or more, and then cooled quickly. But what could have caused the heating? Scientists have proposed an embarrassingly large number of hypotheses to explain chondrule formation, including shock wave heating caused by disturbances in the protoplanetary disk and lightning discharges through dust balls, but there is as yet no commonly accepted explanation.

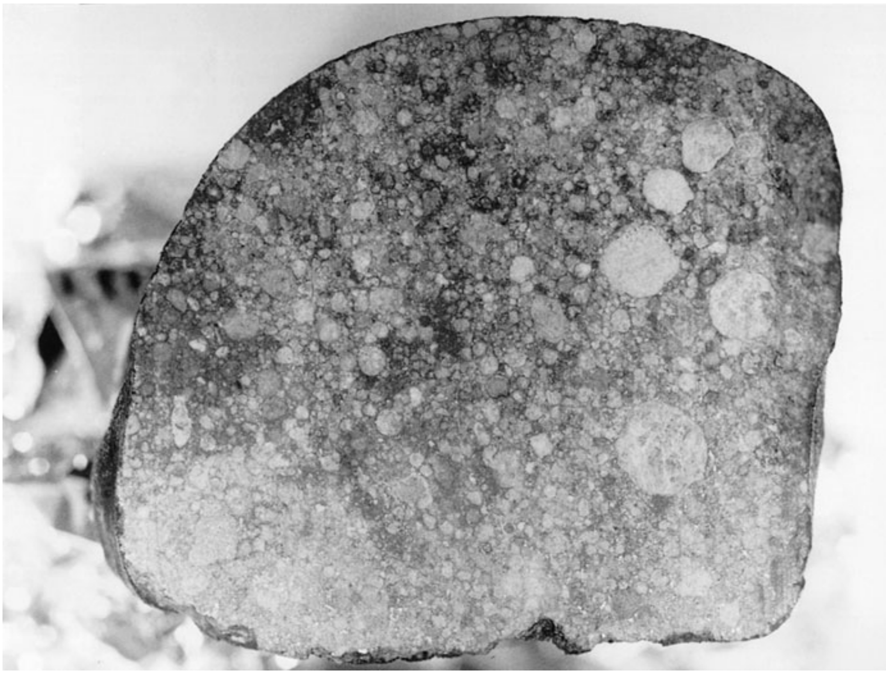


**Fig. 5.4** Chondrules are spherical inclusions of silicate in chondrites; their origins remain a matter for debate. Chondrules are clearly visible in this cut surface of the AH 77278 chondrite. This specimen, which is 8 cm across, was found at Allan Hills—a group of mainly ice-free hills in Antarctica. Since the Hills were first mapped in 1957, many interesting meteorites have been found there. (Credit: NASA)

(That's not too surprising. After all, chondrules formed a *long* time ago and, since they appear in no other type of rock, geologists have no other specimens with which to compare them.) Another suggestion is that 4.567 billion years ago a brief flash of heat spread through the Solar System, fusing dust and forming chondrules. The Irish astronomers Brian McBreen and Lorraine Hanlon propose that a nearby gamma-ray burst[306] (GRB) might have provided the heat. Suppose a GRB occurred within 300 light years of the nascent Solar System. It would have pumped enough energy into the protoplanetary ring of dust and gas to fuse up to $6 \times 10^{26}$ kg of material (100 times the Earth's mass) into iron-rich droplets, which would quickly cool to form chondrules. The chondrules would then absorb gamma-rays and X-rays from the GRB.

In the McBreen–Hanlon scenario the Solar System could be a rarity in possessing chondrules: the formation of chondrules would require a GRB to be in relatively close proximity to a protoplanetary disk at a critical moment in its development. The significance of this is that high-density chondrules, which could have settled quickly into the plane of the protoplanetary disk, might have aided the formation of rocky planets in the Solar System. In other words, in this scenario planetary systems similar to our own—complete with rocky terrestrial planets—would be scarce. And, with only a small number of Earth-like planets on which to develop, ETCs might be rare.

The idea that chondrule formation was initiated by a GRB is interesting. However, other suggestions seem to offer more plausible mechanisms for making chondrules. Moreover, the most accurate dating[307] of radioisotopes within chondrites suggest that CAIs were formed over a brief interval some 4567.3 billion years ago, and that the chondrules were formed over a three million year period beginning at the time of CAI formation. The three million year timescale is similar to the lifetime of protoplanetary disks, and so it seems likely that CAI and chondrule formation is due to some intrinsic process in the development of disks. The results of this research, if confirmed, imply that there was nothing particularly special about the birth of our Solar System. So, as a solution to the Fermi paradox, this perhaps isn't high among the list of contenders.

## Solution 56 A Water Based Solution

*Thousands have lived without love, not one without water.*

W. H. Auden, *First things first*

Life requires water. (At least, "life as we know it" requires water.) It's an almost magical liquid. For a start, pretty much everything is soluble in water: the liquid can transport the substances dissolved within it and thus convey materials around cells, organisms and ecosystems. It has the unusual property of expanding when it freezes, meaning that ice floats on water; if water instead contracted upon freezing then the seas and lakes in cold climates would gradually fill with ice that had fallen to the bottom—a scenario that would cause problems for aquatic life. The large temperature range over which water remains liquid, combined with water's large heat capacity, mean that the oceans moderate Earth's climate. Enzymes—proteins that catalyze chemical reactions and, without which, certain biological processes would occur on timescales measured in millennia rather than milliseconds—require water in their structures. One could go on and on: water is necessary for terrestrial life—and it's not too much of a stretch to say that it's a fundamental requirement of all life. Earth of course has oceans of the stuff. But the Moon doesn't have oceans; rivers might once have flowed on Mars, but it's a rather desiccated place nowadays; both Venus and Mercury are arid planets. Could it be that

Earth is exceptional in possessing so much liquid water? If it turns out that a rocky planet is unlikely to be home to water oceans then we might have a part solution to the Fermi paradox.

How did Earth get its water? This remains a controversial question. One leading suggestion is that 3.85 billion years ago Earth suffered an intense cometary barrage; it was Oort Cloud comets that delivered water to our planet—water that we still drink every day. At first glance this suggestion makes sense. Some planetologists argue that the very early Earth would have been too hot to hold on to large oceans of water, so the water we have now must have been delivered from space; and since we know that cometary nuclei contain ice, and that the Solar System contains trillions of comets, it's not too difficult to imagine how a cometary bombardment could have watered Earth. If such a water-bearing bombardment did indeed occur, the question arises: what could have caused it? If the bombardment arose from some sort of one-off cataclysmic event then the presence of water on Earth would be a fluke. Replay the tape of planetary evolution and Earth might end up dry. Rocky planets with water might be the exception.

However, before we conclude that ours is the only planetary home with running water, we need to address a couple of problems with the notion that comets watered Earth.

The first problem is that cometary water seems to be different to water here on Earth. A water molecule consists of one oxygen atom and two hydrogen atoms—$H_2O$. Now, the nucleus of a hydrogen atom usually contains a single proton; it's possible, though, for a hydrogen nucleus to contain one proton and one neutron. This form of hydrogen is called deuterium. The ratio of normal hydrogen to deuterium in a water sample acts as a "fingerprint" of that water. It turns out that the deuterium abundance in comets such as Hale–Bopp, Halley and Hyakutake is about twice the abundance we see in Earth's oceans. If these three bodies are typical of Oort Cloud comets then it's difficult to see how they could have delivered Earth its oceans. However, the deuterium abundance in asteroids and planetesimals—small objects that were abundant early in the history of the Solar System and that would have collided and adhered to make the proto-Earth—is the *same* as we see in our oceans. Earth and planetesimals contain the same sort of water. Perhaps planetesimals are a more likely source of water than comets?

The second problem is that geologists now have evidence for the presence of water at very early times. The chronology of the early Solar System is becoming increasingly refined. We know that the first solids in the protoplanetary disk, the pebbles and boulders that would collide to form Earth, condensed 4.568 billion years ago. Just 164 million years after that, at a time 4.404 billion years ago, a mineral called zircon[308] crystallized in Earth's crust. A detailed

**Fig. 5.5** The oldest fragment of Earth's crust: a speck of zircon, extracted in 2001 from sandstone from the Jack Hills region of Western Australia. The speck is only about 200 by 400 microns in size—about the size of the period at the end of this sentence. Uranium atoms in the zircon decay into lead atoms a rate that is known precisely. If researchers can measure the amount of uranium and lead in the zircon they can determine the crystal's age. This one is 4.404 billion years old. (Credit: John Valley, University of Wisconsin-Madison)

analysis of those zircons shows they were created in the presence of water. So at the very earliest times in Earth's history—hundreds of millions of years before a cometary bombardment event, and soon after the Moon-forming impact—there appears to have been continental crust and water.

A picture is emerging, then, of water-containing planetesimals giving birth to a wet Earth. The young Earth suffered many giant impacts, but it seems these collisions didn't boil the water off into space. The water went into the atmosphere and later, as the atmosphere cooled, it condensed to form oceans. A cycle of boiling and condensing might have happened several times. Nevertheless, this picture is subject to debate and revision, as are most of the interesting questions in science. In 2011, for example, astronomers used the *Herschel* Space Telescope to measure the deuterium abundance in the comet Hartley 2; they found the same ratio of deuterium to hydrogen as water here on Earth. In 2013, they followed this with a similar measurement of the comet Honda–Mrkos–Pajdušáková; they saw the same abundance.[309] These are both comets from the Kuiper Belt, so it raises the possibility that it was *these* objects, rather than Oort Cloud comets, that brought water to Earth (or, as perhaps is more likely, delivered some fraction of Earth's water). Geologists will surely learn more about the origin of our oceans in the next few years.

At present, however, one can plausibly argue that water oceans are a natural outcome of the process that forms rocky planets. It's premature to conclude that Earth is unique in possessing oceans of life-giving water.

## Solution 57 Continuously Habitable Zones Are Narrow

*Give me more love or more disdain;*
*the torrid or the frozen zone.*

Thomas Carew, *Mediocrity in Love Rejected*

Even if rocky planets form readily around stars, and even if those planets have plentiful amounts of $H_2O$, one could argue that another condition must be met before life as we know it can survive for the billions of years needed for a technological civilization to develop: a terrestrial planet must be in a system's habitable zone (HZ)[310]—the region around a star in which an Earth-like planet could support *liquid* water. It's often called the Goldilocks zone, for obvious reasons. The location of the inner edge of the HZ is set by the point at which a planet loses water due to the high temperatures close to a star; the outer edge is set by the point at which water freezes. This definition of the habitable zone excludes objects of legitimate astrobiological interest. For example, a planet's internal heat might maintain subsurface liquid water far away from the HZ; tidal heating might allow the existence of liquid oceans on the moons of large planets; a "tilted" terrestrial world[311] whose obliquity fluctuated because of the gravitational influence of its star and nearby gas giant planets might a possess climate that prevent glaciation even at large distances from the star. Just as life is not restricted to Earth's surface so might life be possible in these unusual environments. Nevertheless, if we are interested in the existence of technologically advanced civilizations then it would seem to make sense to concentrate on the conventional habitable zone. Modern thinking is that we should also concentrate on planets that have less than about 1.5 times the radius of Earth. If a planet is much larger than this it tends to accumulate a thick atmosphere of hydrogen and helium, which means it resembles a gas giant rather than a terrestrial planet.

It's far from straightforward to calculate the precise location of the HZ boundaries: the inner boundary depends on a runaway greenhouse effect while the outer boundary is determined by the formation of $CO_2$ clouds that act as a sort of "blanket" blocking stellar radiation. Thus calculations of the HZ width, and particularly calculations of the location the outer boundary, require the use of sophisticated climate models. Various estimates have been made

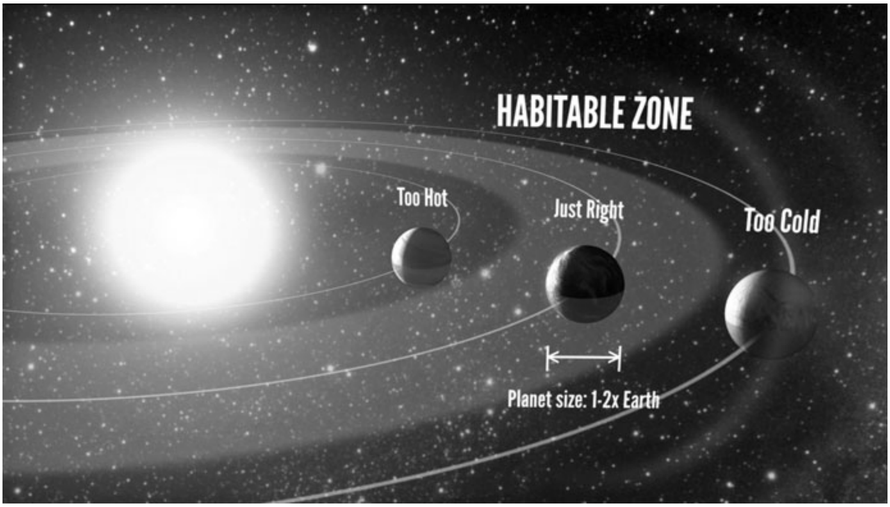



**Fig. 5.6** If a planet orbits too close to a star then it will be too hot to possess liquid water. If a planet orbits too far from a star then it will be too cold to possess liquid water. A planet (with a size that's not too small and not too large) has to be in that "just right" Goldilocks zone for it to have a chance of retaining liquid oceans, and thus to have a chance of possessing life as we know it. (Credit: Petigura/UC Berkeley; Howard/UH-Manoa; Marcy/UC Berkeley)

for the HZ of our Solar System; one recent study[312] gave a range of 0.77–0.87 AU for the inner boundary and 1.02–1.18 AU for the outer boundary, but different estimates exist. If one accepts this particular estimate then our neighbor Venus, with a mean distance of 0.723 AU from the Sun, lies slightly outside the habitable zone; Mars, with a mean distance of 1.524 AU from the Sun, lies well outside the habitable zone. Only Earth, the Goldilocks planet, occupies just the right place.[313]

There's more to the story. Michael Hart pointed out that the habitable zone around a star varies with time. Main-sequence stars become brighter and hotter as they grow older, so the HZ moves outward as a star ages. What's important, according to Hart, is the *continuously* habitable zone (CHZ).

Typically, the CHZ is defined as the region in which an Earth-like planet can support liquid water for 1 billion years—the sort of timescale presumably required by evolution to develop complex lifeforms. In the case of the Solar System, the CHZ has existed for 4.5 billion years, and Earth has been fortunate enough to be in the middle of the zone. Clearly, though, the CHZ must be narrower than the HZ. In the late 1970s, Hart published the results of computer models[314] that seemed to show that the CHZ is extremely narrow. In Hart's models the CHZ was widest around G0 main-sequence stars (the Sun is a G2 star) and shrank to zero at K1 stars (which are cooler than the Sun) and

F7 stars (which are hotter). A K1 star would typically be 0.8 times as massive as the Sun and an F7 star might typically possess 1.2 times the Sun's mass, so according to Hart there was only a limited range of stars that possess a CHZ at all. Furthermore, where a CHZ did exist it was always narrower than 0.1 AU. For the Solar System, for example, he calculated an inner edge of the CHZ at 0.95 AU and an outer edge at 1.01 AU. With such a restricted amount of CHZ "real estate" one might expect Earth-like planets—those that can support life over billions of years—to be much rarer than is commonly supposed.

While Hart's finding didn't *prove* that there were no ETCs, it clearly had a bearing on the Fermi paradox. If the number of potential life-bearing planets is much smaller than most estimates suppose, then the number of potential ETCs out there must also be smaller. Depending upon the values of the other factors in the Drake equation, the total number of communicating civilizations might be reduced to one: us.

**An Exoplanet in the Habitable Zone?** As I write this section, astronomers have announced the detection of the most Earth-like planet[315] yet discovered in a habitable zone. Kepler-186f has a radius just 10% larger than Earth's and, although it's composition is unknown, it seems likely to be a rocky world. The planet receives about a third of the heat energy that Earth gets from the Sun, and it orbits its star once every 130 days. Four other planets in the system are all too close to the star for liquid water to exist. Kepler-186f is in the habitable zone, but is it habitable? The star is of spectral class M, so the planet is probably on the receiving end of violent flare activity. It's also quite possible tidally locked. Personally, I'd bet against this being a home for advanced lifeforms.

SETI astronomers using the Allen Telescope Array have already searched for radio communication from the Kepler-186f system. They heard nothing.

The situation might not be as bleak as Hart argued, however. If one wants to investigate habitable zones today one has access to more powerful computers than Hart's; one can employ more sophisticated models of the Earth's early atmosphere; and one can take account of phenomena not known to Hart, such as the recycling of $CO_2$ by plate tectonics. The results are encouraging for those who would believe in the existence of ETCs (or at least in the existence of planetary homes for ETCs). For example, models developed by James Kasting[316] and co-workers suggest that the 4.5-billion-year CHZ for the Solar System extends from 0.95 AU to 1.15 AU—about four times larger than the range calculated by Hart. Other scientists believe the Solar System's CHZ might be even wider. The CHZ around other stars, too, could be wider than Hart thought.[317]

So: how likely is it that a given planetary system will have a planet that lies within the CHZ? It wasn't so long ago that such a question was purely theoretical, to be answered purely on the basis of computer models. As mentioned in Solution 54, one of the great advances in astronomy in recent decades has been the development of techniques to detect exoplanets and so now we can add observational data to the mix. The answer appears to be that, for Sun-like stars, finding a planet in the continuously habitable zone will be not in the least unusual. Indeed, an analysis of data from the *Kepler* mission and the Keck Observatory suggests that about one-in-five Sun-like stars[318] will have an Earth-size planet in the habitable zone; that means the Galaxy might contain *billions* of Earth-sized planets in the habitable zone of Sun-like stars. Note that just because a planet lies in the habitable zone it doesn't necessarily follow that it's habitable: there are many reasons why a planet in the Goldilocks zone might lack liquid water. But the finding does suggest that ours is unlikely to be the only Sun-like star orbited by a planet on which water can remain in the liquid state.

What about stars that aren't Sun-like? Planets around hot stars of type O, B and A won't stay long in the habitable zone because the stars themselves evolve in luminosity too quickly. But the vast majority of stars in the Galaxy are small, cool stars of type K and M; what about them? Hart argued that such stars won't possess habitable planets because the HZ lies so close to the star that any planets in the zone will become tidally locked. (One side of a tidally locked planet always faces the heat of its star while the other side always faces the cold of open space.) The assumption was that conditions on a tidally locked planet wouldn't permit the existence of large amounts of liquid water and thus the planet would be uninhabitable. Furthermore, the early stages in the life of a small star are marked by huge variability: sometimes they become dim, at other times they emit violent flares. That variability is thought to be inimical to life. However, some climate studies show that oceans or wind currents might mitigate the temperature extremes of a tidally locked planets, and the flare activity might not be the show-stopper we think it is. There are so many small stars, and they shine for such a long time, that perhaps the total amount of CHZ "real estate" is greater around these stars than it is around Sun-like stars. If that is indeed the case, then there could be vast numbers of planets in the continuously habitable zone.

When discussing habitable zones there's a further point to take into consideration. As we shall see in later Solutions, only certain types of star have sufficient metallicity to possess terrestrial planets, and only certain parts of the Galaxy are sufficiently protected from the violence of the central regions. We perhaps need to define a galactic habitable zone[319] (GHZ)—which is an annulus containing perhaps as little as 20% of the stars in the Galaxy. For complex life to evolve, a CHZ must be within the GHZ—and this reduces the possibilities. Nevertheless, it's difficult to see how the numbers can be reduced by a factor that would help address the Fermi paradox. The assumption has to be that the Galaxy contains plenty of planetary homes for life.

## Solution 58 Earth is the First

*'tis not the king's stamp can make the metal better or heavier.*
William Wycherly, *The Plaindealer*

Soon after the Big Bang, the universe contained essentially only hydrogen and helium (in the ratio 75% to 25%). There were small amounts of lithium, and even smaller traces of beryllium and boron, but that was all. To an astronomer, then, the universe consists of hydrogen, helium and everything else; all the elements heavier than hydrogen and helium—the "everything else"—are called metals. Now, the biochemistry of terrestrial organisms, and the biochemistry of any extraterrestrial organisms we can plausibly imagine, depends crucially on six elements: hydrogen (H), sulfur (S), phosphorus (P), oxygen (O), nitrogen (N) and carbon (C). In astronomical terminology, therefore, life depends upon hydrogen and the five metals SPONC. However, none of those metals essential to life were there at the beginning of the universe. Where did they come from? The heavier elements were all cooked in nuclear reactions inside stars, and became part of the interstellar medium only when stars reached the end of their energy-producing life. As time goes by, the concentration of metals in the universe slowly increases.

One resolution of the paradox—often proposed and similar in spirit to Livio's suggestion in Solution 53—is that the heavier elements only recently became sufficiently concentrated in the interstellar medium to allow life to form. Planets around older stars, it is suggested, lack the metals SPONC. Only around quite young stars—stars such as the Sun—can life arise. So humankind is inevitably among the first technological civilizations to arise. Perhaps it's *the* first.

The suggestion that the chemical enrichment of the Galaxy resolves the Fermi paradox *by itself* is surely too strong. As with many suggestions, this one might play a role—but it's unlikely to stand alone as a resolution of the paradox.

One issue with the suggestion is that we don't know what metallicity might be required of a star for it to possess viable planets. (A star's metallicity simply refers to the amount of heavier elements in its chemical make-up.) Would an abundance of heavy elements three quarters of that present in the Sun suffice? One half? One quarter? We don't really know. An analysis of the exoplanets[320] discovered by the *Kepler* mission implies that the formation of small, terrestrial planets doesn't require an environment rich in metals: such planets are as likely to form around low-metallicity stars as they are around high-metallicity stars. If life can develop on planets possessing a smaller abundance of heavy elements than is present in our Solar System, then ancient stars could have been spawning grounds for civilizations.

A second issue is that the relationship between age and metallicity in stars is rather more complicated than at first appears. It's possible for a star to be much older than the Sun and yet possess the same abundance of heavy elements. Consider, for example, the star HIP 102152.[321] It lies about 250 light-years away. The star is of stellar class G3V and it has a surface temperature of 5723 K; for comparison, the Sun is of stellar class G2V and it has a surface temperature of 5778 K. Put them side by side and these stars would look like twins. Furthermore, astronomers have detected the presence of 21 chemical elements within HIP 102152 and found their abundances to be similar to those found in the Sun. These really are stellar twins. And yet HIP 102152 is about 3.6 billion years older than the Sun. So even if high metallicities are a requirement for life as we know it, those conditions have been available for a long time. Our Sun is not the first.

It's not yet known whether terrestrial planets orbit HIP 102152, but an Earth-twin might be there. And intelligent creatures might have evolved on it. If those creatures looked up during the day, they'd see pretty much what we see: a yellow sun dominating the sky. Those creatures could be much older than us; they could have been enjoying that view for a billion years or more. During all that time, those oceans of years, wouldn't they have moved out and sought a different view? Wouldn't the creatures of HIP 102152 at least let others know of their existence?

## Solution 59 Earth has an Optimal "Pump of Evolution"

*When resonance occurs, a small input force can produce large deflections in a system.*

Report on the collapse of the Tacoma Narrows Bridge

Jupiter plays a role in various proposed resolutions of the Fermi paradox. The particular suggestion I discuss here is due to the physicist John Cramer.[322]

We know that large rocks sometimes hit Earth. But where do they come from? One idea is that they fall in from the Asteroid Belt and occasionally happen to strike Earth—but for this idea to work, large numbers of asteroids must be perturbed from their stable orbits and then fall towards the inner part of the Solar System. Why should asteroids be pushed away from their stable orbits? No plausible mechanism was known until, in 1985, George Wetherill highlighted the importance of the gap in the Asteroid Belt[323] at a distance of 2.5 AU.

Kirkwood gaps—regions in the Asteroid Belt in which relatively few asteroids can be seen—were already well known. The gaps occur because of resonance effects. In the case of the gap at 2.5 AU, the resonance occurs because any asteroid at that distance orbits in precisely 1/3 of the time Jupiter takes to orbit the Sun. Therefore, every third occasion a 2.5-AU asteroid reaches a particular position, Jupiter is in the same relative position. Jupiter gives the asteroid a gentle gravitational nudge, always in the same direction, and the effect is cumulative. It's like pushing a swing at precisely the right frequency: the effects build up and the amplitude of the swing increases. Over time, therefore, the orbit of an asteroid at 2.5 AU becomes unstable, and it moves away—and the Asteroid Belt is eventually cleared of objects in this region. Any asteroid wandering into this region from elsewhere is eventually ejected by the same mechanism. The Kirkwood gap at 2.5 AU is due to a 3:1 resonance; other gaps, based on other resonances with Jupiter, also exist.

Where do the asteroids go after they are ejected from the Kirkwood gap at 2.5 AU? Calculations show there's a high probability of their orbits crossing the orbit of Earth. In other words, there's a chance that these asteroids hit Earth—with catastrophic consequences.

However, although the effects of an asteroid impact can be disastrous for any creatures that happen to be around, in the *long* run the impacts might be beneficial for some species. After all, if the meteor impact of 65 million years ago hadn't happened, mammals might still be scraping a living at the margins of a lizard-dominated world. Cramer points out that there might be geological periods when nothing much happens to species; evolution appears to take the commonsense attitude of "if it ain't broke, don't fix it". It's primarily at crises points, when for some reason the environment changes, that evolution works quickly and new species arise to take advantage of altered conditions. Evolution, in Cramer's words, seems to be "pumped" by cycles of crises and stability. And, he suggests, an ideal pump is one that drives evolution through major crises every 20 to 30 million years. Asteroids from the 3:1 Kirkwood gap might be pumping at exactly the right rate.

If Cramer's idea is correct—and he'd be the first to admit that the idea is speculative—it constitutes another reason why life on Earth might be special. Not only might life require an Earth-like environment, the environment might need to occur in a system with planetary masses and orbits that produce a resonance in an Asteroid Belt at just the right rate. If the "pump of evolution" runs too fast—and asteroids hit a life-bearing planet too often—then life never has a chance to evolve intelligence. If the pump runs too slow—and asteroids hit a life-bearing planet too rarely—then life becomes stuck in a rut. The result is a planet full of trilobites or cockroaches or dinosaurs (or, more likely, creatures differing from terrestrial creatures in a myriad of fascinating ways). As

**Fig. 5.7** A montage of images of Eros. The images were taken over three weeks as the NEAR spacecraft approached the asteroid. Near-Earth asteroids such as Eros are relatively few in number. Most asteroids are in the "main belt", orbiting the Sun in a torus between Mars and Jupiter. It's these "belt" asteroids that can be perturbed from their orbits by the gravitational influence of Jupiter—with potentially devastating results. (Credit: NASA)

long as these creatures were successful, in an unchanging environment there'd be no "need" for them to adopt new modes of behavior, no "need" for them to develop intelligence and thence radio telescopes or starships.

The 3:1 resonance in the Asteroid Belt is due to Jupiter. The very existence of the Belt is also due to Jupiter: the asteroids are the remnants of a protoplanet whose formation was aborted because of Jupiter's own formation. If there is such a thing as a "pump of evolution", and if it's tuned to just the right frequency in our planetary system, then we have Jupiter to thank for it.

## Solution 60 The Galaxy is a Dangerous Place

*I am become death, the destroyer of worlds.*

Bhagavadgita

Violent phenomena are common in the universe and pose a variety of threats to civilization. For example, it has been estimated that a million black holes could be wandering through interstellar space. If one of them were to wander into a planetary system it might devour the planets (see Figure 5.8). A magnetar (a type of neutron star) would pose an interesting threat[324] if it came too close. For example, in the summer of 1998, several orbiting detectors recorded radiation from the magnetar SGR1900+14. The radiation was so intense it forced the shutdown of some satellites; the radiation came within 30 miles

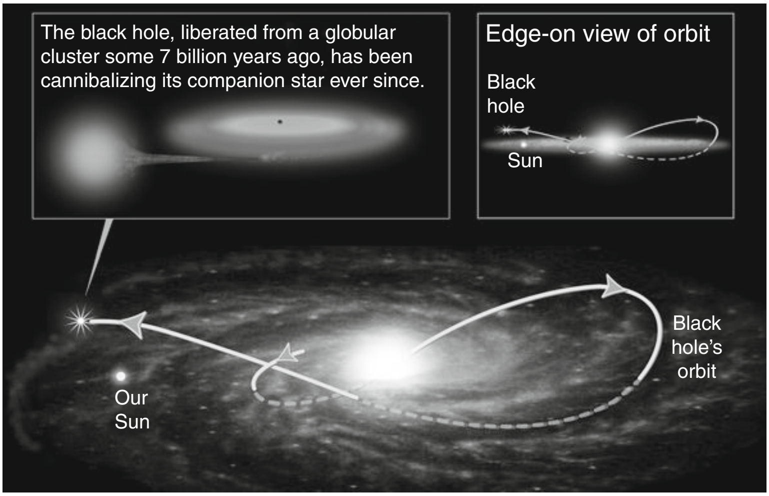



**Fig. 5.8** An artist's conception of the orbital path through the Milky Way followed by the microquasar XTE J1118+480 over the past 7 billion years. A microquasar such as this is powered by a black hole. If its path had taken it close to the Sun, life on Earth would have been affected. (Credit: I. Rodrigues and I. F. Mirabel, Space Telescope Science Institute, NRAO/AUI/NSF)

of Earth's surface. Fortunately, our atmosphere shielded us, as it does from a variety of forms of cosmic radiation. But SGR1900+14 is tens of thousands of light years away—would our atmosphere have saved us had the magnetar been nearby? A galactic nucleus poses yet another threat. Any civilizations living close to the central region of its galaxy must contend with a variety of risks, but the main threat comes if the galaxy possesses an active nucleus. Even the central region of our own Galaxy, which isn't particularly active, is nevertheless quite inhospitable. Close to the center, stars are so crowded that the night sky would be bright enough to read by; closer still, and you meet the accretion disk of a million-solar-mass black hole (see Figure 5.9). This is why the inner edge of the GHZ is defined by the point where the violent central regions are no longer a threat.

Could this be the explanation of the Fermi paradox? Might the random violence of an uncaring universe explain the silence? Are civilizations destroyed before they can reach us?

The three mechanisms mentioned above—stray black holes, magnetars and active galactic nuclei—don't by themselves, or as a group, explain why our Galaxy is silent. Black holes and magnetars might pose a threat to individual stars or stellar groups over the course of the Galaxy's lifetime, but they are too localized to act as a Galaxy-wide sterilizing agent; and while the center of the Galaxy is probably a place to avoid, it isn't a threat to life way out here in the

**Fig. 5.9** An artist's impression of an active galactic nucleus. The central region of any galaxy is thought to play host to a supermassive black hole. Sometimes these holes consume surrounding matter at prodigious rates and in so doing they radiate across the electromagnetic spectrum. Some active galactic nuclei are so bright that astronomers have detected them at the farthest reaches of the observable universe. (Credit: ESA/NASA, the AVO project and Paolo Padovani)

spiral arms, some 30,000 light years away from the action. On the other hand, two other types of astronomical object—supernovae and gamma-ray bursts—might resolve the Fermi paradox.

## Supernovae

A supernova is the cataclysmic explosion of an aging star. Such explosions are powerful and occur rather frequently on an astronomical timescale: the Galaxy on average is host to one or two supernovae per century.

There are two types of supernova. A Type Ia supernova results when a white dwarf in a binary system reaches a critical mass after sucking material from its companion. A violent thermonuclear explosion ignites and blows the star to bits. A Type II supernova occurs in the later stages of the life of massive stars. When the core of a massive star no longer produces enough energy to support itself against the relentless force of gravity, the star collapses under its own weight. The core forms a dense neutron star or even a black hole; the outer layers of the star rebound from the core at high velocity and head off into space, where they become part of the interstellar medium. Such explosions might be lethal, but they are also necessary for life: we wouldn't exist were it not for an ancient Type II supernova that seeded space with heavy elements cooked up in its core. The details of the two types of explosion are different, but both types radiate large amounts of energy: over the course of a few weeks, a supernova can release as much as $10^{44}$ J in a variety of forms.

A nearby supernova might be disastrous for life on Earth. One estimate suggests that a supernova exploding anywhere within 30 light years of Earth[325] could destroy surface life on our planet. The mechanism of destruction is subtle. The threat arises from the enormous amount of gamma-radiation that a nearby supernova would dump into Earth's atmosphere. Direct gamma-radiation from the explosion probably wouldn't harm us, because the upper atmosphere provides an effective shield. However, the gamma-rays would cause atmospheric nitrogen to dissociate, the nitrogen would then react with oxygen to form nitric oxide, and the nitric oxide would react with ozone—thus rapidly depleting the ozone layer. Ozone levels could be reduced by as much as 95% for several years. With Earth's ozone layer down, surface life would have nothing to protect it from lethal UV rays from the Sun. Death would come from a classic one–two punch: first the gamma-radiation from the supernova would lower our defense, then UV radiation from the Sun would devastate multicellular life.

As we shall discuss later, there have been several mass-extinction events since multicellular life took to land. Can any of these be blamed upon the effects of a local supernova? It's difficult to say with certainty. It seems increasingly probable that the last mass extinction—the one in which dinosaurs perished—was due in large part to the effects of a meteor impact. Perhaps the other great die-offs were caused by similar impacts; or perhaps they were due to climate change; or perhaps they were just chaotic events that can happen in complex systems. We see no obvious evidence linking mass extinctions to the after-effects of supernovae. Even if supernovae *can* cause mass extinctions, it's not known whether the extinctions pose a long-term threat to the emergence of intelligence. Perhaps supernovae are *necessary* for intelligent life: perhaps, to use Cramer's phrase, they constitute another "pump of evolution". For the moment, though, let's assume that a nearby supernova can cause a mass-extinction event, and that such an event slows the development of intelligent life.

Since all stars move through space, over the course of aeons random stellar motions will bring the Sun close to a supernova. Eventually, a supernova *will* explode close to Earth. (In case any readers are worried, no star presently within 60 light years of us will go supernova within the next few million years.) The

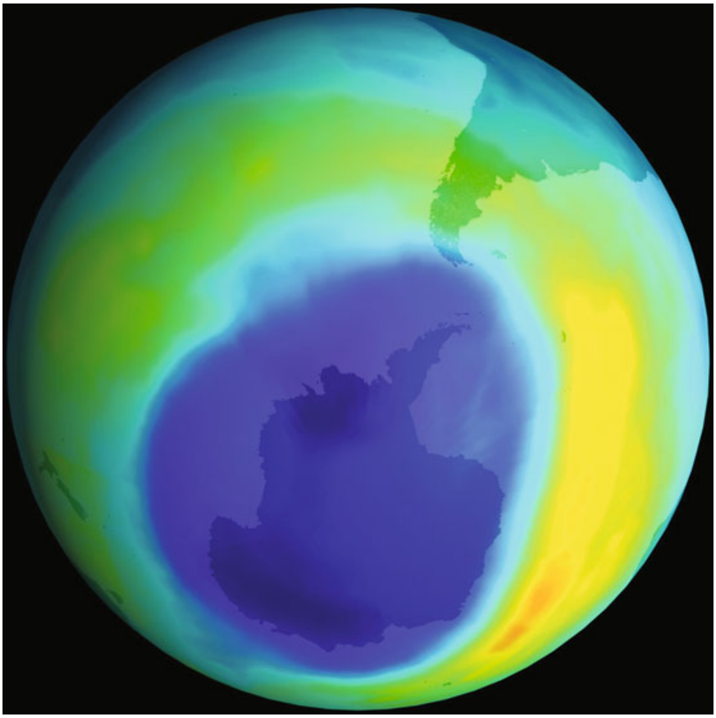


**Fig. 5.10** The dark patch over Antarctica shows a region of ozone depletion in September 2000. The ozone "hole" was caused by a build-up of ozone-depleting chlorofluorocarbons; fortunately the use of these chemicals has been regulated, but full recovery of the Antarctic ozone layer is not expected until 2050. A nearby supernova could reduce ozone levels over the entire globe. (Credit: NASA)

critical question is: how often is a supernova event likely to occur close enough to Earth to cause a mass-extinction event? Well, estimates vary but a mid-range estimate is that a supernova event will occur within 30 light years of Earth on average every 200 million years or so. If that estimate is accurate we have another question to ask: why are *we* here?

One answer to this question could be simply that the calculations of the frequency of supernovae are wrong; or (which is quite likely) that we perhaps don't fully understand the effects on Earth's biosphere of a nearby supernova. In this case, there's no implication for the Fermi paradox. But perhaps we are here because Earth has been extremely lucky; perhaps Earth hasn't seen a really close supernova since the emergence of life on land. If this is true, then we could resolve the Fermi paradox by saying that *every* other life-bearing planet has been less fortunate than Earth.

One difficulty with this suggestion is that there's no astrophysical evidence to support the notion that Earth has been particularly fortunate with regard to supernovae. Furthermore, if we accept that intelligent life is common then supernovae simply aren't effective enough to explain the Fermi paradox. Once an ETC has colonized just a small fraction of its stellar neighborhood, no supernova can stop it. (The threat from supernovae is thus another motivating factor for ETCs to engage in interstellar colonization. Once a civilization has colonized stars within a radius of about 30 light years of the home world, it could survive the effects of a supernova.)

What we need if we want to explain the Fermi paradox is a mechanism that can affect life on *every* planet in the Galaxy, without exception. If there were some mechanism that generated a sufficiently powerful Galaxy-wide sterilizing event it could operate fairly infrequently (every few hundred million years, say) and remain an explanation for the Fermi paradox. Multicellular life would be eradicated before intelligence had the chance to arise; a civilization could never advance to the stage where it might develop effective countermeasures to the threat. Putative ETCs wouldn't have had billions of years to colonize the Galaxy; instead, they 'd have the few hundred million years since the last sterilizing event. In essence, the "Universal Clock" would be reset every time a sterilizing event took place.

It seems almost unbelievable that any natural phenomenon could cause such widespread devastation. Unfortunately, however, astronomers now know of a potential Galaxy-wide sterilizing mechanism: the devastating power of a gamma-ray burst (GRB).

## Gamma-Ray Bursts

Gamma-ray bursts were discovered by accident more than 40 years ago, but until relatively recently their origin was completely unknown.[326] Even now, the precise physical origin of GRBs is a matter for debate. Whatever the progenitor event might be, the important fact is this: the GRB fireball is the most powerful phenomenon in the known universe. A GRB pours out more energy in a few seconds than the Sun will generate in its entire lifetime. A GRB shines so brightly that our detectors can see them from halfway across the universe. All the GRBs we have detected so far seem to have occurred in distant galaxies; if one occurred in our Galaxy, it would be bad news. We need to ask two questions. First, how frequently do GRBs occur in our Galaxy? Second, if our Galaxy hosted a GRB event, just how bad would things be?

Calculating the frequency of occurrence of GRBs is a typical Fermi problem! Very roughly, we can say that a galaxy hosts a visible GRB about once every 100 million years. Interestingly, this rough timescale is pretty much the timescale between mass-extinction events on Earth. People have suggested, therefore, that GRBs might be responsible for mass extinctions.[327]

**The Frequency of Gamma-Ray Bursts** In the 1990s, the orbiting Compton Gamma Ray Observatory detected about one GRB every day. The *Swift* satellite was launched in 2004 in order to make a detailed study of GRBs; it observes the bursts and their afterglows at gamma-ray, X-ray, UV and optical wavelengths. At the time of writing, it has detected 866 GRBs since it began its observations. The *Fermi* Gamma-ray Space Telescope has discovered some bursts not picked up by *Swift*, as have other missions. Between them, they detect about 100 GRBs each year. So our satellites observe somewhere between 100 to 365 GRBs each year. Let's round it up and say that each year in the universe roughly 1000 GRBs are pointing our way. As a rough estimate, let's suppose there are $10^{11}$ galaxies in the universe, so on average there are $10^{-8}$ GRBs pointing our way per galaxy per year. In other words, to a first approximation with which Fermi might be happy, a typical galaxy will host a GRB that we can detect about once every 100 million years. (The true rate is likely to be much higher, since GRBs presumably emit their energy in a beam. The total number of GRBs that happen each year depends on the degree of collimation, but it's likely to be 100–1000 times the observed rate.)

The awesome power released by GRBs means that, even if one occurred at a large distance from Earth, our planet would still be bathed in radiation (assuming the burst was pointing our way). A distant GRB could inflict the same sort of damage to Earth's ozone layer as a nearby supernova. It has to be said that this conclusion is open to debate. GRBs are undeniably more powerful than supernovae, but they are over with much more quickly: they pump out most of their energy in less than a minute. Therefore only half of a planet will be affected directly by a burst; the other half is shielded from the blast by the mass of the planet. Of course, the damage from the affected side of the planet *might* propagate and cause worldwide destruction and secondary effects *might* cause further problems, but with our present state of knowledge one can argue that a planet's ozone layer would protect surface life from the effects of a GRB—unless the GRB occurs *too* close, of course, in which case the planet is toast. If one takes a pessimistic view, however, and accepts the conclusion that GRBs can affect planetary biospheres over large distances, then we have a possible sterilizing agent for the Galaxy.

Suppose, then, that GRB can indeed destroy "higher" life-forms over vast reaches of space; combine this with the prediction from some theories of GRB formation that bursts were more frequent in the past, and you have the resolution of the Fermi paradox proposed by James Annis.[328] The proposal is simple. In the past, GRBs effectively sterilized planets before any life-forms in the Galaxy had the chance to develop intelligence. Only now that the event rate has decreased, and GRBs are less common, has there been time for civilizations to arise. With Annis' proposal, there's nothing necessarily special about Earth; humans aren't here because our planet was lucky in avoiding a catastrophic event. There could be tens of thousands of ETCs in our Galaxy at or near the same stage of development. All of them will have had the same amount of time to develop as life on Earth has had: the time since the last GRB exploded in the Galaxy.

It's undeniable that GRBs occur and are astonishingly powerful; they'd certainly sterilize any nearby planet unlucky enough to be in the line of fire. The SETI optimists—those who argue that technologically advanced ETCs are common—must thus face an unpalatable conclusion: over the course of a Universal Year, countless civilizations will have been within shooting distance of a GRB. Many advanced civilizations must have been consumed by fire.[329] Personally, however, I think it unlikely that GRBs are capable of sterilizing an entire galaxy, and so I don't accept that GRBs by themselves resolve the Fermi paradox.

## Solution 61 A Planetary System is a Dangerous Place

*Man is never watchful enough against dangers that threaten him every hour.*
Horace, *Carmina*, II.13

Destruction can come from the distressingly long list of galactic hazards, but some threats lie much closer to home.[330] We've already mentioned the most obvious worry: meteorite impact. Tiny meteorites hit Earth every day; medium-sized objects land every few years; large objects—say, 20 km wide—strike every few hundred million years. Although large meteorites only hit Earth infrequently, when they *do* hit they cause total devastation. If a 20-km-wide asteroid hit Earth today, it would almost certainly kill every human being. Multiply the small chance of an event occurring by the number of people it would kill, and you arrive at the probability of death per person for the event. It turns out that, averaged over a human lifetime, the chance of being killed by meteorite impact is about the same as dying in an aircraft crash. Strange, then, that we spend vast amounts on air safety and essentially nothing on detecting the near-Earth objects that could destroy our civilization.

Presumably, ETCs also have to contend with the threat posed by meteorite impact, as these objects will be common in planetary systems. But there are many other hazards, and below I discuss a few more.

**Fig. 5.11** An artist's conception of an asteroid impact on Earth. If an object such as this struck our planet today, as we know such objects have struck in the past, then human life would almost certainly be wiped out. (Credit: NASA/Don Davis)

## Snowball Earth

Threats to civilization need not come from space. Recent evidence—particularly the discovery of glacial debris near sea level in the tropics—suggests that, over geological history, Earth has repeatedly been covered in a layer of ice. One event might have happened 2.5 billion years ago, and there might have been four of these so-called Snowball Earth events[331] in the past 800 million years, with each episode lasting for 10 million years or more. Don't mistake these events with the textbook images of the last Ice Age; compared to a Snowball Earth, the last Ice Age was positively tropical. During a Snowball Earth, a kilometer-thick layer of ice covers the oceans, and ice even covers equatorial oceans (though perhaps not to the same depth). Average temperatures drop to −50 °C. Most organisms are unable to cope with such conditions, and life can hang on only by the thinnest of threads—perhaps around volcanoes, or under clear thin ice at the equators.

The mechanism by which our planet can descend into a Snowball Earth is well understood. The ice cover can increase for a variety of reasons, and when it increases the ice reflects an increasing amount of sunlight straight back out into space. This decrease in solar heating of the surface causes the temperature to drop and more ice to form. Once a critical amount of ice cover is reached, a "runaway icehouse" effect takes place and the planet descends into a Snowball Earth event. What's difficult to understand, and what caused scientists to

**Fig. 5.12** Melting ice floes in open water in Antarctica. On a Snowball Earth conditions at the equator would, at best, be like this. The rest of the globe would be covered in thick ice. Complex life would struggle to survive. (Credit: NOAA/Michael van Woert)

dismiss the idea of a Snowball Earth for many years, is how the planet can *escape* from the ice cover. Once Earth is encased in ice, most of the sunlight falling on the planet is reflected into space before it can warm the surface. The solution came with the realization that volcanic activity doesn't stop during a

Snowball Earth event. Volcanoes pump out vast amounts of carbon dioxide—a greenhouse gas. Of course, today volcanoes are still belching out carbon dioxide but under normal conditions this $CO_2$ is absorbed by falling rainwater, which eventually carries it to the ocean where it becomes locked up in solid carbonate deposits on the ocean floor. On a Snowball Earth there's no liquid water to evaporate, and therefore no clouds, and therefore no rain: for 10 million years, maybe more, the $CO_2$ from volcanoes would build up in the atmosphere. Eventually, there'd be about a thousand times more atmospheric $CO_2$ than in today's atmosphere. The temperatures would rise and quickly melt the ice: from icehouse to greenhouse in a geological instant.

The implications of the Snowball Earth hypothesis are profound, and we'll examine some of them later.

## Super-Volcanoes

If volcanoes were life's savior during the Snowball Earth events of the Neoproterozoic era, more recently they proved to be almost disastrous for intelligent life: they almost wiped out *Homo sapiens*. Recent research indicates that humans are genetically remarkably similar. To explain this lack of genetic diversity, some biologists have suggested that our species emerged from a "genetic bottleneck" about 75,000 years ago. A bottleneck occurs when the size of a population reduces dramatically; in the case of our species, the total number of humans alive on Earth might have dropped as low as a few thousand. We almost became extinct.

If this bottleneck really did occur, then we don't have to look far for the smoking gun that could have caused it. The Toba volcano in Sumatra erupted 74,000 years ago; so great was the eruption, it earns the title of a "super-volcano." The eruption was *much* more violent than recent volcanic explosions such as Mount Pinatubo and Mount St. Helens. Climatologists have suggested that a super-volcanic eruption can cause a volcanic winter—similar in effect to a nuclear winter, but without the radiation. It's not implausible that the years of drought and famine following such an explosion could drive a pre-technological human species to the brink of extinction.

## Mass Extinctions

Meteor impact, global glaciation, super-volcanoes. Even on a placid planet such as Earth, life has to contend with a lot. Sometimes, whether the cause is one of the three mechanisms mentioned above, or one of the celestial agents of destruction mentioned earlier, life barely hangs on.

Since animal life became abundant on Earth, in the Cambrian explosion of about 540 million years ago, life on Earth has suffered numerous mass extinctions—a mass-extinction event being defined as a period that sees a significant reduction in biodiversity.[332] The extinction events vary in their severity. In five great mass-extinction events,[333] more than half of all species then alive were killed. These five events are, in chronological order, the Ordovician, Devonian, Permian, Triassic and Cretaceous.

The Ordovician extinction 440 million years ago and the Devonian extinction 370 million years ago both saw more than a fifth of the marine families disappear. The effects on land life are less well known, mainly because the fossil record is so poor for these ages. The cause of these extinction events remains a matter for debate.

The Permian extinction 250 million years ago was by far the largest of all the great mass-extinction events. Perhaps more than 90% of marine species became extinct; eight of the 27 orders of insects were lost (insects survived the other mass extinctions); the loss was devastating. The cause of this catastrophic event is uncertain; several mechanisms, possibly acting in synergy, have been proposed to explain this global catastrophe.

The Triassic extinction 220 million years ago saw significant reductions in the number of marine and land species. Again, scientists debate the cause of this reduction in biodiversity.

The Cretaceous extinction 65 million years ago is the most celebrated and well-known of all the mass extinctions. This event saw the end of the age of the dinosaurs and provided the conditions that led to the rise of the mammals. Almost certainly, the cause of this extinction was the after-effects of a large meteorite impact.[334] There are several reasons for believing in the impact theory of this extinction event. First, the 200-km-wide Chicxulub crater on the Yucatán peninsula in Mexico is of precisely the right age. Second, no matter from where in the world they are drawn, rock samples from the Cretaceous–Tertiary boundary show a high concentration of iridium, which is what one would expect if a large asteroid hit Earth. Third, many of the same sites contain shocked quartz grains—another sign of violent impact. Fourth, geologists often find fine soot particles in clays from the Cretaceous–Tertiary boundary—particles that could have come only from burning vegetation; the implication is that much of Earth's plant matter was on fire. The immediate aftermath of the impact would clearly have killed large numbers of organisms. The precise mechanism for eradicating large numbers of species is less clear; it could have been atmospheric change, a nuclear winter, large-scale long-term fires, acid rain, a combination of these effects . . . or something else entirely. The effects were also dependent upon when and where the meteorite struck Earth, and also on the mass and velocity of the meteorite. Had the meteorite struck just a few hours later, the effects might have been less deadly; had the meteorite been just twice as large, the extinction of life might have been total.

## Extinctions and the Fermi Paradox

It's difficult to say what we can learn from these extinction events. They seem to be different in character, cause and severity. Only in the case of the Cretaceous event is there a definite and identified causal mechanism. The other extinctions might have been caused by something quite different; after all, in this book we've looked at many potential threats. Life-forms on other planets presumably face the same hazards, and they might face additional risks that life on Earth has been spared. For example, some planetary systems might have life-bearing planets in orbits that become chaotic—and a mass extinction would be probable. Or a change in the rotational rate of a planet might trigger a mass extinction. Anything that causes extensive climate change—either a global cooling or warming outside of temperatures that are tolerable for animal life—might induce a mass extinction. Perhaps the lesson is simply that planetary systems are dangerous: over the course of billions of years, mass extinctions are *inevitable*.

It's a short step from arguing that mass extinctions are inevitable to arguing that they play a role in resolving the Fermi paradox. In fact, people have used the idea of mass extinctions to suggest two quite antithetical solutions to the paradox. The straightforward suggestion is that mass-extinction events have impeded the development of intelligent life on other planets. The more subtle suggestion is that, in the immortal capitalization of Sellars and Yeatman, mass extinctions are a Good Thing that occur too infrequently on other planets! (At least, the right *sort* of extinction events happen too infrequently.)

It's easy to understand why mass extinctions might be a Bad Thing. Many people would argue that life—at least life as we know it—has only two defenses against mass extinction. The first defense is simplicity: this is the approach taken by prokaryotes (see page 291), which have survived for billions of years. Bacteria have essentially kept their single-cell body plan over the aeons; indeed, it's possible, though difficult to prove conclusively, that modern bacteria are genetically identical to the earliest living cells of 3.7 billion years ago. Their ability to evolve biochemical responses to new environmental challenges enables prokaryotes to take most things Nature can throw at them. Only a catastrophe on a massive scale would remove all prokaryotic life from Earth. On the other hand, we can't communicate with bacteria. When considering Fermi's question, we're interested in *complex* multicellular life-forms. How do *they* survive the slings and arrows of a billion years of fortune?

