# Chapter 9

# Communication with Extraterrestrial Life

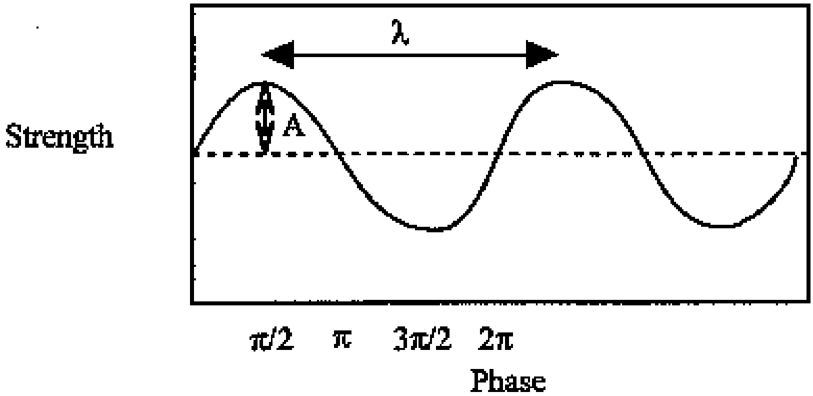
In this chapter, we discuss searches for extraterrestrial intelligence (or SETI). What is the most reasonable way to communicate with extraterrestrial life? In asking this question, we are also asking what means another civilization would choose because both civilizations must choose the same medium of communication. Consequently, it is important to use general considerations in order to avoid too much Earth chauvinism.

Any form of communication involves sending something between the two civilizations. In the next chapter, we will consider the possibility of sending machines or even humans to other stars. We will find that it is extremely difficult and expensive. As a more modest goal, let us consider using any particle with mass to communicate. Using the equation for the kinetic energy given in Chapter 2, we find that the energy required to accelerate a particle of mass m to a velocity v is at least $1/2\ mv^2$ (if v is an appreciable fraction of the speed of light, more energy is required). Suppose we wanted to send the lightest particle, an electron, to the nearest star and get an electron back in a human lifetime. If we allow 80 years for the round-trip, we could accelerate the electron quickly to $v = 0.1\ c$ (1/10 of the speed of light) and get to the nearest star (4 ly away) in 40 years. The energy required to send one electron is $4.1 \times 10^{-9}$ ergs. For comparison, a photon always travels at the speed of light, so we could send a photon to the nearest star in 4 years; the energy of this photon is $h\nu$, where $\nu$ is the frequency of the photon. For a photon with a frequency of $1 \times 10^9$ Hz, a likely frequency for communication, the energy required is $6.6 \times 10^{-18}$ ergs, nearly a billion times less than the energy needed to send an electron. Clearly photons are much cheaper.

Of course we would have to send more than one particle, whether massive or, like the photon, massless, in order to communicate a message. Also, most photons would not fall on the target. Consequently, we would expend quite a lot of energy. Consider a very powerful transmitter, which might produce 100 Megawatts of power. To run this transmitter for a year would take $8 \times 10^8$ kilowatt-hours; a typical cost for electricity is \$0.05 per kilowatt-hour, giving us a total cost per year of 40 million dollars, a substantial sum, but not impossible to imagine. In contrast, it has been estimated that a round trip journey for a small spacecraft to the nearest star, if traveling at 0.7 c, would consume at least $10^{18}$ kilowatt-hours, enough to power the entire United States for 100,000 years, and cost $\$5 \times 10^{16}$ (that's 50,000 trillion dollars!). Whatever arguments may be made for interstellar travel, it is clear that communication by sending photons is cheaper!

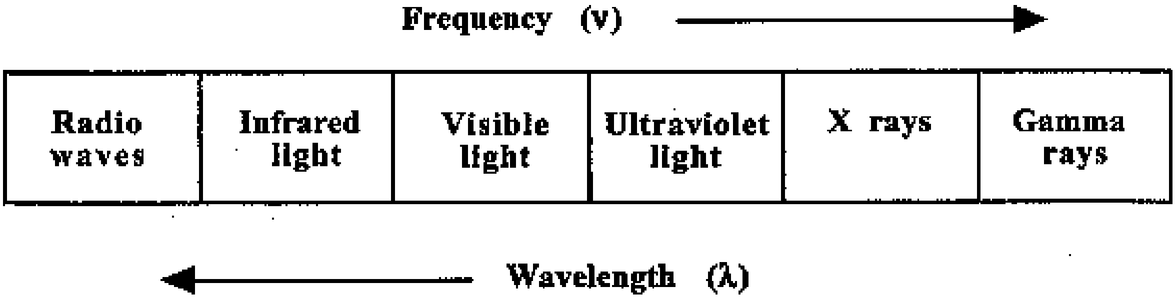
## A. Choice of Wavelength Region

When we speak of photons, we are also talking about electromagnetic radiation. Electromagnetic radiation can be thought of either as a stream of photons or as a wave in which electric and magnetic fields oscillate as they move through space. The wave picture will be more useful for the following discussion. A wave can be characterized by four quantities: its amplitude, its speed, its wavelength, and its frequency. The amplitude (denoted Strength in the figure below) measures how strong the wave is. In the case of an electromagnetic wave, the amplitude measures the strength of the electric field. At right angles to the electric field, a magnetic field also oscillates. Both the fields are perpendicular to the direction that the wave is traveling. An electromagnetic wave travels at the speed of light, $c = 3 \times 10^{10}$ cm $s^{-1}$. The wavelength (denoted $\lambda$ in the figure) is the distance between adjacent crests of the wave. Another useful concept is the phase, which measures how far along the pattern the amplitude has changed. If we define the phase to be zero at the left side of this figure, it will increase to $\pi/2$ at the first peak, $\pi$ at the next point where it crosses zero, and $2\pi$ at the third point where it crosses zero, after one full wavelength.