The second defense against mass extinction is diversity—an approach taken by animals and plants. If a phylum contains many different species, if it has different ways of earning a living, then there's a chance that one or two of the species will survive the extinction event. Later, the diversity of the phylum can be replenished. So even though animal and plant life is less hardy than bacterial life, and is much more susceptible to extinction, in the long run it can survive. (It's something of a theme of this book: don't put all your eggs in one basket.)

We've no idea how evolution has proceeded on other planets, but perhaps Earth is rare in having phyla with many different species. (See Solution 62 for one reason why this might be the case.) Complex life on other worlds might be less likely to survive the inevitable extinction events. We can imagine worlds that are home to many different, weird-looking, truly *alien* creatures—creatures possessing a variety of peculiar body plans. There might be a large number of phyla on such worlds, phyla that took aeons to evolve to their present state. But if those phyla are represented by only a few species . . . well, when the meteorite strikes, or the climate heats up, or the planet's obliquity changes, those phyla might well die out. Maybe Earth has just been lucky (there's that word "lucky" again). This is a gloomy resolution of the Fermi paradox.

We've encountered the more subtle suggestion regarding mass extinctions—namely, that they might be *necessary* for the development of intelligent life—when we discussed the suggestion of a "pump of evolution". Of course, it would be no fun being around when a 20-km-wide asteroid smashes into Earth or global temperatures plummet. But in the long run—a run measured in tens of millions of years—life might benefit from such catastrophes. After the deluge, new and radically different forms have a chance to evolve; Nature can use the changed environment to create and experiment with different species, and perhaps even different body plans. Certainly, following mass-extinction events, biodiversity has always eventually regained the pre-extinction level and then exceeded it.

One controversial suggestion is that two key events in the history of life on Earth—the development of the eukaryotic cell and the Cambrian explosion (more of which in later sections)—were a direct result of the escape from Snowball Earth events. The event itself would case a mass extinction. But the escape? The chemical changes a Snowball Earth would cause in the oceans, the genetic isolation of species, the great environmental pressure on life, the rise in temperature and the rapid melting of ice—all these factors might combine to produce a time of rapid evolutionary activity. According to some scientists, neither animals nor higher plants would exist today if it weren't for past Snowball Earth events.

Perhaps the “right” global glaciation events are uncommon on other planets. A planet must be in the CHZ, it must have oceans of water, it must descend into an icehouse, and it must possess active volcanoes spewing out greenhouse gases to remove the ice. Perhaps the norm for most water-planets is a descent into a Snowball with no means of escape. The mass extinctions would be total.

### The Holocene Extinction

It would be wrong to discuss past mass-extinction events without mentioning the Holocene extinction. The Holocene epoch encompasses the last ten millennia, up to the present day. In other words, we’re living through a mass-extinction event. In this case the cause is clear: human activity. We hunt species to extinction; we cause havoc by introducing alien species into ecosystems; and, most importantly, we destroy habitats. It doesn’t *feel* as if we’re in the midst of a mass extinction because, on an individual scale, 10,000 years is a long time. On a geological scale, though, it’s an instant. Under some estimates, the rate at which species are becoming extinct[335] is now 120,000 times the “normal” or “background” rate. Many of the species made extinct by our destruction of rain forests have never even been documented. If the current rate of extinction is maintained, and the destruction of the rain forests continues, then global atmospheric and climatic effects seem certain to occur. It’s then quite possible that ours will be one of the species that joins the extinction. Harking back to a previous solution discussed in the book, perhaps a general evolutionary law is that intelligence extinguishes itself.

## Solution 62 Earth’s System of Plate Tectonics is Unique

*What we want is a story that starts with an earthquake and builds to a climax.*
Samuel Goldwyn

In the period 2000–2008, an average of 50,184 people were killed each year because of earthquakes.[336] The 2004 Boxing Day tsunami, which was triggered by an undersea earthquake off the west coast of Sumatra, alone claimed almost a quarter of a million victims. It therefore seems strange that some geologists consider the existence of plate tectonics—the process that gives rise to earthquakes and volcanic eruptions—to be necessary for the existence of complex life. But there's a serious reason for believing that three phenomena—life, water oceans, and plate tectonics—are linked. And this linkage might be unique to Earth. The argument goes as follows.

The various planets of the Solar System have different methods of disposing of internal heat. In Earth's case, the heat generated by radioactive decay in the interior is transported by a convective method that gives rise to plate tectonics[337] (or, in more colloquial language, continental drift). Consider what happens near a mid-ocean ridge—an underwater mountain range where new crust forms. Hot material from the deep mantle region of Earth is brought to the surface in a convection cell, and at the surface it spreads out and solidifies into ocean crust—it becomes part of the lithosphere. Over geological timescales, the new material floats on the hot mantle underneath it and moves away from where it was born. During this process it cools and collects masses of igneous rocks. The material becomes heavier, and after many tens of millions of years it sinks back, under its own weight, deep into the mantle at places called subduction zones. Eventually, the cycle repeats. On geological timescales, the outer regions of our planet resemble one of those kitsch lava lamps.

Some scientists have argued that plate tectonics could be the most important requirement for the development of animal life. There are several reasons why plate tectonics might be vital. Let's look at just three of them. (In Solution 67 we consider a fourth possibility.)

First, the mechanism of plate tectonics seems to play an important role in the creation of Earth's magnetic field. The theory of planetary magnetism is formidably complicated but, in essence, planets generate a magnetic field by means of an internal dynamo. Such a dynamo requires three things: the planet must rotate, it must contain a region with an electrically conducting fluid, and it must maintain convection within the conducting fluid region. It's difficult to be sure, but in Earth's case it seems likely that without plate tectonics the convective cells would cease to export heat to the surface, the dynamo would cease to function, and Earth's magnetic field would be a tiny fraction of its present value. The relevance of all this is clear: Earth's magnetic field helps prevent high-energy particles in the solar wind from scattering atmospheric particles into space; over time, such sputtering could cause the Earth's atmosphere to dissipate. In short, without Earth's magnetic field surface life might not have evolved.

Second, plate tectonics created Earth's continents—and continues to refresh them. Continents are important. A world with a mixture of oceans, islands and continents is more likely to offer evolutionary challenges than is a world dominated solely by water or land. Furthermore, plate tectonics causes environmental conditions to alter, and thus helps promote speciation. For example, suppose the splitting of a piece of land from a continental land-mass results in a particular species of bird living on both the new island and the original continent. Over time, the environment on the island will differ from the continental environment; the birds will face different challenges and will evolve in different ways. Over time, there'll be two species where before there was one. Plate tectonics thus promotes biodiversity, which, as we've seen, is important during mass-extinction events.

Third, and perhaps most important, for a billion years or more plate tectonics has played a key role in regulating Earth's surface temperature. The climate on our planet has long been balanced on a razor's edge. If the temperature drops too much, and the icecaps begin to increase in size, then a runaway icehouse effect can occur: Earth freezes. If the temperature increases too much, and the oceans start to simmer, then the extra water vapor in the atmosphere can cause a runaway greenhouse effect: Earth boils. Some prokaryotes might survive these temperature extremes, but complex life-forms flourish only over a much narrower range of temperatures. Plate tectonics, some scientists argue, has a fine-tuning mechanism that keeps the planetary thermostat set "just right" for animal life.

The way in which plate tectonics controls temperature[338] is rather complicated, and more than one mechanism is involved. The key role it plays, however, is in its regulation of atmospheric carbon dioxide. $CO_2$ is an effective greenhouse gas: if the atmosphere contains too much $CO_2$, then global temperatures can rise—as mankind seems hell-bent on demonstrating experimentally. On the other hand, if there's too little atmospheric $CO_2$, then Earth fails to benefit from the greenhouse effect, and the planet cools.

Now, $CO_2$ doesn't remain in the atmosphere indefinitely. Carbon dioxide reacts with water to form carbonic acid; rainfall thus "washes" it out of the atmosphere. This carbonic acid weathers the rocks on Earth's surface and the chemical products of this weathering get transported by rivers to the ocean. The products end up as calcium carbonate ($CaCO_3$) and quartz ($SiO_2$) on the seafloor, both through the formation of rocks and through the formation of the shells of living organisms. Eventually, the plate tectonics mechanism causes this $CaCO_3$ and $SiO_2$ to be subducted down into the depths of the Earth. Thus, atmospheric $CO_2$ is removed. But that's not the end of the story! The high temperatures and pressures deep within Earth convert the calcium carbonate back into $CO_2$ and CaO. Plate tectonics then recycles the $CO_2$—and lots of other useful materials—by creating volcanoes. (Volcanoes vent *tremendous* amounts of material. In 2010, a tongue-twisting Icelandic volcano wreaked havoc with international air travel. Although the Eyjafjallajökull eruption was comparatively small, it still ejected about 250 million cubic meters of ash and cinder and of the order of a million tons of $CO_2$.)

**Fig. 5.13** A minor eruption of Sakurajima in 2009. In the foreground is Kagoshima city. Sakurajima is one of the most active volcanoes in the world; as I write this, in April 2014, Sakurajima is the only volcano under a Level 3 alert—the activity is such that people are being warned not to visit it. (Credit: Kimon Berlin)

If the atmospheric $CO_2$ were not replaced, Earth would undergo a global cooling. But what if too much $CO_2$ is put back into the atmosphere? Do we not run the risk of a runaway greenhouse effect? It turns out that, as the planet warms, the chemical weathering of rocks increases—which causes more $CO_2$ to be removed from the atmosphere, which causes the planet to cool (thus slowing the rate at which $CO_2$ is removed from the system, which causes the planet to warm ... and so on, in a classical feedback mechanism). This $CO_2$–silicate cycle is rather complicated, and the details are still not fully understood, but the cycle seems to be crucially important for the long-term stabilization of global temperature.

One can argue that the development of animal life here on Earth *required* plate tectonics—to promote biodiversity, to generate a magnetic field, to stabilize global temperature, and so on. And yet there's nothing inevitable about plate tectonics. Only Earth, as far as we know for sure, uses this mechanism to dispose of its internal heat. Perhaps the process is rare, and other planets lack animal life because they lack plate tectonics.

We don't know how frequently plate tectonics will occur because we lack a good general theory of the process. The type of questions one might ask—How does the existence of plate tectonics depend on a planet's mass? How does it depend on the chemical composition of the mantle?—can't be answered with present models, so it's not easy to provide a good estimate of how many planets might develop, and maintain, plate tectonics. In the absence of hard facts, from either experiment or theory, one can argue either way. Some scientists believe the titanic collision that formed the Moon laid the seeds from which plate tectonics developed; in this case, plate tectonics might be rare. On the other hand, the basic conditions for plate tectonics seem relatively simple: a planet should have a thin crust floating on top of a hot, fluid region undergoing convection due to rising heat from the core. Perhaps water oceans are also necessary to "soften" the crust and allow subduction. Such conditions are probably not rare. Scarce, perhaps, but not rare. In other words, we simply don't know whether plate tectonics is a common phenomenon.

Even if plate tectonics *is* rare, does it necessarily follow that animal life is rare? Although plate tectonics seems to have played (and continues to play) a beneficial role for the development of life on Earth, is it the *only* mechanism that can provide these benefits? Plate tectonics is an extremely complicated process; the very existence of the $CO_2$–silicate cycle has only been known about for a few decades. In cases such as this, where scientific understanding is still in its relative infancy, it often turns out that there's more than one way to skin a cat. Perhaps right now the scientists of a planet orbiting some anonymous M-class star are marveling at the cooling mechanism of their world and how it almost miraculously stabilizes their global environment.

My guess is that—as with so many factors we have discussed—the possible rarity of plate tectonics is by itself insufficient to provide an answer to the Fermi paradox. But it might be yet another factor making it less probable that ETCs will develop on other planets.

## Solution 63 The Moon is Unique

*How like a queen comes forth the lonely Moon.*

George Croly, *Diana*

The last time I checked, astronomers had found 173 natural satellites orbiting the eight planets of the Solar System. (Since I wrote the first edition of this book, more than one hundred moons have been discovered. On the other hand, the number of planets has dropped by one: in 2006 Pluto was reclassified as a trans-Neptunian dwarf planet or plutoid.) Given the substantial number of planetary satellites that exist in the Solar System it seems absurd to suggest that our Moon is unique, much less that it has anything to do with the Fermi paradox. Yet for decades people have had a nagging suspicion that the Moon is what makes Earth special.

Three questions are relevant here. First, in what way is the Moon unusual? Second, how probable is it that satellites similar to Earth's exist in other planetary systems? Third, in what way might the existence of the Moon have been necessary for the development of intelligent life?

Well, starting with the first question, the Moon is unusual in being large. Indeed, Earth is unique in possessing such a large satellite. Note that our Moon is *not* the largest moon in the Solar System. That honor belongs to Ganymede, which is one of the moons of Jupiter. Two other Jovian satellites—Callisto and Io—are also slightly larger than the Moon; and so is Titan, one of Saturn's moons. But Ganymede, Callisto, Io and Titan orbit giant planets. Compared to their parent bodies these satellites are as grains of dust. Our Moon, on the other hand, is large compared to Earth: it has 1/81 of the mass of our planet. The Earth–Moon system has rightly been called a "double planet". And, moving on to the second question, double planets might be rare.

In order to estimate the scarcity of "double planets" we need to understand how the Moon formed. For many years, the formation of the Moon was one of the longstanding problems of planetary science. Several mechanisms were proposed, including co-accretion (in which Earth and Moon formed at the same time from the gas and dust in the solar nebula), fission (in which Earth formed first but was spinning so quickly a large piece of material tore away and formed the Moon) and capture (in which the two objects formed at different places in the solar nebula, and then the Moon became trapped in orbit after straying too close to Earth). All of these mechanisms had difficulties in explaining several important features of the Earth–Moon system, but it was hoped that analysis of lunar rocks brought back from the Apollo missions would vindicate one of them. Instead, it became clear that none of these ideas worked. A new theory of lunar formation was needed.

In 1975, two groups independently proposed the impact hypothesis[339] for the Moon's origin. They postulated that a Mars-sized object, which has since been given the name Theia, struck the infant Earth in an off-center impact. The unimaginably violent collision ejected a mixture of terrestrial and impactor material into orbit around Earth, and this material quickly coalesced to form the Moon.

Now, scientists dislike having to resort to cataclysmic events to explain their observations, but we know that Earth *has* been hit by a variety of objects throughout its history; and the axial tilts of the planets suggest that really violent collisions were not uncommon in the early Solar System. A collision with an object such as Theia would certainly have been possible.

**Fig. 5.14** Earthrise as seen from the Mare Smythii region of the Moon. The photograph was taken on 20 July 1969, on the Apollo 11 mission. (Credit: NASA)

It must be conceded that the details of the impact are still in dispute. Consider, for example, the fact that Moon rocks brought back from Apollo missions have precisely the same proportion of the three different oxygen isotopes ($^{16}$O, $^{17}$O and $^{18}$O) as one finds in Earth rocks; Martian rocks and meteorites show a different isotopic ratio. Similarly, the ratio of two titanium isotopes ($^{47}$Ti and $^{50}$Ti) is identical in Earth and Moon rocks[340] and different to everywhere else in the Solar System. This is rather puzzling in the giant impact scenario because much of the Moon's substance would have come from Theia, which is unlikely to have had the same isotopic composition as Earth. Another problem with the Theia scenario is that the impact would have created a surface ocean of magma—yet there's no evidence that Earth has ever possessed a magma ocean. Nevertheless, an impact between Earth and Theia is the currently accepted hypothesis for the origin of the Moon.

If our Moon was indeed the consequence of a giant impact, then the uniqueness of the Earth–Moon double planet within our Solar System need not be too surprising. Although violent collisions in the early Solar System were common, cataclysmic Moon-forming collisions might have been scarce. Perhaps the infant Mercury, Venus and Mars were simply fortunate enough to dodge the larger missiles. Or perhaps they were hit but suffered the "wrong sort" of collision or at the wrong stage of development.[341] The Moon-forming collision occurred at a critical time. Had it happened much earlier, when Earth was less massive, then most of the debris from the collision would have ended up in space and the Moon would have been a small object. Had the collision occurred much later, then Earth would have been more massive and its greater

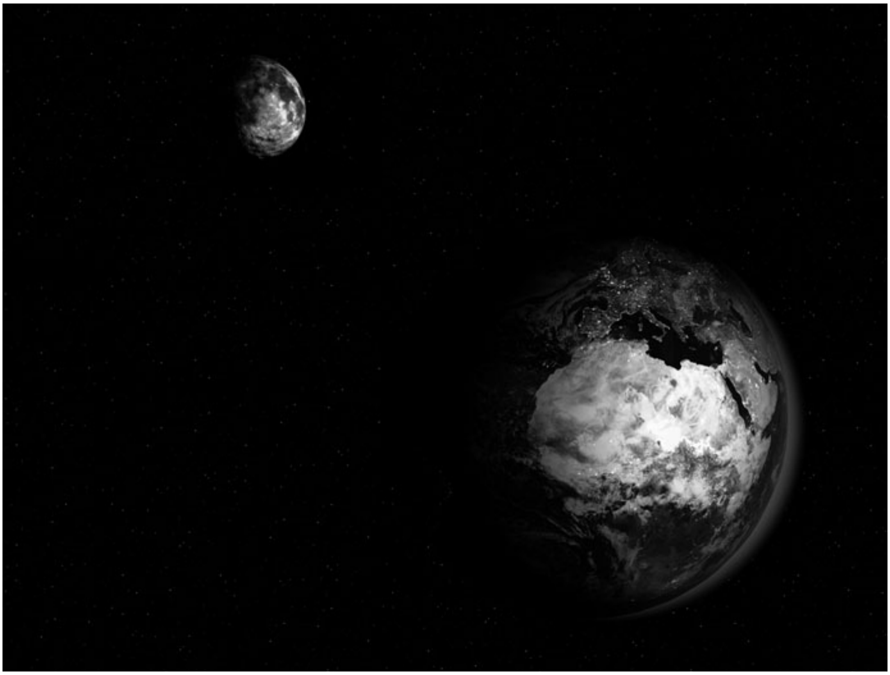


**Fig. 5.15** Earth and Moon: a double planet. (Credit: ESA/AOES Medialab)

surface gravity would have prevented the ejection of enough mass to form a large Moon.

Whereas the original scenarios for lunar formation implied that our Moon was almost a natural by-product of planetary formation, the impact hypothesis hints that the Earth–Moon system might be exceptional. Imagine a collection of primordial stellar nebulae, each identical to the nebula from which our Solar System formed. Perhaps only 1 in 10, or 1 in 100, or 1 in 1000, would generate an Earth-like planet with a Moon as large as ours. Perhaps the figure is 1 in 1,000,000. We have no idea—and huge advances will have to be made in observational astronomy before we can discover whether extrasolar terrestrial planets possess satellites as large as the Moon. With our present knowledge, it's entirely possible to believe Earth is unusual in possessing such a large satellite.

Even if the Moon *is* rare, so what? If Earth were Moonless, then poets through the ages would have lost a source of inspiration. Perhaps humankind's scientific development would have been affected, since historically the Moon has played a large role in advancing our understanding of astronomy. But would life itself really have been any different?[342]

**A Moon for Venus?** It has been suggested that Venus once had a large satellite, which was formed in the same way as the Moon, but which followed a retrograde orbit: in other words, it orbited Venus in the "wrong" direction. Such an orbit could certainly occur if the satellite was created through an impact event. However, whereas tidal forces are causing our Moon to move *away* from Earth, in the case of a retrograde orbit those forces would act in the opposite direction. A satellite in a retrograde orbit moves toward the planet and is eventually destroyed. This is the fate of Triton, the largest of Neptune's satellites.

There are several ways in which the Moon exerted, or indeed continues to exert, an influence on Earth. For example, the Moon raises ocean tides. Soon after the Moon formed it was much closer to Earth than it is now, so the tides of 4 billion years ago would have been huge—a surfer's paradise. It has been suggested that these large tides were a factor in getting life started, perhaps by acting as a giant mixer of the primordial soup and causing nutrient-rich pools where life could have started. Even without the Moon we'd still have ocean tides: the Sun raises tides about half as large as the present lunar tides. We would, however, miss the spring and neap tides, which depend upon the relative positions of Sun and Moon.

A more subtle lunar tidal effect arises from its influence on Earth's crust. The effect of the Moon's gravity might have amplified volcanic activity on Earth and increased continental drift. So it's possible, though not certain, that a Moonless Earth would have been less geologically active; Earth's atmosphere, which formed by volcanic outgassing, might have taken much longer to reach the stage where life could arise. We discussed the importance of plate tectonics in Solution 62.

The most important effect to consider, however, is the way in which the Moon influences Earth's *obliquity*. The eight planets all orbit the Sun in or near a single plane in space; the obliquity—or axial tilt—of a planet is the angle of inclination of its equator to this orbital plane. Earth's obliquity of 23.5° gives rise to the pleasant seasons we enjoy. Other planets aren't so lucky. Mercury has an obliquity of 0°, so its equatorial regions resemble Hell. Life as we know it couldn't survive. (Interestingly, an observer at either of Mercury's poles would see the Sun always on the horizon. Relatively little solar energy can be absorbed at the poles, and indeed the polar regions of Mercury are ice-covered.) Uranus, which has an obliquity of 98°, is almost lying on its side. One pole receives sunlight for half of the Uranian year, while the other pole is in darkness. Again, these are less than ideal conditions for life. Earth—from our biased point of view—seems to be "just right".

The impact event that formed the Moon would have caused Earth's axis of rotation to shift from its initial position. More importantly, as computer simulations have shown, the Moon plays a role in stabilizing Earth's axial tilt over a period of many millions of years. This is important because even

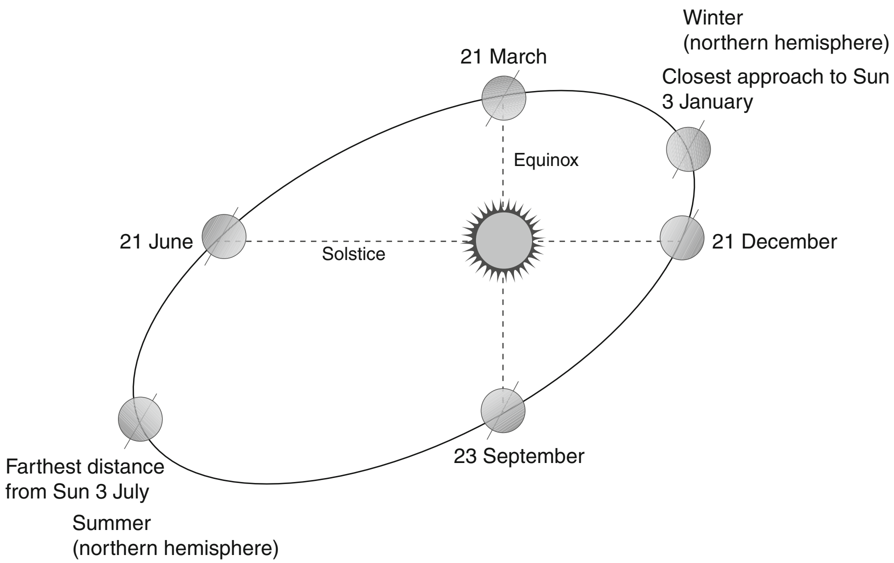



**Fig. 5.16** Earth's obliquity—its tilt relative to the ecliptic plane (in other words, the plane of its orbit around the Sun)—produces the seasons. For planets such as Earth, which has a "moderate" obliquity, most of the solar energy falls in the equatorial regions where the midday Sun is always high in the sky. The polar regions are in constant illumination for 6 months, but in constant darkness for 6 months too; even when the Sun is in the sky, it's never higher in the sky than the obliquity allows (23.5° in the case of the Earth) so the ground is never heated really strongly by sunlight. Thus, the polar regions are cold and the equatorial regions are hot. (The diagram is not to scale.)

small changes in obliquity can cause dramatic changes in planetary climate. For example, Earth's obliquity oscillates by about ±1.5° with a period of oscillation of 41,000 years. This is only a small variation, yet it seems to be linked to the succession of ice ages that Earth has experienced over the past few million years. Mars has no stabilizing influence on its obliquity (Phobos and Deimos being merely boulders, with insufficient mass to have any influence) so although the axial tilt of Mars is currently 25° this value ranges between 15° and 35°, with a period of 100,000 years. Calculations indicate that, over longer timescales, the obliquity of Mars changes chaotically: over the last 10 million years it might have ranged from 0° to 60°. Earth's obliquity, without a Moon stabilize it, would also wander chaotically—to values as large as 90°. Even an object with half the mass of the Moon, which would be a relatively large satellite, would have insufficient heft to stabilize Earth's tilt. Our home planet requires a *large* satellite to prevent its obliquity from wandering and its climate shifting from one extreme to another. (The situation is more subtle than this, however, as we shall discuss in Solution 74.)

Life on Earth has adapted well to climate change in the past, but it's difficult to see how advanced land animals could have prospered had the Martian pattern of obliquity shifts been repeated here. Life on Earth would surely not have evolved into the forms we see today.

There are many "ifs", "buts" and "maybes" in the above discussion. We don't know whether a large satellite is *necessary* for a planet to provide a suitable home for complex life-forms. Our Earth-centered view is necessarily biased. We believe the Moon has been beneficial for the development of life here, but we don't know that life would have been impossible without it. Perhaps if we lived on a Moonless world we'd be grateful to look up and not see a huge chunk of rock hanging in the sky.

And yet that nagging suspicion remains. Perhaps double planets such as our Earth–Moon system *are* necessary for life, for a variety of reasons. And yet they seem to form in chance events. Perhaps the uniqueness of our satellite explains why we are alone. Perhaps that is the tragedy of the Moon.

## Solution 64 Life's Genesis is Rare

*The solution of the problem of life is seen in the vanishing of the problem.*
Ludwig Wittgenstein, *Tractatus Logico-Philosophicus*

Hart's answer to Fermi's question (see Solution 50) was that life's genesis is almost miraculously rare. For practical purposes, we are alone: Earth possesses the only intelligent life—the only *life*—in the visible part of an infinite universe. This miracle loses some of its gloss in an infinite universe, since in that case an infinite number of planets possess intelligent life-forms. However, it's rather difficult to entertain the notion of an infinite universe with an infinite number of habitable planets, not least because there'd be infinite numbers of you and me pondering Fermi's question. That's hard to take in. Can we not instead accept *part* of Hart's idea? Can we dispense with the astronomical notion of an infinite universe and argue solely from biology? Perhaps life isn't a miracle but nevertheless arises only rarely. Perhaps the universe appears sterile because—with the exception of Earth's island of life—it *is* sterile.

Abiogenesis—the process by which non-living matter gives rise to life—might be rare; it might not. Scientists at present don't know how life came into being and so no-one can put a reliable figure on the probability of matter taking the step from inanimate to animate. The sheer unlikelihood of abiogenesis occurring might indeed solve the Fermi paradox; or it might turn out that

Earth-like worlds almost always develop life. Biologists have made strides in recent decades in trying to understand the origin of life, so although there remain two diametrically opposed opinions (as is usual with any aspect of the Fermi paradox), with one group arguing that Nature finds it difficult to create life and another group arguing that life is almost certain to appear as soon as planetary conditions allow, we can hope that the question will be resolved before too many more decades have passed. In the meantime it's worth looking the merits of both positions in order to see what light it sheds on the Fermi paradox. First, though, we need to take a lengthy detour and review what it is we mean by "life" and consider how life on Earth might have arisen.

## What is Life?

At school, my teacher could always drive holes through the attempts of our science class to provide a definition of life. He pointed out that, by some of our definitions, fire is alive (since it grows, it reproduces itself, and so on). On the other hand, by our definitions a mule is not alive (since it can't reproduce itself). For the purposes of this section I'll try my hand at presenting another definition of terrestrial life. My old teacher could probably still drive several holes through the definition, and in any case the definition might be inappropriate in the future. (In ten years, perhaps, scientists might develop a self-aware computer. Will the computer be alive? Or a century hence, perhaps, an explorer on the Altair mission will discover an evil-smelling pink crystal that every morning turns into a goo, clinging to the sides of the spaceship and eating the metal. Is the goo alive? In both cases, under my definition the answer would be "no"—even though the answer should perhaps be "yes". We have to begin somewhere, though, and the definition given below at least constitutes a framework for the discussion.)

I define something to be alive if it possesses the following four properties.

First, *a living object must be made of cells.* Every living creature on Earth consists either of a single cell or a collection of cells. If we knew how cells originated, then we might well be on the road to understanding how life itself originated.

Two quite different types of cell exist: *prokaryotes* and *eukaryotes*. Prokaryotic cells lack a central nucleus. They are simple, small and exist in variety of types. Prokaryotic organisms are hugely successful, in large measure because their simplicity means they can reproduce themselves quickly. A relatively recent and profound discovery is that two different types of prokaryote exist:[343] archaea and eubacteria—or "true" bacteria (or, as I'll write for simplicity, just bacteria). Figure 5.17 illustrates some typical archaea. The two types of prokaryotic cell seem to bear no significantly closer relationship to each other than they do to

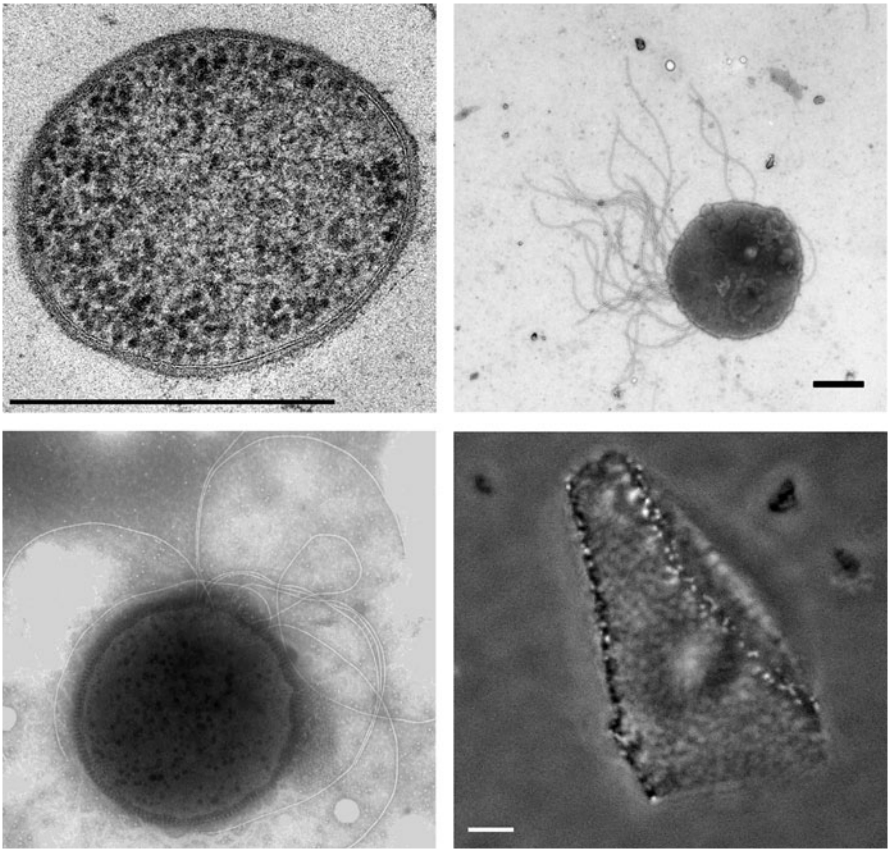


**Fig. 5.17** Four different types of archaea. *Top left*: *Nanoarcheum equitans*. This organism was discovered in a hydrothermal vent off the coast of Iceland, and thrives in temperatures of 80°. Its cells are incredibly small—only 400 nm in diameter. (Credit: R. Rachel and H. Huber) *Top right*: *Methanococcus maripaludis*. This archaeon thrives in relatively moderate conditions, but oxygen is poisonous to it. (Credit: S.I. Aizawa and K. Uchida) *Bottom left*: *Thermococcus gammatolerans*. This is the most radiation-resistant organism known to science. It thrives in temperatures between 55–95°C. (Credit: A. Tapias) *Bottom right*: *Haloquadratum walsbyi*. This archaeon thrives in extremely salty environments, and is unique in possessing a squarish cell shape. (Credit: M.A.F Noor, R.S. Parnell and B.S. Grant)

eukaryotic cells. Eukaryotic cells are much more complicated than prokaryotic cells; within an outer membrane lies a formidable array of biochemical machinery and a nucleus enclosed within its own nuclear membranes. This complexity requires eukaryotic cells typically to possess 10,000 times more volume than prokaryotic cells. Eukaryotes are able to assemble to form complex, multicellular organisms—plants, fungi and animals.

Thus, within the living world there are three domains: archaea, bacteria and eukarya. By this definition, viruses and prions are non-living.

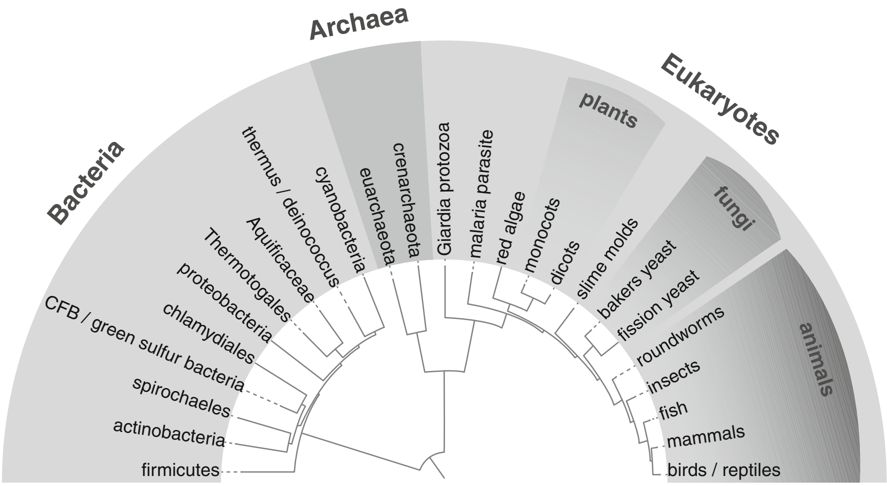



**Fig. 5.18** A highly simplified sketch of the tree of life. The tree contains three domains: archaea, bacteria and eukarya. The domain of eukarya contains the familiar kingdoms of animals, plants and fungi. On a diagram such as this, *Homo sapiens* would appear as a single leaf amongst countless other leaves. The diagram shouldn't be taken too literally—but it does show that life on Earth possesses tremendous unity. (Credit: Madeleine Price Ball)

Second, *a living object must have a metabolism*. Metabolism is what we call the variety of processes enabling a cell, or a collection of cells, to take in energy and materials, convert them for its own ends, and excrete waste products. In other words, all living organisms require food of some description and all living organisms create waste. (Fire has a metabolism, as my old science teacher would point out, but we don't have to regard fire as living since it fails to meet the other criteria.) Metabolism takes place through the catalytic action of *enzymes*: without enzymes, the various biochemical reactions that take place in cells simply wouldn't happen. In turn, enzymes are made of *proteins*. Proteins are therefore a vital constituent of life—at least here on Earth. As we shall see later, the instructions for creating the various proteins necessary for a cell's existence are contained in its deoxyribonucleic acid (DNA), while the biochemical machinery of protein synthesis is based on its ribonucleic acid (RNA). In shorthand form: DNA makes RNA makes proteins.

Third, *a living object can reproduce*—or else it derives from objects that could reproduce. Cells can reproduce either individually or in sexual pairs, and the mechanism of reproduction is DNA. Clearly, then, DNA plays a central role in living organisms—just how central we will come to shortly. Note that crystal structures can reproduce; however, they lack the variation that occurs when living organisms reproduce. Replication, rather than reproduction, is a better term for crystal growth, and certainly we needn't consider crystals as being alive. On the other hand, mules and other sterile organisms came from creatures that *could* reproduce, and thus we needn't classify mules as being non-living.

Fourth, *life evolves*. Darwinian evolution—natural selection acting on heritable variation—is a key aspect of life.

These four properties—cells, metabolism, reproduction and evolution—are enough on which to base a discussion of life, even if the definition itself could be improved. We are now in a position to ask: how did life begin?

## How Did Life Begin?

It's worth stating at the outset that nobody knows how life started. Nevertheless, in recent years tremendous progress has been made in two directions: on the one hand, tracing life's ancestry back as far as possible, and on the other hand attempting to understand the chemical pathways that might have led to the earliest forms of life. (There are various other promising approaches to the problem of abiogenesis, but lack of space prevents us from discussing them.)

The "top-down" method of looking for the origin of life is the search for LUCA—the Last Universal Common Ancestor, the most recent organism from which all present life must have inherited its common biochemical structures. That life stems from LUCA rather than multiple sources seems overwhelmingly likely, since there's a tremendous unity of terrestrial life: all organisms, with a few minor exceptions, use the same genetic code, which enables a sequence of DNA to specify a polypeptide; all organisms use DNA to carry genetic information; and so on. If LUCA was sufficiently simple, if it existed at a very early stage in the history of Earth—and if we can understand LUCA in detail—then we might deduce how it came to be. Unfortunately, biologists can push this approach only so far. One commonly drawn picture is that LUCA was already a sophisticated organism, which had evolved considerably from the time when life first arose, before it branched into the domains of archaea and bacteria. Later, in this picture, the eukaryotic domain branched off from the archaea. The formation of the complex eukaryotic cell probably arose when one prokaryote "ate" another (or, depending on your point of view, one prokaryote "infected" another—it's difficult from this distance to distinguish those two cases). The arrangement must have benefitted both parties and the internal bacteria (whether initially they were food or parasites) were passed on through the generations. This picture is complicated enough, but as the world's many biochemical laboratories discover new information on an almost daily basis, the picture is becoming even more convoluted. We usually think of genetic information as passing only vertically—from parent to child. Early in the history of life, however, *horizontal* transfer of genes between different types of organism seems to have occurred frequently. This horizontal transfer of genetic information means that simple lineages become tangled.

Rather than become bogged down in the details of LUCA, we can consider the "bottom-up" approach to the question of the origin of life. We can ask: how did the universal chemicals of life—nucleic acids and proteins—come into existence? If we can understand that, then we might be able to fill in the gap between the bottom-up and top-down approaches. We might be able to understand how inanimate matter became alive.

## Nucleic Acids

If any molecule deserves the title "molecule of life" it must surely be deoxyribonucleic acid[344]—DNA. According to the definition presented earlier, two of the key aspects of life are that it has a metabolism and that it passes on information through a reproductive process. The DNA molecule is central to both of these aspects. The role it plays in synthesizing proteins, which in turn allow metabolism, is described below. Here, let's concentrate on the reproductive aspect and briefly consider how DNA can replicate itself while providing enough variation upon which natural selection can work.

The DNA molecule is a polymer of *nucleotides*. A nucleotide has three parts.

First, it possesses a deoxyribose sugar. The sugar contains five carbon atoms, conventionally numbered with primes—$1'$ through to $5'$ (pronounced "one prime," "two prime" and so on). The sugar is similar to ribose, but lacks a hydroxyl molecule at the $2'$ position.

Second, it possesses a phosphate group. The nucleotides can link together to form long chains through so-called phosphate ester bonds—bonds between the phosphate group of one nucleotide and the sugar component of the next nucleotide. The sugar–phosphate chains form the backbone of DNA; in the familiar picture of DNA as a "ladder-like" molecule, the sugar–phosphate chains form the "rails" of the ladder. A chain can be indefinitely lengthened simply by attaching more nucleotides through more ester bonds; a DNA molecule can be anywhere between about 100 to a few million nucleotides in length. No matter how long the chain becomes, there are always two ends. One end has a free –OH group at the $3'$ carbon (the $3'$ end) and the other end has a phosphoric acid group at the $5'$ carbon (the $5'$ end).

Third, it possesses a pair of nitrogenous *bases*. These form the "rungs" of the DNA ladder. A base is linked to the deoxyribose sugar at the $1'$ carbon. A base can be either one of the purines, adenine (A) or guanine (G), or one of the pyrimidines, cytosine (C) or thymine (T). Biochemists present the nucleotide

**Fig. 5.19** The double helix structure of DNA is shown here in a computer-generated illustration. (Credit: National Human Genome Research Institute)

sequence in a chain by starting at the 5′ end and identifying the bases in the order in which they are linked; a typical sequence of DNA can thus be written as –G–C–T–T–A–G–G–.

One of the key developments in 20$^{th}$ century science was the realization that DNA in the nuclear material of cells has two strands, twisted around each other to form a double helix, such that one strand is always associated with a complementary strand. The base G is always opposite the base C, the base T is always opposite the base A. This complementarity occurs because only these combinations of base pairs can form hydrogen bonds between them and hold the two strands together. An individual hydrogen bond is weak, but a normal DNA molecule contains so many base pairs that the two strands are held tightly together. This complementarity also means all the information is held in a single strand of DNA—and allows for the possibility of replication and reproduction. (Until recently, essentially all Earth life that has ever existed has had its biological information encoded in four letters, two base pairs: G and C; T and A. As I write this, biologists have announced the production of semi-synthetic bacteria whose engineered DNA contain two extra letters,[345] X and Y. In other words, these modified *E. coli* cells have a third base pair—these cells are a new type of life. Who knows where advances in synthetic biology will take us?)

The process of DNA replication begins when an enzyme called DNA helicase partially unzips the double helix at a region known as the *replication fork*.

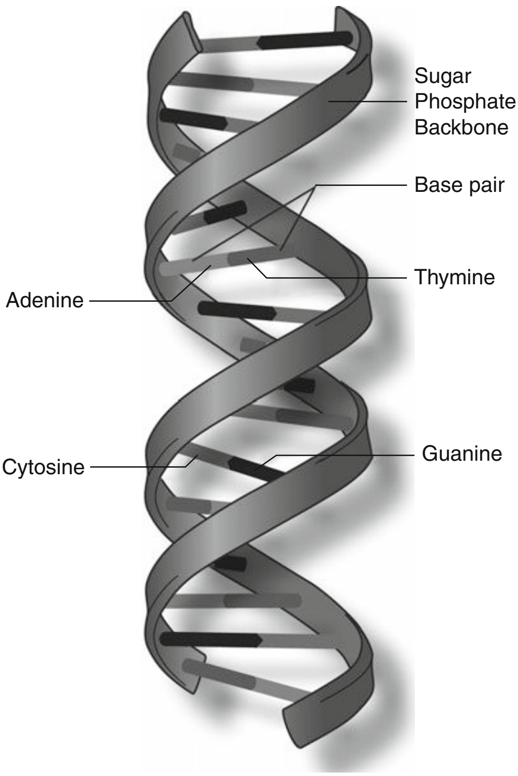



**Fig. 5.20** The backbone of a DNA molecule consists of long chains of deoxyribose sugar and phosphate groups. Nitrogenous bases in each helix form bonds, but they must obey the pairing rules: adenine opposite thymine, and cytosine opposite guanine. (Credit: National Human Genome Research Institute)

At the replication fork there are two strands of DNA—one of which is the *template strand*. With the bases now exposed, an enzyme called DNA polymerase moves into position and begins the synthesis of a DNA strand complementary to the template. The enzyme reads the sequence of bases on the template strand, in the direction from the 3′ end to the 5′ end, and adds the nucleotides to the complementary strand one at a time—always G to C and A to T. (So a sequence on the template strand of –G–C–T–T–A–G–G– would become –C–G–A–A–T–C–C– on the synthesized complementary strand, which grows in the direction from 5′ to 3′.) Eventually, a complete complementary strand is formed; the DNA polymerase catalyzes the formation of the hydrogen bonds between the nucleotides on the two strands, and a new double helix can form. While this whole process takes place, a rather more complicated process manufactures a new strand that is complementary to the other original strand (or *lagging strand*). The net result is the creation

**Fig. 5.21** The specific pairing of nucleotide bases—A with T, C with G—enables DNA to replicate; it's the basis of heredity. When the twin-stranded DNA molecule replicates, the two strands separate at the replication fork. Enzymes (not shown) then add new bases to the two strands while following the pairing rules. The result is two molecules, both of which are identical to the original. (Credit: Madeleine Price Ball)

of two identical copies of the original DNA double helix, and each new helix contains one strand of the original. We have a replication mechanism.

The process outlined above is a simplified version of what actually occurs. One of the aspects I omitted is the role RNA plays in the replication of DNA. Ribonucleic acid is the other major type of nucleic acid and it, too, fulfills key functions for life on Earth. There are several differences between DNA and RNA. A *structural* difference is that RNA usually appears in cells as a single chain of nucleotides, rather than as a double helix of DNA; RNA molecules are also typically smaller than DNA molecules. There are also two *chemical* differences between the molecules. First, the RNA nucleotides contain the sugar ribose rather than deoxyribose (hence the difference in names between the two molecules). Second, RNA employs the base uracil (U) rather than thymine. There's a major *functional* difference between the two acids, too: DNA exists solely to store genetic information in the sequence of its nucleotide bases, whereas RNA molecules *do* things. There are several types of RNA, each performing different tasks, and we shall meet three of them—messenger RNA (mRNA), ribosomal RNA (rRNA) and transfer RNA (tRNA)—below.

The ability of DNA to replicate is the secret of life's ability to reproduce. This ability explains why offspring look like parents—snakes beget snakes, woodpeckers beget woodpeckers, humans beget humans. But for life to *evolve*, and for species to change into other species, heredity must be imperfect. There must be some variation among offspring: natural selection can't adapt things that don't vary. Fortunately, variation *can* arise when DNA replicates. From time to time, a *mutation* occurs: there's a change in the sequence of nucleotide bases. These mutations occur randomly from radiation damage, from chemical agents and simply from errors in the DNA replication process. (The rate of mutation is remarkably small, due to various checks that take place when DNA replicates. After the first stage of replication there are two error-correcting stages: *proofreading* and *mismatch repair*. These extra stages minimize the error rate to 1 in $10^9$.) If an error occurs in a part of DNA that codes for a protein (more on this below), then the mutated DNA will produce a different protein. Usually, the mutation will be harmful or at least neutral. Occasionally, though, the new protein will be better at performing a task than the original protein and the mutation will be beneficial for the organism (and perhaps increase the probability of the organism's survival and thus, through increased numbers of offspring, of its own continued existence). Mutations give natural selection something on which to work.

If all that nucleic acids did was replicate, then they'd be only marginally more interesting than self-replicating crystals. While DNA can *store* genetic information, it would be of little use if the information was not retrieved and put to use. It would be like having a public library stacked full of books, but with no one allowed to read any of the volumes. What makes nucleic acids so fascinating is that they code for and construct proteins. And proteins are what make life so interesting. Proteins enable life to *do* things.

## Proteins

Proteins are complex macromolecules that exhibit tremendous versatility. They function as enzymes (which make possible a cell's metabolism), they act as hormones (thus providing a regulatory function; insulin is a common example), and they provide structure (our fingernails, hair, muscles and the lenses in our eyes are all proteins).

A protein is a long sequence of *amino acids* folded into a three-dimensional structure. A particular sequence of amino acids folds into a particular structure. Change the sequence and you change the way the protein folds up—and thus the task that the protein can fulfill, since the biochemical task that a protein can carry out depends critically upon its shape in three dimensions. Proteins make use of twenty different amino acids. Nature contains many other amino acids, and several of them are important in biology, but proteins employ only twenty. All the amino acids have a common structure: an amino group ($H_2N$), a residue or R group (CHR) and a carboxyl group (COOH). The general structure is written $H_2N$—CHR—COOH, and the chain forms by linking the amino end to the carboxyl end by peptide bonds. (A chain of amino acids is thus called a *polypeptide*; a protein is simply one or more polypeptides.) What makes each amino acid unique is the R side chain: different amino acids have different R groups and thus possess different properties. For example, some side chains create an amino acid that's hydrophobic; such amino acids tend to cluster on the inside of a protein and thus play a factor in determining the three-dimensional structure of the molecule. Other side chains make an amino acid that's hydrophilic—in other words, it reacts readily with water.

Each amino acid is coded for by a set of three RNA nucleotide bases called a *codon*. Since there are four bases (A, C, G, U) there are $4 \times 4 \times 4 = 64$ codons. In theory, then, codons could code for 64 amino acids—and yet only 20 different amino acids are used in protein synthesis. The *genetic code* is thus degenerate: 3 of the codons represent an "end of chain" command, and the other 61 codons code for the 20 amino acids. In other words, nearly all amino acids are coded for by several codons. For example, the amino acid cysteine is coded for by the codons UGU and UGC; isoleucine is coded for by the codons AUU, AUC and AUA; and so on. The genetic code is essentially universal: with only minor exceptions, and disregarding the recent advances in synthetic biology mentioned earlier, all organisms on Earth use it. (Does the universality of the genetic code imply that it's the only possible code? Perhaps there were originally several different codes, and this one just happened to win out over the others? If the present uniqueness of the code means that it arose only once in the history of life, then perhaps the development of an effective code represents a barrier for evolution to overcome–one of Carter's "difficult steps"? We'd learn something about the possibility of extraterrestrial life if we could find examples of the development of different genetic codes here on Earth.)

The way a cell goes about synthesizing a protein is at once both wonderfully straightforward and marvelously intricate. A highly simplified version of the process proceeds as follows.

The information on how to build an organism's proteins—and thus the organism itself—is contained in its DNA. First, then, when a cell receives a signal asking for it to produce a certain protein (and let's suppose the protein is a single polypeptide), the double-helix of DNA unzips in the region of the *coding strand*. This is similar to the template strand mentioned above and contains information for that particular protein. A region of DNA that codes for a polypeptide (or, more accurately, that codes for some form of RNA) is known as a *gene*.

An mRNA copy of the gene is fabricated in a *transcription* process—so called because each triplet in the DNA strand gets transcribed into the corresponding codon in mRNA. The mRNA then moves from the nuclear material to the

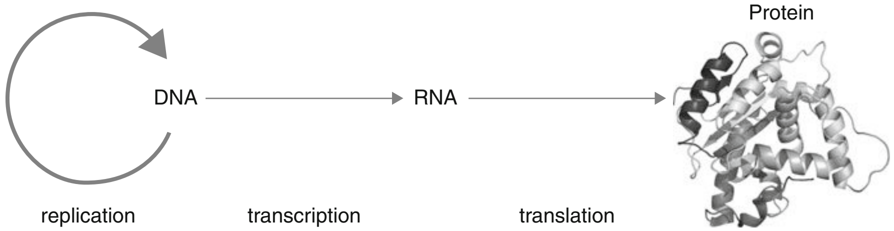



**Fig. 5.22** The DNA molecule stores genetic information, and replicates that information when a cell divides. The expression of that genetic information doesn't take place directly. Instead, DNA is first transcribed into RNA. Information stored in the "four-letter" alphabet of nucleotides (the alphabet used by RNA) is then translated into the "twenty-letter" alphabet of amino acids (which are used to construct proteins). The Central Dogma of biology, first stated by Francis Crick, is that the information flow follows the direction of the arrows in this diagram. In particular, RNA can synthesize proteins through translation, but reverse translation never occurs

cytoplasm of the cell, taking with it its information on amino acid sequences. Within the cytoplasm, organelles called *ribosomes* take the mRNA and use the information contained in the codon sequence to synthesize the protein, adding amino acids onto the growing chain. This process is called *translation*, since a ribosome uses the genetic code to translate from the sequence of codons into a sequence of amino acids. A key ingredient here is tRNA—small molecules, each of which can bind only to a particular amino acid. A series of enzymes is required to catalyze the binding process; each enzyme recognizes one particular tRNA molecule and the corresponding amino acid.

Protein synthesis always begins with methionine (with codon AUG) and it continues until the ribosome encounters one of the stop codons (UAA, UAG or UGA), at which point the protein is released and the synthesis is over. This provides an outline sketch of protein synthesis, at least for prokaryotic cells. In eukaryotic cells, the process is further complicated by the presence of DNA sequences that don't code for anything. A further step is required in eukaryotic cells to remove this seemingly useless information. Space is too limited here to go any further into the details of protein synthesis, but there are many excellent sources available[346] for further reading, and fortunately we don't need extra detail to continue the discussion.

To recap: DNA stores genetic information and replicates the information when a cell divides. The messy business of actually *expressing* the information is left to the more versatile RNA; using the universal genetic code, information is transcribed from DNA into RNA and then translated into protein synthesis.

## How did the Ingredients of Life Arise?

Let's assume, for the moment, that the many intricate steps leading from the first proteins and early nucleic acids through to LUCA are, if not inevitable, at least capable of being understood using well-known physical and chemical processes. We are still left with the question: how did the first proteins and nucleic acids come into existence? If the step from inorganic chemistry to DNA and proteins is a rare phenomenon then we have a resolution of the Fermi paradox, for without these large molecules evolution can't begin the step to LUCA and then to the variety of life we see around us. Without proteins and nucleic acids life, at least as we know it, can't exist.

The basic building blocks of the vital macromolecules appear to be easily synthesized. We find amino acids, for example, both in interstellar space[347] and in laboratory experiments that attempt to mimic the chemistry of early Earth.[348] In 1953, Stanley Miller performed a classic experiment in which he passed an electric discharge through a vessel containing a mixture of water, methane and ammonia. The experiment was intended to investigate the effects of electric currents passing through the atmosphere of the early Earth. At the end of his experiment, Miller found many organic compounds in the vessel. Other scientists have disagreed with Miller's choice of model atmosphere, but the results were unarguably dramatic. It seems probable that amino acids could have formed on Earth soon after our planet cooled; amino acids are almost an inevitability of organic chemistry and the marvelous associative properties of carbon. Similarly, sugars, purines and pyrimidines—the components from which nucleic acids develop—can form in Miller-type experiments (although it must be admitted that yields are often low).

Although the details have yet to be determined, we have no good reason to suppose that the basic chemical building blocks required for life are in any way exceptionally rare. We can be less confident, however, about the probability of natural processes successfully linking these components into the molecules of life—nucleic acids and proteins. Indeed, it's at this point that many creationists (and a few scientists) claim life on Earth is unique: they argue the probability of random processes creating a nucleic acid or a protein is tiny.

Consider, for example, serum albumin (an average-sized protein produced in the liver and secreted into the bloodstream, where it performs several necessary tasks). Serum albumin contains a chain of 584 amino acids, which are curled up into a sphere. In our bodies, the synthesis of the molecule is under the direction of nucleic acids. But imagine a time before DNA existed, so that a molecule of serum albumin had to be synthesized by adding one amino acid at random to the end of a growing chain. The chances are negligible—just 1 in $20^{584}$—that random processes would produce the protein. Similarly, "genesis DNA"—a primitive chain of nucleotides that some scientists have proposed as being necessary for life to start—has a low probability of being created by chance.[349]

**Making a Protein Through Random Processes** Since there are 20 amino acids from which to choose, at each step the probability that the correct amino acid is chosen to add to the end of a growing chain is 1 in 20. Therefore, for serum albumin, which has 584 amino acids, the probability that every amino acid is chosen in the correct order is 1 in $20^{584}$—which is the same as 1 in $10^{760}$. This is an incredibly small probability. There's essentially no chance that this protein can be synthesized by the random process outlined above. Even a small protein such as cytochrome *c*, which consists of just over 100 amino acids, has only a 1-in-$10^{130}$ chance of being synthesized at random. Again, for practical purposes, this number is indistinguishable from zero.

The beginning of life seems to suffer from a "chicken and egg" paradox: DNA contains the instructions necessary for the assembly of amino acids into proteins, but every DNA molecule requires the help of enzymes (in other words, proteins) to exist. DNA makes proteins makes DNA makes proteins . . . which came first?

Although at first glance these criticisms might seem to be fatal to the claim that life arose by chance, biochemists have made great progress in countering them. The details are far from complete, but there's no reason to suppose the problems are insurmountable. Begin with the combinatoric arguments against the primordial synthesis of proteins: there is indeed essentially no chance of cytochrome *c*, for example, somehow coming together by accident. But if we allow for a period of prebiotic *molecular* evolution, then proteins could be synthesized through the workings of chance.

For example, imagine a lake somewhere on the still-young Earth. Suppose that in this lake there were only 10 different amino acids capable of forming peptides; and suppose that a peptide with a length of 20 amino acids displayed some catalytic function making it favored by natural selection. Then Nature only needed to try out $10^{20}$ combinations to hit on this peptide—still an enormous number, but a number that *could* comfortably be accommodated in the timescales available. Once the peptide was created, natural selection would ensure the amount of peptide in the lake increased in volume. Suppose that 1000 different "useful" peptides, each 20 amino acids in length, were created in the lake. If two such peptides could join to form a single chain, then 1 million different peptides with a length of 40 amino acids could be formed. Again, Nature would have plenty of time to try out all the combinations. In the same way, peptides containing 60 amino acids could be synthesized, and 80, and 100 . . . in short, there *was* time for proteins to arise in that ancient lake.

And there were many millions of lakes on the early Earth. (The particular proteins that arose would surely have been an historical accident. Replay the tape of history, and the proteins we use might be very different.)

Similar sorts of argument involving prebiotic molecular evolution can be used to counter the claim that "genesis DNA" was a miraculous fluke. However, such arguments might be unnecessary. It seems plausible that the original self-replicating molecule wasn't DNA, but one of the varieties of the much simpler RNA molecule. Furthermore, RNA provides an answer to the "chicken and egg" paradox. In the early 1980s, Sidney Altman and Thomas Cech demonstrated that some types of RNA molecule could also act as catalysts; they could play the role of enzymes. These RNA enzymes—or ribozymes—led to the idea of the "RNA" world—a time in the early history of life when catalytic RNA enabled all the chemical reactions to take place that are necessary for primitive cellular structures. In a sense, neither the chicken nor the egg came first: catalytic RNA acted both as genetic material and as enzymes.[350]

There seems to be no fundamental reason to suppose that the basic molecules of life could not arise through natural processes that had a reasonable chance of occurring. (Although, in all honesty, one has to concede that the chemical pathways leading to the first RNA molecules are still murky. The subsequent evolution of cellular structures up to LUCA is just as unclear. There are several competing scenarios, each with their advantages and drawbacks. Furthermore, several questions—such as why life uses only the left-hand form of amino acids, and whether the genetic code is inevitable or simply one of a whole raft of possible codes—are outstanding. But progress in these fields is rapid,[351] and we can expect the picture to have more clarity within a few years. Even if life turns out to have a completely different origin from that sketched above—and there are several other competing hypotheses—we aren't yet driven to the hypothesis that life was some bizarre fluke.) There is, however, one last argument to consider regarding the probability of the early Earth being the site of the genesis of life: paradoxically, life seems to have arisen here *too* easily!

## When did Life Arise on Earth?

Life appears to have had little trouble in emerging on Earth. We know our planet formed about 4.55 billion years ago. A maximum of 700 million years after the formation of Earth—3.85 billion years ago—it seems that life had evolved. We believe this to be the case because certain sedimentary rocks in Isua, Greenland—rocks that are among the oldest on this planet—contain isotopes of carbon in a ratio that is a sign of biological processes. The interpretation of these measurements is not without controversy. It's possible that non-biological

**Fig. 5.23** A stromatolite formation in the Bahamas. Stromatolites, similar to the one pictured here, are the oldest known fossils. The oldest of them, found in Western Australia, are 3.5 billion years old. (Credit: Vincent Poirier)

processes can generate a similar isotopic ratio of carbon. Nevertheless, many biologists accept that life was in existence at this time.[352] The earliest fossils are not much younger than the Isua rocks; stromatolites—mounds built up of layers of cyanobacteria and trapped sediment—are preserved as fossils in Western Australia. These stromatolites are 3.5 billion years old.

The haste with which life arose is almost too quick for comfort. The timespan mentioned above for the emergence of life, namely 700 million years, is an upper limit: that time span is squeezed from both ends. On the one hand, there was presumably some evolutionary process leading to the life-forms that might have existed in those ancient Greenland rocks; certainly the cyanobacteria of ancient Western Australia had a biochemistry as sophisticated as more recent forms of life and that sophistication must have taken time to develop. (In other words, if it were possible to find even older rocks then we might well find evidence for life in those rocks—perhaps simpler forms of life, but life nonetheless. Life might have emerged *before* Earth was 700 million years old.) On the other hand, life presumably could not have survived the conditions that were present on the very early Earth. The initial period after formation of Earth, some 4.55 to 3.9 billion years ago, is called the Hadean era. Recent research suggests that Earth's crust formed[353] just 160 million years after the formation of the Solar System itself. However, although the existence of a crust presumably means that conditions on Earth weren't too harsh, the early part of the Hadean era saw our planet peppered by large fast-moving rocks, and some of those impacts would have packed a tremendous punch. It's difficult to comprehend the violence of the literally Earth-shattering impact that gouged out the material that became our Moon. Surely the impact must have sterilized the Hadean Earth—if any form of life was in existence before the impact, it's hard to imagine how it could have survived. So the period of 700 million years postulated for the emergence of life is an upper limit: the actual period was probably much less than this.

Although several hundred million years might seem to offer plenty of time for life to evolve, it's worth remembering that the gap between life and non-life is huge, and that evolution can be a slow process. As the biologist Lynn Margulis famously put it: "The gap between non-life and a bacterium is much greater than the gap between a bacterium and man." And yet this gap was bridged relatively quickly. Some scientists find it difficult to accept that life could have begun so early on Earth without help, and have resorted to the panspermia hypothesis (see page 59). If life came to Earth from the depths of interstellar space, then presumably countless numbers of planets in the Galaxy would have been similarly seeded; life would be everywhere and the Fermi paradox remains as strong as ever. If life came to Earth from Mars, however, then this might hint that life is rare: this possibility is discussed further in Solution 66.

## Finding Life on Other Worlds

There is, of course, a direct way of determining whether life can arise under natural conditions: we could try to find it on other planets. The SETI activity is one way of doing this, but the modern discipline of astrobiology[354] is highlighting another possibility: we could look for primitive life elsewhere in the Solar System or try to observe biosignatures—molecules or phenomena that indicate the past or present presence of life—on distant exoplanets. If we were to find life elsewhere, even the simplest microbe, then we would at least know that life is not unique to Earth. Finding life on just one other world would almost certainly tell us something about how it arose on this one. It would also tell us something about the likely prevalence of life in the Galaxy.

The key ingredient of life seems to be water: find water and there's a chance of finding life. We know that, in the past, Mars almost certainly possessed water; so there's a chance—no matter how remote—of finding fossil remnants of past Martian life. Enceladus, the sixth largest of Saturn's moons, has been observed by NASA's Cassini spacecraft to possess a large subsurface ocean[355] of

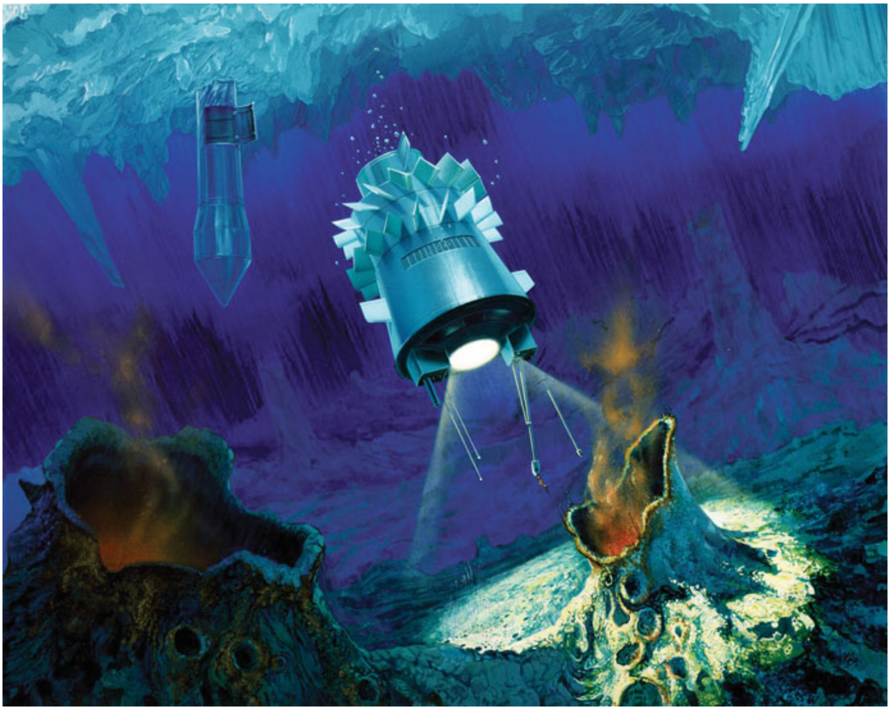


**Fig. 5.24** If there is an ocean beneath the ice of Europa, then a hydrobot similar to this artist's impression will probably be used to explore it. NASA scientists are currently examining the details of how to send a hydrobot to Europa, have it penetrate the ice and reach the ocean without introducing contamination, and then have it send back information to Earth. (Credit: NASA)

liquid water—and the moon is likely to possess an energy source and nutrients as well. It's a good place to look for life. Saturn's largest moon, Titan, might also possess a subsurface ocean of ammonia-water. And two of the moons of Jupiter—Europa and Callisto—might potentially possess liquid water. These bodies are far from the warmth of the Sun, of course, and on the surface of these moons there are thick sheets of ice, but geothermal and tidal heating might be enough to maintain liquid water deep beneath the surface. These four bodies perhaps—just perhaps—are home to alien life. It wouldn't be life with which we could communicate, but if we knew that life arose *independently* in our Solar System more than once then how could we reasonably argue that life is rare throughout the Galaxy? Surely, then, a mission to explore these moons—and particularly Enceladus—should be a priority. Astronomers, meanwhile, are pushing for the construction of telescopes that can search for biosignatures on planets far beyond our Solar System. If life's genesis is common then one day, perhaps in the not too distant future, science will find an example of alien life.

## Solution 65 Life's Genesis is Rare (Revisited)

*The laws of probability, so true in general, so fallacious in particular ...*
Edward Gibbon, *Memoirs of my Life and Writings*

The best way of finding out whether life is abundant in the universe is to go out and look. If we discovered alien life-forms on a variety of exoplanets then we could be fairly sure that abiogenesis—the development of life from prebiotic environments—is a common occurrence. The prevalence of intelligent life-forms would remain unknown, but at least we'd know that the Fermi paradox can't be resolved by claiming abiogenesis is rare. It's difficult to make the relevant observations, however, and it's unclear how quickly astrobiologists will make progress in this regard. Given the observational difficulties, can't we try a theoretical approach? Unfortunately for the theoreticians, we're lacking critical information: we don't know the rate of abiogenesis per unit time and per unit volume as a function of prebiotic chemical and physical conditions. In the absence of this information, one way to procede might be to use our knowledge that life arose at least once on the early Earth and use this knowledge to try and estimate the probability of abiogenesis on an Earth-like planet.

If abiogenesis is improbable then—by definition—there'll be a long time between a planet achieving the conditions suitable for life and life actually developing. On our planet, however, there was a relatively short period of time between Earth cooling and the appearance of life. Does the haste with which cells appeared on Earth hint that the generation of life from inanimate matter is a straightforward process? Can we conclude from Earth's example that the probability of abiogenesis is unlikely to be small[356]—and thus that life is likely to be common in the universe? I have to admit that for a long time I believed this to be almost certainly the case, but is that a reasonable view given the evidence we have?

If we're going to discuss the probability of abiogenesis, a concept about which we have extremely little information, then we must use the correct approach to probability. There are two frameworks for thinking about probability.

The first is to interpret probability as the frequency with which an outcome occurs when you repeat an experiment many times. If you toss an ideal fair coin a billion times then, give or take a few tosses, the coin will land heads-up half a billion times. The probability of throwing heads is thus 0.5. Everyone can accept that. The problem with this approach is that in most situations you can't repeat your experiment. If you are asked to serve as a jury member and must decide the guilt of a defendant beyond reasonable doubt then probability becomes a matter of "degree of belief" rather than the frequency of occurrence. This second approach to probability—one's degree of belief in an outcome rather than the frequency with which an outcome occurs—deals with the messy realities of the world in which we live. It quantifies the degree of belief you should have in a hypothesis given some evidence; as the evidence changes so should the degree of belief. (The famous economist John Maynard Keynes was once berated for changing his mind on key topics. He quite reasonably responded: "When my information changes, I alter my conclusions. What do you do, sir?")

The equation we need to use when discussing probability is the following:

$$P(\mathrm{H}|\mathrm{E}) = \frac{P(\mathrm{E}|\mathrm{H})P(\mathrm{H})}{P(\mathrm{E})}$$

This is one of the most important equations in science. It's quite possibly of even more utility than $F = ma$ or $E = mc^2$. However, unlike Newton's equation for the second law of motion or Einstein's equation showing the equivalence of mass and energy, this equation—despite its importance—is generally unknown to a general audience. Even some scientists don't fully appreciate the formula or apply it correctly, and yet the approach to probability embodied by the formula is indispensable[357] in all branches of experimental science and in medicine, technology, business, warfare ... indeed, in any field where one has to take decisions based on less than perfect knowledge. People are in jail right now because judges and lawyers failed to understand this equation; people have died from cancer because their doctors failed to correctly apply probabilistic reasoning; this formula *matters*.

The formula above is the most common representation of Bayes' theorem, a piece of mathematics named after the English clergyman Thomas Bayes,[358] who wrote down a particular statement of a more general theorem in an essay published after his death in 1761. The formula allows you to calculate $P(\mathrm{H}|\mathrm{E})$, the so-called *posterior probability* of a hypothesis given some evidence. In order to calculate this probability you need to know or be able to estimate the *prior probability* $P(\mathrm{H})$, the *likelihood* $P(\mathrm{E}|\mathrm{H})$ and the probability of the evidence $P(\mathrm{E})$. Before I discuss what this formula has to do with the Fermi paradox (or rather, as it turns out, what it *doesn't* have to do with the Fermi paradox) I need to explain a little bit more about Bayes' theorem. If you already understand what the Reverend Bayes had to say about probability, then feel free to skip the next couple of pages.

Let's look at an example of Bayes-type reasoning in a science fictional context. Suppose government agents have uncovered a plot by extraterrestrials to take over the world: alien shapeshifters, who can assume the form of man or woman, are mingling with the general population. There aren't too many of them just yet—the agents have good reason to believe that only 1% of the general population are, in fact, aliens in disguise. There's an app for telling whether a given individual is an alien, and it's effective: 80% of aliens will show up correctly as possessing alien DNA. But the test isn't perfect: 9.6% of humans will show up falsely as being alien impostors. An agent uses the app on a random passerby and the app flashes a positive result. Given this scenario, what's the probability that the passerby really is an alien?

Before reading on, think about the scenario and make an estimate of the probability.

If you estimated the probability to be between 70%–80% you're in good company. Study after study shows that[359] when doctors are given this problem (with the language of breast cancer and mammograms replacing the language of aliens and shapeshifter-detecting apps) about 6 in 7 of them give an estimate in the 70%–80% range. The correct answer is 7.76%. In other words, even though the accuracy of the test is 80% a positive test result means only a 7.76% chance of the individual being an alien (or of having breast cancer, if the problem is couched in that language). If you want to see how the numbers stack up, the box below applies the Bayes formula with the numbers given in the scenario.

**Bayes and the Shapeshifter Problem** We want to know *P*(H|E)—the probability of the alien shapeshifter hypothesis being true given the evidence of a positive test result.

What we are told is the likelihood *P*(E|H)—the probability of a positive test occurring given the individual is an alien, which in this case is 80%. Bayes tells us that we also need to take into account *P*(H), which is the chance that an individual chosen at random is an alien—1% in this case. Furthermore we also need to take into account *P*(E), which is the chance of getting *any* positive test—either a real positive or a false positive. In this case the chance of a real positive is 1% × 80% or 0.008. The chance of a false positive is 99% × 9.6% or 0.09504. So *P*(E) is equal to 0.008 + 0.09504 or 0.10304.

Put these numbers into the Bayes formula and you find that *P*(H|E) = 7.76%.

Why do so many people get these sorts of problem wrong? The ultimate reason seems to be that they mentally replace the *question* ("what is the probability of an individual being an alien if he or she has a positive test") with the given *information* ("what is the probability that an individual alien will return a positive result on the test"). One of the key benefits of the Bayes theorem, apart from it being the correct approach to reasoning, is that it reminds us to take into account all relevant information when calculating probabilities. Bayes keeps us honest.

For one more example of how Bayes can help make sense of probabilities and tell us how we need to revise our estimates in the light of changing information, consider the infamous Monty Hall problem.[360] The problem was inspired by an American television game show called *Let's Make a Deal*, which ran from 1963 to 1976. The show's host was Monty Hall.

Monty shows you three closed doors behind one of which is a shiny new Bugatti Veyron Grand Sport Vitesse; the other two doors hide lemons. You're given the opportunity to select one of the doors and win whatever is behind it. Clearly, unless you *really* like lemons, you'll want to win the Bugatti. You pick a door. Monty, who knows what is behind each of the doors, then opens one of the remaining doors and reveals a lemon. He then offers you a choice: you can either stick with your original selection or you can select the other unopened door. Should you switch or should you stick with your original selection? Does it make any difference if you switch?

As with the shapeshifter problem above, think about the scenario and commit to an answer before reading on.

When I first heard the question my reaction[361] was that it can't possibly matter whether you stick or twist. The car is equally likely to be behind either door, so my chance of winning would be 50:50. Might as well stick. It turns out that you are twice as likely to win if you change your door. If you want to see how a Bayesian approach leads to the correct answer, and how it demonstrates that we should alter our conclusions when our information changes, see the box below.

**Bayes and the Monty Hall Problem** Let's label the three doors A, B and C and let those letters stand for the event that the Bugatti is behind those doors. It doesn't matter which door you choose, but let's suppose that you choose A. Since Monty won't reveal the door that has the Bugatti behind it then, if the car is behind door A, Monty will randomly pick door B or door C.

The prior probabilities are easy to understand, because at the start of the game you can have equal confidence in the Bugatti being behind any of the three doors:

$$P(A) = P(B) = P(C) = 1/3.$$

Now let's look at the likelihood. You should be able to work out why the probabilities take the values they do.

The probability that Monty opens door B if the prize is behind door A is:

$$P(\text{Monty opens B}|A) = 1/2.$$

The probability that Monty opens door B if the prize is behind door B is:

$$P(\text{Monty opens B}|B) = 0.$$

The probability that Monty opens door B if the prize is behind door C is:

$$P(\text{Monty opens B}|C) = 1.$$

We can now calculate the probability that Monty opens door B:

$$P(A) \times P(\text{Monty opens B}|A) + P(B) \times P(\text{Monty opens B}|B)$$
$$+ P(C) \times P(\text{Monty opens B}|C) = 1/6 + 0 + 1/3 = 1/2.$$

Finally, apply Bayes theorem:

$$P(A|\text{Monty opens B}) = 1/6 \div 1/2 = 1/3.$$
$$P(C|\text{Monty opens B}) = 1/3 \div 1/2 = 2/3.$$

In plain English: if you happen to choose door A and Monty opens door B to reveal a lemon then the probability of the Bugatti being behind door C is 2/3. If you ever find yourself in this situation, you'll double your chances by switching.

These examples show that if we want to talk about the probability of abiogenesis we need to use Bayesian language. We observe that life arose quickly on Earth, but we can't just conclude from this observation that abiogenesis is easy. It *might* be easy—but only a Bayesian analysis can quantify the degree of confidence we should have when we say abiogenesis is easy. Two astrophysicists, David Spiegel and Edwin Turner, undertook just such a Bayesian analysis.[362]

To develop the analysis, Spiegel and Turner proposed a simple model (or hypothesis, using the language above) of abiogenesis. In their model, conditions on a young planet preclude the creation of life; at a certain point life becomes possible and there is then a constant probability per unit time of life developing; and after a certain point, perhaps because of the evolution of its star, the planet again becomes inhospitable to life's creation. The model is simplistic; one could argue that abiogenesis isn't a single event that happens at a particular instant, and it's possible that the probability of abiogenesis per unit time varies with time rather than being constant. Nevertheless, we have no good grounds for proposing some other, more complicated model—so the Spiegel and Turner model is as good a starting point as any other. That's the hypothesis. The evidence we need to take into account is that life arose on Earth at least once, about 3.8 billion years ago, and that this gave sufficient time for cosmologically curious creatures to emerge and be able to contemplate matters such as the Bayes formula and the possible existence of cosmologically curious creatures elsewhere in the universe.

Bayes tells us we must also specify the prior probabilities of the various terms in the model. (Those terms are the rate of abiogenesis, the times before and after which a planet cannot give rise to life, and the minimum time required for intelligence to develop.) There's no theory that predicts values for the various timings in the model so Spiegel and Turner simply choose several different interesting cases. Similarly, there's no underpinning theory that gives us prior information on the rate of abiogenesis, so Spiegel and Turner investigate three different forms for this rate.

The mathematics involved in the analysis is rather more involved than the "alien shapeshifter" and Monty Hall problems discussed above, so I won't repeat it here. But the logic is the same in all three cases: a probability is computed by making use of all available information. The result? Well, it turns out the results are much more heavily influenced by the choices of prior probabilities than they are by the evidence that life arose early on Earth. Choose one set of model parameters and life is common; choose a different, equally plausible set of parameters and life is exceptionally rare. In other words, with the evidence we have available to us, the fact that we are here is entirely consistent with a low probability of abiogenesis. The fact that life arose early *here* gives us little confidence in the belief that life must be common *elsewhere*. It's important to emphasize the following point: *the analysis doesn't show that life is rare*. "Life is common" remains our best guess position; it's just that we can't be confident of the position.

When the Spiegel and Turner paper appeared, some commentators offered it as a resolution of the Fermi paradox. However, this was based on a misreading of the paper. It didn't prove that life, and therefore intelligent life, must be rare. To repeat: the analysis just showed that we can't be confident, given the information we currently possess, that abiogenesis is common. So this isn't a resolution of the paradox. It does, however, highlight the importance of looking for extraterrestrial life: the discovery of one instance of life arising independently of Earth would give us much stronger grounds for believing that the universe is teeming with life—and that in turn might increase our hopes of finding intelligence.

## Solution 66 Goldilocks Twins are Rare

*It takes two to tango.*

Al Hoffman and Dick Manning

Every summer scientists travel from international conference to international conference, like flocks of some exotic migratory bird. The news offices assigned to these conferences scour the abstracts searching for work they think might garner publicity. They issue press releases based on those conference talks; in turn these releases are often picked up by the media; and then those stories appear on Twitter and in the blogosphere, with social media momentarily magnifying their visibility. (I'm as guilty as anyone when it comes to this. You're not the only one, Gentle Reader, whose attention span has been shortened by the internet.) In August 2013 an interesting piece of astrobiological research underwent this treatment: for a few days, newspapers and websites were asking the question: are we all Martians? Despite all the comment, the honest answer to the question is: we don't know. But possibly we are. And if we are, that might have implications for the Fermi paradox.

The conference talk that led to this flutter of interest[363] was given by Steven Benner, a distinguished chemist who has undertaken important research in several fields, including synthetic biology. Benner took as his starting point the fact that, as discussed on page 278, we still don't know how atoms first combined to make those key elements of life on Earth—RNA, DNA and proteins. Suppose it was indeed RNA that came first: well, as chemists have discovered in the decades since Miller's pioneering experiment, zapping the primordial "soup" of organic chemicals that existed on the early Earth generates not ribonucleic acid but gloopy tar-like substances instead. One suggestion is that there was a catalyst—an inorganic, mineral surface—that provided a scaffold on which basic building blocks could assemble into RNA structures. The best scaffolds would have contained boron and oxidized molybdenum: boron-containing minerals assist the formation of prebiotic chemicals from carbohydrate rings and then molybdenum-containing minerals help rearrange those prebiotic chemicals to form ribose, from which RNA follows. It's an excellent suggestion, but there are at least two difficulties with it. First, boron compounds would have dissolved in the early Earth's oceans. Second, the molybdenum would need to be highly oxidized in order to perform its scaffolding function but, back then, Earth's surface contained very little oxygen. So one seemingly critical element was missing and another was in the wrong form. How, then, could the scaffolds have come into being?

Benner pointed out that the early Martian surface might have had those elements that Earth's surface was missing: Mars was drier and had more oxygen. The scaffolding chemistry might have had a better chance of taking place on the red planet than here. The basic building blocks of life, in other words, might have been created on Mars. Benner's suggestion isn't the only approach—it's possible, for example, that the precursors to RNA were indeed made on Earth but that a quite different catalyst was involved—but *if* his suggestion is correct then the idea is that life-bearing rocks were sent into space following asteroid impacts. Those rocks then found their way to Earth. As conditions on Mars changed, life presumably died there; and as conditions on Earth changed, life prospered.

The idea that life could be transferred from one planet to another is far from fanciful; indeed, we've already looked at the possibility in Solution 6 when discussing panspermia. We know that rocks *do* move from planet to planet following a meteor strike. Of the tens of thousands of meteorites that have been found on Earth, scientists have identified just over one hundred of them as having a Martian origin.[364] Over the ages, perhaps as much as a billion tons of rock has travelled from Mars to Earth. Some rocks have even gone the other way; although orbital dynamics informs us that the Mars–Earth trip is about a hundred times more energetically favorable than the Earth–Mars trip, calculations suggest that the dinosaur-killing Chicxulub impact was strong enough to eject[365] 360,000 potentially life-bearing rocks to Mars. (A handful of rocks thrown up by the impact could even have reached Jupiter's moon Europa!)

To repeat: we don't know whether Earth life originated on Mars. We don't know whether we are all Martians. But it's possible. And the implication for the Fermi paradox? Well, perhaps *two* planets are required for life to thrive: one to provide the starting conditions and one to provide a long-term home. It's even possible that the stirring of those prebiotic soups requires *multiple* transfer of material between two planets. Earth is sometimes called the "Goldilocks" planet because conditions have been "just right" for so long. Could it be that life requires "Goldilocks twins"? If that is the case, the number of life-bearing planets might be small. Perhaps that explains the absence of ETCs?

As far as I'm aware, this solution to the Fermi paradox is original to me. However, I wouldn't buy it if I were you—we need to understand much more about the origin of life before this solution needs to be taken seriously.

## Solution 67 The Prokaryote–Eukaryote Transition is Rare

*Life may change.*

Percy Bysshe Shelley, *Hellas*

For a significant fraction of Earth's history the only organisms that existed were single-celled prokaryotes. It took at least a billion years for the byzantine biochemical machinery of the eukaryotic cell to emerge. The development of large multicellular organisms took longer still. This isn't surprising: eukaryotic cells are immensely more complex than prokaryotic cells, and several evolutionary developments had to be made before different eukaryotic cells could learn to cooperate and function effectively in groups. But perhaps the aeons-long wait for eukaryotes to arise here on Earth implies that the development of a sophisticated grade of life follows a tortuous, difficult path. Since complex life of whatever form must presumably evolve from simpler single-celled microbial life, perhaps elaborate multicellular life—and thus eventually life capable of communicating over interstellar distances—has yet to emerge on other planets. Perhaps the prokaryote–eukaryote transition is one of Carter's "difficult steps". Perhaps this explains the silence we observe: the Galaxy is filled with planets on which life has stalled at the prokaryotic stage.

What led to the change from the prokaryotic grade of life, which dominated life on Earth for so long, to the eukaryotic grade of life we see all around us today? To answer that—and to attempt to understand whether the eukaryotic grade of life might be a rare phenomenon—we need to understand something of the differences between two types of cell.

## Differences Between Prokaryotic and Eukaryotic Cells

Whichever way you look at it, bacteria have always been the most successful life-forms on Earth. Even the human body contains many more microbial cells than it does human cells; bacteria swarm all over our skin and in our gut (and in many cases are necessary for good health). The simplicity of bacteria, combined with their capacity to reproduce quickly, is almost a guarantee of success. They evolve biochemical responses to environmental challenges, so even though they all tend to look alike, different bacterial species possess different metabolisms and can inhabit a wide variety of niches. They are also extremely hardy, and some species seem to have survived unchanged for billions of years.

Complex eukaryotic life-forms such as animals are much less robust. They are prone to mass extinctions, and even in the natural run of things the typical lifespan of an animal species is measured in millions rather than billions of years. Nevertheless, the eukaryotic grade of life is much more interesting than the prokaryotic grade. Eukaryotes evolve morphological responses to environmental challenges—in other words, they develop new body shapes and body parts—which leads to a variety and freshness that is absent in the prokaryotes.

A major difference between eukaryotic and prokaryotic cells is that the latter have rigid cell walls or very rigid cell membranes whereas eukaryotic cells either lack cell walls or possess very flexible walls. This flexibility allows eukaryotic cells to change shape, and also to engage in *cytosis*—a process wherein the cell membrane pushes inward to form an intracellular vacuole. Many cellular processes employ cytosis, but perhaps its main role is in *phagocytosis*. In phagocytosis, a eukaryotic cell engulfs a particle of food into a food vacuole, where enzymes then digest it. Obtaining nourishment in this way, by predation, is a much more efficient process than that employed by bacteria, which secrete digestive enzymes into the surrounding medium and then absorb the resulting molecules.

Another distinguishing characteristic is that a eukaryotic cell has a *nucleus*, which contains the cell's DNA. Two membranes separate the nucleus from the *cytoplasm*—the place where most cellular activities take place. Eukaryotic cells also contain *organelles*—"little organs"—which are separated from the rest of the cytoplasm by membranes. The organelles include the *mitochondria* (which play a vital role in energy metabolism) and the *plastids* (which play a role in photosynthesis in plants and algae). In the early 1970s, Lynn Margulis argued that organelles must have arisen by symbiosis. She reasoned that, billions of years ago, very primitive eukaryotic cells would have used phagocytosis to ingest smaller prokaryotic cells for food. Some prokaryotic cells might have been indigestible and would have stayed in the larger eukaryotic cells for some time. And some of those prokaryotes would have performed functions—such as the transformation of energy—more efficiently than their hosts. Both cells would benefit from partnership—and both would have a selective advantage when it came to passing on their genes. An initially indigestible bit of food would become indispensable to the smooth running of a eukaryotic cell. Margulis had to fight hard for her idea to be accepted, but support came from DNA sequencing. Mitochondria and plastids have their own DNA, which is different from the DNA in a cell's nucleus. It turns out that mitochondrial DNA and plastid DNA are much closer to prokaryotic than eukaryotic DNA. The mitochondria, for example, probably share a closest common ancestor with present-day symbiotic purple non-sulfur bacteria.

Another major difference exists between the two cell types: unlike prokaryotes, new eukaryotes can form through the fusion of gametes from two parents—in other words, sex can occur. Furthermore, the amount of genetic information stored by eukaryotes (and passed on either through sex or through parthogenesis) is far greater than that stored by prokaryotes.

Finally, eukaryotes possess a *cytoskeleton*. The cytoskeleton consists of actin filaments, which resist any pulling forces that might act on a cell, and microtubules, which resist any shearing or compression forces that might act on a cell. Thus, even in the absence of a rigid cell wall, a eukaryotic cell can maintain its shape and integrity. But the cytoskeleton can do much more: it can draw the cell into a variety of temporary shapes, it marshals the organelles into various positions, and it allows the eukaryotic cell to increase in size. Actin and tubulin—the structural proteins from which the cytoskeleton forms—are thus among the most important of all proteins for the development of complex life.

So how probable was the emergence of the eukaryotic cell? Was it inevitable, this transition from a primitive prokaryotic cell to the awesome complexity of a modern eukaryotic cell, or was it a fluke? These are difficult questions to answer, not least because the many steps involved in the transition occurred so long ago. One of the first steps must have been loss of the rigid cell wall, even though this would have been fatal to most organisms that attempted it. (Penicillin, for example, works by blocking the formation of bacterial cell walls. Without a rigid wall to protect them, most single-celled organisms are vulnerable to attacks from the environment.) Disposing of the cell wall was *ultimately* useful, because it enabled phagocytosis to occur, but phagocytosis evolved at a later date and thus could have provided no *immediate* benefit to the organism that lost the wall. Evolution has no foresight; unless an organism can survive in the here-and-now and pass its genes on to offspring, any potential it might possess will be lost. Somehow, in ways not yet fully understood, some organism managed to employ new structural proteins—actin and tubulin—and develop a cytoskeleton that helped mitigate the loss of the wall. How likely was this occurrence? We don't know, but it's certainly possible that the eukaryotic cell arose due to a rare and random event—a freak of nature. And what about the origin of what might be the most important innovation of all: cooperation between cells?

## Multicellular Organisms

A few prokaryotes have adopted a multicellular way of life. Stromatolites, for example, consist of bacterial colonies. In general, though, prokaryotic cells live a solitary life (and even in the case of stromatolites it's debatable whether the term "organism" is warranted). For most of Earth's history, eukaryotic cells also lived isolated lives. Then a remarkable transformation occurred. Some eukaryotic cells discovered the benefits of joining together. Because the cells had no external walls isolating them from the environment and from each other, they were free to exchange information and to share materials. The result was the world we see today: three kingdoms of organisms that are hugely complex and various—fungi, plants and, most complex of all, animals.

What caused eukaryotic cells to pool their resources isn't known. It's not even entirely clear *when* the switch to multicellularity occurred. The Cambrian explosion of 540 million years ago, which saw the various animal body plans laid down, was surely a crucial event in the history of life on this planet and it seems to have been a key step on the path to intelligent life. But the details are far from clear. We do know that there's little evidence for animal fossils in rocks older than 540 million years, whereas the Cambrian explosion saw the fossilization of a broad assortment of animals. However, all we can firmly deduce from this observation is that large animals with hard body-parts became common in the Cambrian period; it's entirely possible that small soft-bodied animals were in existence before the Cambrian period and died leaving no trace. (Nematodes are perhaps the most abundant type of animal in the world today. They must have existed since at least the Cambrian explosion, yet they've left no trace in the fossil record.) Gene sequencing leads some biologists to believe animals originated about one billion years ago, which, if true, means the fossil record relates to only half of the history of animal life on Earth. However, whether animals originated a billion years ago, half a billion years ago or some time in between, the fact remains they are johnny-come-latelies in the history of our planet. Single-celled creatures had been around since soon after the Earth cooled; it took 3 billion years for evolution to produce complex creatures. Why the long wait for multicellularity?

One possibility is that a rise in the oxygen content of the atmosphere ignited the Cambrian explosion.[366] Early in Earth's history there was essentially no free oxygen, a situation that posed no hardship for primitive prokaryotes because for those first living organisms exposure to oxygen meant certain death. (Even some present-day bacteria find oxygen to be a lethal poison.) However, organisms such as cyanobacteria produced oxygen as a by-product of their metabolism. For 2 billion years—from about 3.7 billion years ago to about 1.7 billion years ago—these organisms pumped oxygen into the environment. For most of that time there were enough sinks, such as iron dissolved in the oceans, to trap the oxygen. Eventually, though, the sinks became full—and the oxygen content of the atmosphere began to rise. For many organisms, this event spelled doom; the "oxygen crisis" must have created the biggest of all mass extinctions, with many prokaryotic species simply failing to adapt to the large-scale release of such poison. Some organisms, though, prospered: they evolved a metabolism based on oxygen, breaking down food into carbon dioxide and water. This oxygen metabolism generated more energy than did the anaerobic metabolisms, and the organisms prospered; the eukaryotes prospered most of all. Even until about 550 million years ago, however, the amount of oxygen in the atmosphere and dissolved in the oceans was far less than today. Any animals existing before this period must have obtained oxygen for their tissues by diffusion, which is a slow process. Those animals would have had no heart—at least, no pump—nor would they have possessed a circulatory system. They would have been tiny, gossamer-like creatures, so it's little wonder they left no trace in the fossil record. But then, for some reason that's not entirely clear, the atmospheric oxygen level rose yet again in the Cambrian period. Several key evolutionary developments took place—gills, hearts, haemoglobin in blood—allowing marine animals to make much more efficient use of oxygen and to transport the gas to different tissues. Animals became bigger and bulkier and were able to develop various specialized organs. Perhaps the emergence of a predator caused other species to evolve protection in the form of hard shells—and finally animals could become fossils.

The suggestion, then, is that the Cambrian explosion was caused by a rise in the level of oxygen in the atmosphere. And perhaps this was a less-than-inevitable occurrence. Perhaps on most planets the development of large multicellular organisms doesn't take place.

Another suggestion, not necessarily in contradiction to the ideas outlined above, is that evolution "stalled" for a billion years or so because of a long period of tectonic stability.[367] About 1.8 billion years ago, most of Earth's land mass is hypothesized to have been in the form of a supercontinent called Rodinia. However, rather than breaking up after a few tens of millions of years (which is the timescale for substantial change due to continental drift) Rodinia remained in the mid-latitudes until it broke up about 750 million years ago. Rodinia's stability seems to have arisen because at that time Earth's mantle was still so hot that it softened the ocean crust; subducting zones were unable to pull down large areas of Rodinian crust as they would do today. By about 750 million years ago the mantle had cooled sufficiently for modern-style tectonic activity to commence, at which point Rodinia's days (or at least its epochs) were numbered: it was ripped apart. The geologists Peter Cawood and Chris Hawkesworth found that oxygen levels varied before Rodinia formed and after it broke up, but were stable during the billion years of the supercontinent's existence; major glaciation events took place before Rodinia was born and after it died, but not during its lifetime; for a billion years, Earth was a boring place. Perhaps the emergence of the eukaryotic cell, which is a marvel of biochemical sophistication, required that long period of stability? Perhaps the development of complex animal life was a response to the new environmental challenges posed by the breaking up of the long-stable supercontinent?

Another suggestion, then, is that complex life is the product of geological conditions that are "just right". Perhaps on most planets the development of complex multicellular organisms is inhibited by a geology that's too active or too static.

## Energy Considerations

Even if the emergence of the eukaryotic cell is a rare and chance happening—and we've looked at several reasons why this might be the case—who's to say that billions of years of evolution won't produce bigger and more complex prokaryotes? Admittedly, four billion years of evolution has failed to produce large, complex prokaryotic-based life-forms here on Earth, but perhaps things could turn out differently on other worlds? Well, that's unlikely: as we'll see below, energy considerations mean that prokaryotes will tend to stay small and simple. Before we move on to other topics, then, here's one more argument why the possibly freakish nature of the prokaryote–eukaryote transition might explain the Fermi paradox.

Earlier we looked at the various chemical and biochemical requirements for life, but we omitted to mention one other factor that's vital for life: energy.

All living organisms require a supply of energy in order to carry out the various processes that keep the organism alive. Organisms use a variety of methods for obtaining their energy, but before they can use the energy it has to be transformed into a form it can handle. All life on Earth employs the same fuel: the molecule adenosine triphosphate (ATP). Life requires vast amounts of energy. The average human body contains about 250 g of ATP but our insatiable need for energy means that we constantly recycle the molecule; we each turn over the equivalent of our bodyweight in ATP every day. But how can cells make this fuel? Well, in 1961 the British biochemist Peter Mitchell[368] proposed that cells are powered by an electrical potential difference that can exist between two sides of a membrane; the potential difference arises because certain proteins act as "proton pumps" that create different proton concentrations on either side of the membrane. Mitchell's proposal was eventually proved correct: a cell is like a tiny battery. The difference in proton concentration can produce a potential difference of 150 mV across the membrane and, because this operates over a distance of only 5 nm, the field strength is about 30 million volts per meter. It's as if living cells have bolts of lightning at work inside them. This electric potential is employed by cells to make the fuel ATP.

The universality of this proton concentration mechanism in living cells suggests that it arose early; the details of *how* it arose remain unclear, but there's no reason yet to suppose that it involved some miraculous event. What we do know, however, is that there was no gradual progression from simple to complex life: as discussed above, it took a long time before the eukaryotic cell arose—an event that seems to have occurred only once in our planet's history (although it's possible that the first occurrence precludes subsequent occurrences). Furthermore, there's no evidence of the intermediate cells that would have existed had simple prokaryotic cells gradually evolved into complex eukaryotic cells. Instead, we see a great divide in life on Earth. On one side of the divide are the prokaryotes: small in cell volume and small in genome size. On the other side of the divide are the eukaryotes: thousand of times larger in both size and genome.

So why have prokaryotes stayed small and simple? The biochemists Nick Lane and William Martin examined this question in terms of the energy requirements of cells of different size.[369] They found that, if cells are powered by a gradient across their membranes, then in terms of energy it is overwhelmingly favorable for the cell to stay small. Suppose you could take a typical prokaryote and blow it up to the size of a typical eukaryote: each gene in the now-giant prokaryote would have tens of thousands of times less energy available than each gene in a eukaryote. The giant prokaryote couldn't function because genes need energy—lots of energy—in order to make proteins (and it's the proteins that perform the various activities of life). The fact that some truly giant bacteria *do* exist just reinforces the point: these giant bacteria all have thousands of copies of their complete genomes, so each copy of a gene has about the same available energy as in normal-sized bacteria. Why is this the case? Well, the problem for cells is the huge potential difference that exists across their membranes. The potential allows the cell to do useful work, but if the potential gets out of control then it can kill the cell—it's as if the cell is struck by a lightning bolt. It seems that genomes govern the membrane potential by directing the production of proteins; the genomes, which are located near the membrane, respond appropriately if the potential looks like it's getting out of control. This, then, is the bind in which simple cells powered by a membrane potential find themselves. In order to get bigger and acquire more genes and thus more complexity requires the generation of more energy; more energy can only come from increasing the area of the membrane; but to control the larger membrane requires further copies of the genome. The energy available to each copy of the gene remains about the same. Nothing is gained. Bacteria work just fine at a small size, but there's an energy barrier that prevents them from getting bigger. If we ever discover extraterrestrial life, and find that it's powered by a straightforward membrane potential, then the chances seem good that the life will consist of small, simple forms: those cells won't be able to evolve the complexity that permit animals and, eventually, intelligence.

Eukaryotes don't have the same limit imposed upon them. Why not? Well, they possess mitochondria—small structures whose function now is to act as power generators; mitochondria contain the ATP-making membrane and the genome to control the potential across the membrane. With its energy demands taken care of by mitochondria, the rest of the eukaryotic cell is free to grow in complexity. That's how things are now, but it wasn't the case for much of Earth's history. Cells were small. One fateful day in the distant past, however, one simple cell ingested another simple cell (or one cell infected another). Rather than one of the cells dying, however, they somehow managed to coexist and have offspring. The smaller cell got ever-smaller and became the mitochondria we see today; that allowed the host cell to accumulate ever-more DNA. The eukaryotic cell was born. But the birth of the eukaryotic cell appears to have been a fluke, a freak event that happened once here on Earth. There's no guarantee it will happen anywhere else.

Could it be that only Earth experienced the right sequence of biological and environmental events that make possible the evolution of animal life? It seems at least a plausible resolution of the Fermi paradox that life elsewhere in the Galaxy has stalled at the unicellular stage. We might one day visit distant planets and find everywhere oceans teeming with strange, microscopic organisms—lots of life, but not life with which we can communicate.

## Solution 68 Toolmaking Species are Rare

*Man is a toolmaking animal.*

Benjamin Franklin (attributed by James Boswell, *Life of Johnson*)

Let's suppose that, once the eukaryotic cell has developed, complex animal life will eventually make an appearance. Does it follow that an animal species capable of building a radio telescope will develop?

People have long sought to identify one defining characteristic of humankind—one attribute distinguishing *Homo sapiens* from other species. A trait often proposed for this role is tool use and toolmaking. "Man the Toolmaker" is a powerful image. If toolmaking is unique to humans, if among the billions of species that have ever lived on Earth *H. sapiens* alone has mastered the intricacies of tools, then we might have a resolution of the Fermi paradox: perhaps tool use and toolmaking are rare *anywhere* in the Galaxy. Without tools to build spacecraft or construct beacons, it's presumably impossible for a biological species to make its presence known across the depths of space.

There's a major difficulty with this suggestion: many species use tools and some species make them.[370]

For example, several types of bird use twigs to pry out grubs from the bark of trees. Sea otters place anvil stones on their chests and use them to smash open crab shells. Wasps use small pebbles to help hide the entrances to burrows where they've laid eggs. Egyptian vultures pick up rocks in their talons and drop them on ostrich nests to crack open the eggs. The list of tool use among animals is a long one. Of course, one might question whether these examples constitute true tool use. These animal behaviors are all highly stereotyped; they are specific, repetitive responses to particular problems. Change the nature of the problem and these creatures are lost. Nowhere do these animals display insight. Those elaborate displays are the intelligent result of brainless evolution.

If we require more sophisticated examples of tool use then we are forced to look to the primates. At this point *H. sapiens* begins to appear somewhat special, for even among the primates there are relatively few "real" examples of tool use. Apart from the great apes, which we'll come to in a moment, the only primate that spontaneously uses tools in the wild is the capuchin monkey (the type of monkey employed by organ-grinders). Field workers have observed capuchins put stones and sticks to a variety of uses; among other things, the monkeys use them to obtain food and repel predators. In laboratory settings, capuchins learn to use sticks to obtain nuts from different experimental setups. However, it seems that capuchins have no real understanding of the *principles* of tool use, nor any comprehension of why a particular technique might work or fail. Watch them, and it's clear they engage in trial-and-error prodding and poking.

Of all the animals, it's the chimpanzee that seems to make the most creative use of tools in the wild. The chimpanzees of West Africa, for example, use a hammer stone and an anvil stone to crack open nuts (and they make a better job of cracking nuts than I do at Christmas). Suitable stones can be in short supply, and the chimpanzees often have to carry them over long distances to a source of nuts. These chimps plan ahead. The chimpanzees of Tanzania use a variety of twigs for a variety of purposes, and the twigs are modified

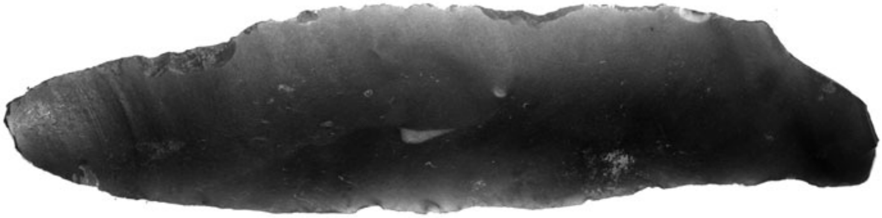


**Fig. 5.25** This serrated stone blade, which was made by some unknown person many millennia ago, is about 9 cm long. It's a simple tool, but the construction of such a blade or scraper is quite beyond the abilities of animals. (Credit: Derby County Council)

beforehand if necessary. These chimps are *making* tools. They also employ various items of foliage for a variety of functions—banana leaves are used as umbrellas, smaller leaves are used to wipe off dirt, and chewed leaves are used as sponges. Even more impressive is the achievements of Kanzi, a bonobo.[371] (The bonobo, along with its sibling species, the chimpanzee, is our nearest relative in the animal kingdom.) Among many other accomplishments, Kanzi appears to have mastered the rudiments of stone tool production.

**Kanzi: An Edison of the Animal Kingdom?** In the early 1990s the archaeologists Nick Toth and Kathy Schick began researching the tool-making capability of captive bonobos. They showed Kanzi how to use sharp-edged flakes of stone to access food kept in a variety of containers. Kanzi picked up this skill in no time. They then showed him the basics of knapping: they showed him how to make a sharp-edged flake by hitting a block with a hammer stone. It took Kanzi a month to make his first tool; within a year, he'd spontaneously made several improvements and advances to his initial flake-making technique.

However, Kanzi's skill as a stone worker shouldn't be oversold. First, bonobos don't make tools such as this in the wild; Kanzi, on the other hand, had the benefit of intensive coaching. Second, Kanzi's stone flakes were small items and it seems he had no insight in how best to fracture rock to obtain large, useful flakes. Finally, although Kanzi is the world's best bonobo knapper, his efforts are crude when compared with tools made by hominin species some 2.5 million years ago.

The lesson to take from these examples is perhaps this: animals use tools because they can. Tool use is less an indicator of the natural "intelligence" of an animal than a reflection of manipulative abilities (and the evolutionary adaptations its species has made to fit a particular ecological niche). A bird can use its beak for a variety of purposes, an elephant can use its trunk, and a chimpanzee is fortunate in possessing a hand that can manipulate objects in several ways. However, a camel, or a cow, or a cheetah, is never going to be a natural tool user—not because these creatures are inherently inferior to birds or less intelligent than chimpanzees, but simply because they lack the requisite manipulative ability. Presumably if they *could* use tools, they would.

Our species is fortunate: we possesses hands that permit a quite astonishing range of actions. (Count how many different ways you configure your hand to carry out tasks during a typical day. You'll be surprised.) We have to ask, then: what is the chance that an extraterrestrial species will follow the same sort of evolutionary path that humans followed? It's not that extraterrestrials will have to possess hands with four fingers and an opposable thumb; the course of evolution doesn't have to be identical. But any intelligent species will need *some* sort of precision-manipulative ability (whether using claws, tentacles or something beyond our imagination) in order to make tools and thus begin the journey to high technology. Even if this isn't the sole solution to the Fermi paradox, the development of toolmaking is perhaps one more hurdle that has to be overcome before a species can communicate. This is yet one more way in which a world full of life can fail to produce a life-form capable of communicating with us.

## Solution 69 High Technology is Not Inevitable

*Any sufficiently advanced technology is indistinguishable from magic.*

Arthur C. Clarke, *Profiles of the Future*

We considered the possibility, in Solution 41, that ETCs might all be stuck at our own present technological level. Communicating civilizations would thus exist, but they'd have the same capacity to communicate over interstellar distances as we do—which is to say they'd have such a limited capacity that we'd be most unlikely to hear from them. A slightly different take on this idea is that even "advanced" life-forms don't necessarily develop sophisticated technology. Perhaps they get stuck at a technological level that's far *less* sophisticated than ours. Could it be that intelligent life-forms of a sort are out there, but that they "work with equipment which is hardly very far ahead of stone knives and bearskins" (as Mr Spock once complained to Captain Kirk). Could it be that we are the only species that has a technology, crude as it might currently be, that permits interstellar communication?

We touched on the emergence of technology in the previous Solution, but we need to look at the development of technology in more detail if we hope to use this as a solution to the Fermi paradox.

More than 2.5 million years ago, members of *Australopithecus garhi* learned how to take a stone in one hand and chip the end off a stone held in the other hand. By repeatedly chipping or knapping stones, our forebears were able to create sharp edges that were extremely useful in hunting. Fabricating sharp stone edges in this way was an impressive feat—as mentioned earlier, even with intensive training our closest living relatives can't match this simple toolmaking exploit of our long-dead ancestors. Not only did the fabrication of stone edges require insight (to realize that one object can be used to manufacture a second, more useful object) it required a large degree of dexterity and bodily control. Brains played a large role in making this happen. The discovery of stone toolmaking by *A. garhi* might even have kick-started the evolution of intelligence and consciousness in our ancestors. (Could it be that a prerequisite for intelligence is a planet with lots of knappable stone?)

But technological advance halted for about a million years. It was only with the advent of *Homo ergaster* ("working man") that more sophisticated stone tools were developed: *H. ergaster* invented Acheulean hand axe technology. The manufacture of a good hand axe requires preparation and planning; the maker must understand how much force to apply to different parts of the stone, and in what order, in order to create a useful tool. Intriguingly, studies have shown that the parts of the modern human brain that "light up" during this sort of toolmaking activity are those areas required for the fine motor control of lips and tongue—the parts of the brain involved in vocalization. When constructing finer types of tool, among the regions that "light up" is Broca's area—a part of the brain associated with language production and verb recognition. Certainly the neuronal circuitry required to control the precision manipulations of the human hand, and to govern activities such as the throwing of projectiles at moving prey, is phenomenal—and quite beyond the capacity of any present-day robot. Could it be that technology not only kick-started high intelligence in humans but also drove the continued development of intelligence? Perhaps we aren't better toolmakers than other animals because we are more intelligent; perhaps we are more intelligent than other animals because we were better toolmakers. If alien beings remain stuck at the level of "stone knives and bearskins" then perhaps they never develop high intelligence.

One could argue that there's an inevitable sense of progress associated with technology. After all, over the course of a million years or so, Acheulean technology improved slowly but steadily, with refinements in axe manufacturing, the development of cleavers and the invention of spearheads. And then hominid species caused an explosion in the rate of technological progress . . . right? Well, the story isn't so straightforward.

There's now only one hominid species on Earth, but until recently—about 40,000 years ago—we shared the planet with at least two other hominid species. Of course, 40,000 years is a long time when compared with the history of our civilization, but it's a mere instant in the Universal Year; even in the history of our species it represents only about a quarter of the time we've been in existence. Our species grew up co-existing with *Homo neanderthalensis* and with

**Fig. 5.26** Two faces of a hand axe. The axe was found in Spain, and is made of quartzite. It's about 350,000 years old. (Credit: J-M Benito Álvarez)

*Homo denisova*. (Earth has been home to a dozen or more hominid species at various times, and some of these species must have co-existed. The simple and widely publicized picture of hominid evolution—an ape-like creature gradually evolves into "more advanced" species and culminates triumphantly with Man—is wrong. Rather, *H. sapiens* is the last remaining twig on a convoluted branch of the evolutionary tree. Looked at in this light we appear to be much less of a success story.)

The existence of Neanderthals has been well researched and documented but Denisovans[372] were only "discovered" in 2010. In the Altai mountains in Siberia there's a cave called Denisova, after a Russian hermit named Denis who made it his home in the 18th century. In 2008, scientists found a bone fragment from a juvenile girl's little finger and two years later a team led by Svante Pääbo, who is perhaps the world's leading authority in ancient human DNA, extracted mitochondrial DNA from part of the fragment. They were fortunate that the average annual temperature in the Denisova cave is at freezing point; this helped preserve the DNA. An analysis of the girl's mitochondrial DNA made it clear that she was neither a modern human nor a Neanderthal; she was from a related but different species. (At the time of writing, the bone fragments and a couple of teeth are all that remains of this entire species. Denisovans have been called a genome in search of a fossil.) What's remarkable about the Denisova cave is that there's clear evidence of Neanderthal occupation and there's equally clear evidence for human activity in the cave. The cave is unique in that it's

**Fig. 5.27** The entrance to the Denisova cave in the Altai mountains of Siberia. Tens of thousands of years ago, the cave provided shelter for humans, Neanderthals and Denisovans. (Credit: Novosibirsk Institute of Archaeology and Ethnography)

the only place we know for sure that all three human species lived, but it's a sure bet that humans, Neanderthals and Denisovans were at the very least close neighbors in a variety of places on the globe. Indeed, there's evidence of interbreeding: if you are a non-African then between 2–4% of your genome is likely to have a Neanderthal origin, while between 4–6% of the genome of Melanesians and Australian Aborigines comes from Denisovans.

Our knowledge of the Denisovan lifestyle is almost non-existent, but we know more about our closest relatives, the Neanderthals. It's instructive to remember the abilities and achievements of our cousins. Individual Neanderthals must have lived short, hard lives, but as a species they survived for much longer than humankind has been around; they inhabited a large area of Earth; they coped with severe swings in climate; in short, they successfully filled a biological niche. There's some evidence that Neanderthals buried their dead (though whether this practice was associated with the ritual accompanying modern human burials is a matter for debate) and they must have spent long hours preparing clothing. It's particularly interesting that Neanderthals had a form of tool technology, called *Mousterian* (after the French cave of Le Moustier where such tools were first discovered). Mousterian tools are made of stone and take a variety of basic forms. The Mousterian craftsmen, then, were presumably able to hold several patterns of tool design in their minds and, combined with their deep appreciation of the properties of stone, produce quite beautifully constructed implements. The Neanderthals might not have matched the achievements of humans, but they were no mugs.[373]

However, during their period on Earth, Neanderthals demonstrated little in the way of innovation. Their technology, though effective, wasn't subject to the sort of progress we've come to believe is inevitable. The late Mousterian tools weren't significantly better than those of the early Mousterian. (This is possibly doing the Neanderthal's a disservice.[374] In 2013, a team of archaeologists led by

Marie Soressi published evidence of the discovery of bone tools called lissoirs in a rock shelter called Pech-de-l'Azé in southwestern France. Present-day workers in the leather industry still use lissoirs, but the Pech-de-l'Azé tools date back 41,000–51,000 years. One interpretation of this discovery is that Neanderthals independently invented a tool that we still use today. However, the matter is far from settled. The uncertainty in the dating is so large that it's entirely possible the Neanderthals in that area met the first wave of modern humans to enter Europe. It would be quite a coincidence for Neanderthals to have gone so long without using these tools and then suddenly invent them just as modern humans were arriving with their advanced toolkit.) Although some archaeologists would disagree, I think it's possible to regard the Neanderthals as an intelligent toolmaking species that survived for more than 100,000 years without making significant technological advance. Our knowledge of the Denisovans is fragmentary, but there's no evidence to suggest that their technology was any more developed. Our cousins edged into extinction—for reasons not entirely clear—without inventing the ratchet much less the radio telescope. Perhaps this situation is mirrored on other worlds.

The suggestion, then, is that for some reason (lack of language, lack of a "creative spark", lack of hand–eye coordination, lack of whatever) alien species reach a certain level of toolmaking and then remain at that level. Perhaps the Galaxy abounds with species that are experts at handling wood or stone or bone, but that never develop further. We don't hear from ETCs because none of them have the required technology: in other words, *communicating* ETCs don't exist.

One weakness of this suggestion is that it requires *all* toolmaking species to develop in the same way. It fails to convince, just as "sociological" explanations fail to convince when they assume all ETCs will behave in the same way. After all, even if hominid species in general have been poor technological innovators, one member of the hominid family is exceptionally innovative. That's a ratio of about one in ten—not too shabby. If there are many extraterrestrial toolmaking species out there, and just 10% of them discover the benefits of continual innovation, then the odds of finding ETCs wouldn't seem so bad.

Before rejecting the suggestion completely, however, it's worth remarking that for much of our history we weren't much better than the Neanderthals when it came to technological innovation. It's only 40,000 or so years ago that our technology and art began to dazzle.[375] The cave art of the Cro-Magnons is truly beautiful, recognizably human and capable of speaking to us across millennia. It's unlike anything appearing before that date. Until this explosion of creativity the three extant hominid species appear to have been equally stagnant. Why the sudden change for us humans? There are several possible explanations. Perhaps the development of language triggered the creative explosion. Perhaps the explosion occurred much earlier, but artifacts prior to

40,000 years ago weren't well preserved. Perhaps the humans of more than 40,000 years ago were anatomically modern, but lacked a modern brain. Perhaps cultural knowledge accumulated slowly until, 40,000 years ago, it passed a critical threshold. Perhaps the exceptionally long stage of human child development, and the play that children engage in, enabled humans to envision their environment in new and creative ways. Perhaps a constellation of factors was involved. We don't know. But if whatever caused this explosion of creativity was a fluke, an accident, then we might expect the number of communicating ETCs to be small.

## Solution 70 Intelligence at the Human Level is Rare

> *Many difficulties which nature throws in our way may be smoothed away by the exercise of intelligence.*
>
> Livy, *Histories,* Book XXV, Sec. 11

When Fermi asked "where *is* everybody?", the "everybody" referred to *intelligent* extraterrestrial creatures. While the discovery of any life elsewhere would be profoundly important, it's intelligent life we search for with passion. It is (presumably) only *intelligent* life that can travel between stars and with whom we can communicate, interact and learn from. But perhaps intelligence—the sort that can investigate and understand the laws of physics—is rare in the universe? As many as 50 billion species have lived on Earth, but only one has evolved the sort of intelligence that can determine the existence of the Higgs boson. Could it be that the development of intelligence is a fluke, so that the $f_i$ term in the Drake equation is small?

There are many aspects to this question, and insufficient space here to address them all. Let's discuss just a couple. First, how should we define intelligence? Second, how likely is intelligence—human level intelligence—to evolve?

### What is Intelligence, Anyway?

In terms of SETI activities a good operational definition of intelligence is: it's the ability to operate a radio telescope. Unfortunately, under this definition humans only became intelligent within the last century! Are there other definitions that can better capture the essence of intelligence?

A common approach is to define intelligence in terms of the ability to perform certain mental tasks that we find difficult, such as playing a decent game of chess. However, it's not much more difficult to write a chess-playing program than it is to play the game itself, and few would contend that cheap chess-playing software possesses intelligence. It turns out that the sorts of activity humans and other animals do without thinking are much more difficult to program. No one has yet programmed a robot capable of independently navigating the outside world while coping with the continuous challenges everyday life throws up. If finding sustenance and avoiding danger are any measure of intelligence, then the average rodent is *much* more intelligent than the smartest robot. So if we want to appreciate what intelligence really means, and whether humans are unique in this regard, it might help if we understood something of animal intelligence. Unfortunately, if it's difficult to define intelligence in humans, it's even more difficult to define intelligence in other creatures.[376]

Most people, if asked to rank non-marine animals in terms of intelligence, would probably rate man as the most intelligent animal, followed perhaps by apes, down through dogs and cats, down further to the likes of mice and rats, down even further to birds, and so on. It's a comfortable picture for the human ego: we are at the top of the tree of intelligence, our closest relatives are clever, our pets are quite bright, and the animals we don't particularly like are stupid. Implicit in this picture, though, is the notion of evolution as progress from a "less evolved" state (rats, say) to a "highly evolved" state (us), with intelligence being the scale against which one can measure progress. This is simply wrong.

In the first place, we have no reason for supposing intelligence (however we choose to define it) is the sole criterion by which we can rank animals. Why not instead use visual acuity, or speed, or strength? Indeed, why try to rank animals in this way at all? We shouldn't view evolution as a ladder, with ourselves at the top and all other animals below us because they aren't yet "evolved enough" to possess intelligence. Apes, bears, cats, dogs, mice and men are all equally "evolved" because we share a common ancestor that lived about 65 million years ago.[377] The various species have adapted to their environments in different ways; our species has certain characteristics that make it successful, but so has every other species on the planet. These species are all equally successful, since they've passed the critical test: they've all survived. If we want to assign different levels of intelligence to different animals, then we need a better gauge than our prejudices.

Biologists face an almost impossible task when trying to measure the intelligence of animals. Measuring the IQ of humans in a non-culturally biased way is difficult enough. But if tests on humans are biased, how can we possibly test the intelligence of different animal species? How can we factor out the differences in perceptual ability, manipulative ability, temperament, social behavior, motivation and all the other variations between species? Does a monkey fail to complete a maze because it is brainless or because it is bored? If a cat fails to press a lever that produces a food reward, should we conclude that the cat is stupid or is simply not hungry? Does a rat fail an intelligence test because it is dense or because the test demanded visual discrimination (at which rats are poor) rather than discrimination between smells (at which rats excel)? These sorts of questions make it exceptionally difficult to be sure that we're testing an animal's cognitive ability.

Suppose we try to account for as many cross-species variables as we can think of in these cognitive tests. (For example, biologists might want to investigate how many list items an animal can remember, or whether an animal can recognize a face; either of these tasks might tell us something about cognitive processes in animals. The investigator would have to ensure the details of the test were different for different animals. The tests for pigeons and for chimpanzees would *have* to be different, if only to take into account their different physical abilities.) Suppose further that we define intelligence, *general* intelligence, to be a measure of how well animals score on such fundamental cognitive tests. Then a surprising fact emerges: most animals perform at about the same level! Of course there are some differences between species, but the differences are much smaller than one might expect. Chimpanzees can remember about seven items from a list at one time—but so can pigeons (so no more cracks about "bird brains"). Monkeys can quickly discern whether pile A contains more food treats than pile B—but so can cats. In fact, if intelligence is defined as the ability to perform these basic non-verbal tasks, then one can argue that to a first approximation all birds and mammals, including humans, are about equally intelligent! This conclusion remains controversial, but if it turns out to be true we shouldn't be surprised. After all, every species, including mankind, has to negotiate the same perilous world; we all have to eat and drink and find mates. The basic cognitive skills enabling animals to perform these tasks might well be common to all species.

On the other hand, one can equally take the opposite approach: perhaps intelligence in animals consists precisely in all those factors we deliberately omit in cognitive tests. To use a computing analogy, we shouldn't just consider the processor (the brain) but also the attached input and output devices (the senses and manipulative abilities of an animal). After all, a chimpanzee has hands that enable it to perform tasks that a cow simply can't attempt. From this viewpoint there might be little general intelligence residing in the brain; rather, intelligence should be defined in terms of *specialized* intelligence—adaptations that enable particular species to succeed in their particular ecological niches. Support for this view is that the ability to learn, which is surely a large part of intelligence, seems to be specialized. Many animals can learn a particular task with ease but find it impossible to learn a logically equivalent task. It appears that an animal's learning ability depends upon the hard-wired behaviors already present in its brain. In this view, all animals are *differently* intelligent. It simply makes no sense to ask whether a bonobo is brighter than a homing pigeon: both creatures possess specialized intelligence enabling them to succeed in their particular environments.

These two seemingly opposite views of intelligence—that either *general* intelligence or *specialized* intelligence is the important factor—are perhaps just two faces of the same coin. The lesson is that, cognitively, animals are both similar and yet different. In the case of mankind, much as we might like to think otherwise, our similarities with other animals are clear: we simply aren't much better than many other animals at tasks that investigate fundamental, non-verbal cognition.

Nevertheless, it's impossible to deny the quite profound difference that exists between mankind and every other species. We might not be atop some evolutionary ladder of intelligence, but we *are* the only species capable of contemplating the intricacies of integral calculus. Only a member of our species can reflect upon his or her own thoughts and the thoughts of others of the species. Only humans are in the slightest bit interested in defining the concept of intelligence, trying to measure it or wondering precisely what it means. If that unique intelligence arose from a rare combination of many factors then we might be alone in the Galaxy. Could it be that the Galaxy is full of intelligent alien species—but species that possess intelligence in the same sense that animals possess intelligence? Might they all be *differently* intelligent?

## How Likely is it That Human-Like Levels of Intelligence will Evolve?

Most of the SETI enthusiasts I meet have a physical science background. Almost without exception they argue that over geological timescales evolution has caused a monotonic increase in levels of intelligence, and they further argue that this is only to be expected since, as Carl Sagan once remarked, "all things being equal it's better to be smart than to be stupid". What evidence is there for this generally held supposition that stupid creatures eventually evolve smartness? It's impossible to measure the IQ of long dead creatures, of course, so physical scientists tend to bolster their argument by picking a proxy for intelligence, such as the ratio of cranial capacity to body size, and then drawing a graph that shows this ratio increasing over time. Sure enough, if you draw a graph of relative brain size for the vertebrates, for example, you see that the lower vertebrates had a small brain relative to its body; the archaic mammal branched off from this line and evolved a larger brain; insectivores subsequently split from this line and they evolved an even larger brain; prosimians split from the insectivores and evolved still larger brains; and so on through time until we reach modern humans. It's the story of an unbroken increase in braininess, culminating in us. My SETI friends, who have no wish to appear vainglorious, point out that this increase in relative brain size—and presumably therefore in intelligence—would have been repeated over much longer timescales on planets older than ours. So although humans are at the pinnacle of intelligence here on Earth, compared to some extraterrestrial species we will be intellectual pygmies. It seems to be an overwhelmingly strong argument. Except . . .

As Charles Lineweaver pointed out,[378] when we choose to draw a graph of the evolution of brain size we are choosing to plot a property that defines us humans: there's a selection bias at work here. Any species possesses some unique feature; with humans it just happens to involve brains. But if you choose the extreme feature that characterizes a species, and then plot how that feature developed over time, you'll inevitably see the sort of graph described in the paragraph above. For example, if present-day elephants could think about such things they'd probably choose the ratio of nose length to body length as being an animal's most important feature; sure enough, if you plot such a ratio for various species over time you'll see an ever-increasing trend culminating in elephants. Such a trend says nothing about life in general.

In order to be convincing, the sort of graph mentioned above needs to consider what happened to a line after we diverged from it. If we could show that insectivores, for example, had increased in intelligence after they diverged from the line that led to us, then that might be evidence to support Sagan's view that smart is better than stupid. But there is no evidence that insectivores, after they split from the line leading to us, have increased in intelligence.

Lineweaver points out another source of bias in the plot of brain size over time: the lines on the graph represent quite different things. In some cases a line represents just one species (in particular, the topmost line represents humans) whereas in other cases a line represents many thousands of species. This isn't a reliable way of analyzing the data. If we are looking for evolutionary trends then we need to look at *all* the data. And if we did that we'd see that sometimes one species will get smarter and sometimes a different species will get stupider. There's no directionality. It all depends on how a species responds to the pressures acting on it as time goes by. For many species it has paid off to become more sophisticated; for some species, such as parasitic flukes and tapeworms, it has paid off to become less sophisticated.

Is there then no evidence for the evolutionary convergence of intelligence? After all, birds, dinosaurs, fish, insects, mammals and reptiles all independently evolved the ability to fly—in other words, after their evolutionary lines diverged, different species developed mechanisms for getting airborne (as a means of escaping from predators or for hunting prey). Are there really no examples of species evolving increasing levels of intelligence after they have diverged from our own lineage? Well, here is one example: the last common ancestor of birds and humans lived about 310 million years ago. It was a creature with a small brain-to-body ratio. Independently, that ratio increased in birds and humans. A second example: the last common ancestor of dolphins and humans is more recent, but again we find that brain capacity increased independently in both lines.

The example of birds is particularly interesting. Our last common ancestor would have been about the size of a Border Collie, but considerably stupider. However, inside its small brain was an area called the pallium; this evolved into the prefrontal cortex in mammals and into the nidopallium caudolaterale in birds. Human brains and bird brains are thus quite different, but the brain regions that evolved from that initial pallium are in both cases involved in handling working memory, learning, prediction. In 2013, scientists showed how crows[379] can handle abstract reasoning and the researchers watched neurons firing in the nidopallium caudolaterale as crows performed this reasoning task. This example of crow intelligence is probably the closest scientists have got to examining alien intelligence: birds can reach the same answer to an abstract puzzle as humans do, but they use a completely different brain structure to do it. Doesn't this prove, then, that intelligence is a convergent feature of evolution? Even if it does, I'd make a couple of comments.

First, we believe that all organisms ultimately have a common ancestor. Before two organisms diverged, therefore, they'll have shared billions of years of evolutionary history—giving them a shared biochemical and genetic toolkit with which subsequent evolution can play. For example, eyes have evolved independently many times but the genetic signaling system for eyes can work across species: a particular example of this is the way a mouse gene can control the development of a fly's eye.[380] When an organism responds to environmental pressures, its past evolutionary history limits the choices available to it in the present; should we not regard the independent development of eyes or intelligence as parallel evolution constrained by a long shared history? Evolution plays variations on existing themes.

Second, even if we choose to consider crows and dolphins as being intelligent, neither of these creatures seem inclined to build a radio telescope. Let's look at this another way. Suppose tomorrow a meteorite strikes Earth and causes a mass extinction, wiping out all trace of humanity. Would some other species with human-like intelligence arise over the course of a few tens of millions of years? Lineweaver calls this notion, this idea that stupid things become smarter, the Planet of the Apes hypothesis. It's named, of course, after *Planet of the Apes*—a novel by Pierre Boulle published in 1963 and made famous by the 1968 Hollywood movie of the same name. In the movie, astronauts somehow travel into the future and their spacecraft then crashes onto a strange planet. The survivors encounter a society in which apes have developed speech (English, fortunately) and human-like intelligence. The apes are the dominant species. At the end of the movie, the astronauts discover that the planet is (spoiler alert!) a post-apocalyptic Earth. The famous biologist Ernst Mayr argued that the history of life on Earth refutes this idea. After almost 4 billion years of evolution, intelligence hasn't appeared in archaea or bacteria; within the eukaryotes, intelligence hasn't appeared in fungi or plants; within the animals intelligence hasn't appeared in ... well, you get the idea. Humans represent just one tiny twig on an immense bush of life. Looked at in this way, the development of a species with the intelligence to communicate over interstellar distances seems anything but inevitable.

Billions of years of shared biological evolution have produced just one species—*Homo sapiens*—that can build such a radio telescope. Why should we expect creatures with whom we share *no* evolutionary history to possess the intelligence (and other factors such as symbolic language, tool use and so on) needed to communicate over interstellar distances? Human-like intelligence is as the name suggests: a species-specific characteristic.

## Solution 71 Language is Unique to Humans

> ... *I learn'd the language of another world.*
>
> Lord Byron, *Manfred*, Act III, Scene 4

Ludwig Wittgenstein once famously remarked that "if a lion could talk, we would not understand him". It's easy to see the philosopher's reasoning: lions must perceive the world in ways quite alien to us. They possess drives and senses and abilities that we don't share. On the other hand, the statement is all wrong. If a lion spoke English then presumably English speakers *could* understand it—but the mind of that lion would no longer be a lion's mind. *The lion would no longer be a lion*. Humans talk; lions don't.[381]

Many will argue that humans are unique in being the only species in the history of Earth to have employed language. If language developed in only one of the 50 billion or so species that have ever existed then perhaps the likelihood of language developing is small. Perhaps it developed in humans just through dumb luck—a chance assembly of several unlikely physical and cognitive adaptations. We are unique on Earth and we might be unique in the Galaxy: perhaps humans are the only creatures that can talk. And since language opens up so many possibilities—so much of what we do individually and socially wouldn't otherwise occur—creatures without language would surely be unable to build radio telescopes. No matter how intelligent those creatures might be, if they lacked language,[382] then we wouldn't hear from them.

Could this explain the Fermi paradox—only on Earth has a species learned to talk?

In attempting to answer this question we must first consider whether ours is indeed the only species that possesses language. After all, some birds communicate through complex songs; bees communicate through dance; dolphins communicate through whistles and chirps and clicks. Perhaps all animals have innate language abilities to a greater or lesser extent? One of the difficulties in contemplating this question is our own use of language: we seem compelled to anthropomorphize. Even when describing inanimate objects we anthropomorphize: genes are "selfish", the car is "acting funny", my chess program is "figuring out" the best move to make. There is of course nothing wrong with employing metaphor—assigning intentionality to inanimate objects enables us to convey the appropriate thought quickly—but sometimes we can forget that anthropomorphic statements don't necessarily describe what's really happening. We must be careful when describing an animal's actions in terms of our own conscious thoughts and motives. When we describe an animal as communicating some word or idea—effectively, when we say an animal is "talking"—we could be wrong.

Here is just one example where an initial interpretation of events could be wrong. Some types of ground squirrel living in open country suffer two main predators: hawks, relying on speed, attack from the air, while badgers, relying on stealth, attack from the ground. When a squirrel spots a predator it chooses (there's an anthropomorphic usage!) from one of two defensive strategies. If it spots a badger, the squirrel retreats to the opening of its burrow and maintains an erect posture. A badger, seeing that posture, knows the squirrel has spotted it and thus an attack would be a waste of time and energy. If a squirrel spots a hawk, it runs like hell for the nearest cover. Squirrels also emit two different alarm sounds. If they spot a badger, they make a rough chattering sound; if they spot a hawk, they emit a high-pitched whistle. Other squirrels in the vicinity react when they hear the sounds, retreating to their burrows when they hear the badger alarm or running for cover when they hear the hawk alarm. One's initial and quite natural inclination is to think that squirrels are talking to each other, that they are saying in effect: "Careful, now, there's a badger around; better head for home" or "Oh-oh, hawk; get *out* of here!" But are they?

As its actions upon spotting a predator clearly show, any individual squirrel is interested in saving its own skin. Indeed, evolutionary theory tells us that this must be the case: a squirrel couldn't care less about the fate of its acquaintances.

But if the squirrel alarm calls carry semantic information—if they are calling out "brock!" or "hawk!" in squirrelese—we meet a paradox. Selection will favor those squirrels who keep quiet, sneak off silently, and let the other suckers get eaten; being a non-caller in a group of callers is selectively advantageous, and the squirrel gets to pass on its genes. Soon, though, you end up with a community of silent squirrels. Where does the instinct to cry out arise?

The behavior of the squirrels makes sense only if their calls do *not* convey semantic information. Consider the squirrel's "hawk alarm". First, it's a high-pitched whistle—which, as experiments have shown, hawks find difficult to locate. So the squirrel is revealing little about its position to the hawk. Second, being the only one to run for cover makes a squirrel conspicuous; it's much better to be one of a group of squirrels that are scrambling around because the chances of being singled out by the hawk are reduced. Similarly, squirrels that run for cover when they hear a high-pitched whistle are less likely to be eaten by a hawk than squirrels that stand their ground. So selection will tend to favor squirrels that cry out when they see a hawk and also those that run for cover when they hear a high-pitched whistle. When us humans look at the situation we interpret it as squirrels sharing information. But that's not what's happening. The behavior is simply a trait passed on through the generations because it's effective. The squirrels don't even have to be aware of one another for this sort of behavior to evolve. No words; no language; just the forces of evolution.

Animals certainly *communicate*. But then so do bacteria. Even cells communicate. Communication isn't the same as language, however. Despite years of research there's no evidence that any creature—not even the bonobo, the chimpanzee or the dolphin—possess a communication system that can do all the things that human language can do. Humans can refer to abstract concepts; to objects in their environment; to events that happened in the past and that will or could occur in the future. Humans can combine small meaningful elements into larger meaningful elements, and they can memorize hundreds of thousands of concepts and map them onto specific combinations of vocal patterns. Humans can produce an infinite combination of meanings from a systematic grammar. Only humans possess this sort of symbolic logic.

None of this is to say that because only humans possess language we are somehow "better". Birds can perform feats of navigation that no human can match without aids. Some marine animals can, unlike humans, sense electric currents. Dogs can hear sounds beyond our perception and smell scents to which our noses are dead. Bats use an incredible system of echolocation. Horses have been known to pick up on cues that humans miss completely. And so on. Every species has abilities, forged by evolution, which enable them to scrape a living in a world that cares not whether they survive. This diversity is wonderful, and should be celebrated. It's arrogant to measure the abilities of animals in terms of *our* capabilities. Defining other species in terms of how well or how badly they use *human* traits is to demean those species. Nevertheless, only humans possess language. And language has opened up the world to us.

Language is truly remarkable. A well-educated person, the sort of person who will be reading this book, knows about 75,000 words. That means you, Gentle Reader, must have in your youth learned on average one word per hour for 13 years (assuming 8 hours per day for sleep). When speaking words out loud you are performing mechanical motions of the utmost complexity: you must coordinate and regulate the movements of various organs to within a fraction of a millimeter and get timings right to within tenths of a second. When you listen to the speech of some other person your brain is comprehending information at an impressive rate. The amazing thing about all this is that dealing with the complexities of speech production, of speech comprehension, of a grammar that permits the utterance of an infinite number of different sentences is all so . . . well, effortless. If I ask you to multiply 267 and 384 in your head, your face will contort with concentration. But speech just happens. How did we develop this almost miraculous facility?

The question of the origin of language is surely one of the most difficult in science, not least because relevant observational evidence is so limited. Voices don't leave behind fossils. Fossilized skulls don't tell us what the brains they once housed were capable of. Pretty much the only hard evidence is in the anatomy of the vocal tract; modern speech sounds couldn't have been made by creatures living longer than 100,000 years ago. (An analysis of a fossilized Neanderthal hyoid bone,[383] a structure that supports the roof of the tongue and whose position is crucial for complex vocalization, suggests that Neanderthals might have had the physical capacity for speech—though whether they could communicate in the way we do remains uncertain.) Language could well have began earlier, but it would perforce have employed a restricted range of sounds.

Faced with these difficulties, scientists have proposed a variety of different theories to explain the origin of language and how it is that humans came to possess this amazing asset. Perhaps the most influential approach to the question is due to the philosopher and linguist Noam Chomsky,[384] who argues that language is innate. A child doesn't need to *learn* language; rather, language *grows* in the child's mind. In other words, a child is genetically programmed with a blueprint—a set of process rules and simple procedures that make the acquisition of language inevitable. All of us possess a "language organ"—not something a surgeon can cut out with a knife, but a set of connections in the brain dedicated to language in the same way that parts of the brain are dedicated to vision. In this view, language acquisition happens to a child in much the same way that body hair suddenly sprouts on a pubescent teenager; it's part of growing up. Language is part of our genetic heritage.

Although Chomsky's ideas have been attacked by adherents to the standard social science model (who argue that human practices within a social group are moulded by the culture of the group), by linguists (who have offered a variety of competing models for the evolution of human language) and by computer scientists (who take quite a different approach to the study of language), his theory has framed the terms of the debate for several decades and it does address several puzzles regarding language acquisition.

One puzzle is that, as already mentioned, language is an infinite system: one can construct an infinite number of sentences from a finite number of words. If I were to speak this present sentence out loud, then there's an excellent chance that I'd be the first person ever in the entire history of the universe to utter this particular concoction of words in this particular order; it's a unique combination. In order to cope with this infinite set, the brain *must* be following rules rather than accessing a store of responses. And when one considers what a child hears when its parents and siblings talk to it—just a sequence of sounds, including meaningless "um's", "huh's" and "coochy-coo's" interspersing the poorly formed and incomplete sentences we all inevitably utter—it's remarkable that children develop and employ complex grammars so rapidly, all without the benefit of training, and often without feedback from adults on the errors they make. (Grammar in this sense, it should be noted, refers to the structure of a language rather than the trivial rules that pedants try to enforce. Grammar is about the basic patterns of language, rather than disputes over whether it should be "to boldly go" or "to go boldly".) However, if children are innately equipped with a *language acquisition device* (LAD) that lets them pluck the relevant syntactic patterns from the gobbledygook assaulting their ears then the puzzle disappears. Rather than there being one device for Albanian, another for Basque and yet another for Czech, there's just one LAD common to all humanity. Any child—so long as he or she receives sufficient stimuli to trigger the LAD at the correct age—can learn to speak any language. The stimulus need not even be auditory. If they are exposed to signing at the right age, hearing children of deaf parents can acquire sign language.

If it exists, the operation of the human LAD might be similar to the innate visual acquisition device (VAD) of many animals. Scientists have performed experiments on kittens, blindfolding them immediately after birth. If the blindfold is removed any time before the first 8 weeks, the normal development of the kitten's visual system is resumed and the adult cat will see normally. If the blindfold is kept on for longer than 8 weeks, the cat will suffer permanent visual impairment. It seems, therefore, that there's a critical period in which the VAD must receive external visual stimuli in order to establish the appropriate neuronal connections in *specific* pre-wired locations in the kitten's brain. If the connections aren't established within this period, the chance of developing a fully functioning visual system is lost. Other parts of the brain are unable to act as stand-ins for the visual system. The same effect has been observed to occur in those tragic cases in which linguistic input is withheld from children during the critical period up to puberty: their ability to speak grammatically is severely impaired. The existence of a critical period for language acquisition isn't necessarily mysterious: it's presumably simply part of the same genetically controlled maturation process that causes our sucking reflex to disappear, our baby teeth to erupt, and all the other changes that occur to the human body. It would make evolutionary sense for the LAD to switch on early, as that way we have the maximum time to enjoy the considerable benefits of language. It would also make sense for the LAD to switch off when its job is done, since maintaining such a device would presumably incur considerable costs in terms of energy requirements.

Although different tongues differ in the specifics, there's a universality to language. It's these universal *principles* that Chomsky and his followers argue are innate. When a child develops language, then, the procedure follows an internal, predetermined course. A child who's acquiring Dutch will set the parameters of this predetermined system in one way; the child acquiring English will set the parameters in another way; and the child acquiring French will set the parameters of the system in yet another way. But the underlying principles are the same. To use a software analogy, language acquisition is rather like a macro with arguments—one argument for each language. (Vocabulary, of course, must be learned: if individual words were innate, then a neologism like "pulsar" would have to be assimilated into the gene pool before astronomers could use it! Cultural evolution would move at the same glacial pace as genetic evolution. Certain grammatical constructions must also be learned. For example, although there's a rule for forming the regular past tense of an English verb—namely, add *-ed*—the past tense of irregular verbs must be learned on a case-by-case basis.)

Clinical evidence is at least consistent with the notion that language is innate. In some unfortunate patients, trauma or disease harms particular locations in the brain—locations that appear to be responsible for language processing. The effects can be distressing.

For example, patients in which Wernicke's area is damaged find it difficult to comprehend the speech taking place around them. More bizarrely, they suffer from Wernicke's aphasia: their own speech is rapid, fluent, filled with grammatically correct phrases—yet their utterances contain little or no meaning. They often substitute one word for another, and they coin new words; when asked to name objects, they give semantically related words or words that distort the sound of the correct word. Transcripts of their speech can make for disturbing reading—like reading the ramblings of a psychotic. On the other hand, patients with damage to Broca's area suffer from Broca's aphasia—speech that is slow, halting and ungrammatical. They can often comprehend the speech going on around them, or at least make informed guesses as to the meaning of speech, thanks to their prior knowledge of the world and the built-in redundancy of speech. (These sufferers can understand a sentence such as "the cat chased the mouse" because they know cats chase mice.) Patients in which the *connection* between Wernicke's and Broca's areas is damaged suffer a form of aphasia that renders them incapable of repeating sentences. Even worse is the aphasia affecting patients in which Wernicke's and Broca's areas, and the connection between them, are undamaged but isolated from the rest of the cortex. The patients can repeat what they hear but have no understanding of what they are saying; they never initiate conversation. In yet other cases, damage to specific parts of the brain—often through stroke—causes remarkably specific language problems. Some aphasics can recognize colors but not name them; others can't name food items, though they know what they like to eat; others have no trouble dressing themselves yet they can't name items of clothing. At present, neuroscientists remain unable to map the brain and highlight different areas as handling different aspects of language. However, the evidence is that language is localized. And although localization itself does not mean language is innate, for some researchers it is suggestive of their being a language organ.

If we do possess an innate language faculty, the obvious question is: how did we come by such an intricate and complex organ? The answer is equally obvious: it evolved by natural selection of heritable variations.[385]

Unless we invoke the involvement of a creator, natural selection is the only known process that can generate such wonderful structures. Critics have argued that if our language organ is the result of evolution then we should see traces of it in the apes. After all, we are descendants of apes, aren't we? Well, no, we aren't. Humans and apes are linked by a common ancestor that perhaps lived as long ago as 7 million years. It's entirely possible that a LAD evolved some time within the last 7 million years, so that it's not shared with the evolutionary branch leading to modern apes. Indeed, some scientists have suggested that the minds of early modern humans of about 100,000 years ago contained several separate "modules": a module for language, a module for technical intelligence, a module for social intelligence, a module for natural history, and so on. It might be that these isolated modules began to communicate only 50,000 years ago; and only then could people get together in groups and discuss, for example, the merits of a new tool design for use in hunting. Only then did we become fully human.

Articulate speech is vitally important to the success of our species. It's not unreasonable to suppose that it's impossible for any species to develop the ability to travel or communicate over interstellar distances if it lacks some equally sophisticated method of communication. And yet, in the case of the evolution of human speech, we seem forced to conclude that articulate speech is the result of a series of chance environmental changes and evolutionary responses; it was just good luck. Consider, for example, what happened to the bodies of our ancestors: they underwent a restructuring of the diaphragm, larynx, lips, nasal passages, oral cavity and tongue, all of which were vital for articulate speech to develop, but none of which occurred *in order* for speech to develop. The changes to these organs were initially completely unrelated to the capacity for speech; they were small changes that brought immediate selective benefits. At least one of the changes—the positioning of the larynx deep in the throat—seems bizarre. Having a larynx low in the throat provides the tongue with enough room to move and produce a large number of vowel sounds, but any food and drink we swallow has to pass over the trachea: choking to death becomes a distinct possibility. The benefits are great, but so are the costs. If the tape of life were replayed, perhaps humans wouldn't develop language.

On Earth, of the 50 billion species that have existed, only humans possess language. Language enables us not only to think but to think about the thoughts we have, to try out new patterns of thought, and to record our thoughts. Language is what makes us human. If we ever visit other worlds, perhaps we'll find billions of other species—each well adapted to its particular niche, but none of them with the single adaptive trait we are searching for: language.

## Solution 72 Science is Not Inevitable

*For science is like virtue, its own exceedingly great reward.*

Charles Kingsley, *Health and Education*

An extraterrestrial civilization will presumably need to possess a high level of scientific ability and prowess if it is to communicate with us, for only through science can it gain the knowledge that would enable it to build a radio telescope (or some other device that permits interstellar communication). But even if an intelligent extraterrestrial species *does* learn to make tools, *does* develop technology, and *does* acquire language, will it then inevitably develop the methods of natural science? Perhaps the Galaxy is swarming with species more intelligent than us—creatures excelling in the arts and philosophy—but who lack the techniques of science. We don't hear from these species because they can't make themselves heard over interstellar distances.

This suggestion is implicit in thousands of SF stories, and those who offer it as a solution to the paradox presumably take their cue from the historical development of natural science on Earth. Many civilizations have developed mathematics and medicine, but the origins of natural science have historically been much more restricted. Consider, for example, the Aborigines. They arrived in Australia as far back as 50,000 years ago[386]—a landmark achievement in human exploration that is too often underestimated. The culture of Australia's indigenous peoples is perhaps the oldest continuously maintained culture in the world; their stories and belief systems are the most ancient on Earth. They have survived in a wide range of environments with great success for an unimaginable length of time. Yet in all that time they never created the techniques of modern science. The dawn of modern science began only about 2500 years ago with the Greeks. Despite possessing some of the most brilliant scientists of all time, however, Hellenistic science was limited and shackled by a pervading intellectual snobbery that valued contemplation over experiment. It took almost 2000 years for science as we now understand it to really get underway, with scientists such as Galileo and in particular Newton pioneering a quantitative approach to scientific reasoning. Why did it take so long for the seeds planted by the Greeks to flower into our modern scientific endeavors? And although science is now a global activity, why did the flowering take place in such a restricted geographical area?

After the demise of the ancient Greek civilization, many other civilizations developed sophisticated technologies and systems of mathematics. The Arabic civilizations in North Africa and the Middle East possessed some excellent mathematicians (much of our knowledge of Greek astronomy was preserved by them). The civilizations of South America possessed architects who were capable of building fantastic structures. The Chinese civilization was for many hundreds of years the most advanced on Earth. Yet none of them—nor any of the others civilizations around the world—developed the methods of modern science, and none of them developed the scientific approach to the study of Nature that has proven to be so powerful. Why?

It might be that cultural factors played a role. For example, some authors believe the prevailing philosophy of the Chinese civilization encouraged a "holistic" view of the world, so it was more difficult for them to take a Western-style "analytic" approach to science. Newton was ready to consider a system in isolation from the rest of the universe and apply his techniques to that idealized, simplified system; had he attempted to provide a complete description of Nature in all its messy holistic complexity, he'd surely not have succeeded. And in 1709, while the world was still absorbing the impact of Newton's great scientific books, the industrial revolution began and with it came an increase in the rate at which science could be transformed into technology. The spark that ignited the industrial revolution—Abraham Darby's use of coke rather than charcoal for smelting iron—took place in Ironbridge, England. At the same time in China, a centuries-old iron works was in the process of being closed. The Chinese thought they had no further need of it.

It can be argued, then, that the development of science is far from being inevitable. There is a variety of reasons—cultural preferences, environmental obstacles, philosophical inclination, sheer luck—why ETCs might not hit on the techniques of science.

Yet it's hard to accept this as a plausible explanation of the Fermi paradox. Yes, the gap between the emergence of Hellenistic science and the rise of modern science[387] was almost 2000 years—undoubtedly a long time on the human scale. But this isn't the correct timescale with which to consider these questions. In the Universal Year, 2000 years corresponds to less than 5 seconds. On a cosmic timescale it matters not at all that natural science was developed by a Western European civilization rather than the Incas, Ottomans or Chinese. Whether it takes 2000 years or 20,000 years to invent science makes little difference as far as the paradox is concerned. For humankind, the scientific method had to be invented only once: its effectiveness meant it spread quickly, and it's now the common heritage of our species. Won't the same be true for ETCs?

## Solution 73 Consciousness is Not Inevitable

*What's the difference between being dead, and just not knowing you're alive?*
Peter Watts, *Blindsight*

If you've made it this far—well, I admire your perseverance. I'm sure some readers will at times have wanted to throw the book away in frustration when they encountered lapses in logic (I'm sure those lapses exist; but *you* try distilling such a large amount of technical information for a general audience ... it ain't easy) or a less than elegant turn of phrase (well, tastes differ). If I'm lucky, someone will point out the errors and go on to develop better proposals for addressing the Fermi paradox. If I'm *really* lucky I'll have provoked someone into developing a completely new solution to the paradox. Whatever your reaction to the book happens to be—boredom and frustration at one end of the spectrum, excitement and enjoyment at the other—the mere fact that you spent time weighing up and criticizing the various speculative ideas I've presented, and that you're capable of an emotional reaction to those ideas, is really quite astonishing. We all possess an internal "theatre" in which we not only register feelings and emotions but also contemplate abstruse matters such as the possible existence of extraterrestrial civilizations. Why do we possess this wonderful phenomenon known as consciousness (or sentience or awareness or whatever you prefer to call it; it's difficult to define consciousness, even though we all know from our own subjective experience what it means)?

Consciousness is surely what makes life worth living, and the phenomenon is certainly advantageous to us in our complex modern world because it enables us to perform so many different tasks. But evolution has no foresight. How could consciousness have been an advantage to humans eking out an existence in Africa 50,000 years ago? Indeed, wouldn't consciousness have been a definite *disadvantage* to them? If one of our ancestors spotted a lion then the correct reaction would have been to scarper, not stop and ponder the grace with which big cats stalk their prey. Even today, sports people talk about the importance of being "in the zone", a state where everything just *flows* rather than being the result of conscious effort. Spend time deliberating about how to catch a fast-moving ball and you'll miss it; let the body do its stuff and you'll probably catch it. In many cases consciousness just gets in the way. Perhaps, then, intelligent creatures can get along just fine without consciousness?

When we search for extraterrestrials we're hoping not only to find intelligence but also consciousness; we want to talk to creatures with whom we can share insights into science and art and philosophy. Could it be that this search is doomed because intelligent species generally don't develop the handicap of consciousness? For without consciousness, presumably they'd have no urge to communicate or explore or reach out to other intelligent, conscious species. They wouldn't care.

The notion that the solution to the Fermi paradox lies in the concept of consciousness is due to the Canadian SF author Peter Watts. As we saw in Solution 44, his compatriot Karl Schroeder suggested that intelligence, and by implication consciousness, is a passing phase; Watts on the other hand suggests that consciousness is unlikely to evolve in the first place. It's just not important. According to Watts, intelligence can exist without consciousness. Watts and Schroeder take completely different paths but reach the same ultimate destination: intelligent and conscious beings are rare.

Watts dramatizes his idea in a truly chilling science fiction novel[388] called *Blindsight*. The novel takes its name from a spooky phenomenon involving certain patients suffering from lesions in their primary visual cortex. These patients possess normally functioning eyes but are nevertheless cortically blind: they are declared blind by all the usual tests for blindness. Despite their handicap, some cortically blind patients can see without knowing they can see. They somehow sense the presence of objects and can even catch objects thrown at them but they have no conscious experience of perceiving the objects. In one case a man rendered completely blind by two strokes in different regions of the brain was asked by psychologists to walk down a corridor without the aid of a white cane. The patient was apprehensive, but since the psychologists told him the corridor was empty, and they'd be right behind him if he needed help, he attempted the task. The researchers captured the experiment on film.[389] The video shows the man carefully manoeuvring around a succession of obstacles; he avoided dustbins, trays and various other office paraphernalia even though he had no conscious knowledge that the objects were there or that he was steering his body around them.

So what's going on? How can one see without seeing? It's a difficult question for researchers to investigate. One difficulty is that only a relatively small number of patients suffer cortical blindness in isolation; it's often accompanied by severe injury to other parts of the brain. Furthermore, most patients with cortical blindness are unlikely to realize that they might possess unconscious visual function, precisely because the phenomenon is not accessible to conscious awareness. Finally, from this small sample size, researchers are reliant upon subjective accounts. All this makes it difficult to tease out what might be happening. Nevertheless, one quite widely accepted explanation of the phenomenon is that the human eye sends information to two quite distinct visual systems in the brain: an advanced mammalian system, which is located in the occipital lobe, and an older, more primitive reptilian system located in the midbrain. Damage to the occipital lobe can stop signals from reaching the mammalian visual system, but it wouldn't stop those signals from reaching the reptilian visual system in the midbrain. If consciousness can access the advanced visual system but not the primitive system, and if the primitive system is associated with basic behaviors such as recognizing motion and positioning one's body, then this would explain how blindsight occurs. Perhaps this is the closest we can come to experience what it's like to be a reptile. A lizard doesn't need to recognize a fly, or think about what the fly represents, in order to catch its lunch; it just has to notice movement. That recognition of movement causes the lizard's tongue to shoot out and catch the fly: everything happens automatically. Consciousness is unnecessary for the lizard to survive. In fact, consciousness would be a hindrance for the lizard.

If this is indeed the explanation of blindsight then it has implications for our understanding of consciousness. It implies that consciousness doesn't exist in every part of the brain. More importantly, it raises once again the question posed above: if our brains don't depend upon being conscious, what exactly was it that consciousness brought to the party when humankind was still hunting and gathering in Africa? Why do we possess consciousness? What's the purpose of this first-person narrator that we all carry around with us in our heads?

I've come across no convincing account of the development of human-level consciousness, of how matter can somehow become aware of itself. As far as my reading has taken me,[390] the nature of human consciousness remains a mystery to science. Could it be that consciousness, this facility to speculate and ponder and reflect, which leads, ultimately, to the glories of human civilization, is just a fluky and far from necessary byproduct of evolution? Perhaps one day we'll leave Earth, explore the Galaxy, and find intelligent creatures. But it could be they won't be conscious, won't understand why we might want to talk to them. Lights on, but no one home.

## Solution 74 Gaia, God or Goldilocks?

*It is the fortunate who should extol fortune.*

Johann Wolfgang von Goethe, *Torquato Tasso*

As mentioned in the introduction to chapter 5, Peter Ward (a geologist and paleontologist) and Donald Brownlee (an astronomer and astrobiologist) took a wide-ranging look at why our planet might be special. In *Rare Earth*, Ward and Brownlee described several factors, ranging from the size of a galactic habitable zone to the rate of extinction events, that might restrict the number of planets on which complex life arises. More recently, the British geophysicist David Waltham in his book *Lucky Planet* focuses on one detail that might make Earth a special place for life: its four billion year history of clement weather.[391]

Since life first appeared, Earth has had a relatively stable climate. Average surface temperatures have inevitably fluctuated over the aeons—our planet has suffered episodes of glaciation and episodes of heat—but we can measure those temperature differences in terms of tens rather than hundreds of degrees. And that stability has been important. A frigid Earth, with its water locked up as ice, would be inhospitable. Life, if it existed, would lie dormant and be unable to *do* much—it's the ability of liquid water to carry out various jobs that makes life what it is. A hothouse Earth would be even worse because high temperatures tend to unfold proteins away from their native state. Protein unfolding can be a useful activity for humans—denaturing proteins through the application of heat is a fancy term for "cookery" (at least, it is for a person of my level of culinary ability)—but it does little to promote biodiversity. Organisms can certainly survive in extreme conditions. *Methanopyrus kandleri* strain 116, for example, can survive at 122 °C (at large ocean depths, where the high pressure means the water doesn't boil). In pockets of very salty water surrounded by sea ice, it's possible to find organisms that thrive in the cold.

*Methanogenium frigidum*, for example, lives at the bottom of an Antarctic lake, where it needs neither oxygen nor sunlight. However, if Earth had gone the way of Venus (average surface temperature of about 462 °C) or Mars (average surface temperature of about −55 °C) then life would not have survived. Furthermore, although extremophiles have adapted to temperature extremes, they can't well survive temperature *variations*: *M. kandleri*, for example, struggles in temperatures less than about 80 °C while *M. frigidum* stops growing when the mercury hits room temperature. Evolution has provided more complex organisms with a variety of mechanisms for handling changes in environmental temperature—from feathers and fur to shivering and sweating—but any organism is comfortable only over a restricted range of temperature. Over the past half billion years, Earth's climate has been suitable for complex multicellular life.

Earth's uninterrupted spell of good weather is highly surprising since a variety of factors—astronomical, biological and geological—act independently to control surface temperature and all of them have changed over Earth's lifetime. Ocean make-up has changed, atmospheric composition has changed, the amount of land has changed ... Let's look in a little more detail at a one particular factor that influences surface temperature. When Earth was born the Sun was smaller than it is today. Gradually, our star has expanded—it's due to the helium "ash" generated by the nuclear fusion of hydrogen: the helium sinks to the Sun's core, causing the core to contract and get hotter, which in turn increases the amount of hydrogen burning. In effect, over time the Sun has become a larger radiator capable of pumping out more heat. When life got going here on Earth, the Sun was radiating only about 70% of the heat it does today. If Earth's atmosphere back then was the same as it is now then liquid oceans would have been impossible, and yet we know those oceans existed. If Earth's surface temperature had followed the Sun's increasing heat output then Earth's biodiversity might now extend to a few species of heat-loving extremophiles, and yet here we all are—animals, plants, fungi—enjoying a pleasant climate that, if anything, has exhibited a slight cooling trend. It seems that the various changes have all more or less cancelled each other out. For example, the young Earth's atmosphere contained large amounts of greenhouse gases and they provided a warming effect that compensated for the weak Sun; and as the Sun's luminosity increased, Earth's atmosphere lost enough of its greenhouse gas (through mechanisms described in earlier sections) to maintain the temperature.

Waltham outlines three possible explanations for this happy circumstance, this aeons-long spell of settled climate that has allowed complex multicellular life to develop. He calls the explanations Gaia, God and Goldilocks.

The "God" explanation needs no further explanation. If one believes that a benevolent deity has fine-tuned a set of parameters in order to allow life in general to flourish, and humans in particular, then that's that. Nothing else to say.

The "Gaia" explanation is based on the hypothesis, originated by James Lovelock,[392] that various feedback mechanisms allow life itself to create, maintain and develop the conditions needed for life to survive and thrive: Earth can be seen as a single, living, self-regulating organism. Life has undeniably had a significant impact on Earth—the atmosphere would look quite different if our planet were lifeless, for example—but the Gaia hypothesis is not without its critics. Although Lovelock's idea has been the spur for a great deal of biological research, it still lacks clear observational support. Gaia might exist; it might not.

The "Goldilocks" explanation is another way of saying that we've been incredibly lucky. Consider the discussion of temperature mentioned above: the increasingly luminous Sun has had a warming effect and some other mechanism has had a cooling effect and taken together there's been a slight cooling trend. Proponents of Gaia argue that the cooling is ultimately due to biological feedback loops. However, there's another way of looking at this, a view that doesn't require feedback. Perhaps life did indeed play an important role by removing greenhouse gases from the atmosphere, and perhaps certain geological effects have also produced a cooling effect, but instead of attributing all this to feedback loops we can attribute it to coincidence. The overall warming effect caused by our evolving Sun just happens to be roughly cancelled by the overall cooling effect caused by biology and geology. The net result is a general cooling trend with background fluctuations of a few tens of degrees, and it's all down to chance. Luck. The cancellation on most other Earth-like planets will be less effective, those planets will freeze or boil, and complex life will be impossible. The cancellation on some other Earth-like planets will be better, but result in a net warming trend that will also make the development of complex life unlikely. We just happen to live on a lucky planet where the cancellation is just right to permit complex life—and eventually intelligence—to evolve.

Should a scientific explanation rely on luck? Well, once again we encounter the anthropic principle. As Carter pointed out, Earth's history must be compatible with our presence as intelligent observers. We are here. We can explain that fact by supposing Gaia feedback mechanisms are in place or, as Waltham suggests, by supposing we happen to live on a planet where temperature trends from different sources happened to cancel in a way that permitted live to evolve. We could hardly find ourselves on a planet where the past climate was such that it *didn't* allow life to evolve.

Is there a way of distinguishing between Gaia and Goldilocks? For most of the factors that affect Earth's habitability it's going to be almost impossible to decide, after the fact, whether a given outcome was down to dumb luck or to some particular property of life. Waltham, however, discusses one factor that does allow us to distinguish between Gaia and Goldilocks, between life and luck: the Moon.

Solution 63 looked at ways in which the Moon might be necessary for life and discussed how the Moon seems to play a role in stabilizing Earth's axial tilt: remove the Moon and Earth's obliquity starts to vary wildly. And wild obliquity changes cause the climate to change in ways that are inimical to life. However, Waltham points out that instead of asking "What would happen if we removed the Moon right now?" a more appropriate question to ask is "What would have happened if the Moon-forming collision had created a Moon much larger than the one we actually have?"—and the answer to that is surprising.

The Moon raises tidal bulges in the oceans (and, to a much smaller extent, in the continental land masses) and this acts as a frictional brake on Earth's rotation. Every 50,000 years or so the day gets longer by about 1 second. Furthermore, because the tidal bulges are slightly ahead of the Moon rather than directly under it, the Moon is pulled forward and thus moves into a slightly higher orbit. Every year the Moon moves away from Earth by about 4 centimeters or so. This evolution of the Earth–Moon system, an increasing Earth day and an increasing separation between Earth and Moon, is just the playing out of Newtonian dynamics. One consequence of this orbital evolution is that Earth's precession—the change in the direction of its spin axis, which moves rather like the axis of a gyroscope—will slow. At present Earth precesses once every 26,000 years. This precession period will increase as Earth rotates more slowly and the evermore distant Moon raises smaller tides. Eventually, in about 1.5 billion years, the precession period will have become about 50,000 years. Unfortunately for our successors, the planetary orbits also oscillate with a period of about 50,000 years. Earth enters a "zone of instability". I say that this zone is "unfortunate" because when oscillation periods are aligned resonance occurs: it's like pushing on a swing—push at the right period and the amplitude builds up. (We saw something similar in Solution 59, with resonance effects from Jupiter causing a gap in the Asteroid Belt.) So, about 1.5 billion years in the future, the effects of planetary orbits will cause Earth's axial tilt to begin to wobble chaotically. The resultant temperature extremes will cause life to struggle. (In truth, any Earth life of 1.5 billion years hence will have other concerns. For example, difficulties caused by axial instability will only compound the effects of a much more luminous Sun.)

Waltham developed computer models to study how Earth–Moon systems evolve. The beauty of computer models is that you can investigate what would have happened if the Moon were just smidgen smaller or just a little bit bigger;

or what would have happened if the young Earth's day had been a tad longer or a few minutes shorter. It turns out that the possession of a large Moon is a mixed blessing. A large Moon increases the axial stability of the planet through increased tidal forces, which is good, but a large Moon also increases the rate at which the planet enters that zone of instability. It turns out that our Moon is almost as large as it could possibly have been and not have caused instability for us. Waltham's models suggest that if the Moon-forming collision had produced a satellite with a radius just 10 kilometers bigger than our Moon, and if the young Earth had turned so that its day was just 10 minutes longer than our Earth, then we would be entering the zone of instability right about now. Or imagine an Earth–Moon system that's identical to ours in every way except that tidal drag is increased by a few percent; again, we would be just about to enter the zone of instability. Our days would be numbered.

It's rather a coincidence that we have a Moon that's almost as big as it could possibly be without rushing us to the zone of axial instability. But the possession of such a Moon is what we might expect to see if a big Moon promotes the existence of complex life for some other reason that has nothing to do with axial stability. Waltham likens the situation to average speeds observed on British motorways. The 70 mph limit puts a cap on the allowed speed; but we're all in a hurry and so we tend to drive close to the limit. Pick a car at random on a British motorway and it's likely to be going close to 70 mph—as fast as it could possibly be traveling without entering the zone of illegality. So are there any reasons why having a big Moon is of benefit for complex life? Well, we considered some speculations in Solution 63. Waltham adds his own speculation. Our Moon causes Earth's axis to precess slowly and it causes Earth's day to be relatively long. Waltham argues persuasively—although at present it's still speculation—that these two effects mean that Earth suffers relatively mild and infrequent ice ages.

We can imagine billions of Moon-forming collisions, each of which are subtly different. In most cases the resulting Earth–Moon system descends into glaciation or else wanders into climate chaos; in most cases, life struggles to survive. Our own Earth–Moon system has hit that "sweet spot". The size of the Moon, the length of day, the angle of the obliquity all combine to give us good weather. And the key point is that all this has nothing to do with Gaia. We can't argue that biofeedback loops have somehow been involved, or that the sheer adaptability and robustness of life has been key. This is just Newtonian mechanics acting on Goldilocks.

Is this dose of good fortune too much to swallow? Not necessarily: it's the anthropic argument again. If life depends upon various planetary factors (such as the existence of a magnetic field, the correct amount of rock, the presence of a large but not huge Moon, and so on) and those factors combine in a manner reminiscent of the Drake equation to make complex life a one-in-a-trillion chance ... well, life is surely going to happen somewhere simply because there are many trillions of planets out there. And if that life gives rise to intelligent observers then those observers will inevitably find themselves on a planet in which those factors combined in the right way. Those observers might look for a Gaia (or God) explanation but all that would be required is a Goldilocks explanation.

As planetary scientists learn more about exoplanets, and the different ways of building a planetary system, they'll be better able to understand whether Earth truly is an oddball. At present it's too early to say. But it's certainly possible that we live on the Lucky Planet.

# 6
# Conclusion

I have criticized 74 proposed resolutions of the Fermi paradox,[393] so it's only fair that I give my own. I was unhappy with the treatment I presented in the first edition to this book, so this time round I'm trying a different approach. The conclusion is the same, but the route I take to get there is rather different. This is by no means an original suggestion, but it does sum up what I feel the paradox might be telling us about our universe.

The American science fiction author David Brin, in his superb 1983 analysis of the Great Silence, wrote that "few important subjects are so data-poor, so subject to unwarranted and biased extrapolations—and so caught up in mankind's ultimate destiny—as is this one". More than three decades after Brin published his review, little has changed.

The subject is *still* data-poor. To be sure, we have more relevant knowledge now than we did even at the turn of the century. There have been some tremendous advances in particular areas. Developments in computing and astronomical technology have made possible a variety of powerful SETI programs; astronomers understand more about the formation of planetary systems, while the discovery of exoplanets has become routine; biologists are uncovering the fundamental workings of life on Earth (although, as is usual in science, new discoveries seem to create an expanding shell of ignorance). Nevertheless, we've barely begun to find answers to many of the deep questions in this area.

The subject is *still* liable to unwarranted, biased extrapolations. Given the profound importance of the subject, though, should our lack of hard data force us to remain silent? Surely the best we can do under the circumstances is to be frank about our biases and open about our extrapolations. At least then a debate can take place, even if for the moment such debate will generate more heat than light.

The subject is *still* important. What could be more so? Either we are alone, or else we share the universe with creatures with whom we might one day communicate. Either way, it's a staggering thought.



S. Webb, *If the Universe Is Teeming with Aliens … WHERE IS EVERYBODY?*,
Science and Fiction, DOI 10.1007/978-3-319-13236-5_6

## Solution 75 The Fermi Paradox Resolved . . .

*When facts are few, speculations are most likely to represent individual psychology.*
Carl Gustav Jung

The paradox resolved? Well, no. Of course not. The topic remains so intangible that honest people can reach quite opposite conclusions. The reader is free to choose one or more of the solutions presented earlier, or to originate his or her own. Here, I present the solution that makes most sense to me. Before presenting my own take on the paradox, however, I'd like to discuss briefly why so many people believe that intelligent extraterrestrial beings must exist.

My non-scientist friends tend to defend their belief in extraterrestrial intelligence by giving what one might call the Douglas Adams response:[394] "Space is big. Really big. You just won't believe how vastly, hugely, mindbogglingly big it is." Surely we *can't* be the only intelligent species in such a large universe? When one looks at how insignificant Earth appears in figure 1, in a photograph taken from a neighboring planet, it's difficult to conclude there aren't other civilizations out there in all that vastness. And yet the size argument really has little relevance because it turns out that most of our universe is empty. Well, that's not quite right. The universe appears to be full of "stuff", but it's "stuff"—dark energy and dark matter—of which we have almost zero knowledge except for the fact that it isn't suitable for constructing life. Even the 5% of the mass–energy content of the universe that we understand—atoms and neutrinos and radiation—is spread thinly, and most of it isn't in a form that would permit the existence of life. The universe might be big, but size alone tells us little about whether there are homes for beings such as us.

My physical-scientist friends tend to defend their belief in extraterrestrial intelligence by pointing to the numbers. It's not the size of the universe *per se* that's important, but the fact that it's big enough to contain vast numbers of terrestrial planets. We don't know precisely how many such planets are out there, but one recent estimate[395] suggested (perhaps optimistically) that the Galaxy might contain as many as 100 billion habitable, Earth-like planets. There are about 500 billion galaxies in the universe, and so there might be as many as 50 *sextillion* potential homes for life. That's a 5 followed by 22 zeros. Surely we can't be the only intelligent species when there are so many sites on which intelligent species could evolve? A sextillion is a big number, right?

The trouble with this argument is that we don't know whether a sextillion (or 50 sextillion, or 100 sextillion, or whatever number you believe is appropriate) is large in this context. It might be. It might not. Large numbers arise quite easily in the simplest of contexts. Let me give just one example; it's a problem to ponder the next time you're in some boring committee meeting. List every possible subcommittee that can be formed from the people in your meeting and consider every possible pair of subcommittees. Assign each pair to one of two groups. What's the smallest number of people in the original committee that will guarantee, no matter how the assignment is made, there will be four subcommittees in which all pairs are in the same group and all the people belong to an even number of subcommittees?

Okay, I guess at first glance it's not the most interesting of problems. I'm being unfair, too, because it's a tough problem: it hasn't yet been solved. The mathematician Ronald Graham, however, once proved that there exists a solution to this problem—or, to be precise, to an equivalent problem—and he proved the solution lies between 6 and some number we'll call $G$ (which stands for Graham's number).[396] The point I want to make is that Graham's number, which arises from a simple enough problem, is big. Very, *very* big. $G$ is so big its representation requires a special notation. A commonly used notation to represent very big numbers is due to Don Knuth of TeX fame but, as we'll see, even this notation doesn't work easily with an integer the size of Graham's number.

Knuth introduced the operator $\uparrow$. A single $\uparrow$ is the same as exponentiation:

$$m \uparrow n = m \times m \times \cdots \times m = m^n.$$

So we have $2 \uparrow 2 = 2 \times 2 = 2^2 = 4$ and $3 \uparrow 4 = 3 \times 3 \times 3 \times 3 = 3^4 = 81$ and so on. Things get interesting when you have a pair of arrows, $\uparrow\uparrow$. This represents a tower of exponents:

$$m \uparrow\uparrow n = m^{m^{\cdot^{\cdot^{m}}}}$$

where the tower is $n$ rows high. This lets you generate some big numbers very quickly. For example:

$$3 \uparrow\uparrow 2 = 3^3 = 27$$
$$3 \uparrow\uparrow 3 = 3^{3^3} = 3^{27} = 7625597484987.$$

Play around with the double arrow notation to get a feel for it. See if you can comprehend how big $3 \uparrow\uparrow 4 = 3^{7625597484987}$ is. If you can, you're doing better than me. The number is already *vastly* greater than the number of particles in the known universe. But we haven't even started yet. Consider the operator

$\uparrow\uparrow\uparrow$, which generates a tower of a tower of exponents. Let's look at $3 \uparrow\uparrow\uparrow 3$:

$$3 \uparrow\uparrow\uparrow 3 = 3 \uparrow\uparrow 7625597484987 = 3^{3^{\cdot^{\cdot^{\cdot^{3}}}}}$$

where the total height of the tower contains 7625597484987 levels. It's a crazily large number. But we still haven't scratched at Graham's number. Let's consider the operator $\uparrow\uparrow\uparrow\uparrow$, which generates a tower of a tower of a tower of exponents. Think of the number $3 \uparrow\uparrow\uparrow\uparrow 3$ which is . . . well, it's so big that it's very difficult to write out. Try it and you'll see. When thinking about Graham's number we *start* with this number, which is represented by $g_1$. In other words, $g_1 = 3 \uparrow\uparrow\uparrow\uparrow 3$. The number $g_2$ is *absurdly* vast:

$$g_2 = 3 \uparrow\uparrow \cdots \uparrow\uparrow 3, \qquad \text{with } g_1 \text{ arrows between the 3's.}$$

Just four uparrow operators between the 3's generates a number that's too big to write comfortably. Here we're thinking about a number with $3 \uparrow\uparrow\uparrow\uparrow 3$ uparrow operators between the 3's. That's $g_2$. The number $g_3$ has $g_2$ uparrow operators between the 3's. And so on. Graham's number is $g_{64}$.

It's almost impossible to comprehend the utter humongousness of Graham's number. It dwarfs anything your mind (well, my mind at least) can comprehend. Compared to Graham's number, 50 sextillion—the possible number of habitable, Earth-like planets—is laughably small. So is 50 sextillion a big number when we're discussing the possibility of extraterrestrial intelligence? It might be, if it turned out for example that life is present on most of those planets. But if the development of intelligent life from inanimate material turned out to be a 1-in-$G$-type event then the number of planets would be irrelevant.

**Some Big Numbers** Graham's number is insanely large, and its digits can't possibly be written down (the universe isn't big enough to hold contain a decimal representation of the number, no matter how small you write), but we do know what the final few digits are. For what it's worth, Graham's number ends . . . 2464195387.

Other numbers, even bigger than Graham's number, have appeared in serious mathematics literature. Mathematicians working in the area of combinatorics or computer science work with astoundingly large numbers, which require special notations to represent them. Mathematicians working on Kruskal's tree theorem, for example, encounter a number that makes Graham's number look puny: they employ a function called TREE, which starts off TREE(1) = 1 and TREE(2) = 3 but TREE(3) is so insanely big that the Knuth up-arrow notation struggles to cope. Graham's number is much closer to TREE(2) than it is to TREE(3).

My life-scientist friends, unlike those who studied physical sciences (or indeed those who haven't studied any science), tend to be much more skeptical about the prospects for intelligence—or at least about the prospects for intelligence that develops into a civilization capable of communicating with us.

Biologists tend to agree that other forms of life will exist (the number of planets on which life might start *is* large, after all) but they don't buy the deterministic "high intelligence evolved on Earth so it must eventually evolve on other planets" argument. They tend to see the implausibility, rather than the inevitability, of intelligence.

My own view? Well, I side with my biologist friends.

The debate about extraterrestrial intelligence contains just one gleaming, hard fact: we haven't been visited by ETCs nor have we heard from them. So far, the universe remains silent to us. Those who would deny this fact of course have a ready solution to the Fermi paradox (and presumably stopped reading this book after the first few pages). The job for the rest of us is to interpret this lone fact.

As the quotation that starts this section suggests, when we have just one piece of evidence to play with our biases will come to the fore. My own biases, such as I can identify them, include optimism about our future. I like to think our scientific knowledge will continue to expand and our technology to improve; I like to believe humankind will one day reach the stars—first by sending messages and then later, perhaps, by sending craft. I like to hope that something akin to the Galaxy-spanning civilization described by Asimov in his classic *Foundation* stories might one day come to pass. But these biases collide with the Fermi paradox: if *we* are going to move out into the Galaxy, why have *they* not already done so? They've had the means, motive and opportunity to establish colonies, yet they appear not to have done so. Why? Well, I believe it's because "they"—sentient, intelligent, sapient creatures that build civilizations and with whom we can communicate—don't exist.

I agree that when one looks up at the sky on a clear moonless night, and gazes with the naked eye at the myriads of stars and the vastness of space, it's difficult to believe we might be alone. We are too small and the universe is too big for this to make sense. But appearances can be deceptive: even under ideal observing conditions one is unlikely to see more than about 3000 stars, and few of those would provide conditions hospitable to our form of life. The gut reaction we perhaps all feel when we look at the night sky—that there *must* be intelligent life somewhere out there—isn't a good guide. We have to be guided by reason, not gut reaction. Well . . . reason tells us there are billions of Earth-like planets in our Galaxy, and trillions of such planets in the immediate galactic neighborhood, so aren't the physicists and astronomers right? Doesn't sheer weight of numbers mean that intelligence, and perhaps intelligence vastly superior to our own, is inevitable? I don't think so. I think that argument has a whiff of arrogance about it. Let me explain what I mean.

First, in our search for sapient, intelligent extraterrestrial beings we're assuming that abiogenesis—the creation of life from non-life—is not unlikely.

This assumption might be unwarranted; it's possible that life on Earth arose from some fluke, never-to-be repeated event. However, since many of the chemical ingredients for life are present in cosmic dust, and since there are so many planets out there, let's agree that there are many instances out there of life getting a start.

We're searching, then, for planets on which life has started and on which conditions remained hospitable for billions of years—long enough for evolution to work its magic. But how many planets will possess that level of stability? Conditions have been suitable for life throughout Earth's history, but we can't use this to argue about the likely stability of conditions elsewhere: we are here so we *have* to look back on a history that has been suitable for the development of intelligent life. Other planets might lack a large moon, or be without a protective magnetic field, or orbit a star that's too variable, or possess a climate that runs away into greenhouse or icehouse, or have its atmosphere stripped by a nearby gamma-ray burst, or . . . well, we've seen how dangerous the universe can be. Not every planet that gives birth to life will be able to protect its progeny.

We're searching for those planets that not only provide a long-term home for life but that have seen the appearance of complex, multicellular organisms. But why should we expect life to evolve beyond the prokaryotic grade? There seems to be nothing inevitable about it. On those planets where complex life-forms do come into being, we're searching for life-forms that developed the same or similar sense organs to us, in order for communication to take place. But why should we expect this to be common? Perhaps olfaction, or magnetoception, or thermoception—or, more likely, senses we don't even dream about—are more useful to creatures trying to survive the environments to be found on other planets.

We're searching for life-forms in which high intelligence has developed. But why should we expect intelligence to be widespread? It certainly isn't common on Earth. The archaea and the bacteria didn't evolve intelligence after our line split off from theirs. The fungi and the plants didn't evolve intelligence after the animal line diverged from theirs. Of the various animal phyla, only chordates evolved intelligence; and of the chordates only vertebrates evolved intelligence; and of the vertebrates only mammals; and of the mammals only humans have evolved the high intelligence we're seeking. We look back and see a ladder of intelligence that leads to us at the top. But that's a biased way of looking. If we look around rather than back, we see that high intelligence just isn't that important: millions of species manage perfectly well without it.

We're searching for intelligent life-forms that have also evolved conscious self-awareness. We're searching for intelligent, conscious life-forms that have both the available resources and the need to manipulate raw materials into tools. We're searching for intelligent, conscious, tool-making beings that have developed a language we're capable of understanding. We're searching for intelligent conscious, tool-making, communicative beings that live in social groups (so they can reap the benefits of civilization) and that develop the tools of science and mathematics.

We're searching for ourselves . . .

That's the sense in which I mean the argument for the existence of extraterrestrial intelligence has an air of arrogance about it. When we look up into the night sky, why should we expect to find beings possessing precisely those qualities that define humanity? The millions of species with which we share our planet are all equally as "evolved" as us: they all earn a living in a harsh world that cares not whether they live or die. They manage to survive in a spectacular number of different ways. There's no evolutionary drive towards the sort of intelligence that defines our species. If we don't find intelligence here, why on Earth should we find it out there?

If we're engaged in a search for ourselves, though, then the activity assumes tremendous importance. What would it mean to us if we learned that we really were the only conscious species in the universe? The responsibility would be astounding.

The famous French biologist Jacques Monod[397] once wrote that "Man at last knows he is alone in the unfeeling immensity of the Universe, out of which he has emerged only by chance". It's a melancholy thought. I can think of only one thing sadder: if the only species possessing consciousness, the only species that can light up the universe with acts of love and humor and compassion, were to extinguish itself through acts of stupidity or ignorance. The various "Solutions" discussed in chapter 4 don't, I believe, solve the Fermi paradox; but they do describe a range of possible futures for our descendants. We can choose which future we want. If we survive, we have a Galaxy to explore and make our own. If we destroy ourselves, if we ruin Earth before we're ready to leave our home planet . . . well, it could be a long, long time before a creature from another species looks up at its planet's night sky and asks: "Where *is* everybody?"

# Notes

## Where *Is* Everybody?

[1]**Pg 1 reading the works of Isaac Asimov** The American author Isaac Asimov (1920–1992) was one of the 20th century's most prolific authors. He wrote on a vast number of topics—from the Bible to Shakespeare—but it was his science books, both fiction and non-fiction, that had the most impact on me. For a memoir, written towards the end of his life, see Asimov (1994).

[2]**Pg 1 appeared in successive issues** The "pro-Fermi" article, by the American geologist and science fiction writer Stephen Lee Gillett (1953–), appeared in the August 1984 issue of *Asimov's*. The rebuttal, by the American scientist and author Robert A. Freitas Jr (1952–), appeared in the September issue. A few years later, Gillett expanded upon his original article and pointed out a different interpretation of the "lemming paradox" introduced by Freitas and discussed here on page 2. If Earth were empty except for lemmings then the creatures *would* be everywhere; but Earth teems with other living things, which out-compete lemmings and limit their spread. The correct conclusion to draw from the non-observation of lemmings is that Earth has an abundance of living species competing for resources (which we knew anyway, because we see life all around us). When we look into space, however, we see *nothing* that indicates the presence of life.

[3]**Pg 5 latest cosmological measurements** The WMAP and *Planck* space missions have tied down the key numbers describing our universe. For details see, for example, NASA (2012) and ESA (2014).

## The Physicist Enrico Fermi

[4]**Pg 10 precocious ability in mathematics** For details of Fermi's life I consulted two sources: a biography written by his wife Laura (Fermi 1954); and

 
S. Webb, *If the Universe Is Teeming with Aliens . . . WHERE IS EVERYBODY?*, Science and Fiction, DOI 10.1007/978-3-319-13236-5

a readable account of Fermi's life in physics, written by Emilio Segré (1905–1989), a friend, student and collaborator of Fermi (Segré 1970). Segré himself won the Nobel Prize for physics in 1959. A symposium held in Chicago in 2001 to commemorate the centenary of Fermi's birth highlighted the sheer breadth of his impact on physics; the proceedings were later published (Cronin 2004).

[5]**Pg 10 quickly outstripped his teachers** Luigi Puccianti (1875–1952), Fermi's teacher, was the director of the physics laboratory at the Scuola Normale Superiore in Pisa. According to Laura's account (Fermi 1954) Puccianti asked the young Fermi to teach him relativity. "You are a lucid thinker", Puccianti said, "and I can always understand what you explain".

[6]**Pg 13 pile went critical** The man in overall charge of the project that aimed to achieve the first self-sustaining nuclear reaction was Arthur Holly Compton (1892–1962), an American physicist who won a Nobel prize for his work in subatomic physics. When it was clear Fermi had attained the goal, Compton telephoned James Bryant Conant (1893–1978), the President of Harvard University. The telephone call was cryptic: "Jim, you'll be interested to know that the Italian navigator has just landed in the new world". See Compton (1956) for details of the project.

## Paradox

[7]**Pg 13 word paradox comes from** See Poundstone (1988) for an entertaining and readable book dealing with a variety of paradoxes. As well as those I discuss here, you can read about Russell's barber paradox, Newcomb's psychic paradox and many others—but not the Fermi paradox.

[8]**Pg 14 Rapoport once remarked** The Russian-born biomathematician Anatol Rapoport (1911–2007) is known for his work in a variety of fields, including the analysis of a famous mathematical paradox: the prisoner's dilemma. For a short, readable introduction to this paradox, see Rapoport (1967).

[9]**Pg 15 intentional vagueness** Our word "sorites" comes from the Greek word *soros* meaning "heap", since it was first used in the type of reasoning described in the text. (In other words, one grain of sand doesn't make a heap; if one grain of sand doesn't make a heap, then neither do two grains; and so on *ad infinitum*.) See Williamson (1994) for a comprehensive account of the sorites paradox.

[10]**Pg 15 raven paradox** The raven paradox was developed by the German-born philosopher Carl Gustav Hempel (1905–1997), one of the leaders of the logical positivist movement. The paradox first appeared in Hempel (1945a, b).

[11]**Pg 16 generated a huge literature** The paradox of the unexpected hanging was first noticed by the Swedish mathematician Lennart Ekbom when he heard the following wartime announcement by the Swedish Broadcasting Company: "A civil defense exercise will be held this week. In order to make sure that the civil defense units are properly prepared, no-one will know in advance on what day this exercise will take place." For more details on this paradox, see Gardner (1969). Although Martin Gardner (1914–2010) was best known for his mathematics columns in *Scientific American*, he trained as a philosopher and published scholarly articles on paradox.

[12]**Pg 17 fact of interstellar travel** Although the twin paradox involves Einstein's special theory of relativity, Einstein himself of course understood his own theory well enough not to present this phenomenon as a paradox. However, although Einstein was also one of the founders of quantum theory, he was less sure of his ground in this field. He and his co-workers Boris Podolsky (1896–1966) and Nathan Rosen (1909–1995) constructed a marvelously subtle argument (now called the EPR paradox) intended to prove that quantum physics is incomplete. Again, a full analysis shows there is no paradox—but at the expense of introducing a "spooky" (Einstein's own word) phenomenon called entanglement. The EPR result tells us that everything we have ever touched is invisibly tied to us by the weird rules of quantum theory. Clear accounts of the EPR paradox can be found in Mermin (1990) and Gribbin (1996). The paradox was originally described in Einstein et al. (1935).

[13]**Pg 17 first proposed in 2012** The paper that proposed the firewall paradox was available as a preprint in 2012, and appeared in print the following year. See Almheiri et al. (2013).

[14]**Pg 18 proposed an idea** See for example Webb (2004).

[15]**Pg 19 named after Heinrich Olbers** The dark-sky paradox was named after the German astronomer Heinrich Wilhelm Matthäus Olbers (1758–1840), but several other astronomers, including most notably Johann Kepler (1571–1630) and Edmond Halley (1656–1742), had considered the problem before

Olbers published his analysis in 1826. See Harrison (1987) for a thorough, elegantly written discussion of Olbers' paradox, including the early history of the question of why the sky is dark at night.

## The Fermi Paradox

[16]**Pg 21 whose report I draw heavily upon** Eric Jones, an astronomer who spent most of his career at Los Alamos, contacted Emil John Konopinski (1911–1990), Edward Teller (1908–2003) and Herbert Frank York (1921–2009), Fermi's luncheon companions on the day he asked his famous question, and requested them to record their recollections of the incident. He published their accounts in Jones (1985). During the early 1950s, the Americans Konopinski and York were both involved in theoretical work on the development of nuclear weapons, as was the Hungarian-born Teller (who has been described as "the father of the H-bomb"). All three of them would have enjoyed Fermi's input into their discussions on nuclear physics.

[17] **Pg 24 after the radio astronomer** The American astronomer Frank Donald Drake (1930–) was the first person in history to use a radio telescope to search for ETCs. A fascinating account of what led him to a life in astronomy, and of the prospects for finding ETI, can be found in Drake and Sobel (1991).

[18]**Pg 26 formulate the argument as a paradox** See for example Haqq-Misra and Baum (2009) or Prantzos (2013).

[19]**Pg 26 a scientific visionary** The Russian author and philosopher Konstantin Eduardovich Tsiolkovsky (1857–1935) was born into a poor family in the eastern city of Izhevsk. From the age of nine he suffered almost total deafness following a streptococcus infection. Nevertheless, he educated himself and studied chemistry and physics. As early as 1898 he explained the need for liquid-fueled rockets for spaceflight, and in his 1920 SF novel *Beyond the Earth* he described how people would live in orbiting colonies. He promoted his ideas on extraterrestrial life in two essays entitled "There are also planets around other suns" (dated 1934) and "The planets are occupied by living beings" (dated 1933). For a description of Tsiolkovsky's philosophy and his anticipation of the Fermi paradox, see Lytkin et al. (1995).

[20]**Pg 27 clearly stated the dilemma** See Viewing (1975).

[21]**Pg 27 a 1975 paper** See Hart (1975). It was this paper more than any other, I believe, that generated widespread interest in the Fermi paradox.

[22]**Pg 28 the House of Lords** Lord Douglas of Barloch (1889–1980) suggested (Douglas 1977) that the number of evolutionary steps leading from primitive life to intelligence was so large that the probability of it happening elsewhere was infinitesimal.

[23]**Pg 28 Tipler reasoned** The American mathematical physicist Frank Jennings Tipler III (1947–) has published several popular articles on the use of probes to colonize the Galaxy. See, for example, Tipler (1980).

[24]**Pg 28 coolest and best summary** Glen David Brin (1950–) trained as an astronomer, but is much better known as an award-winning SF writer. His article on the "Great Silence" (Brin 1983) remains one of the clearest treatments of the subject. In a popular article (Brin 1985) he gives a brief treatment of 24 possible solutions to the Fermi paradox.

[25]**Pg 28 proceedings were published** See Zuckerman and Hart (1995). The updated second edition of this very readable book is easier to obtain than the first.

[26]**Pg 28 the probability of extraterrestrial life is 1** See Aczel (1998) for a breezy account suggesting that the sheer number of stars in the universe means there *must* be life elsewhere: give something enough of a chance to happen and eventually it will. However, many readers may find the arguments leading to this conclusion unconvincing.

[27]**Pg 28 Smolin wrote that** See Smolin (1997).

[28]**Pg 28 Gould wrote that** See Gould (1985).

[29]**Pg 28 and the economist** Mention of economists reminds me of a proof of the non-existence of time travelers that employs Fermi paradox-like reasoning (Reinganum 1986–7): if time travelers existed, then interest rates would not be positive! In fact, if people could travel back in time, then interest rates would have to be 0%—otherwise savers could use banks as bottomless ATM machines. Savers could simply travel back in time a few thousand years, deposit

a few dollars, then return to the present; compound interest on even a small sum would guarantee riches.

[30]**Pg 29 the acid test of experiment** A good example of the need for experiment was Tipler's argument that, in the distant future, we will all be resurrected in software by a God-like intelligence (Tipler 1994). His argument rested on the universe possessing certain cosmological properties; modern observations seem to exclude these properties and thus at least the initial version of Tipler's theory. We wouldn't know this, however, unless astronomers had looked.

## They Are Here and They Call Themselves Hungarians

[31]**Pg 32 The joke originated** McPhee (1973) ascribes the "theory" that Hungarians are descended from Martians to Leo Szilard, who would have been one of the Martians. However, a posthumously published letter (Morrison 2011) provides a slightly different—and more likely—account of the tale.

[32]**Pg 32 a formidable array of intellect** The five "Martians" mentioned in the text did indeed constitute an extraordinary grouping of talent. Edward Teller has already been mentioned in a previous note. Leo Szilard (1898–1964) made contributions to molecular biology as well as nuclear physics—and also invented a novel type of home refrigerator; his co-inventor was Einstein! (See Lanoutte (1994) for a good biography of Szilard.) Eugene Paul Wigner (1902–1995) was one of the leading experts in quantum theory. John von Neumann (1903–1957) was hugely influential and made immense contributions in a number of fields. Theodore von Kármán (1881–1963) was one of the world's foremost aeronautical engineers. All five were born in Budapest. Another physicist born in Budapest around the same time, although he never worked at Los Alamos, was Dennis Gabor (1900–1979); he was awarded the Nobel prize for his invention of holography. The radiochemist George de Hevesy (1885–1966) was awarded the 1943 Nobel prize in chemistry; he too was born in Budapest. Such a grouping of talent is rare, but probably not unique. Other pockets of brilliance have occurred from time to time. For example, the 1979 Nobel prize-winning particle theorists Sheldon Lee Glashow (1932–) and Steven Weinberg (1933–), who worked independently on electroweak unification, were in the same class at the Bronx High School of Science. Also in the class was Gerald Feinberg (1933–1992), who developed the idea of the tachyon. In addition to Glashow and Weinberg, the Bronx High School has produced three other Nobel prize-winning physicists! A rather more sinister

constellation of people occurred in 1913 in Vienna, the capital of the Austro-Hungarian Empire: Adolf Hitler, Joseph Stalin, Joseph Tito, Leon Trotsky and Sigmund Freud all lived within a couple of miles of each other. Coincidences happen.

## They Are Here and They Call Themselves Politicians

[33]**Pg 34 According to Icke** See, for example, Icke (1999). Icke's was once a well known face on English TV so when I learned about his beliefs I found myself compelled to read one of his books. The book I chose started out badly, rapidly descended to that curious level where something is so bad it's good, but unfortunately continued its descent so that after a few pages I could take no more.

[34]**Pg 34 members of President Obama's administration** See Citizen Hearing on Disclosure (2013) for details of Hellyer's testimony, along with that of 39 other witnesses.

[35]**Pg 34 local election success** At the time of writing, Parkes represents the Stakesby Ward of Whitby Town Council. For details of the 2012 election results, see Scarborough Borough Council (2012). An internet search for Parkes will provide links to several television appearances in which he discusses his dealings with "Mantid" aliens.

[36]**Pg 35 take them seriously** See Nasar (1994) for a thought-provoking biography of the mathematician John Forbes Nash, Jr (1928–), published at around the same time as Nash was awarded the Nobel prize for economics.

## They Are Throwing Stones at Radivoje Lajic

[37]**Pg 35 book on materials science** See Miodownik (2013). I know of only one other popular book on materials science that is better than Miodownik's *Stuff Matters*, and that's the classic *New Science of Strong Materials or Why You Don't Fall Through the Floor* (Gordon 1991).

[38]**Pg 35 about 100,000 meteorites** See Brown et al. (2002) for an estimate of the rate at which small objects strike Earth. Although any particular square meter of Earth is unlikely to be struck by a meteorite during any given year, there's at least one well documented case of an extraterrestrial object striking

a human. The Sylacauga meteorite fell in Alabama on 30 November 1954; a fragment crashed through a roof, bounced off a wooden cabinet radio, and struck Ann Hodges on the hip while she slept on a couch.

[39]**Pg 35 lost their ticket** See Guardian (2001) for the story of the couple who failed to claim their winnings in time.

[40]**Pg 36 student called Martin Andrews** For one of the earliest mentions of Gorman's fake story see Digital Spy (2013). A quick internet search will be sufficient to demonstrate how the story mutated.

## They Are Watching Us from UFOs

[41]**Pg 37 strange lights in the sky** Ezekiel 1:4–28 contains a description of a wheel in the sky that some have chosen to interpret as a flying saucer. The interpretation of apocalyptic writings is notoriously difficult, but it's probably fair to say that the prophet Ezekiel wasn't describing a physical event. Depending upon one's outlook on these things, he could have been describing a message from God or he might have eaten some funny mushrooms.

[42]**Pg 37 flying his private plane** Kenneth Arnold (1915–1984) wrote about his sighting in *The Coming of the Saucers* (Arnold 1952).

[43]**Pg 37 As surveys consistently show** Many surveys have examined peoples' attitudes to UFOs over the past few decades. Depending on the precise nature of the question asked, the percentage of Americans professing to a belief in the existence of UFOs—which presumably equates to a belief in the existence of extraterrestrial spacecraft—generally ranges between 30 and 50%. For the results of a recent survey see, for example, Harris Interactive (2013).

[44]**Pg 38 coined by Edward Ruppelt** The relatively early death of Edward J. Ruppelt (1922–1959), due to a heart attack, sadly but inevitably sparked more than a few conspiracy theories. A biography of Ruppelt, and a discussion of the 1950s UFO phenomenon from the point of view of "ufologists", is given in Hall and Connors (2000).

[45]**Pg 39 noted skeptic Robert Sheaffer** Many books have been written in support of the thesis that UFOs are alien spacecraft; skeptical approaches

are much less common. One of the clearest skeptical essays on the UFO phenomenon is in Sheaffer (1995).

[46]**Pg 41 we should use Occam's razor** The law of parsimony—the principle which states that entities are not to be multiplied beyond necessity—must have been invoked by numerous philosophers and scientists before the $14^{th}$ century. But William of Occam (1284–1347) applied the principle so frequently and so sharply that it became known as Occam's razor.

## They Were Here and Left Evidence of Their Presence

[47]**Pg 42 footprints of alien technology** See Davies (2012) for a discussion of "astroforensics" and the difficulties involved in searching for traces of past alien activity. In addition to his technical physics writing, Paul Davies is an outstanding writer of popular science; see for example Davies (2010) for some beautifully clear explanations of the Great Silence.

[48]**Pg 42 traces that might yet survive** We can try to get a handle on the present traces of possible past technological activity by asking what elements of our current civilization might survive into the far future. If every person died tomorrow, what evidence that our species had once walked the Earth would survive for a million years? Or ten million years? Or longer? See Weisman (2007) for a popular-level account of the question; a more scientific account, written by a geologist, can be found in Zalasiewicz (2009).

[49]**Pg 42 the Oklo reactor** See Meshik (2005) for a clear, non-technical discussion of the Oklo reactor.

[50]**Pg 44 famous for a series of books** Erich Anton von Däniken (1935–), a Swiss author, wrote his most famous book, *Chariots of the Gods*, when he was working as a hotel manager. He followed it up with titles such as *The Gold of the Gods* and *The Return of the Gods* (see von Däniken 1969, 1972, 1997). For an excellent and entertaining discussion of why these books are wrong-headed, see Story (1976).

[51]**Pg 45 covered them** See Crawford et al. (2008) for a related problem: the survivability and detectability of terrestrial meteorites on the Moon.

[52]**Pg 45 past extraterrestrial visitations** See Davies and Wagner (2013) for a strategy that could be employed to search for alien artefacts on the Moon.

[53]**Pg 45 a bridge** Six decades on it seems strange to us that anyone would claim to have observed a bridge on the Moon, but the Welsh astronomer Hugh Percy Wilkins (1896–1960) was a fine observer. He produced some excellent maps of the near side of the Moon, and was honored in 1961 by having a 57-km diameter lunar crater named after him.

[54]**Pg 46 where might we find them** For a treatment of how we might search for Earth-observing probes, see Freitas and Valdes (1980) and Freitas (1983a, b).

[55]**Pg 46 view the entire planet from space** The idea that a probe might observe Earth over a period of millennia is not so outlandish. Even with our present level of technology, the KEO project plans to put a passive satellite in orbit 1400 km above Earth's surface and have it stay in orbit for 50,000 years. The project was the brainchild of French artist Jean-Marc Phillipe (1939–2008), who came up with the idea in 1994. Phillipe hoped to send a message to our descendants, just as the cave artists of Lascaux sent a message to us. The information was to be encoded on radiation-resistant DVDs, and there would be symbolic instructions in several formats to show any future finders how to build a suitable reader. The current planned launch date is 2015, although at the time of writing it's far from clear that this will be achieved (the launch was initially planned for 2003, but has been delayed several times). See KEO (2014).

[56]**Pg 46 best known are the Lagrangian points** The Italian–French mathematician Joseph-Louis Lagrange (1736–1813) was undoubtedly one of the greatest mathematicians of the 18$^{th}$ century. Perhaps his most important astronomical investigations concerned calculations of the libration of the Moon and of the orbits of the planets. For a brief biography of Lagrange, see Rouse Ball (1908).

[57]**Pg 48 not provide the stable vantage point** Lissauer and Chambers (2008) ran a series of numerical simulations that showed how the gravitational influence of the planets, when combined with the much larger influence of the Sun, are enough to destabilize the orbits on a timescale of a few million years.

[58]**Pg 48 A more prosaic explanation** An explanation of LDEs was given by Lawton and Newton (1974). Their paper responded to the hypothesis presented at length by Lunan (1974) that LDEs were evidence of ETC probes at L4 or L5. See Faizullin (2010) for a different take on the issue.

[59]**Pg 48 long been thought to be home to life** For an excellent account of observations of Mars, see Sheehan (1996).

[60]**Pg 49 in a series of observations beginning in 1877** The Italian astronomer Giovanni Virginio Schiaparelli (1835–1910), director of the observatory at the Brera Palace in Milan, made important observations of meteors and comets before turning his attention to the planets. He wasn't the first to record channels on Mars; the first true map of Mars, published in 1830 by the German astronomers Wilhelm Beer (1797–1850) and Johann Heinrich von Mädler (1794–1874), contains at least one feature that seems to be a channel. Nevertheless, Schiaparelli so popularized the idea of *canali* that they became the defining theme of Mars. Perhaps the most famous of the stories that tapped into the public's subsequent fascination with the red planet was *War of the Worlds* (Wells 1898) by English author Herbert George Wells (1866–1946).

[61]**Pg 49 Lowell also saw** Percival Lowell (1855–1916) came from a wealthy Boston family and only took up astronomy in earnest at the relatively late age of 40. He achieved a lot in science, despite his late start: he had the determination to initiate the search for a planet beyond Neptune, and the Lowell Observatory in Arizona is named after him. However, he'll always be associated with his ideas concerning Mars. For an interesting article about Lowell, see Zahnle (2001).

[62]**Pg 49 in the early 1960s** The Ukrainian astrophysicist Josif Samuelevich Shklovsky (1916–1985) is best known for his explanation of continuum radiation from the Crab Nebula, but he also made important contributions in cosmic ray astronomy and on the distance scale for planetary nebulae. His popular book *Intelligent Life in the Universe*, which Carl Sagan had translated from the Russian and then expanded upon, is a classic in the field (Shklovsky and Sagan 1966). The American astronomer Bevan P. Sharpless (1904–1950), on whose observations Shklovsky based his suggestion regarding Phobos, worked at the US Naval Observatory; poor health hampered his work throughout his career and he died early. The fifth largest crater on Phobos is named after him.

[63]**Pg 49 Salisbury pointed out** The German-born astronomer Heinrich Louis d'Arrest (1822–1875), who became director of the Copenhagen Observatory, mounted a thorough search for Martian moons in 1862. However, it was the American astronomer Asaph Hall (1829–1907) who discovered the moons in 1877 (see Sheehan 1996 for further details). The reason Hall found them and d'Arrest did not is simple: the Martian satellites are much closer to the planet than d'Arrest thought possible. Hall looked in the right place; d'Arrest did not. Thus, the suggestion by American biologist Frank Boyer Salisbury (1926–) that Phobos and Deimos were artificial satellites launched between 1862–1877 is unnecessary.

[64]**Pg 53 but no human face** The Cydonian "face" was first pointed out in 1977 by American electrical engineer Vincent DiPietro. The view that the face is artificial has been championed most strongly by the American writer Richard C. Hoagland (1945–). See, for example, Hoagland (1987). See Hancock et al. (1998) for another book in similar vein. For a refreshingly sane article about the face, see Gardner (1985).

[65]**Pg 54 Papagiannis argued** The Greek–American astronomer Michael Demetrius Papagiannis (1932–1998) was the first president of the International Astronomical Union's commission on bioastronomy. See Papagiannis (1978) for his suggestion regarding hiding places for colonies in the Asteroid Belt. Kecskes (2002) offers reasons why humanity might end up as "asteroid dwellers". Is this another solution to the paradox: ETCs choose to colonize not space, which is difficult, but their home system's Asteroid Belt?

[66]**Pg 54 mine the asteroids for natural resources** There has been discussion about the possibility of mining the asteroids for various minerals; however, it might turn out that such activity is prohibitively expensive. See Elvis (2014).

[67]**Pg 54 the result of an astroengineering project** See Stephenson (1978).

[68]**Pg 55 Loeb and Turner showed** See Loeb and Turner (2012) for a discussion of how it would be possible to search for artificially illuminated objects in the outer Solar System.

[69]**Pg 55 a professor of electrical engineering** The first paper to calculate the minimum distance for the Sun's gravitational lens was von Eshleman (1979).

[70]**Pg 56 Maccone, who perhaps more than anyone** For more on the possibility of exploiting the Sun as a gravitational lens, see Maccone (1994, 2000, 2009, 2011, 2013) and Maccone and Piantà (1997).

[71]**Pg 56 the Belgian astrophysicist Michaël Gillon** For details of the argument that SETI could do worse than focusing on the solar focus, see Gillon (2014).

[72]**Pg 56 spectacular observatories** In Webb (2012) I give an account of the many observatories that have recently come on line or are in the planning stage.

[73]**Pg 57 can't rule out the possibility** See Haqq-Misra and Kopparapu (2012) for an in-depth discussion of why it's difficult to assert that there are no small (say, 1–10 m) probes in the Solar System. They argue that searching the Solar System at the spatial resolution required to detect a 1–10 m probe is analogous to searching for a needle in a 1000-ton haystack.

[74]**Pg 57 has yet been uncovered** See Freitas (1983, 1985).

[75]**Pg 58 embed some sort of signal** See shCherbak and Makukov (2013) for the claim that a signal is embedded in the terrestrial genetic code.

[76]**Pg 58 a few investigations have been performed** See Yokoo and Oshima (1979).

## They Exist and They Are Us—We Are All Aliens!

[77]**Pg 59 dates back to Anaxagoras.** Anaxagoras (c. 500–428 BC), one of the greatest of Greek philosophers and the teacher of Socrates, spoke of the "seeds of life" from which spring all organisms. See O'Leary (2008).

[78]**Pg 59 a book by Arrhenius** The Swedish chemist Svante August Arrhenius (1859–1927) is best known as the man who helped lay the foundations of modern physical chemistry. His book *Worlds in the Making* popularized the notion that life on Earth might have arrived from space. See Arrhenius (1908).

[79]**Pg 60 mass outbreaks of disease** The astronomers Fred Hoyle (1915–2001) and Nalin Chandra Wickramasinghe (1939–) have made exceptional contributions to science, but they have also proposed several hypotheses that go against received wisdom. This is one such hypothesis. Nevertheless Hoyle, Wickramasinghe and collaborators have published widely on the subject. See for example Hoyle and Wickramasinghe (2000) and references therein. The physicist Thomas Gold (1920–2004) was another scientist who liked to propose unorthodox ideas. He jokingly proposed the "garbage" scenario for the origin of terrestrial life: ETCs landed here, dumped their waste, and the contamination from the garbage was the seed for life!

[80]**Pg 60 the ability of some extremophiles** Calculations tend to suggest that life would struggle to survive the radiation environment found in space; see for example Secker, Wesson and Lepock (1996). Nevertheless, Lage (2012) demonstrates the remarkable capacities for survival of extremophiles in conditions that attempt to simulate those found in the space environment.

[81]**Pg 60 inactivated virus-like organisms** See Wesson (2010) for the interesting notion of necropanspermia.

[82]**Pg 60 *directed* panspermia** See Crick and Orgel (1973) and Crick (1981). The English biophysicist Francis Harry Compton Crick (1916–2004) gained fame for his discovery, along with the American biochemist James Dewey Watson (1928–), of the double-helix structure of DNA. The English-born biochemist Leslie Eleazer Orgel (1927–2007) made major contributions to the study of life's origins. The Crick–Orgel idea of directed panspermia originated at the first conference on communication with extraterrestrial intelligence, organized in 1971 by Sagan and Kardashev, and held at the Byurakan Astrophysical Observatory in Armenia. Many of the luminaries in the field of SETI attended this conference.

## The Zoo Scenario

[83]**Pg 61 zoo scenario was proposed** The American astronomer John Allen Ball (1935–) has written extensively on the Fermi paradox. For the zoo hypothesis, see Ball (1973).

[84]**Pg 61 be in control of the universe** Hair (2011) argues that if the oldest civilization still present in the Galaxy had a hundred million year "head start"

on the next oldest civilization then they could have established a hegemony that guides the development of younger civilizations; in that case, he suggests, a modified zoo scenario is an appealing answer to the Fermi paradox. See Forgan (2011) for a criticism of the idea that a total hegemony could be established that would allow the zoo scenario to occur.

[85]**Pg 62 editorship of John Campbell** Asimov's famous "humans-only" Galaxy was a reaction against Campbell's insistence that humans should always win out against aliens. Asimov thought that human civilization would be less advanced than any extraterrestrial civilizations we might encounter, and he couldn't bring himself to write stories in which primitive Earth technology triumphed over superior alien technology (see Asimov 1979). On the other hand, he wanted to sell stories to Campbell. He therefore removed the potential source of conflict, and his *Foundation* trilogy described a Galaxy containing only humans. If the Fermi paradox implies that we are alone, then perhaps an empire something like Asimov reluctantly described will come to pass.

[86]**Pg 62 to slowly prepare us** The leaky embargo hypothesis was proposed by James Warner Deardorff (1928–2014), a retired atmospheric physicist; for details of the proposal see Deardorff (1986, 1987). Although Deardorff had a scientific background, his leaky embargo hypothesis is unscientific. For a nice introduction to scientific method, which uses Deardorff's hypothesis as an example to be critiqued, see Carey (1997).

## The Interdict Scenario

[87]**Pg 63 expanded form of the zoo scenario** See Fogg (1987) for the original presentation of the interdict hypothesis; Fogg (1988) is a more popular account. Martyn J. Fogg (1960–) originally trained as a dentist. He's now one of the foremost authors on "speculative" engineering techniques, such as terraforming.

[88]**Pg 64 Asimov pointed out** See Asimov (1981) for a dated but still readable introduction to the subject. Asimov was an optimist and argued that half a million planets in our Galaxy are home to technological civilizations.

[89]**Pg 64 a *Codex Galactica* is established** The notion of a *Codex Galactica* is discussed in Newman and Sagan (1981); note, however, that this is yet another idea that appeared in the pages of SF magazines before gaining respectability in the pages of a refereed journal.

## The Planetarium Hypothesis

[90]**Pg 66 Baxter has proposed** The British writer Stephen Baxter (1957–) is known for his "hard" science fiction. For details of his planetarium hypothesis, see Baxter (2000a).

[91]**Pg 67 a fake town** Many examples exist of this paranoid trope in SF. The earliest such story of which I'm aware is "The Earth-Owners" by Edmond Hamilton (1904–1977), which describes an Earth invaded by aliens in disguise; the aliens, of course, are busy manipulating us. Hamilton's story appeared in the August 1931 issue of *Weird Tales*. Historians of science fiction could doubtless point to even earlier examples. The Asimov story was "Ideas Die Hard" (*Galaxy*, October 1957). Weiner's "The News from D Street" appeared in the September 1986 issue of *IASFM*. The philosophical considerations underpinning the planetarium hypothesis are well discussed in Deutsch (1998); see also Tipler (1994).

[92]**Pg 69 Bekenstein showed** The Bekenstein bound is named after the Mexican-born US–Israeli physicist Jacob David Bekenstein (1947–), who introduced the concept in terms of the thermodynamics of black holes.

[93]**Pg 70 most readers would wager is the case** The idea that our universe is a simulation is being debated quite seriously by heavyweight philosophers, so perhaps we shouldn't be too quick to discount the idea. See for example Bostrom (2003) and Bostrom and Kulczycki (2011). A physics paper that takes the proposition seriously (Beane et al. 2012) concludes that in principle there'll always be the possibility for the simulated to discover the simulators.

## God Exists

[94]**Pg 70 think of them as gods** A haunting short story called the "The Last Question" (see Asimov 1959) tells how a pair of drunken technicians one night ask a supercomputer whether there is a way to reverse the increase of entropy and thereby halt the death of the universe. The computer says there is insufficient data for a meaningful answer. The same question is asked of the computer six times over many different epochs. I won't spoil the story by telling you the computer's final answer!

[95]**Pg 72 evolutionary ideas to cosmology** See Smolin (1997) for a discussion of why we might want to apply Darwinian thinking to the problem of the universe as a whole.

[96]**Pg 74 a specific forecast** The Austrian–British philosopher Karl Raimund Popper (1902–1994) propounded the notion that scientific hypotheses must be falsifiable. The drive to falsify hypotheses is the essence of science. If an hypothesis cannot be tested and perhaps found to be false, then it isn't a valid part of the process of science. See for example Popper (1963). Although his views about scientific progress have been attacked, they remain influential. Smolin's idea is certainly falsifiable, since it makes specific testable predictions; the novelty is that it must be tested by calculation rather than experiment.

[97]**Pg 74 speculation one step further** See Harrison (1995). Byl (1996) criticizes Harrison's speculation as being post hoc, unverifiable and essentially a more elaborate version of the theistic or anthropic principles. For further reading about the notion of a multiverse see Gribbin (2010) for a popular account and Carr (2007) for more technical aspects. See Vaidya (2007) for a mention of the Fermi paradox in the multiverse setting.

## They Exist, But We Have Yet to See or Hear from Them

[98]**Pg 77 the aestivation hypothesis** At the time of writing, details are only available as a preprint (Sandberg et al. 2014).

## The Stars Are Far Away

[99]**Pg 79 Voyager will take** For information about Voyagers 1 and 2 see Voyager (2013). For useful material on several of the advanced propulsion concepts discussed in this section see NASA (2013).

[100]**Pg 79 the speed of light** According to the theory of special relativity, massless objects such as photons always travel at light speed *c*, while objects with non-zero mass inevitably travel more slowly. Of course, it's possible to accelerate a slow-moving body to a faster speed by acting upon it with a force. Unfortunately for the prospects for space travel, special relativity tells us that the faster things move the more massive they become. At speeds close to *c*, the accelerating force tends to make the body more massive rather than make it move faster. The speed of light is a barrier that can't be reached by any object with mass—including space ships. For a good introduction to these concepts, see French (1968).

[101]**Pg 79 nearest star to our Sun** See Webb (1999) for an in-depth discussion of astronomical distances.

[102]**Pg 79 Bernal proposed the idea** John Desmond Bernal (1901–1971), an Irish physicist, published the idea of a generation ship in a visionary book (see Bernal 1929). His book contains the following quote, which is relevant to any discussion of the Fermi paradox. "Once acclimatized to space living, it is unlikely that man will stop until he has roamed over and colonized most of the sidereal universe, or that even this will be the end. Man will not ultimately be content to be parasitic on the stars but will invade them and organize them for his own purposes." For "man" read "ETC". So—where *are* they?

[103]**Pg 79 Heinlein's story *Universe*** The short novel *Universe*, written by the American author Robert Anson Heinlein (1907–1988), appeared in the May 1941 issue of *Astounding Science Fiction*. (It can be found more easily in Bova (1973).) The story is one of many SF classics penned by Heinlein.

[104]**Pg 80 things that we can learn** Crawford (2009) makes the science case for interstellar spaceflight. There's only so much one can learn by telescopic observation. In order to make progress in astronomy, astrobiology and planetary science, there's a strong argument that we must develop interstellar spaceflight.

[105]**Pg 80 possible within a human lifetime** This possibility was dramatized by the American writer Poul William Anderson (1926–2001) in his novel *Tau Zero* (Anderson 2000). The novel tells the story of a ramjet that accelerates to speeds so close to *c* that circumnavigation of the universe becomes possible.

[106]**Pg 82 might be able to detect them** For a possible addition to the SETI search strategy, see Garcia-Escartin and Chamorro-Posada (2013). The authors suggest that we should look for reflected light from objects traveling at relativistic speeds.

[107]**Pg 82 a navigation problem** For an interesting discussion of the problems inherent in navigating to a particular star, see Hemry (2000).

[108]**Pg 82 by Eugen Sänger** In addition to as conceiving the idea of an antimatter rocket, the Austrian scientist Eugen Sänger (1905–1964) pioneered several practical ideas in rocketry. For superb introductions to many different proposals for interstellar travel, see Mallove and Matloff (1989) and Crawford (1995).

[109]**Pg 83 a *fusion ramjet*** Bussard's idea for the ramjet appeared over half a century ago (Bussard 1960). Since then, various authors have made proposals and suggestions for the improvement of the initial ramjet design.

[110]**Pg 83 Forward began to consider** Robert Lull Forward (1932–2002), as with many of the scientists mentioned in this book, was also a successful SF writer. For a technical discussion of the laser sail, and how it might be used in a round-trip interstellar mission, see Forward (1984).

[111]**Pg 83 have designed schemes** See Dyson (1982) for a discussion of how laser sails could be used in colonization methods; see Wright (1992) for a general discussion of space sailing.

[112]**Pg 83 sail would be expensive** For a discussion of the costs and required technologies associated with different types of sail, see Andrews (2004).

[113]**Pg 84 a gigantic, massive mirror** Shkadov (1987) introduced the thruster idea. See Forgan (2013) for how we might detect the use of a Shkadov thruster by an ETC. Benford and Niven (2012) give a fictional account of a star thruster.

[114]**Pg 84 near light speed** Stanislaw Marcin Ulam (1909–1984), a Polish-born mathematician, contributed to several fields. His autobiography (Ulam 1976) is fascinating. (Ulam appears in fig. 4.9 on page 116.) The English-born physicist Freeman John Dyson (1923–) is one of the most imaginative physicists of his generation, and has contributed to many topics mentioned in this book. For the papers on gravitational propulsion, see Ulam (1958a) and Dyson (1963).

[115]**Pg 85 negative mass** For a discussion of negative mass, see Forward (1990).

[116]**Pg 85 no evidence such particles exist** In September 2011, the OPERA experiment shocked physicists by announcing they had observed muon neutrinos traveling faster than *c* (OPERA Collaboration 2011). A few months later they retracted their claim, stating the original results were affected by equipment failures.

[117]**Pg 87 novel *Contact*** Carl Edward Sagan (1934–1996) based the science in his novel *Contact* (Sagan 1985) on work by the American theoretician Kip

Stephen Thorne (1940–) who has been prominent in investigating the properties of wormholes. (For a popular account of this work, see Thorne (1994).) In 1997, Sagan's novel was made into a movie of the same name, starring Jodie Foster.

[118]**Pg 87 a certain class of wormhole** For details of the Krasnikov tube, see Krasnikov (1998).

[119]**Pg 88 surfs a spacetime wave** Miguel Alcubierre Moya (1964–), a Mexican theoretical physicist, is now Director of the Nuclear Sciences Institute at the National Autonomous University of Mexico. See Alcubierre (1994) for his paper describing the warp drive.

[120]**Pg 88 unrealistic features** For details on the possibility of using wormholes for transport, see Krasnikov (2000). For details on Van Den Broeck's warp drive, see Van Den Broeck (1999). These matters have been covered in detail, and at a non-mathematical level, in John Cramer's "Alternate View" columns in *Analog*.

[121]**Pg 89 The Casimir effect** In 1948, the Dutch physicist Hendrik Brugt Gerhard Casimir (1909–2000) predicted that quantum fluctuations of the EM field would cause a small attractive force to act between two close parallel uncharged conducting plates. The first measurement of the Casimir force between parallel plates took place in 2002 (see Bressi et al. 2002). The experiment confirmed Casimir's predictions. For articles propounding the idea that mankind might one day mine the zero-point energy see, for example, Haisch et al. (1994) and Puthoff (1996).

[122]**Pg 89 safely to Saturn and back** It might be that the future of human exploration of the Solar System over the next few decades lies in a combination of human and robotic elements. For example, having humans land on Enceladus, a moon of Saturn that provokes interest for a variety of reasons, would be risky and costly. Perhaps a better bet would be to have astronauts orbit Enceladus while they used teleoperation to control rovers and robots on the surface. See Schmidt et al. (2012).

## They Have Not Had Time to Reach Us

[123]**Pg 90 temporal explanation of the paradox** One of the first responses to Hart's paper was by Cox (1976). Cox argued that a temporal explanation of the paradox is indeed valid.

[124]**Pg 91 Several authors have developed** See for example Jones (1975, 1981). In Jones (1995) the author has written a particularly entertaining discussion of various colonization processes, from past human expansions through to possible human settlement of the Solar System and nearby stars. See also Finney and Jones (1985).

[125]**Pg 91 demands of population growth** See Newman and Sagan (1981).

[126]**Pg 92 detailed model of galactic exploration** See Bjørk (2007) for details of his exploration algorithm.

[127]**Pg 92 expanded on Bjørk's model** See Cotta and Morales (2009).

[128]**Pg 93 models have been analyzed** See Crawford (2000) for a well-written account of galactic colonization models and their relation to the Fermi paradox. See Fogg (1987) for details of one particular model of galactic colonization.

[129]**Pg 93 Prantzos reinforced this conclusion** See Prantzos (2013) for an interesting framework in which to think about the Fermi paradox.

## A Percolation Theory Approach

[130]**Pg 93 bases his model** Geoffrey Alan Landis (1955–), an American physicist who works at NASA, is yet another scientist who is perhaps better known as an SF writer. For details of his approach, see Landis (1998).

[131]**Pg 94 key task in a percolation problem** Percolation theory was developed in 1957 by the British mathematician John Michael Hammersley (1920–2004) and his colleagues. See Stauffer (1985) for the best introduction to the ideas of percolation theory; however, although this excellent book is entertaining reading, readers should be aware that it inevitably contains an element of mathematics.

[132]**Pg 97 conclusion is similar** See Kinouch (2001) for details of the "persistence solution" to the Fermi paradox.

[133]**Pg 97 an economist's point of view** See Hanson (1998) for an interesting model of colonization. To fully appreciate the argument requires some mathematics, but the conclusions are clearly expressed in layman's terms. See also Bainbridge (1984).

[134]**Pg 98 simple extensions of the model** See Wiley (2011) for a detailed critique of the percolation model, as well as various other colonization models.

## Wait a Moment

[135]**Pg 99 Martin Gardner popularized** The *Game of Life* was devised by the British mathematician John Horton Conway (1937–) as an offshoot of his thinking on von Neumann's attempt to construct a mathematical model of a self-replicating machine. The game became an immediate hit amongst the public when Martin Gardner (1914–2010) discussed it in his "Mathematical Games" column in *Scientific American* (Gardner 1970).

[136]**Pg 100 Fermi was the first to dabble** See Metropolis (1987) for the early history of the Monte Carlo method, including Fermi's early experimentation and Ulam's.

[137]**Pg 101 to the study of the Fermi paradox** See for example Forgan (2009).

[138]**Pg 101 sophisticated cellular automata** See Vukotić and Ćirković (2012).

## The Light Cage Limit

[139]**Pg 102 a model of migration** See McInnes (2002) for a discussion of how ETCs could be hindered by the light cage limit. The basic idea was briefly mentioned much earlier by von Hoerner (1975).

[140]**Pg 103 Baxter calls this radius** See Baxter (2000b) for an interesting fictional take on one possible solution to the Fermi paradox.

## They Change Their Mind

[141]**Pg 104 planets must be re-engineered** Fogg (1995) is perhaps the most comprehensive resource on terraforming, and how it might be possible to engineer a planet so that it becomes suitable for life.

[142]**Pg 104 some simple equations** See Gros (2005) for details of rate equations that govern the population dynamics of civilizations that are assumed to be able to change character and priorities.

## We Are Solar Chauvinists

[143]**Pg 105 simply inapplicable** This resolution to the Fermi paradox was discussed in Rood and Trefil (1981), a book that is now sadly out of print.

[144]**Pg 105 encloses the star** The concept of the Dyson sphere first appeared in Dyson (1960). (A Dyson sphere is a *loose* collection of bodies moving on independent orbits around a star; a rigid sphere would be unstable.) The idea inspired two great SF novels: *Ringworld* (Niven 1970) and *Orbitsville* (Shaw 1975). Scientists have suggested numerous other mega-engineering projects that technologically advanced ETCs might embark upon. For example, Roy et al. (2013) discuss the possibility of "shell worlds". A shell world is formed by enclosing an airless, sterile body by a shell of material to create a cozy home for life.

[145]**Pg 106 a better analogy for colonization** Kecskes (1998, 2002) outlines a possible "trajectory" for the development of technical civilizations: they move from being planet dwellers to asteroid dwellers to interstellar travelers to interstellar space dwellers. In this picture we don't meet extraterrestrials because our habitats are different.

## Aliens Are Green

[146]**Pg 108 Haqq-Misra and Baum propose** See Haqq-Misra and Baum (2009) for a discussion of the "sustainability solution" to the Fermi paradox.

## They Stay at Home . . .

[147]**Pg 110 happened on 20 July 1969** The American astronauts Neil Alden Armstrong (1930–2012) and Edwin Eugene Aldrin Jr. (1930–) landed at the edge of Mare Tranquillitatis on 20 July 1969; Armstrong walked on the Moon at 22:56 (Eastern Daylight Time). The *last* man to walk on the Moon was Eugene Andrew Cernan (1934–), and unfortunately he seems set to hold this honor for quite some time to come. Cernan recounts his experiences of the Apollo program in Cernan and Davis (1999). Smith (2005) is an evocative account of the Apollo era.

[148]**Pg 111 China expanded her empire** The two emperors mentioned in the text were Hongwu (1328–1398) and Yongle (1359–1424). The incredible voyages of the admiral Zheng He (c. 1371–c. 1435), a court eunuch and diplomat, have only relatively recently come to light. For a readable account of the seven epic voyages made by Zheng He, see Levathes (1997).

[149]**Pg 111 cause us serious problems** "Inconstant Moon", one of the finest stories by the American author Laurance (Larry) van Cott Niven (1938–), describes the events of a single night when the full moon shines brighter than ever before. It's a gem of a tale, and deservedly won the 1972 Hugo award for best short story; it's available in Niven (1973).

[150]**Pg 111 Zuckerman has shown** See Zuckerman (1985).

## . . . and Surf the Net

[151]**Pg 113 plausible future for humankind** Set a billion years in the future, *The City and the Stars* (Clarke 1956) imparts a sense of wonder and magnificent scope few novels can match. In the novel Arthur Clarke presents at least two explanations of the Fermi paradox, including the notion that beings might prefer to stay in the "City"—safe from facing the realities of a harsh universe.

## Against the Empire

[152]**Pg 114 Ćirković points out** See Ćirković (2008) and references therein.

[153]**Pg 115 already being debated** See for example Rummel (2001) for thoughts about the problems of contamination when we engage in planetary exploration.

[154]**Pg 115 Bostrom's term** For a definition of the term "singleton", see Bostrom (2005). See also Caplan (2008) for a discussion of the problems with singletons.

## Bracewell–von Neumann Probes

[155]**Pg 117 As long ago as 1980** For an early discussion of interstellar exploration by probe, see Freitas (1980). As mentioned in chapter 2, the relevance of self-reproducing probe technology to the Fermi paradox was considered by Tipler (1980). One could argue that the starting point for the discussion is even earlier, with Crick's motto for directed panspermia (see page 60): "bacteria go further". Crick and Orgel argued that a small probe filled with a payload of bacteria would be easy to construct, cheap to propel, and would enable an ETC to seed the Galaxy. However, a bacteria-filled probe is of little use to an ETC wanting to explore and learn about the Galaxy. To be successful in that endeavor, a Bracewell–von Neumann probe would be better.

[156]**Pg 118 The first person to suggest** The Australian-born electrical engineer Ronald Newbold Bracewell (1921–2007) was for many years a leading light in SETI. See Bracewell (1960).

[157]**Pg 118 more recent investigations** See for example Forgan et al. (2013) and Nicholson and Forgan (2013) for discussions of how the judicious use of the slingshot effect could reduce the time for galactic exploration by probe; in particular, if self-reproducing probes make use of the slingshot effect then colonization times can be similar to that calculated by Tipler. See Barlow (2013) for yet another analysis of galactic colonization in the context of Bracewell–von Neumann probes. Cartin (2013) discusses a different approach to colonization, which doesn't involve self-reproducing probes.

[158]**Pg 119 probes for interplanetary exploration and exploitation** Mathews (2011) argues that probes are a natural extension of our planetary explorer craft. We'll send out robots, not humans, to explore the Solar System. Perhaps the development of this technology will lead us on a path to the sort of self-reproducing probes discussed in the text.

[159]**Pg 119 not exactly a risk-free technology** For a criticism of galactic exploration via Bracewell-von Neumann probes, and why it might not work, see Chyba and Hand (2005). Wiley (2011), however, concludes that criticisms of the self-reproducing probe approach to galactic colonization possess little merit.

[160]**Pg 120 significantly sharpened the paradox** See Armstrong and Sandberg (2013).

## Information Panspermia

[161]**Pg 121 an interesting hypothesis** For details of the argument that the universe might be full of low-complexity bit strings, see Gurzadyan (2005). See Scheffer (1993) for an earlier and thorough defense of the notion that "information transfer" is a much cheaper option for interstellar travel than physical travel. Scheffer resolves the Fermi paradox by arguing that the first civilization to colonize its galaxy will have done all the hard work; for any emerging society it will be overwhelmingly attractive to join the existing civilization rather than try to physically colonize the galaxy. There will be a single, unified civilization. If that first civilization in our Galaxy didn't bother to contact Earth, for whatever reason, then subsequent societies won't have bothered either.

[162]**Pg 121 *Kolmogorov complexity*** The idea that a measure of the complexity of a system can be the length of an algorithm that produces that system is due to Andreii Nikolaevich Kolmogorov (1903–1987), who was one of the outstanding mathematicians of the twentieth century. For an appreciation of just some of Kolmogorov's output, see for example Parthasarathy (1988).

[163]**Pg 122 our extinct cousins, the Neanderthal** In December 2013, scientists published a high-quality genome sequence from a Neanderthal woman who lived 130,000 years ago in what is now Siberia. The DNA came from one of her toe bones. See Prüfer et al. (2013).

## Berserkers

[164]**Pg 123 famous berserker stories** The American author Fred Thomas Saberhagen (1930–2007) wrote many stories about berserkers, with the first

collection appearing in *Berserker* (Saberhagen 1967). The concept of a Doomsday weapon was brilliantly satirized by Stanley Kubrick in *Dr. Strangelove*, and the original *Star Trek* television series aired an episode called *The Doomsday Machine*, which dramatized the notion of an indestructible world-killing machine (though Kirk & Co. managed to destroy it, of course). The machine in *Star Trek* was a single, large, slow-moving object. My mental picture of berserkers is somewhat different: I imagine swarms of small, fast-moving machines. A novel entitled *The Unreasoning Mask*, by the American author Philip José Farmer (1918–2009), is another that treats the notion of world-killers (Farmer 1981). But perhaps the idea of malignant killing machines has been treated most thoroughly by the American astrophysicist Gregory Benford (1941–), who is also one of the finest modern SF writers; see, for example, Benford (1977).

## They Are Signaling but We Don't Know How to Listen

[165]**Pg 126 no plausible signatures** See Jugaku and Nishimura (1991). They continued their search of the solar neighborhood, but failed to find any candidates; see Jugaku and Nishimura (1997, 2000)

[166]**Pg 126 found nothing unusual** See Mauersberger et al. (1996)

[167]**Pg 126 Carrigan carried out** See Carrigan (2009). For an entertaining essay on whether interstellar archaeology is possible, see Carrigan (2010, 2012).

[168]**Pg 126 Wright and his colleagues** For a discussion of the G-HAT search for Kardashev civilizations see for example Battersby (2013).

[169]**Pg 127 Minsky pointed out** It was at the seminal Byurakan conference on communication with extraterrestrial intelligence that the American computer scientist Marvin Lee Minsky (1927–) pointed out that a truly advanced energy-conscious ETC might radiate at a temperature just above the cosmic background. See Minsky (1973).

[170]**Pg 127 beacons can be transmitted** Whitmire and Wright (1980) was not the first paper to suggest the stars themselves could be used to send signals. Philip Morrison (1915–2005) suggested the "eclipse" method 20 years earlier, and Drake had made similar suggestions before. But their paper is perhaps

the first to give detailed calculations of how to modify stellar spectra to send a signal.

[171]**Pg 127 rule out a natural phenomenon** See page 245 of Sullivan (1964). See also Arnold (2013).

[172]**Pg 128 brainchild of Ray Davis** The American chemist Raymond Davis Jr. (1914–2006) ran his solar neutrino experiment for more than three decades, and was awarded the 2002 Nobel prize for his research. See Bahcall and Davis (2000) for the early history of neutrino astronomy.

[173]**Pg 129 neutrino beams to communicate** For discussions of neutrino-based searches for extraterrestrial intelligence, see for example Learned et al. (1994), Silagadze (2008) and Learned et al. (2009).

[174]**Pg 130 problem of detecting gravitational waves** Einstein's theory of general relativity predicted the existence of gravitational waves—ripples in spacetime. Such waves were demonstrated indirectly by the American physicists Joseph Hooten Taylor Jr. (1941–) and Russell Alan Hulse (1950–) through exquisitely accurate observations of PSR 1913+16. This pulsar is part of a binary system, its partner being another neutron star. As the two stars orbit each other, they lose energy in precisely the manner predicted by general relativity: the binary system is radiating gravitational energy in the form of waves. See Weisberg and Taylor (2005) for more information. The current generation of detectors is typified by LIGO (Laser Interferometer Gravitational-wave Observatory). If LIGO doesn't observe gravitational waves then astronomers will pin their hopes on the next generation of detectors, of which the Einstein Observatory is perhaps most advanced.

## They Are Signaling but We Don't Know at Which Frequency to Listen

[175]**Pg 133 first to consider this question** The Italian physicist Giuseppe Cocconi (1914–2008) worked at Cornell University with Morrison before returning to Europe to work at CERN, where he eventually became Director. Their paper (Cocconi and Morrison 1959) is one of the classics in SETI.

[176]**Pg 134 it will send a *narrowband signal*** Although there are good reasons for concentrating on narrowband signals, increasing attention is being paid to

the possibility of wideband signals. The search for wideband signals is much more challenging than the search for narrowband signals; on the other hand, a wideband beacon can carry vastly more information than a narrowband beacon. For more information on wideband SETI see, for example, papers by Benford, Benford and Benford (2010a, b); Harp et al. (2011); Messerschmitt (2012); Morrison (2012).

[177]**Pg 137 study other frequencies** For suggestions of some other likely SETI frequencies see Kardashev (1979), Mauersberger et al. (1996) and Kuiper and Morris (1977).

[178]**Pg 138 a new search strategy** Hair (2013) considers some of the difficulties in applying statistical techniques to any "long stare" strategy that hopes to construct an archive of provocative radio transients.

[179]**Pg 138 from some unknown terrestrial source** See Gray (2011) for an entertaining and in-depth discussion of the "Wow!" signal, and one man's attempt to understand it better.

[180]**Pg 139 increased in sophistication over time** See Tartar (2001) and Bowyer (2011) for more background on SETI projects.

[181]**Pg 139 developed in 1985 by Paul Horowitz** Paul Horowitz (1942–), a Harvard astronomer, has been at the forefront of SETI research for several years. Much of the funding for META came from Steven Spielberg (1947–), the director of the film *E.T. the Extra-Terrestrial*. See Lazio, Tarter and Backus (2002) for a discussion of Project META

[182]**Pg 139 piggybacks on radio telescopes** The idea for SERENDIP originated with the American astronomers C. Stuart Bowyer (1934–) and Jill Tarter (1944–) in 1978. Tarter, who in 2012 announced her retirement from the position as director of research at the SETI Institute, is an icon in the field. She is widely believed to have been the inspiration for Sagan's heroine in *Contact*. See for example Korpela et al. (2011) for further information about SERENDIP and other SETI-related projects.

[183]**Pg 139 has great potential** For background on and papers about the Allen Telescope Array see, for example, Welch et al. (2009), Siemion et al. (2010) and Tarter et al. (2011).

[184]**Pg 139 also play a role** For contrasting viewpoints on how the SKA might be relevant to SETI see, for example, Penny (2004), Loeb and Zaldarriaga (2007), Forgan and Nichol (2011), Rampadarath et al. (2012).

[185]**Pg 139 Optical SETI is not as advanced** The slow uptake of OSETI is perhaps due to the relative novelty of the technology. Credit for the invention of the laser is a matter of some dispute (see, for example, Hecht (2010). The American physicists Arthur Leonard Schawlow (1921–1999) and Charles Hard Townes (1915–2015) were both awarded the Nobel prize for laser-related work (Townes in 1964 and Schawlow in 1981). Townes was far-seeing in regard to the potential of lasers. The suggestion that SETI should consider optical searches is almost as old as the Cocconi–Morrison paper: see Schwartz and Townes (1961).

[186]**Pg 140 starting to develop large-scale projects** For two early examples of optical searches, see Eichler and Beskin (2001) and Reines and Marcy (2002). See Korpela et al. (2011) for further details of Project SEVENDIP.

[187]**Pg 140 Ball once hypothesized** See Ball (1995).

[188]**Pg 140 play the role of "synchronizers"** See Corbet (1999) for a discussion of the role that gamma-ray bursts might play in synchronizing signals; essentially, they would act as universal timing markers.

[189]**Pg 141 Local Group of galaxies** See LePage (2000).

## They Are Signaling but We Don't Know Where to Look

[190]**Pg 142 taken this approach** See Turnbull and Tarter (2003a, b) for details of the *Hipparcos* habstars.

[191]**Pg 142 judged to be most amenable** Siemion et al. (2013) discuss a targeted search of 86 *Kepler* objects of interest; they looked for radio emission from ETCs, but found none.

[192]**Pg 142 more readily discovered** See Nussinov (2009) for an interesting suggestion about preferred directions for SETI.

[193]**Pg 142 straight-line alignments** For details of this suggestion, and for one way in which pulsars might be used as beacons, see Edmondson and Stevens (2003) and Edmondson (2010).

[194]**Pg 143 as many stars as possible** See Hohlfeld and Cohen (2000) and Cohen and Hohlfeld (2001).

[195]**Pg 144 a universal frequency** The "universal" frequency standard was first discussed by Drake and Sagan (1973). See also Gott (1995).

## The Signal is Already There in the Data

[196]**Pg 145 logged several pulses** From a total of about 60 trillion events, META researchers found only 11 good candidate signals. If these signals were really attempts at communication, however, why could astronomers not observe them again? One suggestion was that interstellar plasmas or gravitational microlenses, passing between the sources and Earth, caused what were steady beacon-like signals to "twinkle"—and temporarily become strong enough for us to detect. A detailed analysis of the data ruled out this possibility, however, and the result seemed to indicate that the Galaxy contains at most one other civilization with a comparable level of technology to ours that is deliberately trying to contact us. See Lazio, Tarter and Backus (2002).

## We Haven't Listened Long Enough

[197]**Pg 146 profoundly change the world** Drake wrote this in the Preface to *Is Anyone Out There?* (Drake and Sobel 1991).

[198]**Pg 146 have to be patient** At the turn of the millennium, 39% of almost 75,000 respondents to an online poll stated they believed that the discovery of an ET signal would happen within 10 years (SETI@home 2000). Fourteen years later, we're still waiting.

## They Are Signaling but We Aren't Receiving

[199]**Pg 147 an amateur scientist** Although Smith is an "amateur" scientist, he has published in a variety of reputable and peer-reviewed journals across a range of fields. Regarding his contribution to the Fermi paradox debate, see Smith (2009).

## Everyone is Listening, No One is Transmitting

[200]**Pg 149 no one is transmitting** This idea, that we might live in a universe where there are lots of searchers but no senders, has been dubbed the "SETI" paradox by Zaitsev (2006).

[201]**Pg 149 detect our inadvertent** If ETCs could detect our television transmissions, then they could deduce a great deal about our planet even without decoding the programs. Astronomers have shown how an ETC could deduce the rotational speed of Earth, estimate its size, the length of our year, the distance of Earth from the Sun, and the Earth's surface temperature! See Sullivan, Brown and Wetherill (1978).

[202]**Pg 149 been some deliberate transmissions** Denning (2010) gives a partial list of deliberate broadcasts to the sky, but this reference is of more interest for its treatment of the debate about whether we *should* transmit to the sky.

[203]**Pg 150 more cost-effective to listen** Billingham and Benford (2011) discuss the costs of traditional SETI compared to active SETI.

[204]**Pg 151 *Hipparcos* mission** For more information on the ESA *Hipparcos* mission, see Webb (1999).

[205]**Pg 151 thinkers are opposed** Not everyone is convinced that active SETI is a good idea. Billingham and Benford (2011) call for a moratorium on active SETI and Haqq-Misra et al. (2013) urge caution. Denning (2010) and Musso (2012) give good overviews of the "to transmit or not to transmit" debate. Vakoch (2011) is more upbeat about active SETI. He argues that if we transmit then the burden of decoding and interpreting the message is placed on them; since they are likely to be older, and presumably more advanced, the task will be easier for them and thus communication will be facilitated. Penny (2012) makes the point that transmitting might be dangerous but then

so might listening (as dramatized in "A for Andromeda" by Hoyle and Eliot (1963)); indeed, it's even possible that in some cases even *not* listening could be dangerous. We just don't know.

[206]**Pg 151 ways of signaling** The idea that we could send a signal to extraterrestrial civilizations is almost 200 years old. In 1820 the German mathematician Johann Karl Friedrich Gauss (1777–1855), one of the greatest of all mathematicians, suggested planting forests of pine trees in such a way that they illustrated the Pythagorean theorem. The idea was expanded upon by Joseph Johann von Littrow (1781–1840), director of the Vienna Observatory, who suggested digging large ditches with geometrical shapes, filling them with kerosene, and setting them ablaze. He believed that light from these plainly artificial fires would be visible throughout the Solar System. In 1869, the French physicist Charles Cros (1842–1888) suggested that reflecting sunlight toward Mars using suitably arranged mirrors would be the best way to signal our presence to Martian astronomers. See Cerceau and Bilodeau (2012) for a comparison of old and new attempts at communication.

[207]**Pg 152 other communications** See Zaitsev (2012) for a list of all cosmic messages sent up to that date.

[208]**Pg 152 content of the signal** See Atri et al. (2011) for a proposed protocol for active SETI.

[209]**Pg 152 ethical difficulties** For a discussion of this suggestion, as well as for general SETI questions, see SetiLeague (2013).

[210]**Pg 152 game-theory analysis** For a game-theory approach to the problem of passive and active SETI, see de Vladar (2013).

## They Have No Desire to Communicate

[211]**Pg 153 caution is a general trait** Drake tells the story of how the English astronomer Martin Ryle (1918–1984), an Astronomer Royal who was awarded the Nobel prize for physics, was distraught upon learning of the 1974 Arecibo transmission toward M13. Ryle was worried that advanced ETCs might prey upon us. More recently, Stephen Hawking has warned against humanity trying to initiate contact with alien intelligences; see Hawking (2010). Korhonen

(2013) analyses the risk of ETCs initiating an attack by drawing inferences from the Cold War and mutually assured destruction scenarios. My favorite fictional description of a species whose defining trait is extreme caution — taken to the point of cowardice—is that of "Puppeteers". They occur in Larry Niven's "Known Space" stories, including the award-winning *Ringworld* (Niven 1970).

[212]**Pg 154 taking place in the Galactic Club** Kuiper and Morris (1977) argue that "Complete contact with a superior civilization (in which their store of knowledge is made available to us) would abort [our] further development".

[213]**Pg 154 different for societies** See page 210 of Drake and Sobel (1991).

## They Develop a Different Mathematics

[214]**Pg 155 as Wigner put it** See Wigner (1960) for the source of this quotation.

[215]**Pg 155 an anti-Platonic stance** For a critique of the Platonic view of mathematics, see for example Chaitin (1997), Dehaene (1997), Hersh (1997), Davies (2007) and Abbott (2013).

[216]**Pg 156 rudimentary numerical judgments** For a critique of what animals might be doing when we say they are counting, see Budiansky (1998). Budiansky gives a superb introductory account of animal cognitive processes.

[217]**Pg 157 Why should they?** For a powerful argument as to why we *should* be able to converse with aliens using our system of mathematics, and perhaps a language such as LINCOS, see Minsky (1985).

[218]**Pg 157 different systems can't exist** One author who might have been able to imagine alien mathematics was Jorge Luis Borges (1899–1986), perhaps the greatest Spanish-language writer of the last century. Borges (1998) contains several mathematical-based stories; Bloch (2008) examines the mathematical ideas in one of Borges most famous stories.

[219]**Pg 157 if mathematics itself is universal** Lemarchand (2008) suggests that the golden section $\phi$, which arises in the problem $a/b = b/(a+b)$, might

be a cognitive universal and possess the potential to be used for interstellar communication codes, semantics and interstellar artistic works. However, a great deal of nonsense has been written about the golden section. It isn't the universal it's claimed to be in the human sphere, let alone the extraterrestrial one; see for example Devlin (2007).

## They Are Calling but We Don't Recognize the Signal

[220]**Pg 158 one can imagine various options** One could imagine trying to communicate with extraterrestrials using icons, for example. As mentioned in Solution 31, Gauss suggested this approach: for example, giant geometrical figures, drawn on the Siberian tundra and constructed from pine forest and crops such as wheat, would signal our intelligence to observers on Mars. Perhaps something more sophisticated could be attempted for interstellar communication. Musso (2011) suggests something more interesting: a cosmic language based on analogy.

[221]**Pg 158 Hogben's *Astraglossa*** In *Astraglossa*, which was developed by the British mathematician Lancelot Hogben (1895–1975), the counting numbers are represented by radio pulses. For example, three pulses would represent the number three. A mathematical concept such as "equals" would be represented by a *radioglyph*—a pattern of longer pulses. The scheme was outlined in Hogben (1963). Philip Morrison expanded upon the radioglyph idea; see Morrison (1962).

[222]**Pg 158 or Freudenthal's** The LINCOS language was developed by the German mathematician Hans Freudenthal (1905–1990). There are a few websites devoted to LINCOS, but if you really want to learn the language I believe there is only one source: the original, but out of print, book (Freudenthal 1960). Freudenthal's book dealt only with mathematics. Although he planned a second part that would consider the problem of communicating non-mathematical concepts, he lost interest in the topic. His colleague Alexander Ollongren (1928–) took up the challenge and has developed LINCOS in a number of ways; see for example Ollongren (2011, 2013).

[223]**Pg 158 the Voynich Manuscript** The best print resource for the mysterious Voynich Manuscript is a small-press book (D'Imperio 1978), which is difficult to find. However, many websites describe the various tantalizing aspects of the Voynich Manuscript puzzle.

[224]**Pg 159 the early 15$^{th}$ century.** See Hodgins (2012).

[225]**Pg 159 a medieval hoax** There have been many suggestions about who might have created a hoax manuscript and why they might have done it. And the hoax theory explains why we haven't found meaning in the Voynich Manuscript: there *is* no meaning to be found. On the other hand, a variety of scientists believe they have found patterns in the Voynich Manuscript that suggest the words aren't random, that there's meaning contained in the sentences. See, for example, Amancio et al. (2013).

[226]**Pg 160 would be frustration** Elliott (2011) discusses a protocol for how, after a signal has been detected but not yet deciphered, scientists might disseminate of timely and accurate information to an expectant world. See also Elliott and Baxter (2012) and Elliott (2012).

[227]**Pg 160 indistinguishable from blackbody radiation** If EM radiation is used to transmit information, the most efficient format for a given message is indistinguishable from blackbody radiation (to a receiver who is unfamiliar with the format). This was first shown by Caves and Drummond (1994). The same result, using different arguments, was derived by Lachman et al. (2004).

## Message in a Bottle

[228]**Pg 161 clear but counter-intuitive** Their work (Rose and Wright 2004) appeared as a letter in *Nature* and caused quite a stir in the SETI community. For a theoretical paper it's remarkably easy to follow.

## Oops . . . Apocalypse!

[229]**Pg 165 "home of the next supernova"** Fermilab's management became so exasperated with Dixon's protests that they discussed the matter in their newsletter *FermiNews* (FNAL 1998).

[230]**Pg 165 collapse of the quantum vacuum state** Kurt Vonnegut (1963), in his novel *Cat's Cradle*, gives a fictional account of the effects of a phase transition (albeit a phase transition involving not the quantum vacuum state but the imaginary "ice-nine"—a form of $H_2O$ that's more stable than ordinary water at room temperature.)

[231]**Pg 165 suggesting this could be the case** The idea that our universe might not be in the "true" vacuum didn't originate from cranks! Martin John Rees (1942–), an English astrophysicist, was appointed Astronomer Royal in 1995 and between 2005 and 2010 was the President of the Royal Society. Lord Rees is one of Britain's foremost scientists. His Dutch colleague Piet Hut (1952–) works at the Princeton Institute for Advanced Studies. See Hut and Rees (1983) for details of their suggestion.

[232]**Pg 165 higher than anything physicists can achieve** On 15 October 1991 the Fly's Eye detector in Utah detected a cosmic ray with an energy of 320 EeV. (This energy is so large that the rarely-used SI prefix "Exo" was pressed into action; the prefix represents a factor of $10^{18}$.) The particle detected by Fly's Eye packed a *staggering* amount of energy: about 50 J. In other words, this single subatomic particle carried more kinetic energy than a tennis ball traveling at 180 mph. Its energy was more than 10 million times greater than the maximum achievable energy of the largest accelerator ever been planned. How this particle acquired so much energy is something of a mystery. No obvious process can produce a particle with this much kinetic energy; yet *whatever* produced it must have been relatively nearby, because if it had traveled cosmological distances its interactions with the microwave background would have slowed it down. See Bird (1995).

[233]**Pg 166 usual arrangement of quarks** The existence of strange quarks has been known for decades (see Webb 2004). Their key properties were first highlighted by George Zweig (1937–) and Murray Gell-Mann (1929–) in 1964. However, their presence was first evident in cosmic-ray experiments performed by Clifford Charles Butler (1922–1999) and George Rochester (1909–2001) in 1947; it's an injustice they weren't awarded a Nobel prize for their work.

[234]**Pg 166 surrounding electron cloud** These calculations were the work of the American physicist Robert Loren Jaffe (1946–) and others. For a non-technical account, see Matthews (1999). For a more in-depth analysis, see Jaffe et al. (2000).

[235]**Pg 167 a piece by two lawyers** See Johnson and Baram (2014).

[236]**Pg 167 patiently answered the worries** See for example Ellis et al. (2008).

[237]**Pg 168 investigate earth's core** See Stevenson (2003).

[238]**Pg 168 a rather dangerous activity** See Ćirković and Cathcart (2004).

[239]**Pg 169 different subject areas** The term "nanotechnology" was popularized by the American physicist K. Eric Drexler. In an influential book (Drexler 1986) he presented his vision of a forthcoming revolution in nanoscale engineering. Drexler introduced the term "nanotechnology" to refer to molecular manufacturing (the construction of objects to complex, atomic specifications using sequences of chemical reactions directed by non-biological molecular machinery) together with its techniques, its products, and their design and analysis. Recently, the term has come to denote any technology that has nanoscale effects—submicron lithography (or etching) for example. To distinguish his original concept from the work currently taking place in laboratories, Drexler now refers to "molecular nanotechnology". The field of nanotechnology itself might be said to have started with a lecture given by Feynman (1959), in which he considered the direct manipulation of individual atoms.

[240]**Pg 169 potential to improve health care** For a collection of SF stories that deal with medicine, as well as a discussion of the science behind the stories, see Aiken (2014). Many of the stories touch in some way on nanotechnology.

[241]**Pg 169 a self-replicating machine** A Royal Society (2004) report discussed the potential of nanotechnology and concluded that regulators need not concern themselves with self-replicating machines, for a while at least. Their development lies too far in the future.

[242]**Pg 170 the gray goo problem** One of the best fictional treatments of the grey goo problem is Greg Bear's wonderful short story "Blood Music", which was published in 1983—three years before Drexler's book. The story is available in a collection (Bear 1989).

[243]**Pg 170 less than three hours** See Freitas (2000) for a detailed mathematical assessment of the environmental risks of nanotechnology.

## Ouch . . . Apocalypse!

[244]**Pg 171 on the verge of demonstrating** Drake and Sobel (1991) report how Shklovsky, who as we saw earlier was one of the first to publicize the Fermi paradox, lost heart in the SETI enterprise in the years before his death. Shklovksy was convinced that nuclear war was inescapable, and the same inevitable holocaust would occur with other technological civilizations.

[245]**Pg 171 ruinous for our species** See Turco et al. (1983) for a discussion of the consequences of a nuclear winter.

[246]**Pg 171 knowledge is preserved** Walter Michael Miller Jr. (1923–1996) was an American radioman and tailgunner on 53 bombing raids over Italy and the Balkans in World War II. His award-winning *A Canticle for Liebowitz* (Miller 1960) is one of the classic post-apocalyptic SF novels. He wrote the novel in response to the Allied attack on Monte Cassino—a raid in which he took part and which almost certainly affected him psychologically. (The detailed effects of a nuclear winter were only determined quite recently so, although Miller's post-holocaust world is vividly described, it necessarily lacks scientific accuracy. Nevertheless, the novel is highly recommended.)

[247]**Pg 173 Cooper offers bioterrorism** See Cooper (2013) for a discussion of bioterrorism and its link to the Fermi paradox.

## Heat Wave

[248]**Pg 176 Charles Keeling began measuring** The American chemist Charles David Keeling (1928–2005) worked at the Scripps Institution of Oceanography for more than four decades, and throughout that period maintained beautiful observations of atmospheric carbon dioxide. For substantial biographies of Keeling, see Weart (2008) or Bowen (2006).

[249]**Pg 177 average surface temperature** IPCC (2013) contains details of the increases in Earth's surface temperature averaged over land and oceans.

[250]**Pg 178 latest research suggests** Goldblatt and Watson (2012) argue that it's probably impossible for humanity to trigger a runaway greenhouse by burning fossil fuels. They also point out that their work offers no comfort to the climate change deniers: they clearly state that anthropogenic greenhouse

gas emissions are a major threat to human civilization. They also point out that, even if their work is correct and a runaway greenhouse is not possible, nothing in their models exclude an abrupt change to a "hot, moist greenhouse" state: this wouldn't be a runaway process, but it would be a truly dire outcome.

## Apocalypse When?

[251]**Pg 179 reasoned in the following way** J. Richard Gott III (1947–) is a professor of astrophysics at Princeton University. His original paper on the Doomsday argument (Gott 1993) purported to show, among other things, that mankind is unlikely to colonize the Galaxy; see Gott (1997) for a simplified account of the argument. The article generated an extremely interesting correspondence (Buch et al. 1994). The philosopher John Leslie independently developed the Doomsday argument (Leslie 1996). Perhaps the first person to appreciate the power of this type of reasoning was the Australian physicist Brandon Carter (1942–); Carter's anthropic arguments are outlined in chapter 5.

[252]**Pg 183 an ingenious manner** See Wells (2009) for a fascinating look at the question of human survival, by way of the recorded lifetimes of stage shows and businesses! Wells was one of the few students that Feynman mentored, and I find something of Feynman's irreverence and fearless questioning in this book.

## Cloudy Skies Are Common

[253]**Pg 183 a planet in a system of six stars** At the time of writing, we have yet to find a planetary system as extreme as that in *Nightfall*. In 2012, however, astronomers discovered an example of a planet in a four-star system; see Schwamb et al. (2013). An artist's representation of the planet appears in fig. 4.25.

[254]**Pg 184 a wonderful story** *Nightfall*, written in 1941, is routinely voted as the best SF short story of all time. It can be found in many collections, including Asimov (1969).

## As Good as it Gets

[255]**Pg 186 final element of the standard model** For a lucid description of the discovery of the Higgs boson, and why it was so important, see Carroll (2013).

[256]**Pg 187 telescopes of quite astounding capability** See Webb (2012) for a discussion of new and planned observatories.

## They Are Distance Learners

[257]**Pg 188 Lampton, a scientist** Lampton is involved in SETI activities at the University of California, Berkeley, and in particular the optical SETI program that I outlined in Solution 26. For further details of his proposed solution to the paradox, see Lampton (2013).

[258]**Pg 189 we wouldn't need to send astronauts** The notion that we could replicate Martian life on Earth by having a genome-sequencing probe on Mars transmit genetic information back here, and then use bioprinters to "build" them, is discussed in Venter (2013)

## They Are Somewhere but the Universe is Stranger Than We Imagine

[259]**Pg 190 universe A and universe B** The American physicist Hugh Everett III (1930–1982) developed the many-worlds interpretation of quantum mechanics for his PhD thesis at Princeton. See Everett (1957) for a summary of the thesis. Unfortunately his ideas weren't taken seriously at the time of publication, and he became dispirited and left academia. See Byrne (2010) for well researched account of Everett's rather sad life story.

[260]**Pg 191 *really* interesting places** Alfred Bester (1913–1987) first published his famous novel *The Stars My Destination* under the title *Tiger! Tiger!* (Bester 1956). Arthur Clarke's most ambitious work is perhaps *Childhood's End* (Clarke 1953). Seemingly *outré* speculations aren't limited to science fiction, however. Theoretical physicists also delight in dreaming up wild ideas; see, for example, Tegmark and Wheeler (2001).

[261]**Pg 191 move through the "bulk"** This idea appears in Gato-Rivera (2006); it's a seemingly sincere suggestion, but I find it difficult to take it seriously.

## Intelligence Isn't Permanent

[262]**Pg 192 philosophical speculations** See Schroeder (2002).

[263]**Pg 192 an "adaptationist" solution** See Ćirković (2005) and Ćirković, Dragićević and Berić-Bjedov (2005).

## We Live in a Postbiological Universe

[264]**Pg 194 noted historian of science** See Dick (2003, 2008) for lucid explanations of the implications for SETI if we live in a postbiological universe. His book *The Biological Universe* (Dick 1996) is also highly recommended.

[265]**Pg 194 Stapledon was a British philosopher** Stapledon's science fiction novels influenced writers such as Brian Aldiss, Arthur C. Clarke, Stanislaw Lem and Vernor Vinge. In addition to the novels *Last and First Men* and *Star Maker* mentioned here (Stapledon 1930, 1937), he wrote other influential novels including *Sirius* and *Odd John*.

[266]**Pg 194 give or take 37 million years** The best estimate of the age of the universe comes from a combination of data from the ESA *Planck* satellite and previous missions such as the NASA WMAP satellite; both *Planck* and WMAP worked by measuring the cosmic microwave background radiation. I find it incredible that astronomers can specify fundamental cosmological parameters with such accuracy. When I was a student, realistic estimates for the age of the universe differed by billions of years! See Webb (2012) for a discussion of these space-based missions.

[267]**Pg 195 using an argument based upon stellar evolution** See Norris (2000). Norris's paper appears in a very interesting volume edited by Allen Tough.

[268]**Pg 196 Martin once wrote** *A Song For Lya* appeared in *Analog* magazine in 1974 and went on to win the Hugo Award for Best Novella. It appears in a story collection of the same name (Martin 1976).

## They Are Hanging Out Around Black Holes

[269]**Pg 197 a scale of inward manipulation** See Barrow (1998).

[270]**Pg 197 plenty of room at the bottom** See Feynman (1959). He gave the lecture entitled "There's Plenty of Room at the Bottom" to a meeting of the American Physical Society at Caltech on 29 December 1959. In it, Feynman considered the possibility of directly manipulating individual atoms—it's a lecture that in many ways prefigured the field of nanotechnology.

[271]**Pg 197 Vidal argues** Vidal's PhD thesis is entitled *The Beginning and the End: the Meaning of Life in a Cosmological Perspective* (Vidal 2013).

[272]**Pg 197 nothing can escape** We can't look inside a black hole—not even light can reach us from beyond the event horizon that cloaks a hole—but if we *could* look inside a particular type of black hole might we see an extraterrestrial civilization living there? In 2011, a Russian physicist showed that stable periodic orbits can exist inside a black hole and he hypothesized that KIII civilizations could live safely inside a supermassive black hole. Such a civilization would by definition be invisible to our telescopes. Could *that* be the resolution to the paradox? That ETCs choose to live inside black holes and thus are unable to communicate with us? See Dokuchaev (2011).

[273]**Pg 198 store or extract energy** Inoue and Yokoo (2011) suggest that KIII civilizations might construct what would essentially be a Dyson sphere around a supermassive black hole. However, they make no reference to the Barrow scale: this is essentially a souped-up version of a "traditional" Dyson sphere.

## They Hit the Singularity

[274]**Pg 200 Back in 1965** Gordon Earle Moore (1929–) co-founded Intel in 1968 and quickly became one of the world's richest individuals. See Moore (1965) for the first statement of his "law".

[275]**Pg 201 some time before 2030** The American mathematician Vernor Steffen Vinge (1944–) has explored the idea of the Singularity in several SF novels and short stories. A non-fictional account of the idea can be found in Vinge (1993). A discussion of the seemingly inexorable development of computing power can be found in Moravec (1988).

[276]**Pg 201 calls such an event** The term "singularity" was used in the 1950s by von Neumann, who is quoted as saying: "The ever accelerating progress of technology . . . gives the appearance of approaching some essential singularity in the history of the race beyond which human affairs, as we know them, could not continue". See Ulam (1958).

[277]**Pg 201 transcendental event** Vinge wasn't the first to explore the idea that mankind's intellectual development might profoundly change our global society. The French Jesuit priest Pierre Teilhard de Chardin (1881–1955) thought individual minds would somehow merge to form the noösphere—an expanding sphere of human knowledge and wisdom; spiritual and material would eventually merge to form a new state of consciousness he called the Omega point. His argument, although mystical and woolly, reaches a conclusion that seems similar to Vinge's Singularity. There are two main differences between Vinge and Teilhard de Chardin. First, Vinge has extrapolated real-world trends to suggest specific mechanisms that might get us to the Singularity. Second, organic evolution requires millions of years to construct the noösphere; we (and our successors) construct the Singularity in a few decades. For an insight into this sort of thinking, see for example Teilhard de Chardin (2004).

[278]**Pg 202 a non-biological substrate** See Searle (1984) and Penrose (1989) for two stimulating books criticizing the idea that human-level "artificial" intelligence can exist. I happen to disagree with the conclusions of these highly distinguished thinkers, but the two references here make for extremely interesting reading.

[279]**Pg 202 typeset this book** TeX was developed by the American computer scientist Donald Ervin Knuth (1938–). See Knuth (1984). He wrote the TeX (along with a program for designing typefaces) just so that he could typeset his multi-volume *Art of Computer Programming* to his own satisfaction!

## The Transcension Hypothesis

[280]**Pg 203 a paper published in 2012** See Smart (2012). In this paper Smart builds on a decade of thinking about transcension and its relationship to the Fermi paradox.

[281]**Pg 204 the World Health Organization** For details on urban population growth, see WHO (2013).

[282]**Pg 206 new ideas of evolutionary developmental biology** For a readable account of evolutionary developmental biology, see Carroll (2006).

## The Migration Hypothesis

[283]**Pg 207 futurologist Robert Bradbury** See Ćirković and Bradbury (2006) for details of the migration hypothesis. Robert J. Bradbury (1956–2011) was interested in a variety of unorthodox scientific pursuits, including options for radical life extension. Sadly, he did not live to benefit personally from the life-extending technologies in which he was interested.

## Infinitely Many Civilizations Exist but Only One Within Our Particle Horizon: Us

[284]**Pg 209 done so much to promote** Hart is a particularly clear and forceful writer. For a description of his proposal of how an infinite number of life-bearing planets exist, yet we are alone in the observable universe, see Hart (1995). An equally clear treatment of the subject, by a cosmologist, appears in Wesson (1990).

[285]**Pg 210 Guth has presented** See Guth (2007).

[286]**Pg 210 one of the key underpinning concepts** See Webb (2014) for a discussion of inflation, and of how it's possible that observational results made public in 2014 might provide confirmation of inflation.

## They Don't Exist

[287]**Pg 213 a stimulating and thought-provoking book** The book *Rare Earth* (Ward and Brownlee 1999) articulated the growing suspicion of a number of astrobiologists that Earth is unusual, perhaps unique, in harboring complex forms of life.

[288]**Pg 214 will seem narrow-minded** For an imaginative, unorthodox and challenging book on the possible forms that life may take, see Feinberg and Shapiro (1980). The authors discuss the notions of plasma life in stars, radiant life in interstellar clouds, silicate life, low-temperature life and many other possibilities. One of the earliest and most delightful SF stories about alien biochemistries was *A Martian Odyssey* by Stanley G. Weinbaum (in *Wonder Stories*, July 1934). You can find the story in several anthologies, including Asimov (1971).

## The Universe is Here for Us

[289]**Pg 214 a dozen such steps** See, for example, Mayr (1995).

[290]**Pg 215 some astronomers believe** The Sun's luminosity has increased by about 25% since the formation of the Solar System. Earth's surface temperature has been quite stable over that time, however, mainly thanks to negative feedback loops that reduce the $CO_2$ greenhouse effect. These loops will be unable to maintain Earth's surface temperature at a level that's suitable for complex life beyond another billion years or so. See for example Bergman et al. (2004).

[291]**Pg 217 presented by Brandon Carter** See Carter (1974).

[292]**Pg 217 recent analysis** See Watson (2008) for an extension of Carter's work. See also McCabe and Lucas (2010).

[293]**Pg 217 observational bias** See Bostrom (2002) for a thorough discussion of anthropic bias.

[294]**Pg 218 will never die out** See Barrow and Tipler (1986)—a remarkable and stimulating book, which covers the various types of anthropic principle in detail.

[295]**Pg 218 expanded upon the notion** See Tipler (1994)

## The Canonical Artefact

[296]**Pg 219 searching for a "theory of everything"** For a beautiful treatment of the motivations behind the hunt, see Weinberg (1993).

[297]**Pg 219 addressed by Gerard Foschini** Foschini has won numerous awards for his contributions to communications engineering. See Foschini (1994) for the intriguing notion of the canonical artefact.

## Life Can Have Emerged Only Recently

[298]**Pg 223 takes issue with** See Livio (1999).

[299]**Pg 225 occurred billions of years ago** Work by Sobral et al. (2013) suggests that the rate of star formation peaked about 11 billion years ago, which is rather earlier in the history of the universe than previously thought.

## Planetary Systems Are Rare

[300]**Pg 226 originating in more exotic locales** The novels mentioned in the text are *Integral Trees* (Niven 1984) and *Dragon's Egg* (Forward 1980).

[301]**Pg 227 competing scenarios for planetary formation** The French naturalist George-Louis Le Clerc, Comte de Buffon (1707–1788), proposed in 1749 that the planets formed when a comet collided with the Sun. The German philosopher Immanuel Kant (1724–1804) proposed the nebular theory of planetary formation in 1754. See Williams and Cremin (1968) for a comparative survey of the various different ideas that had been proposed to explain the origin of the Solar System.

[302]**Pg 227 condensed to form the planets** The first models of planetary formation through stellar collisions were developed by the American scientists Thomas Chrowder Chamberlin (1843–1928) and Forest Ray Moulton (1872–1952). The models were improved by the British mathematicians James Hopwood Jeans (1887–1946) and Harold Jeffreys (1891–1989). See Taylor (1998) for a fascinating tour of the Solar System, including its formation. Taylor reaches the conclusion that life on Earth might be the result of chance; and perhaps this means that life is unlikely to occur elsewhere.

[303]**Pg 229 exoplanetary discovery** For more details on the newest planetary discoveries, visit *The Extrasolar Planets Encyclopædia* (Exoplanet Team 2014). For a haunting, beautifully written account of the scientists who search for exoplanets, see Billings (2013).

## Rocky Planets Are Rare

[304]**Pg 230 formation of Earth itself** The accepted age of the Earth, as calculated by geochemists using radioisotopic dating techniques, is 4.54 ± 0.05 billion years. A value close to this was first presented in 1956 by the American geochemist Clair Cameron Patterson (1922–1995); research since then has refined Patterson's value, but not substantially revised it. For more details on how scientists determined the age of the Earth, see for example Dalrymple (2001).

[305]**Pg 230 precise nature of the chondrules** References to what we now know are chondrules were made in the scientific literature as far back as 1802. They were named in 1864, by the German mineralogist Gustav Rose (1798–1873). The English geologist Henry Clifton Sorby (1826–1908), one of the great amateur scientists, used a petrographic microscope—a device he invented—to carry out the first detailed study of chondrules. He suggested that chondrules, which he described as being "like drops of a fiery rain", might be pieces of the Sun that had been ejected in solar prominences. See Sorby (1877).

[306]**Pg 231 a nearby gamma-ray burst** See McBreen and Hanlon (1999). See also Duggan et al. (2003).

[307]**Pg 232 most accurate dating** For further details, see Connelly et al. (2012).

## A Water Based Solution

[308]**Pg 233 a mineral called zircon** Uranium (U) decays to lead (Pb) through two different chains ($^{238}$U decays to $^{206}$Pb with a half-life of 4.47 billion years; $^{235}$U decays to $^{207}$Pb with a half-life of 0.704 billion years). Zircon strongly rejects lead, so any lead that's detected in the mineral must have come from radioactive decay. This gives rise to the possibility of a uranium–lead dating mechanism, and Valley et al. (2014) have shown that the uranium–lead "clock" in zircon is reliable. They confirmed that a scrap of zircon from the Jack Hills region of Western Australia formed 4.4 billion years ago.

[309]**Pg 234 saw the same abundance** See Hartogh *et al.* (2011) for details of the Hartley 2 observations; see Lis et al. (2013) for details of the Honda–Mrkos–Pajdušáková observations.

## Continuously Habitable Zones Are Narrow

[310]**Pg 235 a system's habitable zone** One of the first books to discuss the conditions that might be required to make a planet habitable was Dole (1964). Although now extremely dated, it remains a good guide. The book was the outcome of a RAND study and is rather technical. A popular version, also recommended, is Dole and Asimov (1964). Seager (2013), published almost half a century after the Dole study, provides a detailed summary of the factors that might affect an exoplanet's habitability.

[311]**Pg 235 a "tilted" terrestrial world** See Armstrong et al. (2014) for a discussion of how a fluctuating obliquity does not necessarily preclude the existence of life, and in some cases might actually be a boon for life.

[312]**Pg 236 one recent study** See Vladilo et al. (2013), which considers the effect of atmospheric pressure on the habitable zone.

[313]**Pg 236 just the right place** In several calculations of the boundaries of the habitable zone, Earth can be seen pushing the limits. It's easy to take an "Earth-centric" view of the possibilities for life, but increasingly scientists are discovering that liquid water could exist in a wide variety of situations. Heller and Armstrong (2014) point out that some planets might be *more* suitable for life than Earth is.

[314]**Pg 236 results of computer models** See Hart (1978, 1979).

[315]**Pg 237 most Earth-like planet** For details of the discovery of Kepler-186f, see Quintana et al. (2014).

[316]**Pg 237 models developed by James Kasting** The American geologist James Fraser Kasting (1953–) has made several contributions to our understanding of the long-term stability of Earth's climate. The models he and his colleagues use are much more detailed than Hart's original model. See for example Kasting, Reynolds and Whitmire (1992) and Selsis et al. (2007) for further details.

[317]**Pg 237 could be wider than Hart thought** Rushby et al. (2013) consider a simple model of how the habitable zone evolves over time, and show that some exoplanets can spend many billions of years in their star's habitable zone.

[318]**Pg 238 one-in-five Sun-like stars** Petigura, Howard and Marcy (2013) analyzed *Kepler* and Keck data on exoplanets to conclude that 22% of Sun-like stars harbor Earth-size planets orbiting in their habitable zones.

[319]**Pg 238 galactic habitable zone** See Gonzalez, Brownlee and Ward (2001) for an initial definition of the galactic habitable zone, and Lineweaver, Fenner and Gibson (2004) for a detailed discussion of the size and time evolution of the zone. Gowanlock, Patton and McConnell (2011) describe a model of the GHZ in terms of the spatial and temporal dimensions of the Galaxy that may favor the development of complex life.

## Earth is the First

[320]**Pg 239 analysis of the exoplanets** See Buchhave et al. (2012).

[321]**Pg 240 the star HIP 102152** A detailed study of this solar twin is given in Monroe (2013).

## Earth has an Optimal "Pump of Evolution"

[322]**Pg 240 physicist John Cramer** See Cramer (1986) for a popular account of the idea that Jupiter might affect evolution on Earth.

[323]**Pg 240 gap in the Asteroid Belt** The American geologist George West Wetherill (1925–2006) was well known for his research into the role that Jupiter plays in the Solar System. That resonance effects should cause gaps to exist in the Asteroid Belt was first suggested in 1866 by the American astronomer Daniel Kirkwood (1814–1895). Jack Leach Wisdom (1953–), an American physicist, was one of the first scientists to apply the modern techniques of nonlinear dynamics to the study orbits in the Solar System. Wisdom looked at the Asteroid Belt's 3:1 resonance in detail. For an authoritative and up-to-date account of many of these ideas, as well as a more general discussion of the origin and evolution of the Solar System, see Yeomans (2012).

## The Galaxy is a Dangerous Place

[324]**Pg 242 pose an interesting threat** Magnetars are neutron stars with exceptionally strong magnetic fields. The field of SGR1900+14 is estimated to be $5 \times 10^{10}$ tesla—compare that with the strongest non-destructive magnetic field scientists have made, which is only a little over 100 tesla. The magnetic field of a magnetar is so strong that it could suck the keys from your pocket at a distance of more than 100,000 miles. Of course, if you were standing that close to a magnetar, then the radiation and charged-particle wind that it spews out would kill you instantly. At the time of writing, 21 magnetars have been discovered. See Mereghetti (2008) for more information.

[325]**Pg 245 30 light years of Earth** Gehrels et al. (2003), for example, calculate that a Type II supernova occurring within 8 pc could double the "biologically active" ultraviolet flux at Earth's surface.

[326]**Pg 247 origin was completely unknown** Astronomers first detected gamma-ray bursts in 1969 using data from the VELA satellites, which were in orbit to look for gamma-rays from possible nuclear explosions, but it wasn't until 1997 that astronomers obtained proof that bursts occur at cosmological distances. Even now, the detailed nature of the progenitor events is a matter for debate. See Vedrenne and Atteia (2009).

[327]**Pg 248 responsible for mass extinctions** Melott et al. (2004) suggest that a GRB might have initiated the late Ordovician mass extinction some 440 million years ago. For further details on this suggestion see Thomas (2009).

[328]**Pg 248 proposed by James Annis** See Annis (1999).

[329]**Pg 249 consumed by fire** Arthur Clarke's short story "The Star" describes how humans find the remains of a civilization destroyed by an astronomical explosion. Light from the explosion would have reached Earth about two thousand years ago—a fact that gives the story its haunting quality. I find it poignant that within a few hours of Clarke's death in 2008, the Swift satellite detected GRB 080319B—an explosion so tremendously powerful that, despite occurring 7.5 billion years ago, was potentially visible to the naked eye for half a minute."The Star" appears in many anthologies. See, for example, Asimov (1972).

## A Planetary System is a Dangerous Place

[330]**Pg 249 much closer to home** For an in-depth look at planetary threats, see Bostrom and Ćirković (2008).

[331]**Pg 250 Snowball Earth events** The notion that Earth experienced a global glaciation in the Neoproterozoic age is not new: the English geologist Walter Brian Harland (1917–2003) postulated precisely this as long ago as 1964. At the same time, the Russian geologist Mikhail Budyko (1920–2001) showed how a runaway icehouse effect could take place. Only recently, however, has the notion been taken seriously—largely due to the work of groups led by the American geologists Joseph Kirschvink and James Kasting, who have investigated the escape route from "Snowball Earth". For an early introduction, see Harland and Rudwick (1964). A clearly written introduction to Snowball Earth theories appears in Hoffman and Schrag (2000). More technical papers include Hoffman et al. (1998) and Kirschvink (1992).

[332]**Pg 253 reduction in biodiversity** There might well have been many more extinctions earlier in Earth's history, particularly in Snowball Earth events, but only in the past half billion years have creatures with hard skeletons become common; only relatively recently could creatures become fossils. Indeed, we are now living in the geological eon known as the Phanerozoic era, the name

coming from Greek words meaning "visible life". Nature began experimenting with the present animal phyla in the Cambrian explosion, 540 million years ago; the 4 billion years before the Cambrian explosion is known as the Cryptozoic era, from Greek words meaning "hidden life". For most of Earth's history, virtually all organisms lived and died without leaving traces. For more information about the explosion of animal life in the Cambrian, see Gould (1986).

[333]**Pg 253 great mass-extinction events** See Raup (1990).

[334]**Pg 253 large meteorite impact** The idea that a meteorite impact killed the dinosaurs is an old one. The key paper is Alvarez et al. (1980). Years before that paper appeared, however, a remarkably prescient article was published in an SF magazine (see Enever 1966). It described the consequences of a large meteor hitting Earth. An entertaining look at the evidence for a meteorite impact causing the Cretaceous–Tertiary extinction appears in Alvarez (1997); the book is as good as its title!

[335]**Pg 256 species are becoming extinct** See Leakey and Lewin (1995).

## Earth's System of Plate Tectonics is Unique

[336]**Pg 256 killed each year because of earthquakes** See McClean (2010).

[337]**Pg 257 gives rise to plate tectonics** The first to marshal evidence for the suggestion that continents move was the German meteorologist Alfred Lothar Wegener (1880–1930). He published his ideas on continental drift in 1915, but they were met with ridicule. One of the seeming flaws in his theory was that no known mechanism could account for the drift of continents. Wegener died in a blizzard on an Arctic expedition, shortly before the British geologist Arthur Holmes (1890–1965) suggested that convection might provide a suitable mechanism to explain continental drift. Holmes was a respected geologist; he was the first, for example, to suggest a reasonable timescale for geological processes—his 1913 estimate of 4 billion years for the age of the Earth was far better than any previous estimate. But it was to be almost another 20 years before the idea of continental drift became established. In 1960, the American geologist Harry Hammond Hess (1906–1969) showed that the seafloor was spreading from vents in mid-ocean rifts. As magma welled up and cooled, it pushed the existing seafloor away from both sides of the rifts. It was this force

that moved the continents. See Oreskes (2003) for an in-depth account of how the theory of plate tectonics came into being. Marshak (2009) is a superb textbook that explains the details behind the concepts discussed in this section.

[338]**Pg 258 plate tectonics controls temperature** The first description of Earth's geological-timescale carbon dioxide thermostat appeared in Walker, Hays and Kasting (1981). This mechanism doesn't take into account the effect that biological organisms might have had on stabilizing global surface temperature. Several prominent scientists take the view that life itself has played the key role in keeping temperature at an equable level.

## The Moon is Unique

[339]**Pg 261 the impact hypothesis** Two groups of American scientists independently arrived at the idea of lunar formation by a Mars-sized impactor. One group was led by the American astronomers William Kenneth Hartmann (1939–) and Donald Ray Davis (1939–), who work at the Planetary Science Institute in Arizona. The other group was led by the Canadian-American astronomer Alastair Graham Walter Cameron (1925–2005) of Harvard University. See Hartmann and Davis (1975) and Cameron and Ward (1976).

[340]**Pg 262 identical in Earth and Moon rocks** For details of the oxygen isotope ratios in Moon rock samples, see Wiechert et al. (2001). For details of the titanium isotope ratios in Moon rock samples, see Zhang et al. (2012).

[341]**Pg 262 wrong stage of development** Jacobson (2014) pins down the Moon-forming event to 95 million years (give or take 32 million years) after the Solar System formed. This is rather later than many previous estimates, but a high-energy collision that occurred relatively late in the development of the Solar System is consistent with the observation that Moon and Earth have an identical isotopic composition (see text).

[342]**Pg 263 been any different** For an entertaining treatment of the importance of the Moon, which is aimed at non-scientists, see Comins (1993).

## Life's Genesis is Rare

[343]**Pg 267 two different types of prokaryote exist** The classification of living organisms into the domains of archaea, bacteria and eukarya is relatively recent. The proposal originated in the late 1980s and early 1990s with the American biophysicist Carl Richard Woese (1928–2012), who discovered micro-organisms living in extreme environments (extremes of heat, salinity, acidity—places previously thought to be hostile to life). At first it was thought that these organisms were bacteria that had managed to adapt to extreme conditions; certainly, the cell nucleus of these organisms was not enclosed within a nuclear membrane, which made them look like bacteria. However, Woese and co-workers embarked on a study of the ribosomal RNA of these extremophiles. (In cells, ribosomal RNA is the site of protein synthesis—the place where amino acids are assembled into proteins. It is thus found in all living cells, and a study of the nucleotide sequence of rRNA provides an ideal "evolutionary chronometer".) They found that the rRNA of extremophiles differs quite radically from the rRNA of bacteria. These and other fundamental differences made it clear to Woese that life consists of *three* domains. The landmark paper is Woese, Kandler and Wheelis (1990).

[344]**Pg 271 deoxyribonucleic acid** The story of nucleic acids goes back a long way. The first to investigate the chemical structure of the nucleic acid molecule was Albrecht Kossel (1853–1927), a German biochemist. Kossel isolated the nitrogen bases and named them adenine, guanine, cytosine and thymine. He was awarded the 1910 Nobel prize for his work. Forty years later, the role that DNA might play in heredity was one of the burning issues of biology. In 1953, Francis Crick and James Watson made one of the key breakthroughs in all of science when they proposed the double-helix model of the DNA molecule. For details of the story, and the personalities involved, see Watson (2010) and Ridley (2011).

[345]**Pg 272 two extra letters** The work on expanding the genetic "alphabet" is described in Malyshev et al. (2014).

[346]**Pg 277 many excellent sources available** If you have access to a good library, Brooker (2011) is a popular introductory textbook on genetics.

[347]**Pg 278 in interstellar space** Elisa, Glavin and Dworkin (2009), for example, report the existence of the amino acid glycine in material brought back to Earth from comet Wild 2 by the Stardust spacecraft. A number of polycyclic

aromatic hydrocarbons—molecules that might be important as starting material for life—have been detected in the interstellar medium. The basic building blocks that form complex organics are common in space.

[348]**Pg 278 chemistry of early Earth** The story of scientific research into the question of life's origin is long and fascinating. It began in 1924 with the Russian biologist Alexander Ivanovich Oparin (1894–1980), who suggested that small lumps of organic matter might have formed naturally and become the precursor of modern proteins. Along with the British biologist John Burdon Sanderson Haldane (1892–1964), he produced the evocative idea of the primordial soup, from which living material arose. It wasn't until 1953 that the American biologist Stanley Lloyd Miller (1930–2007), a graduate student working in the laboratory of the Nobel prize-winning chemist Harold Clayton Urey (1893–1981), put these ideas to an experimental test. The results of Miller's experiments suggested that at least the basic building blocks of life could form naturally on a primordial Earth. Nevertheless, there are many steps leading from these building blocks to life itself, and the route remains shrouded in fog. This is a fascinating and active area of research. See Deamer (2012) for an account by someone working in the field.

[349]**Pg 279 being created by chance** For an argument as to why the emergence of life might be a rare occurrence, see Hart (1980). I believe the arguments in the paper are wrong, but as usual Hart states his case clearly and forcefully.

[350]**Pg 280 as genetic material and as enzymes** The first ribozymes—enzymes made of RNA—were discovered independently in 1983 by the American biochemist Thomas Robert Cech (1947–) and the Canadian biochemist Sidney Altman (1939–), who shared the 1989 Nobel prize for chemistry for this work. A good overview of the RNA world is given by Bernhardt (2012).

[351]**Pg 280 progress in these fields is rapid** There are numerous proposals regarding the genesis of life. The following references, which give only a flavor of the wide range of thinking on offer, all appeared within the timescale of the writing of this book. Sharov and Gordon (2013) take what I believe is a hugely speculative approach, and argue that life's origin lies 9.7 billion years ago; compare that with Earth's age of 4.5 billion years. Quite a claim! England (2013) takes a much more traditional approach, but nevertheless arrives at an equally stunning claim: he believes he has identified fundamental physical principles

that drive the origin of life. If England is correct, life arises quite naturally. Deacon (2013) talks about "autogenesis"—a physical process of reciprocal catalysis and self-assembly that can not only *create* order, but also *preserve* order and *reproduce* it; those are the sorts of properties we look for when we talk about life. Martins et al. (2013) discuss the possibility that the chemicals necessary for life were created in shocks when icy comets struck rocky bodies, or rocks impacted on icy surfaces. As you might conclude from this brief sample of papers, the fascinating question of the origin of life is a subject of continuing debate. Indeed, Gollihar, Levy and Ellington (2014) point out that the origin of life remains mysterious in part, paradoxically, because scientists are aware of lots of possible mechanisms that could have led to self-replication of nucleic acids and the creation of cells!

[352]**Pg 281 in existence at this time** See Pons et al. (2011) for the suggestion that life began in mud volcanoes in Isua, Greenland about 3.85 billion years ago.

[353]**Pg 281 Earth's crust formed** As mentioned in the discussion of Solution 56, researchers have dated a tiny zircon crystal from Western Australia to 4.4 billion years. The speck is the oldest known part of our planet. See Valley et al. (2014).

[354]**Pg 282 modern discipline of astrobiology** There are now many introductions to and textbooks on the relatively new science of astrobiology. Three that I can recommend are Dartnell (2007), Sullivan and Baross (2007) and Catling (2014).

[355]**Pg 282 large subsurface ocean** See Witze (2014).

## Life's Genesis is Rare (Revisited)

[356]**Pg 284 unlikely to be small** See Lineweaver and Davis (2002).

[357]**Pg 285 formula is indispensable** For a discuss of the history behind the Bayes formula, and its importance in the modern world, see McGrayne (2011).

[358]**Pg 285 clergyman Thomas Bayes** Not a huge amount is known about the life of Thomas Bayes. His formula appears in Bayes (1763).

[359]**Pg 286 study shows that** For research on how medical professionals often fail to use Bayesian reasoning see, for example, Casscells, Schoenberger and Graboys (1978); Eddy (1982); Gigerenzer and Hoffrage (1995).

[360]**Pg 286 infamous Monty Hall problem** The Monty Hall problem rose to prominence in 1990, when a columnist in *Parade* magazine (see vos Savant, 1990) argued that it pays to switch. The columnist was Marilyn vos Savant, who is clearly a very bright woman: from 1986 to 1989 she was listed in the *Guinness Book of World Records* as possessor of the "Highest IQ (women)"; she ceased to appear not because some other woman was deemed to possess a higher IQ, but because the editors at *Guinness* saw sense and realized that attaching a number to intelligence in this way is essentially meaningless. Her proposed solution to the Monty Hall problem nevertheless provoked outrage from several mathematics professors; at least one academic argued that by publishing such nonsense she was doing a disservice to the public understanding of mathematics. And yet her analysis was perfectly correct.

[361]**Pg 287 my reaction** In failing to spot the answer to the Monty Hall problem I was in good company. Paul Erdös was one of the most prolific mathematicians of the twentieth century. Mathematicians and scientists like to boast of their "Erdös number". If you co-authored a paper with him you have an Erdös number of 1; if you co-authored a paper with someone who has an Erdös number of 1 then you have an Erdös number of 2; and so on. (See Hoffman (1998) for a biography of Erdös.) My own Erdös number is a rather poor 5. Anyway, even the great Paul Erdös only accepted the correct conclusion after he saw computer simulations.

[362]**Pg 288 just such a Bayesian analysis** For the full technical details of the analysis, see Spiegel and Turner (2012).

## Goldilocks Twins are Rare

[363]**Pg 289 flutter of interest** The research in question was presented at the Goldschmidt Conference in Florence; see Benner (2013).

[364]**Pg 290 having a Martian origin** See for example Belbruno et al. (2012).

[365]**Pg 290 strong enough to eject** See Worth, Sigurdsson and House (2013).

## The Prokaryote–Eukaryote Transition is Rare

[366]**Pg 294 ignited the Cambrian explosion** See Knoll and Carroll (1999).

[367]**Pg 295 period of tectonic stability** For a billion year period, tectonic activity on Earth was minimal. Cawood and Hawkesworth (2014) describe the timescales on which the mechanism of plate tectonics has operated.

[368]**Pg 296 biochemist Peter Mitchell** Peter Dennis Mitchell (1920–1992) was awarded the 1978 Nobel prize in chemistry for his proposal of the chemiosmotic hypothesis—the notion that ATP synthesis occurs thanks to a potential difference across a membrane. Mitchell's idea was met with huge skepticism when he proposed it (Mitchell 1961); it took many years before the weight of experimental observation proved the correctness of his hypothesis.

[369]**Pg 297 cells of different size** For a beautifully clear discussion of the development of the eukaryotic cell, and of a variety of other topics in evolutionary biology, see Lane (2010).

## Toolmaking Species are Rare

[370]**Pg 299 some species make them** There is a wide literature on animal tool use, though there's no single definition of what constitutes tool use—is a dog using a wall as a tool when it scratches its back? Depending upon one's definition, many animals have been observed to use tools. With regard to chimps, for example, see Boesch and Boesch (1984, 1990). With regard to capuchin monkeys, see Visalberghi and Trinca (1989). With regard to elephants, see Chevalier-Skolnikoff and Liska (1993). Three good general books on the subject of tool use (including the development of human tool use) are Calvin (1996), Gibson and Ingold (1993) and Griffin (1992).

[371]**Pg 300 Kanzi, a bonobo** For the story of this remarkable bonobo, Kanzi (1980–), see Savage-Rumbaugh and Lewin (1996).

## High Technology is Not Inevitable

[372]**Pg 303 but Denisovans** The story of our Denisovan cousins is still being written. The discovery of *Homo denisova* was announced in Krause (2010). Since then, a mitochondrial genome sequence of an ancient hominim (Meyer,

2013) has led to the suggestion that Denisovans interbred with an as-yet unidentified hominid species. At the time of writing, the chronicle of human evolution is difficult to read, but the incredible advances being made by geneticists will surely bring about some clarity.

[373]**Pg 304 they were no mugs** An introductory article describing how various hominid species must once have co-existed is given in Tattersall (2000). For four excellent books on early-human tool use, see Tattersall (1998), Schick and Toth (1993), Leakey (1994) and Kohn (1999). A modern synthesis of these ideas, and of what might distinguish modern humans from Neanderthals, is Stringer (2012). Svante Pääbo is the "master of Neanderthal DNA"; see Pääbo (2014) for the fascinating story of how modern technology is transforming out understanding of both humans and Neanderthals.

[374]**Pg 304 the Neanderthal's a disservice** See Soressi et al. (2013) for details of bone lissoirs found in Neanderthal sites in present-day Dordogne. Appenzeller (2013) gives two sides of the debate surrounding the putative achievements of Neanderthals.

[375]**Pg 305 began to dazzle** For a discussion of cave art see, for example, Sieveking (1979).

## Intelligence at the Human Level is Rare

[376]**Pg 307 intelligence in other creatures** Herzing (2014) offers an attempt to assess and compare various non-human intelligences, as part of the larger goal of preparing for the assessment of intelligence in life on other planets. We might need to take a flexible approach if we ever encounter extraterrestrial species. For example, if we came across a species that could build a structure complete with cultured gardens, internal temperature control and ventilation would we consider the species to be intelligent? Well, termites build such structures and we generally don't attribute an individual termite with a high level of intelligence. Or is intelligence to be found in the termite "hive mind"? This possibility has been discussed in many science fiction stories; perhaps scientists and philosophers will one day have to grapple with the question for real.

[377]**Pg 307 lived about 65 million years ago** See O'Leary et al. (2013) for details of research on the likely appearance of the ancestor of all mammals.

[378]**Pg 310 Lineweaver pointed out** See Lineweaver (2008) for a strong and beautifully reasoned argument suggesting that human-level intelligence is not a convergent feature of evolution.

[379]**Pg 311 scientists showed how crows** This research appeared in Viet and Nieder (2013).

[380]**Pg 311 development of a fly's eye** In 1993, Walter Gehring and Rebecca Quiring found a gene called *eyeless* that seemed to act as a master control gene for the formation of an eye in fruit flies (see Quiring et al. (1994) and Halder et al. (1995) for more information). By suitable manipulation, they could "turn the gene on" in different places and have a fly sprout an ectopic eye on its wing or its leg or its antenna. *Eyeless* was not the gene "for" an eye—the way genes work is much more subtle—but it seemed, among other functions, to orchestrate the action of thousands of other genes that form an eye in the early development of an embryo. It soon became clear that the fly *eyeless* gene was similar to a mouse gene called *small eye*. A mouse with a defective *small eye* gene develops shrunken eyes. Furthermore, the gene is similar to a human gene responsible for the condition Aniridia, sufferers of which can have defects of the iris, lens, cornea and retina. When geneticists made a detailed comparison it was discovered that the "eye genes" in these three quite different species—fruit fly, mouse and man—were essentially identical in two crucial locations. Georg Halder and Patrick Callaerts decided to implant the mouse *small eye* gene into a fruit fly. The gene worked. It caused the fly to develop ectopic eyes—fruit fly eyes, not mouse eyes. The eyes were not wired to the brain, but they looked like normal insect compound eyes and they responded to light. So although eyes take on different designs across the animal kingdom, the biochemical pathways that allow eyes to function seem to have been laid down very early in history.

## Language is Unique to Humans

[381]**Pg 312 lions do not** Budiansky (1998) is an accessible account of research into animal cognition. For a different take on the question of animal consciousness and intelligence, see Rogers (1997).

[382]**Pg 313 if they lacked language** See Olson (1988) for a discussion of the relevance of human linguistic abilities to the Fermi paradox.

[383]**Pg 315 fossilized Neanderthal hyoid bone** See D'Anastasio et al. (2013).

[384]**Pg 315 philosopher and linguist Noam Chomsky** The American linguist Avram Noam Chomsky (1928–), one of the world's most respected intellectuals, writes widely on political and social issues as well as on linguistics. His linguistic work is highly abstruse, but for an introduction to the revolution that he sparked in 1959—and to the advances made by others in the intervening decades—look no further than Pinker (1994), which is a superbly readable book.

[385]**Pg 318 natural selection of heritable variations** Half the members of a British family known as KE suffer from severe language difficulties: not only do they struggle with grammar, writing and comprehension, they can't properly coordinate the complex mechanical motor sequences required for fluid speech. Geneticists (Lai et al. 2001) discovered that the source of the problem was a mutation in the gene Forkhead box protein P2—FOXP2, for short. Normally, FOXP2 coordinates the expression of other genes, but in affected members of the KE family it was broken. This was the first time that scientists had implicated a specific gene in a speech and language disorder, and so it's not surprising that journalists began calling it "the language gene". This was taking the interpretation much too far: FOXP2 isn't a language or a grammar gene. But it *is* an interesting gene, and ongoing research will clarify the role it seems to play in language.

## Science is Not Inevitable

[386]**Pg 320 50,000 years ago** Genetic studies suggest that Aboriginal people are descended from the first humans to migrate out of Africa. They migrated to Asia about 70,000 years ago and somehow made the journey to Australia about 50,000 years ago. See Rasmussen et al. (2011).

[387]**Pg 321 rise of modern science** There are many good accounts of the historical development of science. See, for example, Asimov (1984).

## Consciousness is Not Inevitable

[388]**Pg 322 chilling science fiction novel** Watts (2006) packs his novel *Blindsight* with speculations based on hard science. By making the case for separating intelligence from consciousness he succeeds in describing creatures that appear

to be truly alien. The novel has a thoroughly bleak outlook on life, but is well worth reading—particularly since the author has been kind enough to make the novel available for free online.

[389]**Pg 323 captured the experiment on film** For a good discussion of the blindsight phenomenon, see de Gelder (2010); the article also links to video of the experiment mentioned in the text. The film shows a patient TN successfully navigating a litter-strewn corridor; also shown in the film, walking behind TN, is the British psychologist Lawrence Weiskrantz (1926–) who in the 1970s discovered and named the phenomenon of blindsight.

[390]**Pg 324 far as my reading has taken me** On the basis of Watts' recommendation in *Blindsight*, I'm employing Metziger (2003) as a tour guide to the phenomenon of consciousness and subjectivity. It's tough going (I find most philosophy books tough going) but Metzinger is clearly a brilliant thinker and his arguments are compelling.

## Gaia, God or Goldilocks?

[391]**Pg 324 history of clement weather** For details of the argument in this section, and of the various ways in which Earth might be special, see Waltham (2014).

[392]**Pg 326 originated by James Lovelock** Although he is best known for developing the Gaia hypothesis, Lovelock (1919–) has several inventions to his name and a number of contributions to science, even though he's an unaffiliated, independent scientist. For more on Gaia, and humankind's possible future, see for example Lovelock (2009, 2014).

## Conclusion

[393]**Pg 331 resolutions of the Fermi paradox** Note that new solutions to the paradox, and new work that is inspired by the paradox, appear frequently in the scientific and science fictional literature. Whates (2014), for example, is an anthology of original science fiction stories inspired by Fermi's question. It was published just weeks before this book went to print.

## The Fermi Paradox Resolved . . .

[394]**Pg 332 the Douglas Adams response** The quote appears, of course, in *The Hitchhiker's Guide* (Adams 1979).

[395]**Pg 332 one recent estimate** The estimate of 100 billion for the number of habitable, Earth-like planets is larger than previous estimates, but is not unreasonable. The estimate appears in Abe et al. (2013).

[396]**Pg 333 stands for Graham's number** The story of Graham's number first appeared in Martin Gardner's *Scientific American* column (Gardner 1977), in which it was called "the largest number ever used in a serious mathematical proof". Gardner's column referred to a number used by Graham in an unpublished proof. In 1971, Graham co-published a paper that discussed the problem mentioned in the text (although the problem was couched in terms of coloring the lines connecting vertex pairs of an *n*-dimensional hypercube, rather than in terms of committees and subcommittees); see Graham and Rothschild (1971). The upper bound calculated by Graham and Rothschild was much smaller than Graham's number, but still vast. The lower bound has been improved, and now stands at 13. The upper bound has been improved, too, and now stands at $2 \uparrow\uparrow 2 \uparrow\uparrow 2 \uparrow\uparrow 9$.

[397]**Pg 337 biologist Jacques Monod** See Monod (1971). This translation from the French original is by A. Whitehouse.

# References

Abbott D (2013) The reasonable ineffectiveness of mathematics. Proc IEEE 101:2147–2153

Abe F et al (2013) Extending the planetary mass function to earth mass by microlensing at moderately high magnification. Mon Not R Astro Soc 431:2975–2985

Aczel A (1998) Probability 1: why there must be intelligent life in the universe. Harcourt Brace, New York

Adams D (1979) The Hitchhiker's guide to the galaxy. Pan, London

Aiken B (2014) Small doses of the future: a collection of medical science fiction stories. Springer, Berlin

Alcubierre M (1994) The warp drive: hyper-fast travel within general relativity. Class Quantum Gravity 11:L73–L77

Almheiri A, Marolf D, Polchinski J, Sully J (2013) Black holes: complementarity or firewalls? J High Energy Phys 2013(2):1–20

Alvarez L et al (1980) Extra-terrestrial cause for the Cretaceous–Tertiary extinction. Science 208:1094–1108

Alvarez Q (1997) T-Rex and the crater of doom. Princeton University, Princeton

Amancio DR, Altmann EG, Rybski D, Oliveira ON Jr, Costa L da F (2013) Probing the statistical properties of unknown texts: application to the Voynich Manuscript. PLoS ONE 8(7):e67310

Anderson P (2000) Tau zero (SF Collector's Edition). Orion, London.

Andrews DG (2004) Interstellar propulsion opportunities using near-term technologies. Acta Astronaut 55:443–451

Annis J (1999) An astrophysical explanation of the great silence. J Br Interplanet Soc 52:19

Appenzeller T (2013) Neanderthal culture: old masters. Nature 497:302–304

Armstrong JC, Barnes R, Domagal-Goldman S, Breiner J, Quinn TR, Meadows VS (2014) Effects of extreme obliquity variations on the habitability of exoplanets. Astrobiology 14:277–291

Armstrong S, Sandberg A (2013) Eternity in six hours: intergalactic spreading of intelligent life and sharpening the Fermi paradox. Acta Astronaut 89:1–13

Arnold K (1952) The coming of the saucers (Privately published)

Arnold L (2013) Transmitting signals over interstellar distances: three approaches compared in the context of the Drake equation. Int J Astrobiol 12:212–217

Arrhenius SA (1908) Worlds in the making. Harper and Row, New York

S. Webb, *If the Universe Is Teeming with Aliens . . . WHERE IS EVERYBODY?*, Science and Fiction, DOI 10.1007/978-3-319-13236-5

Asimov I (1959) Nine tomorrows. Doubleday, New York
Asimov I (1969) Nightfall and other stories. Doubleday, New York
Asimov I (ed) (1971) Where do we go from here? Doubleday, New York
Asimov I (ed) (1972) The Hugo winners, volumes 1 and 2. Doubleday, New York
Asimov I (1979) In memory yet dreen. Doubleday, New York
Asimov I (1981) Extraterrestrial civilizations. Pan, London
Asimov I (1984) Asimov's new guide to science. Basic Books, New York
Asimov I (1994) I, Asimov: a memoir. Doubleday, New York
Atri D, DeMarines J, Haqq-Misra J (2011) A protocol for messaging to extraterrestrial intelligence. Space Policy 27:165–9
Bahcall JN, Davis R (2000) The evolution of neutrino astronomy. CERN Cour 40(6):17–21
Bainbridge WS (1984) Computer simulation of cultural drift: limitations on interstellar colonization. J Br Interplanet Soc 37:420–429
Ball JA (1973) The zoo hypothesis. Icarus 19:347–349
Ball JA (1995) Gamma-ray bursts: the ETI hypothesis. www.haystack.mit.edu/hay/staff/jball/grbeti.ps
Barlow MT (2013) Galactic exploration by directed self-replicating probes, and its implications for the Fermi paradox. Int J Astrobiol 12:63–68
Barrow JD (1998) Impossibility: the limits of science and the science of limits. OUP, Oxford
Barrow JD, Tipler FJ (1986) The anthropic cosmological principle. OUP, Oxford
Battersby S (2013) Alien megaprojects: the hunt has begun. New Sci 2911:42–45
Baxter S (2000a) The planetarium hypothesis: a resolution of the Fermi paradox. J Br Interplanet Soc 54:210–216
Baxter S (2000b) Manifold: space. Voyager, London
Bayes T (1763) An essay towards solving a problem in the doctrine of chances. Phil Trans R Soc 53:370–418
Beane SR, Davoudi Z, Savage MJ (2012) Constraints on the universe as a numerical simulation. arXiv:1210.1847v2
Bear G (1989) Tangents. Warner, New York
Belbruno E, Moro-Martín A, Malhotra R, Savransky D (2012) Chaotic exchange of solid material between planetary systems: implications for lithopanspermia. Astrobiology 12:754–74
Ben-Bassat A, Ben-David-Zaslow R, Schocken S, Vardi Y (2005) Sluggish data transport is faster than ADSL. Ann Improbable Res 11:4–8
Benford G (1977) In the ocean of night. Dial, New York
Benford G, Niven L (2012) Bowl of heaven. Tor, New York
Benford J, Benford G, Benford D (2010a) Messaging with cost-optimized interstellar beacons. Astrobiology 10:475–90
Benford J, Benford G, Benford D (2010b) Searching for cost-optimized interstellar beacons. Astrobiology 10:491–8
Benner SA (2013) Planets, minerals and life's origin. Mineral Mag 77:686

Bergman NM, Lenton TM, Watson AJ (2004) COPSE: a new model of biogeochemical cycling over Phanerozoic time. Am J Sci 304:397–437
Bernal JD (1929) The world, the flesh and the devil. Cape, London
Bernhardt HS (2012) The RNA world hypothesis: the worst theory of the early evolution of life (except for all the others). Biol Direct 7:23. doi:10.1186/1745-6150-7-23
Bester A (1956) The stars my destination. Sidgwick and Jackson, London
Bezsudnov I, Snarskii A (2010) Where is everybody?—Wait a moment . . . New approach to the Fermi paradox. arXiv:1007.2774v1
Billingham J, Benford J (2011) Costs and difficulties of large-scale "messaging", and the need for international debate on potential risks. arXiv:1102.1938v2
Billings L (2013) Five billion years of solitude: the search for life among the stars. Current, New York
Bird DJ (1995) Detection of a cosmic ray with a measured energy well beyond the expected spectral cutoff due to cosmic microwave radiation. Astrophys J 441:144–151
Bjørk R (2007) Exploring the galaxy using space probes. Int J Astrobiol 6:89–93
Bloch WG (2008) The unimaginable mathematics of Borges' library of Babel. OUP, Oxford
Boesch C, Boesch H (1984) Mental map in wild chimpanzees: an analysis of hammer transports for nut cracking. Primates 25:160–170
Boesch C, Boesch H (1990) Tool use and tool making in wild chimpanzees. Filia Primatol 54:86–99
Borges JL (1998) Collected fictions. Viking, London
Bostrom N (2002) Anthropic bias: observer self-selection effects in science and philosophy. Routledge, New York
Bostrom N (2003) Are you living in a computer simulation? Phil Q 53(211):243–255
Bostrom N (2006) What is a singleton? Ling Phil Investig 5:48–54
Bostrom N, Ćirković MM (2008) Global catastrophic risks. OUP, Oxford
Bostrom N, Kulczycki M (2011) A patch for the simulation argument. Analysis 71:54–61
Bova B (ed) (1973) The science fiction hall of fame, volume 2A. Doubleday, New York
Bowen M (2006) Thin ice: unlocking the secrets of climate in the world's highest mountains. Holt, New York
Bowyer S (2011) A brief history of the search for extraterrestrial intelligence and an appraisal of the future of this endeavor. Proc. SPIE: Instruments, Methods, and Missions for Astrobiology XIV 8152 Ed. R B Hoover, P C W Davies, G V Levin and A Y Rozanov
Bracewell RN (1960) Communication from superior galactic communities. Nature 186:670–1
Bressi G, Carugno G, Onofrio R, Ruoso G (2002) Measurement of the Casimir force between parallel metallic surfaces. Phys Rev Lett 88:041804
Brin GD (1983) The "great silence": the controversy concerning extraterrestrial intelligent life. Q J R Astro Soc 24:283–309
Brin GD (1985) Just how dangerous is the Galaxy? Analog 105(7):80–95
Brooker RJ (1998) Genetics: analysis and principles, 4th edn. McGraw Hill, New York

Brown P, Spalding RE, ReVelle DO, Tagliaferri E, Worden SP (2002) The flux of small near-earth objects colliding with the earth. Nature 420:294–296
Buch P, Mackay AL, Goodman SN (1994) Future prospects discussed. Nature 358:106–108
Buchhave LA et al (2012) An abundance of small exoplanets around stars with a wide range of metallicities. Nature 486:375–377
Budiansky S (1998) If a lion could talk. Weidenfeld and Nicolson, London
Bussard RW (1960) Galactic matter and interstellar flight. Acta Astronaut 6:179–194
Byl J (1996) On the natural selection of universes. Q J R Astro Soc 37:369–371
Byrne P (2010) The many worlds of hugh Everett III. OUP, Oxford
Calvin WH (1996) How brains think. Basic Books, New York
Cameron AGW, Ward WR (1976) The origin of the moon. Abstr Lunar Planet Sci Conf 7:120–122
Caplan B (2008) The totalitarian threat. In: Bostrom N, Ćirković MM (eds) Global catastrophic risks. OUP, Oxford, pp 504–519
Carey SS (1997) A beginner's guide to scientific method. Wadsworth, Stamford
Carr B (ed) (2007) Universe or multiverse? CUP, Cambridge
Carrigan RA Jr (2009) IRAS-based whole-sky upper limit on Dyson spheres. Astrophys J 698:2075–2086
Carrigan RA Jr (2010) Starry messages: searching for signatures of interstellar archaeology. J Br Interplanet Soc 63:90–103
Carrigan RA Jr (2012) Is interstellar archeology possible? Acta Astronaut 78:121–126
Carroll S (2013) The particle at the end of the universe. Oneworld, London
Carroll SB (2006) Endless forms most beautiful: the new science of Evo Devo. Norton, New York
Carter B (1974) Large number coincidences and the anthropic principle in cosmology. In: Longair MS (ed) Confrontation of cosmological theories with observation. Reidel, Dordrecht
Cartin D (2013) Exploration of the local solar neighbourhood I: fixed number of probes. Int J Astrobiol 12:271–281
Casscells W, Schoenberger A, Graboys TB (1978) Interpretation by physicians of clinical laboratory results. N Engl J Med 299:999–1001
Catling DC (2014) Astrobiology: a very short introduction. OUP, Oxford
Caves CM, Drummond PD (1994) Quantum limits on bosonic communication rates. Rev Mod Phys 66:481–537
Cawood PA, Hawkesworth C (2014) Earth's middle age. Geology 42:503–506
Cerceau FR, Bilodeau B (2012) A comparison between the 19th century early proposals and the 20th–21st centuries realized projects intended to contact other planets. Acta Astronaut 78:72–9
Cernan E, Davis D (1999) The last man on the moon. St Martin's, New York
Chaitin GJ (1997) The limits of mathematics. Springer, Berlin
Chevalier-Skolnikoff S, Liska J (1993) Tool use by wild and captive elephants. Anim Behav 46:209–219

Chyba CF, Hand KP (2005) Astrobiology: the study of the living universe. Ann Rev Astron Astrophys 45:31–74

Ćirković MM (2005) Permanence—an adaptationist solution to Fermi's paradox? J Br Interplanet Soc 58:62–70

Ćirković MM (2008) Against the empire. J Br Interplanet Soc 61:246–254

Ćircović MM, Bradbury RJ (2006) Galactic gradients, postbiological evolution and the apparent failure of SETI. New Astron 11:628–39

Ćirković MM, Cathcart RB (2004) Geo-engineering gone awry: a new partial solution of Fermi's paradox. J Br Interplanet Soc 57:209–215

Ćirković MM, Dragićević I, Berić-Bjedov T (2005) Adaptationism fails to resolve Fermi's paradox. Serb Astron J 170:89–100

Citizen Hearing on Disclosure (2013) The citizen hearing on disclosure homepage http://citizenhearing.org. Accessed 17 Jan 2014

Clarke AC (1953) Childhood's end. Del Rey, New York

Clarke (1956) The City and the Stars. New American Library, New York

Cocconi G, Morrison P (1959) Searching for interstellar communications. Nature 184:844–6

Cohen N, Hohlfeld R (2001) A newer, smarter SETI strategy. Sky Telesc 101(4):50–51

Comins NF (1993) What if the moon didn't exist? Harper Collins, New York

Compton AH (1956) Atomic quest. OUP, Oxford

Connelly JN, Bizzarro M, Krot AN, Nordlund A, Wielandt D, Ivanova MA (2012) The absolute chronology and thermal processing of solids in the solar protoplanetary disk. Science 338:651–655

Cooper J (2013) Bioterrorism and the Fermi paradox. Int J Astrobiol 12:144–148

Corbet R, H D (1999) The use of gamma-ray bursts as direction and time markers in SETI strategies. Pub Astron Soc Pacific 111:881–885

Cotta C, Á M (2009) A computational analysis of galactic exploration with space probes: implications for the Fermi paradox. J Br Interplanet Soc 62:82–88

Cox LJ (1976) An explanation for the absence of extraterrestrials on earth. Q J R Astro Soc 17:201–208

Cramer JG (1986) The pump of evolution. Analog 106(1):124–127

Crawford IA (1995) Interstellar travel: a review. In: Zuckerman B, Hart MH (eds) Extraterrestrials: where are they? CUP, Cambridge

Crawford IA (2000) Where are they? Sci Am 283(7):28–33

Crawford IA (2009) The astronomical, astrobiological and planetary science case for interstellar spaceflight. J Br Interplanet Soc 62:415–421

Crawford IA, Baldwin EC, Taylor EA, Bailey J, Tsembelis K (2008) On the survivability and detectability of terrestrial meteorites on the moon. Astrobiology 8:242–252

Crick F, H C (1981) Life itself. Simon and Schuster, New York

Crick FHC, Orgel LE (1973) Directed panspermia. Icarus 19:341–6

Cronin JW (2004) (Ed) Fermi remembered. UCP, Chicago

Dalrymple GB (2001) The age of the earth in the twentieth century: a problem (mostly) solved. Geol Soc London, Special Publ 190:205–221

D'Anastasio R et al (2013) Micro-biomechanics of the Kebara 2 hyoid and its implications for speech in Neanderthals. PLOS ONE doi:10.1371/journal.pone.0082261
Dartnell L (2007) Life in the universe: a beginner's guide. Oneworld, London
Davies EB (2007) Let platonism die. Euro Math Soc Newsletter (June) 64:24–25
Davies P, C W (2010) The Eerie silence. Allen Lane, London
Davies P, C W (2012) Footprints of alien technology. Acta Astronaut 73:250–7
Davies PCW, Wagner RV (2013) Searching for alien artifacts on the moon. Acta Astronaut 89:261–5
Deacon T (2013) Life before genetics: autogenesis, and the outer solar system. https://www.youtube.com/watch?v=jeMwy3xuEs8. Accessed 8 May 2014
Deamer D (2012) First life: discovering the connections between stars, cells, and how life began. University of California, Oakland
Deardorf JW (1986) Possible extraterrestrial strategy for earth. Q J R Astro Soc 27:94–101
Deardorf JW (1987) Examination of the embargo hypothesis as an explanation for the Great silence. J Br Interplanet Soc 40:373–379
de Gelder B (2010) Uncanny sight in the blind. Sci Am 302:60–65
Dehaene S (1997) The number sense: how the mind creates mathematics. OUP, Oxford
Denning K (2010) Unpacking the great transmission debate. Acta Astronaut 67:1399–1405
Deutsch D (1998) The fabric of reality. Penguin, London
de Vladar HP (2013) The game of active search for extra-terrestrial intelligence: breaking the "Great Silence". Int J Astrobiol 12:53–62
Devlin K (2007) The myth that will not go away. MAA Online (May). www.maa.org/external_archive/devlin/devangle.html. Accessed 21 Feb 2014
Dick SJ (1996) The biological universe: the twentieth century extraterrestrial life debate and the limits of science. CUP, Cambridge
Dick SJ (2003) Cultural evolution, the postbiological universe and SETI. Int J Astrobiol 2:65–74
Dick SJ (2008) The postbiological universe. Acta Astronaut 62:499–504
Digital S (2013) "Jesus Christ image" found in fabric conditioner www.digitalspy.co.uk/fun/news/a474585/jesus-christ-image- found-in-fabric-conditioner.html. Accessed 20 Jan 2014
D'Imperio ME (1978) The Voynich manuscript—an elegant enigma. Aegean Park Press, Laguna Hills
Dokuchaev VI (2011) Is there life inside black holes? Class Quantum Gravity 28:235015
Dole SH (1964) Habitable planets for man. Blaisdell, New York
Dole SH, Asimov I (1964) Planets for man. Random House, New York
Douglas F (1977) The absence of extraterrestrials on earth. Q J R Astro Soc 18: 157–158
Drake FD, Sagan C (1973) Interstellar radio communication and the frequency selection problem. Nature 245:257–8
Drake FD, Sobel D (1991) Is anyone out there? Simon and Schuster, London
Drexler KE (1986) Engines of creation: the coming era of nanotechnology. Doubleday, New York

Duggan P, McBreen B, Carr AJ, Winston E, Vaughan G, Hanlon L, McBreen S, Metcalfe L, Kvick AA, Terry AE (2003) Gamma-ray bursts and X-ray melting of material to form chondrules and planets. Astron Astrophys 409:L9–L12

Dyson FJ (1960) Search for artificial sources of infrared radiation. Science 131:1667

Dyson FJ (1963) Gravitational machines. In: G W Cameron A (ed) Interstellar communication. Benjamin, New York

Dyson FJ (1982) Interstellar propulsion systems. In Zuckerman B, Hart MH (eds) Extraterrestrials: where are they? CUP, Cambridge

Eddy DM (1982) Probabilistic reasoning in clinical medicine: problems and opportunities. In Kahneman D, Slovic P, Tversky A (eds) Judgment under uncertainty: heuristics and biases. CUP, Cambridge, pp 249–267

Edmondson WH (2010) Targets and SETI: shared motivations, life signatures and asymmetric SETI. Acta Astronaut 67:1410–1418

Edmondson WH, Stevens IR (2003) The utilization of pulsars as SETI beacons. Int J Astrobiol 2:231–271

Eichler D, Beskin G (2001) Optical SETI with air Cerenkov telescopes. Astrobiology 1:489–493

Einstein A, Podolsky B, Rosen N (1935) Can a quantum-mechanical description of physical reality be considered complete? Phys Rev 41:777–780

Elliott JR (2011) A post-detection decipherment strategy. Acta Astronaut 68:441–444

Elliott JR (2012) Constructing the matrix. Acta Astronaut 78:26–30

Elliott JR, Baxter S (2012) The DISC quotient. Acta Astronaut 78:20–25

Ellis J, Giudice G, Mangano ML, Tkachev T, Wiedemann U (2008) Review of the safety of LHC collisions. J Phys G Nucl Part Phys 35:115004

Elsila JE, Glavin DP, Dworkin JP (2009) Cometary glycine detected in samples returned by Stardust. Meteor Planet Sci 44:1323–30

Elvis M (2014) How many ore-bearing asteroids? Planet Space Sci 91:20–6

Enever JG (1966) Giant meteor impact. Analog 77(3):62–84

England JL (2013) Statistical physics of self replication. J Chem Phys 139:121923

ESA (2013) ESA home page http://sci.esa.int/planck/

Everett H (1957) "Relative state" formulation of quantum mechanics. Rev Mod Phys 29:454–462

Exoplanet Team (2014) http://exoplanet.eu. Accessed 1 April 2014

Faizullin RT (2010) Geometrical joke(r?)s for SETI. arxiv.org/abs/1007.4054

Farmer PJ (1981) The unreasoning mask. Putnam, New York

Feinberg G, Shapiro R (1980) Life beyond earth. Morrow, New York

Fermi L (1954) Atoms in the family. UCP, Chicago

Feynman RP (1959) There's plenty of room at the bottom. Lecture given to the American Physical Society at Caltech, 29 Dec

Finney BR, Jones EM (eds) (1985) Interstellar migration and the human experience. University of California, Berkeley

Fischer DA et al. (2002) A second planet orbiting 47 Ursae Majoris. Astrophys J 564:1028–1034

FNAL (1998) The universe lives on and rumors of its imminent demise have been greatly exaggerated. www.fnal.gov/pub/ferminews/FermiNews98-06-19.pdf
Fogg MJ (1987) Temporal aspects of the interaction among the first galactic civilizations: the "interdict hypothesis". Icarus 69:370–384
Fogg MJ (1988) Extraterrestrial intelligence and the interdict hypothesis. Analog 108(10):62–72
Fogg MJ (1995) Terraforming: engineering planetary environments. SAE, New York
Forgan DH (2009) A numerical testbed for hypotheses of extraterrestrial life and intelligence. Int J Astrobiol 8:121–131
Forgan DH (2011) Spatio-temporal constraints on the zoo hypothesis, and the breakdown of total hegemony. Int J Astrobiol 10:341–347
Forgan DH (2013) On the possibility of detecting class A stellar engines using exoplanet transit curves. J Br Interplanet Soc 66:144–154
Forgan DH, Nichol RC (2011) A failure of serendipity: the square kilometre array will struggle to eavesdrop on human-like ETI. Int J Astrobiol 10:77–81
Forgan DH, Papadogiannakis S, Kitching T (2013) The effects of probe dynamics on galactic exploration timescales. J Br Interplanet Soc 66:171–178
Forward RL (1980) Dragon's egg. Del Rey, New York
Forward RL (1984) Roundtrip interstellar travel using laser-pushed lightsails. J Spacecraft 21:187–195
Forward RL (1990) The negative matter space drive. Analog 110(9):59–71
Foschini GJ (1994) The canonical artefact and its cosmological interpretations. Proc Math Phys Sci 444:3–16
Freitas RA Jr (1980) A self-reproducing interstellar probe. J Br Interplanet Soc 33:251–264
Freitas RA Jr (1983a) If they are here, where are they? Observational and search considerations. Icarus 55:337–343
Freitas RA Jr (1983b) The search for extraterrestrial artifacts (SETA). J Brit Interplanet Soc 36:501–506
Freitas RA Jr (1985) There is no Fermi paradox. Icarus 62:518–520
Freitas RA Jr (2000) Some limits to global ecophagy by biovorous nanoreplicators, with public policy recommendations. Available from www.foresight.org/nano/Ecophagy.html
Freitas RA Jr, Valdes F (1980) A search for natural or artificial objects located at the earth–moon libration points. Icarus 42:442–447
French AP (1968) Special relativity. Norton, San Francisco
Freudenthal H (1960) Design of a language for cosmic intercourse. North Holland, Amsterdam
Garcia-Escartin JC, Chamorro-Posada P (2013) Scouting the spectrum for interstellar travellers. Acta Astronaut 85:12–18
Gardner M (1969) The unexpected hanging and other mathematical diversions. Simon and Schuster, New York
Gardner M (1970) The fantastic combinations of John Conway's new solitaire game "Life". Sci Am 223:120–123
Gardner M (1977) Mathematical games. Sci Am 237:18–28

Gardner M (1985) The great stone face and other nonmysteries. Skept Inquirer 10(2): 14–18

Gato-Rivera B (2006) The Fermi paradox in the light of the inflationary and brane world cosmologies. Trends in general relativity and quantum cosmology. Nova, New York

Gehrels N, Laird CM, Jackman CH, Canizzo JK, Mattson BJ, Chen W (2003) Ozone depletion from nearby supernovae. Astrophys J 585:1169–1176

Gibson KR, Ingold T (eds) (1993) Tools, language and cognition in human evolution. CUP, Cambridge

Gigerenzer G, Hoffrage U (1995) How to improve Bayesian reasoning without instruction: frequency formats. Psych Rev 102:684–704

Gillon M (2014) A novel SETI strategy targeting the solar focal regions of the most nearby stars. Acta Astronaut 94:629–633

Goldblatt C, Watson AJ (2012) The runaway greenhouse: implications for future climate change, geoengineering and planetary atmospheres. Phil Trans R Soc A 370:4197–4216

Gonzalez G, Brownlee D, Ward PD (2001) Refuges for life in a hostile universe. Sci Am 285(4):60–67

Gordon JE (1991) The new science of strong materials or why you don't fall through the floor (2nd revised ed.). Penguin, London

Gott JR III (1993) Implications of the Copernican principle for our future prospects. Nature 363:315–319

Gott JR III (1995) Cosmological SETI frequency standards. In Zuckerman B, Hart MH (eds) Extraterrestrials: where are they? CUP, Cambridge

Gott JR III (1997) A grim reckoning. New Scientist 15 Nov pp 36–39

Gould SJ (1985) SETI and the wisdom of Casey Stengel. In: The Flamingo's smile. Penguin, London

Gould SJ (1986) Wonderful life. Norton, New York

Gowanlock MG, Patton DR, McConnell SM (2011) A model of habitability within the Milky Way galaxy. Astrobiology 11:855–873

Graham RL, Rothschild BL (1971) Ramsey's theorem for *n*-parameter sets. Trans American Math Soc 159:257–292

Gray RH (2011) The elusive wow: searching for extraterrestrial intelligence. Palmer Square, Chicago

Gribbin J (1996) Schrödinger's kittens. Phoenix Press, London

Gribbin J (2010) In search of the multiverse. Penguin, London

Griffin DR (1992) Animal minds. Chicago University, Chicago

Gros C (2005) Expanding advanced civilizations in the universe. J Br Interplanet Soc 58:108–110

Guardian (2001) It could've been you. www.theguardian.com/society/2001/may/02/lottery.g2. Accessed 20 Jan 2014

Gurzadyan VG (2005) Kolmogorov complexity, string information, panspermia and the Fermi paradox. Observatory 125:352–355

Guth AH (2007) Eternal inflation and its implications. J Phys A Math General 40:6811

Hair TW (2011) Temporal dispersion of the emergence of intelligence: an inter-arrival time analysis. Int J Astrobiol 10:131–135

Hair TW (2013) Provocative radio transients and base rate bias: a Bayesian argument for conservatism. Acta Astronaut 91:194–197
Haisch B, Rueda A, Puthoff HE (1994) Beyond $E = mc^2$. Science 34(6):26–31
Halder G et al (1995) Induction of ectopic eyes by targeted expression of the eyeless gene in *Drosophila*. Science 267:1788–1792
Hall MD, Connors WA (2000) Captain Edward J. Ruppelt: summer of the saucers. Rose, Albuquerque
Hancock G, Bauval R, Grigsby J (1998) The Mars mystery. Michael Joseph, London
Hanson R (1998) Burning the cosmic commons: evolutionary strategies for interstellar colonization. http://hanson.gmu.edu/filluniv.pdf
Haqq-Misra J, Baum S (2009) The sustainability solution to the Fermi paradox. J Br Interplanet Soc 62:47–51
Haqq-Misra J, Kopparapu RK (2012) On the likelihood of non-terrestrial artifacts in the solar system. Acta Astronaut 72:15–20
Haqq-Misra J, Busch M, Som S, Baum S (2013) The benefits and harms of transmitting into space. Space Policy 29:40–48
Harland WB, Rudwick MJS (1964) The great infra-Cambrian glaciation. Sci Am 211(2):28–36
Harp GR, Ackermann RF, Blair SK, Arbunich J, Backus PR, Tarter JC, ATA Team (2011) A new class of SETI beacons that contain information. In: Vakoch DA (ed) Communication with extraterrestrial intelligence (CETI). State University of New York, Albany, pp 45–70
Harris I (2013) Americans' belief in God, miracles and heaven declines. www.harrisinteractive.com. Accessed 23 Jan 2014
Harrison E (1987) Darkness at night. Harvard University, Cambridge
Harrison E (1995) The natural selection of universes containing intelligent life. Q J R Astro Soc 36:193–203
Hart MH (1975) An explanation for the absence of extraterrestrials on earth. Q J R Astro Soc 16:128–135
Hart MH (1978) The evolution of the atmosphere of the earth. Icarus 33:23–39
Hart MH (1979) Habitable zones about main sequence stars. Icarus 37:351–357
Hart MH (1980) *N* is very small. In strategies for the search for life in the universe. Reidel, Boston, pp 19–25
Hart MH (1995) Atmospheric evolution, the Drake equation and DNA: sparse life in an infinite universe. In Zuckerman B, Hart MH (eds) Extraterrestrials: where are they? CUP, Cambridge
Hartmann WK, Davis DR (1975) Satellite-sized planetesimals and lunar origin. Icarus 24:504–14
Hartogh P et al. (2011) Ocean-like water in the Jupiter-family comet 103P/Hartley 2. Nature 478:218–220
Hawking SW (2010) Stephen Hawking warns over making contact with aliens. BBC News (25 April). http://news.bbc.co.uk/1/hi/8642558.stm. Accessed 14 March 2014
Hecht J (2010) Beam: the race to make the laser. OUP, Oxford
Heller R, Armstrong J (2014) Superhabitable worlds. Astrobiology 14:50–66

Hempel CG (1945a) Studies in the logic of confirmation I. Mind 54(213):1–26
Hempel CG (1945b) Studies in the logic of confirmation II. Mind 54(214):97–121
Hemry JG (2000) Interstellar navigation or getting where you want to go and back again (in one piece). Analog 121(11):30–37
Hersh R (1997) What is mathematics really? OUP, Oxford
Herzing DL (2014) Profiling nonhuman intelligence: an exercise in developing unbiased tools for describing other "types" of intelligence on earth. Acta Astronaut 94:676–80
Hoagland RC (1987) The monuments of Mars. North Atlantic, Berkeley
Hodgins G (2012) Forensic investigations of the Voynich MS. Voynich 100 Conference www.voynich.nu/mon2012/index.html. Accessed 4 March 2014
Hoffman P (1998) The man who loved only numbers. Hyperion, New York
Hoffman PF, Schrag DP (2000) Snowball earth. Sci Am 282(1):68–75
Hoffman PF, Kaufman AJ, Halverson GP, Schrag DP (1998) A neoproterozoic snowball earth. Science 281:1342–6
Hogben LT (1963) Science in authority. Norton, New York
Hohlfeld R, Cohen N (2000) Optimum SETI search strategy based on properties of a flux-limited catalogue. SETI beyond Ozma. SETI Press, Mountain View
Höss M (2000) Ancient DNA: neanderthal population genetics. Nature 404:453–4
Hoyle F, Eliot J (1963) A for Andromeda. Corgi, London
Hoyle F, Wickramasinghe NC (2000) Astronomical origins of life. Kluwer, Dordrecht
Hut P, Rees MJ (1983) How stable is our vacuum? Nature 302:508–509
Icke D (1999) The biggest secret. Bridge of Love, Newport
Inoue M, Yokoo H (2011) Type III Dyson sphere of highly advanced civilisations around a super massive black hole. J Br Interplanet Soc 64:58–62
IPCC (2013) Climate change 2013: the physical science basis. WMO, Geneva
Jacobson SA, Morbidelli A, Raymond SN, O'Brien DP, Walsh KJ, Rubie DC (2014) Highly siderophile elements in earth's mantle as a clock for the moon-forming impact. Nature 508:84–7
Jaffe RC et al. (2000) Review of speculative "disaster scenarios" at RHIC. Rev Mod Phys 72:1125–1140
Johnson EE, Baram M (2014) New US Science Commission should look at experiment's risk of destroying the earth. Int. Business Times www.ibtimes.com. Accessed 7 March 2014
Jones EM (1975) Colonization of the Galaxy. Icarus 28:421–2
Jones EM (1981) Discrete calculations of interstellar migration and settlement. Icarus 46:328–36
Jones EM (1985) Where is everybody? An account of Fermi's question. Physics Today (Aug) pp 11–13
Jones EM (1995) Estimates of expansion timescales. In Zuckerman B, Hart MH (eds) Extraterrestrials: where are they? CUP, Cambridge
Jugaku J, Nishimura SE (1991) A search for Dyson spheres around late-type stars in the IRAS catalog. In: Heidemann J, Klein MJ (eds) Bioastronomy: the search for extraterrestrial life (lecture notes in physics) 390. Springer, Berlin

Jugaku J, Nishimura SE (1997) A search for Dyson spheres around late-type stars in the solar neighborhood II. In: B Cosmovici C, Bowyer S, Wertheimer D (eds) Astronomical and biochemical origins and the search for life in the universe. Editrice Compositori, Bologna, pp 707–10

Jugaku J, Nishimura SE (2000) A search for Dyson spheres around late-type stars in the solar neighbourhood. III. In: Lemarchand G, Meech K (eds) Bioastronomy: a new era in the search for life. UCP, Chicago

Kardashev NS (1979) Optimal wavelength region for communication with extraterrestrial intelligence—$\lambda = 1.5$mm. Nature 278:28–30

Kasting JF, Reynolds RT, Whitmire DP (1992) Habitable zones around main sequence stars. Icarus 101:108–128

Kecskes C (1998) The possibility of finding traces of extraterrestrial intelligence on asteroids. J Br Interplanet Soc 51:175–180

Kecskes C (2002) Scenarios which may lead to the rise of an asteroid-based technical civilisation. Acta Astronaut 50:569–577

KEO (2014) Welcome to KEO. www.keo.org

Kinouchi O (2001) Persistence solves Fermi paradox but challenges SETI projects. arXiv:cond-mat/0112137v1

Kirschvink JL (1992) Late proterozoic low-latitude global glaciation: the snowball earth. In Schopf JW, Klein C (eds) The Proterozoic biosphere. CUP, Cambridge

Knoll AH, Carroll S (1999) Early animal evolution: emerging views from comparative biology and geology. Science 284:2129–2137

Knuth DE (1984) The TEX book. Addison Wesley, Reading

Kohn M (1999) As we know it. Granta, London

Korhonen JM (2013) MAD with aliens? Interstellar deterrence and its implications. Acta Astronaut 86:201–210

Korpela EJ et al (2011) Status of the UC-Berkeley SETI efforts. Proc. SPIE: Instruments, Methods, and Missions for Astrobiology XIV 8152 Ed. R B Hoover, P C W Davies, G V Levin and A Y Rozanov

Krasnikov SV (2000) A traversible wormhole. Phys Rev D 62:084028

Krause J, Fu Q, Good JM, Viola B, Shunkov MV, Derevianko AP, Pääbo S (2010) The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. Nature 464:894–897

Kuiper TBH, Morris M (1977) Searching for extraterrestrial civilizations. Science 196:616–621

Lachman M, Newman MEJ, Moore C (2004) The physical limits of communication, or why any sufficiently advanced technology is indistinguishable from noise. Am J Phys 72:1290–1293

Lage C (2012) Probing the limits of extremophilic life in extraterrestrial environment-simulated experiments. Int J Astrobiol 11:251–256

Lai C, S L et al. (2001) A forkhead-domain gene is mutated in a severe speech and language disorder. Nature 413:519–523

Lampton M (2013) Information-driven societies and Fermi's paradox. Int J Astrobiol 12:312–3

Landis GA (1998) The Fermi paradox: an approach based on percolation theory. J Br Interplanet Soc 51:163–166

Lane N (2010) Life ascending. Profile, London

Lanoutte W, Szilard B (1994) Genius in the shadows: a biography of Leo Szilard, the man behind the bomb. UCP, Chicago

Lawton AT, Newton SJ (1974) Long delayed echoes: the search for a solution. Spaceflight 6:181–187

Lazio TJW, Tarter J, Backus PR (2002) Megachannel extraterrestrial assay candidates: no transmissions from intrinsically steady sources. Astron J 124:560–564

Leakey R (1994) The origin of humankind. Weidenfeld and Nicolson, London

Leakey R, Lewin R (1995) The sixth extinction. Doubleday, New York

Learned JG, Pakvasa S, Simmons WA, Tata X (1994) Timing data communications with neutrinos: a new approach to SETI. Q J R Astro Soc 35:321–329

Learned JG, Pakvasa S, Zee A (2009) Galactic neutrino communication. Phys Lett B 671:15–19

Lemarchand GA (2008) Counting on beauty: the role of aesthetic, ethical, and physical universal principles for interstellar communication. arXiv:0807.4518

LePage AJ (2000) Where they could hide. Sci Am 283(7):30–31

Leslie J (1996) The end of the world. Routledge, London

Levathes L (1997) When China ruled the seas: the treasure fleet of the dragon throne. OUP, Oxford, pp 1405–1433

Lineweaver CH (2008) Paleontological tests: human-like intelligence is not a convergent feature of evolution. In: Seckbach J, Walsh M (eds) From fossils to astrobiology. Springer, Berlin, pp 355–68

Lineweaver CH, Davis TM (2002) Does the rapid appearance of life on earth suggest that life is common in the universe? Astrobiology 2:293–304

Lineweaver CH, Fenner Y, Gibson BK (2004) The galactic habitable zone and the age distribution of complex life in the Milky Way. Science 303:59–62

Lis DC et al (2013) A Herschel study of D/H in water in the Jupiter-family comet 45P/Honda–Mrkos–Pajdušáková and prospects for D/H measurements with CCAT. Astrophys J Lett 774:L3–L8

Lissauer JJ, Chambers JE (2008) Solar and planetary destabilization of the earth–moon triangular Lagrangian points. Icarus 195:16–27

Livio M (1999) How rare are extraterrestrial civilizations, and when did they emerge? Astrophys J 511:429–431

Loeb A, Turner EL (2012) Detection technique for artificially illuminated objects in the outer solar system and beyond. Astrobiology 12:290–294

Loeb A, Zaldarriaga M (2007) Eavesdropping on radio broadcasts from galactic civilizations with upcoming observatories for redshifted 21 cm radiation. J Cosmol Astropart Phys. doi:10.1088/1475-7516/2007/01/020

Lovelock JE (2009) The vanishing face of Gaia: a final warning—enjoy it while you can. Allen Lane, London

Lovelock JE (2014) A rough ride to the future. Allen Lane, London

Lunan D (1974) Man and the stars. Souvenir, London

Lytkin V, Finney B, Alepko L (1995) Tsiolkovsky, Russian cosmism and extraterrestrial intelligence. Q J R Astro Soc 36:369–376
Maccone C (1994) Space missions outside the solar system to exploit the gravitational lens of the Sun. J Br Interplanet Soc 47:45–52
Maccone C (2000) The gravitational lens of Alpha Centauri a, b, c and of Barnard's star. Acta Astronaut 47:885–897
Maccone C (2009) Deep space flight and communications—exploiting the Sun as a gravitational lens. Springer, Berlin
Maccone C (2011) Focusing the galactic internet. In: Paul Schuch H (ed) Searching for extraterrestrial intelligence: SETI past, present and future. Springer, Berlin, pp 325–49
Maccone C (2013) Sun focus comes first, interstellar comes second. J Br Interplanet Soc 66:25–37
Maccone C, Piantà M (1997) Magnifying the nearby stellar systems by FOCAL space missions to 550 AU. Part I. J Br Interplanet Soc 50:277–280
Mallove EF, Matloff GL (1989) The starflight handbook. Wiley, New York
Malyshev DA, Dhami K, Lavergne T, Chen T, Dai N, Foster JM, Corrêa IR, Romesberg FE (2014) A semi-synthetic organism with an expanded genetic alphabet. Nature. doi:10.1038/nature13314
Marshak S (2009) Essentials of geology, 3rd edn. Norton, New York
Martin G, R R (1976) A song for Lya. Avon, New York
Martins Z, Price MC, Goldman N, Sephton MA, Burchell MJ (2013) Shock synthesis of amino acids from impacting cometary and icy planet surface analogues. Nat Geosci 6:1045–1049
Mathews JD (2011) From here to ET. J Br Interplanet Soc 64:234–241
Matthews R (1999) A black hole ate my planet. New Scientist 28 Aug pp 24–27
Mauersberger R et al (1996) SETI at the spin-flip line frequency of positronium. Astron Astrophys 306:141–144
Mayr E (1995) A critique of the search for extraterrestrial intelligence. Bioastronomy News 7(3)
McBreen B, Hanlon L (1999) Gamma-ray bursts and the origin of chondrules and planets. Astron Astrophys 351:759–765
McCabe M, Lucas H (2010) On the origin and evolution of life in the galaxy. Int J Astrobiol 9:217–226
McClean D (ed) (2010) World disaster report 2010. International Federation of Red Cross and Red Crescent Societies, Geneva
McGrayne SB (2011) The theory that would not die. YUP, Yale
McInnes CR (2002) The light cage limit to interstellar expansion. J Br Interplanet Soc 55:279–284
McPhee J (1973) The curve of binding energy. Farrar, Straus and Giroux, New York
Melott AL, Lieberman BS, Laird CM, Martin LD, Medvedev MV, Thomas BC, Cannizzo JK, Gehrels N, Jackman CH (2004) Did a gamma-ray burst initiate the late Ordovician mass extinction? Int J Astrobiol 3:55–61

Mereghetti S (2008) The strongest cosmic magnets: soft gamma-ray repeaters and anomalous x-ray pulsars. Astron Astrophys Rev 15:225–287

Mermin ND (1990) Boojums all the way through. CUP, Cambridge

Meshik AP (2005) The workings of an ancient nuclear reactor. Sci Am 293(5):83–91

Messerschmitt DG (2012) Interstellar communication: the case for spread spectrum. Acta Astronaut 81:227–38

Metropolis N (1987) The beginning of the Monte Carlo method. Los Alamos Science (Special issue dedicated to Stanislaw Ulam). pp 125–130

Metzinger T (2003) Being no one: the self-model theory of subjectivity. MIT Press, Cambridge

Meyer M et al. (2013) A mitochondrial genome sequence of a hominin from Sima de los Huesos. Nature 505:403–406

Miller WM Jr (1960) A canticle for Liebowitz. Lippincott, Philadelphia

Minsky M (1973) Talk given at the Communication With Extraterrestrial Intelligence (CETI) conference. In C Sagan (ed) Proceedings of a Conference Held at Byurakan Astrophysical Observatory, Yerevan, USSR, 5–11 Sept 1971 p ix (Cambridge, MA: MIT Press)

Minsky M (1985) Communication with alien intelligence. In Regis E (ed) Extraterrestrials: science and alien intelligence. CUP, Cambridge

Miodownik M (2013) Stuff matters. Viking, London

Mitchell P (1961) Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. Nature 191:144–148

Monod J (1971) Chance and necessity. Collins, London

Monroe T, W R et al. (2013) High precision abundances of the old solar twin HIP 102152: insights on Li depletion from the oldest Sun. Astrophys J Lett 774:L32(6pp)

Moore GE (1965) Cramming more components onto integrated circuits. Electronics 38(8):114–117

Moravec H (1988) Mind children. Harvard University, Cambridge

Morbidelli A, Chambers J, Lunine JI, Petit JM, Robert F, Valsecchi GB, Cyr KE (2000) Source regions and timescales for the delivery of water on earth. Meteor Planet Sci 35:1309–1320

Morrison IS (2012) Detection of antipodal signalling and its application to sideband SETI. Acta Astronaut 78:90–98

Morrison P (1962) Interstellar communication. Bull Phil Soc Wash 16:68–81

Morrison P (2011) Hungarians as Martians: the truth behind the legend. In: Schuch HP (ed) Searching for extraterrestrial intelligence: SETI past, present and future. Springer, Berlin, pp 515–517

Musso P (2011) A language based on analogy to communicate cultural concepts in SETI. Acta Astronaut 68:489–499

Musso P (2012) The problem of active SETI: an overview. Acta Astronaut 78:43–54

NASA (2012) Wilkinson Microwave Anisotropy Probe home page. http://wmap.gsfc.nasa.gov

NASA (2013) Voyager home page. http://voyager.jpl.nasa.gov

Nasar S (1994) A beautiful mind. Simon and Schuster, New York

Newman WI, Sagan C (1981) Galactic civilizations: population dynamics and interstellar diffusion. Icarus 46:293–327

Nicholson A, Forgan D (2013) Slingshot dynamics for self-replicating probes and the effect on exploration timescales. Int J Astrobiol 12:337–344

Niven L (1970) Ringworld. Ballantine, New York

Niven L (1973) Inconstant moon. Gollancz, London

Niven L (1984) Integral trees. Del Rey, New York

Norris RP (2000) How old is ET? In Tough A (ed) When SETI succeeds: the impact of high-information contact. Foundation for the Future, Bellevue, pp 103–105.

Nussinov S (2009) Some comments on possible preferred directions for the SETI search. arXiv:0903.1628v1

O'Leary M (2008) Anaxagoras and the origin of panspermia theory. iUniverse, Bloomington

O'Leary MA et al (2013) The placental mammal ancestor and the post-K-Pg radiation of placentals. Science 339:662–667

Ollongren A (2011) Recursivity in Lingua Cosmica. Acta Astronaut 68:544–8

Ollongren A (2013) Astrolinguistics. Springer, New York

Olson EC (1988) N and the rise of cognitive intelligence on earth. Q J R Astro Soc 29:503–509

OPERA Collaboration (2011) Measurement of the neutrino velocity with the OPERA detector in the CNGS beam. arXiv:1109.4897v1

Oreskes N (2003) Plate tectonics: an insider's history of the modern theory of the earth. Westview, Boulder

Pääbo S (2014) Neanderthal man: in search of lost genomes. Basic, London

Papagiannis MD (1978) Are we all alone, or could they be in the asteroid belt? Q J R Astro Soc 19:236–251

Parthasarathy KR (1988) Obituary: Andreii Nikolaevich Kolmogorov. J Appl Prob 25:445–450

Penny AJ (2004) SETI with SKA. The scientific promise of the SKA (SKA Workshop, Oxford)

Penny AJ (2012) Transmitting (and listening) may be good (or bad). Acta Astronaut 78:69–71

Penrose R (1989) The emperor's new mind. OUP, Oxford

Petigura EA, Howard AW, Marcy GW (2013) Prevalence of earth-size planets orbiting Sun-like stars. Proc Natl Acad Sci. doi:10.1073/pnas.1319909110

Pinker S (1994) The language instinct. Allen Lane, London

Pons M-L, Quitte G, Fujii T, Rosing MT, Reynard F, Moynier F, Douchet C, Albarede F (2011) Early Archean serpentine mud volcanoes at Isua, Greenland, as a niche for early life. Proc Natl Acad Sci. doi:10.1073/pnas.1108061108

Popper K (1963) Conjectures and refutations: the growth of scientific knowledge. Routledge, London

Poundstone W (1988) Labyrinths of reason. Penguin, London

Prantzos N (2013) A joint analysis of the Drake equation and the Fermi paradox. Int J Astrobiol 12:246–253

Prüfer K et al (2013) The complete genome sequence of a Neanderthal from the Altai mountains. Nature 505:43–49

Puthoff HE (1996) SETI, the velocity of light limitation, and the Alcubierre warp drive: an integrating overview. Phys Essays 9:156

Quintana EV (2014) An earth-sized planet in the habitable zone of a cool star. Science 344:277–280

Quiring R et al (1994) Homology of the eyeless gene of Drosophila to the small eye in mice and aniridia in humans. Science 265:785–789

Rampadarath H, Morgan JS, Tingay SJ, Trott CM (2012) The first very long baseline interferometric SETI experiment. Astron J 144(2):38

Rapoport A (1967) Escape from paradox. Sci Am 217(1):50–56

Rasmussen M et al (2011) An Aboriginal Australian genome reveals separate human dispersals into Asia. Science 334:94–98

Reines AE, Marcy GW (2002) Optical SETI: a spectroscopic search for laser emission from nearby stars. Pub Astron Soc Pacific 114:416–426

Reinganum MR (1986–1987) Is time travel impossible? A financial proof. J Portfolio Manage 13(1):10–12

Ridley M (2011) Francis Crick. Harper Perennial, London

Rogers LJ (1997) Minds of their own. Westview, Boulder

Rood RT, Trefil JS (1981) Are we alone? Charles Scribner's, New York

Rose C, Wright G (2004) Inscribed matter as an energy-efficient means of communication with an extraterrestrial civilization. Nature 431:47–9

Rouse Ball WW (1908) A short account of the history of mathematics. Dover, New York

Roy KI, Kennedy RG III, Fields DE (2013) Shell worlds. Acta Astronaut 82:238-245

Royal S (2004) Nanoscience and nanotechnologies. Royal Society, London

Rummel JD (2001) Planetary exploration in the time of astrobiology: protecting against biological contamination. Pub Natl Acad Sci 98:2128–2131

Rushby AJ, Claire MW, Osborn H, Watson AJ (2013) Habitable zone lifetimes of exoplanets around main sequence stars. Astrobiology 13:833–849

Saberhagen F (1967) Berserker. Ballantine, New York

Sagan CE (1985) Contact. Simon and Schuster, New York

Sandberg A, Armstrong S, Ćirković MM (2014) That is not dead which can eternal lie: what are the physical constraints for the aestivation hypothesis? Preprint

Savage-Rumbaugh S, Lewin R (1996) Kanzi: the ape at the brink of the human mind. Wiley, New York

Scarborough Borough Council (2012) Election results, Stakesby Ward of Whitby Town Council. http://democracy.scarborough.gov.uk/mgElectionAreaResults.aspx?ID=91&RPID=0. Accessed 20 Jan 2014

Scheffer LK (1993) Machine intelligence, the cost of interstellar travel and Fermi's paradox. Q J R Astro Soc 35:157–175

Schick KD, Toth N (1993) Making silent stones speak: human evolution and the dawn of technology. Simon and Schuster, New York

Schmidt GR, Landis GA, Oleson SR (2012) Human exploration using real-time robotic operations (HERRO): a space exploration strategy for the 21st century. Acta Astronaut 80:105–113

Schroeder K (2002) Permanence. Tor, New York

Schwamb ME et al (2013) Planet hunters: a transiting circumbinary planet in a quadruple star system. Astrophys J 768:127 (21pp)

Schwartz RN, Townes CH (1961) Interstellar and interplanetary communication by optical masers. Nature 190:205–208

Seager S (2013) Exoplanet habitability. Science 340:577–581

Searle JR (1984) Minds, brains and programs. Harvard University, Cambridge

Secker J, Wesson PS, Lepock JR (1996) Astrophysical and biological constraints on radiopanspermia. J R Astro Soc Canada 90:184–192

Segré E (1970) Enrico Fermi: physicist. UCP, Chicago

Selsis F, Kasting JF, Levard B, Paillet J, Ribas I, Delfosse X (2007) Habitable planets around the star Gliese 581? Astron Astrophys 476:1373–1387

SETI@home (2000) SETI@home poll results. http://boinc.berkeley.edu/slides/xerox/polls.html. Accessed 14 Feb 2014

SETI@home (2013) SETI@home homepage. www.SetiAtHome.ssl.berkeley.edu. Accessed 14 Feb 2014

SetiLeague (2013) Ask Dr SETI www.setileague.org/askdr/index.html. Accesssed 14 Feb 2014

Sharov AA, Gordon R (2013) Life before earth. arXiv:1304.3381

Shaw B (1975) Orbitsville. Gollancz, London

shCherbak VI, Makukov MA (2013) The "Wow! signal" of the terrestrial genetic code. Icarus 224:228–242

Sheaffer R (1995) An examination of claims that extraterrestrial visitors to earth are being observed. In Zuckerman B, Hart MH (eds) Extraterrestrials: where are they? Cambridge, CUP

Sheehan W (1996) The planet Mars: a history of observation and discovery. University of Arizona, Tucson

Shklovsky IS, Sagan C (1966) Intelligent life in the universe. Holden-Day, San Francisco

Siemion APV, von Korff J, McMahond P, Korpela E, Werthimer D, Anderson D, Bowera G, Cobb J, Foster G, Lebofsky M, van Leeuwen J, Wagner M (2010) New SETI sky surveys for radio pulses. Acta Astronaut 67:1342–1349

Siemion APV, Demorest P, Korpela E, Maddalena RJ, Werthimer D, Cobb J, Howard AW, Langston G, Lebofsky M, Marcy GW, Tarter J (2013) A 1.1 to 1.9 GHz SETI survey of the Kepler field: I. A search for narrow-band emission from select targets. Astrophys J 767:94

Sieveking A (1979) The cave artists. Thames and Hudson, London

Silagadze ZK (2008) SETI and muon collider. Acta Phys Polonica B 39:2943–2948

Smart JM (2012) The transcension hypothesis: sufficiently advanced civilizations invariably leave our universe, and implications for METI and SETI. Acta Astronaut 78:55–68

Smith A (2005) Moondust: in search of the men who fell to earth. Bloomsbury, London

Smith RD (2009) Broadcasting but not receiving: density dependence considerations for SETI signals. Int J Astrobiol 8:101–105

Smolin L (1997) The life of the cosmos. Weidenfeld and Nicolson, London

Sobral D, Smail I, Best PN, Geach JE, Matsuda Y, Stott JP, Cirasuolo M, Kurk J (2013) A large, multi-epoch H$\alpha$ survey at $z = 2.23$, 1. 47, 0.84 & 0.40: the 11 Gyr evolution of star-forming galaxies from HiZELS. Mon Not R Astro Soc 428:1128–1146

Sorby HC (1877) On the structure and origin of meteorites. Nature 15:495–498

Soressi M et al. (2013) Neandertals made the first specialized bone tools in Europe. Proc Natl Acad Sci 110:14186–14190

Spiegel DS, Turner EL (2012) Bayesian analysis of the astrobiological implications of life's early emergence on earth. Proc Natl Acad Sci 109:395–400

Stapledon O (1930) Last and first men. Methuen, London

Stapledon O (1937) Star maker. Methuen, London

Stauffer D (1985) Introduction to percolation theory. Taylor and Francis, London

Stephenson DG (1978) Extraterrestrial cultures within the solar system? Q J R Astro Soc 19:277–281

Stevenson DJ (2003) Mission to earth's core—a modest proposal. Nature 423:239–240

Story R (1976) The space gods revealed. Barnes and Noble, New York

Stringer C (2012) The origin of our species. Penguin, London

Sullivan WS (1964) We are not alone. Pelican, London

Sullivan WT III, Baross J (eds) (2007) Planets and life: the emerging science of astrobiology. CUP, Cambridge

Sullivan WT III, Brown S, Wetherill C (1978) Eavesdropping: the radio signature of the earth. Science 199:377–388

Tarter J (2001) The search for extraterrestrial intelligence (SETI). Ann Rev Astron Astrophys 39:511–548

Tarter J et al. (2011) The first SETI observations with the Allen telescope array. Acta Astronaut 68:340–346

Tattersall I (1998) Becoming human. OUP, Oxford

Tattersall I (2000) Once we were not alone. Sci Am 282(1):56–62

Taylor SR (1998) Destiny or chance. CUP, Cambridge

Tegmark M, Wheeler JA (2001) 100 years of the quantum. Sci Am 284(2):68–75

Teilhard deCP (2004) The future of man. Image, London

Thomas B (2009) Gamma-ray bursts as a threat to life on earth. Int J Astrobiol 8:183–186

Thorne K (1994) Black holes and time warps. Norton, New York

Tipler FJ (1980) Extraterrestial intelligent beings do not exist. Q J R Astro Soc 21:267–281

Tipler FJ (1994) The physics of immortality. Anchor, New York

Turco RP et al (1983) Nuclear winter: global consequences of multiple nuclear explosions. Science 222:1283–1297

Turnbull MC, Tarter J (2003a) Target selection for SETI. I. A catalog of nearby habitable stellar systems. Astrophys J Supp 145:181–198

Turnbull MC, Tarter J (2003b) Target selection for SETI. II. Tycho-2 dwarfs, old open clusters, and the nearest 100 stars. Astrophys J Supp 149:423–436

Ulam SM (1958a) On the possibility of extracting energy from gravitational systems by navigating space vehicles. Report LA-2219-MS. (Los Alamos, NM: Los Alamos National Laboratory)

Ulam SM (1958b) Tribute to John von Neumann, 1903–57. Bull Am Math Soc 64:1–49

Ulam SM (1976) Adventures of a mathematician. University of California, Berkeley

Vaidya PG (2007) Are we alone in the multiverse? arXiv:0706.0317v1

Vakoch DA (2011) Asymmetry in active SETI: a case for transmissions from earth. Acta Astronaut 68:476–488

Valley JW et al. (2014) Hadean age for a post-magma-ocean zircon confirmed by atom-probe tomography. Nat Geosci 7:219–223

Van Den Broeck C (1999) A "warp drive" with more reasonable total energy requirements. Class Quantum Gravity 16:3973–3979

Vedrenne G, Atteia J-L (2009) Gamma-ray bursts: the brightest explosions in the universe. Springer, Berlin

Venter C (2013) Life at the speed of light: from the double helix to the dawn of digital life. Viking, New York

Vidal C (2013) The beginning and the end: the meaning of life in a cosmological perspective. PhD thesis, Vrije Universiteit Brussel.

Viet L, Nieder A (2013) Abstract rule neurons in the endbrain support intelligent behaviour in corvid songbirds. Nat Commun 4. doi:10.1038/ncomms3878

Viewing D (1975) Directly interacting extra-terrestrial technological communities. J Br Interplanet Soc 28:735–744

Vinge V (1993) VISION-21 Symposium (NASA Lewis Research Center)

Visalberghi E, Trinca L (1989) Tool use in capuchin monkeys: distinguishing between performing and understanding. Primates 30:511–521

Vladilo G, Murante G, Silva L, Provenzale A, Ferri G, Ragazzini G (2013). The habitable zone of earth-like planets with different levels of atmospheric pressure. Astrophys J 76:65 (23 pp)

von Däniken E (1969) Chariots of the gods. Souvenir, London

von Däniken E (1972) The gold of the gods. Bantam, New York

von Däniken E (1997) The return of the gods. Element, London

von Eshleman R (1979) Gravitational lens of the Sun: its potential for observations and communications over interstellar distances. Science 205:1133–1135

von Hoerner S (1975) Population explosion and interstellar expansion. J Br Interplanet Soc 28:691–712

Vonnegut K (1963) Cat's cradle. Holt, Rinehart and Winston, New York

vos Savant M (1990) "Game show problem". marilynvossavant.com. Accessed 30 May 2014

Voyager (2013) HPL home page. www.jpl.nasa.gov/index.cfm

Vukotić B, Ćirković MM (2012) Astrobiological complexity with probabilistic cellular automata. Orig Life Evol Biosph 42:347–371

Walker J, Hays P, Kasting J (1981) A negative feedback mechanism for the long-term stabilization of the earth's surface temperature. J Geophys Res 86:9776–9782

Waltham D (2014) Lucky planet. Icon, London

Ward PD, Brownlee D (1999) Rare earth. Copernicus, New York
Watson AJ (2008) Implications of an anthropic model of evolution for emergence of complex life and intelligence. Astrobiology 8:175–185
Watson JD (2010) The double helix: a personal account of the discovery of the structure of DNA. Phoenix, London
Watts P (2006) Blindsight. Tor, New York
Weart SR (2008) The discovery of global warming: revised and expanded edition. Harvard University, Cambridge
Webb S (1999) Measuring the universe. Springer, Berlin
Webb S (2004) Out of this world. Praxis, Chichester
Webb S (2012) New eyes on the universe. Springer, New York
Webb S (2014) Ripples from the start of time? In: Mason J (ed) Patrick Moore's yearbook of astronomy 2015. Macmillan, London, pp 243–265
Weinberg S (1993) Dreams of a final theory: the search for the fundamental laws of nature. Hutchinson, London
Weisberg JM, Taylor JM (2005) The relativistic binary pulsar B1913+16: thirty years of observations and analysis. In Rasio F A and Stairs I H (eds) Astron Soc Pacific Conf Series 328, p. 25
Weisman A (2007) The world without us. Picador, New York
Welch J et al. (2009) The Allen Telescope Array: the first widefield, panchromatic, snapshot radio camera for radio astronomy and SETI. Proc IEEE Spec Issue Adv Radio Telesc 97:1438–47
Wells HG (1898) War of the worlds. Heinemann, London
Wells W (2009) Apocalypse when? Praxis, Chichester
Wesson PS (1990) Cosmology, extraterrestrial intelligence, and a resolution of the Fermi–Hart paradox. Q J R Astro Soc 31:161–170
Wesson PS (2010) Panspermia, past and present: astrophysical and biophysical conditions for the dissemination of life in space. Space Sci Rev 156:239–252
Whates I (2014) Paradox: stories inspired by the Fermi paradox. NewCon Press, Cambridgeshire
Whitmire DP,Wright DP (1980) Nuclear waste spectrum as evidence of technological extraterrestrial civilizations. Icarus 42:149–156
WHO (2013) Urban population growth. http://www.who.int/gho/urban_health/situation_trends/urban_population_growth_text/en/
Wiechert U, Halliday AN, Lee D-C, Snyder GA, Taylor LA, Rumble D (2001) Oxygen isotopes and the moon-forming giant impact. Science 294:345–358
Wigner E (1960) The unreasonable effectiveness of mathematics in the natural sciences. Commun Pure Appl Math 13(1):1–14
Wiley KB (2011) The Fermi paradox, self-replicating probes, and the interstellar transportation bandwidth. arXiv:1111.6131v1
Williams IP, Cremin AW (1968) A survey of theories relating to the origin of the solar system. Q J R Astro Soc 9:40–62
Williamson T (1994) Vagueness. Routledge, London

Witze A (2014) Icy Enceladus hides a water ocean. Nature. doi:10.1038/nature.2014.14985

Woese CR, Kandler O, Wheelis ML (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci U S A 87:4576–4579

Worth RJ, Sigurdsson S, House CH (2013) Seeding life on the moons of the outer planets via lithopanspermia. Astrobiology 13:1155–1165

Yeomans DK (2012) Near-earth objects: finding them before they find us. Princeton University, Princeton

Yokoo H, Oshima T (1979) Is bacteriophage phi X174 DNA a message from an extraterrestrial intelligence? Icarus 38:148–153

Zahnle K (2001) Decline and fall of the Martian empire. Nature 412:209–213

Zaitsev A (2006) The SETI paradox. Bull Spec Astrophys Obs 60

Zaitsev A (2012) Classification of interstellar radio messages. Acta Astronaut 78:16–19

Zalasiewicz J (2009) The earth after us: what legacy will humans leave in the rocks? OUP, Oxford

Zhang J, Dauphas N, Davis AM, Leya I, Fedkin A (2012) The proto-earth as a significant source of lunar material. Nat Geosci 5:251–255

Zuckerman B (1985) Stellar evolution: motivation for mass interstellar migration. Q J R Astro Soc 6:56–59

Zuckerman B, Hart MH (eds) (1995) Extraterrestrials—where are they? CUP, Cambridge

# Index





S. Webb, *If the Universe Is Teeming with Aliens . . . WHERE IS EVERYBODY?*,
Science and Fiction, DOI 10.1007/978-3-319-13236-5