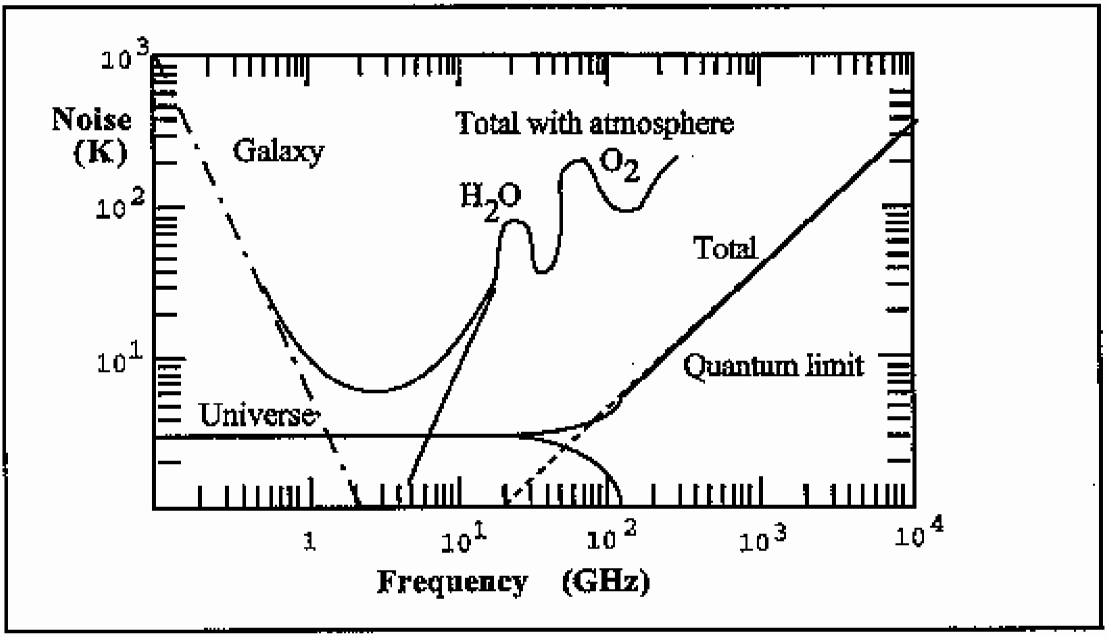
The frequency is the number of times per second the field passes its peak value, measured at a fixed point in space. Frequency is measured in cycles per second, which has been given the name hertz (abbreviated Hz). The unit is named for Heinrich Hertz, who demonstrated the existence of electromagnetic waves in 1887. The current coming out of a wall socket has a frequency of 60 Hz in the U.S. Most of the frequencies we will be discussing are much higher, so we use some convenient units: 1 kilohertz (kHz) is 1000 Hz; 1 megahertz (MHz) is $10^6$ Hz; 1 gigahertz (GHz) is $10^9$ Hz. An important fact about light waves is that the speed is related to the wavelength ($\lambda$) and frequency ($\nu$) through the following equation: $c = \lambda \nu$.

This equation means that higher frequency implies shorter wavelength, since the product must remain constant. Thus the plot of the electromagnetic spectrum below shows frequency increasing to the right, but wavelength increasing to the left. The divisions into different wavelength regions are purely arbitrary, though convenient; all electromagnetic waves are the same phenomenon. The divisions between regions reflect different technologies, or biologies. For example, the visible region just corresponds to the range of wavelengths to which our eyes respond, a very small fraction of the total spectrum.



It would be extremely expensive for a civilization to broadcast over this entire range of wavelengths, so we should consider arguments that would allow us to narrow the range, bearing in mind that another, possibly much more advanced, civilization must come to the same conclusions. A satisfactory communication wavelength should be able to travel large distances without unnecessary loss. All the wavelengths shorter than infrared are somewhat attenuated by interstellar clouds. In addition, the energy required to send a photon increases with frequency. These general arguments would tend to favor low frequencies, or long wavelengths, like infrared or radio.

A second consideration is making the signal detectable in the presence of noise. Even if a signal is not attenuated by interstellar clouds, it will decrease in amplitude by the square of the distance traveled. Consequently, even very strong signals will have become very weak by the time they reach another civilization. Just as it becomes more difficult to pick up a radio station as you drive farther from the transmitter, and it is overwhelmed by the hissing noise that we hear between stations, so the signal from a distant civilization may be difficult to distinguish from noise. Since "noise" has a technical meaning here, we should be clear about what we mean. We do not imply any judgment about the quality or content of the signal, only about whether it is the signal we are looking for. If we are looking for extraterrestrial signals, any terrestrial signal, be it a rock music station, a TV signal, or a defense radar, is noise. These types of noise are usually given the name "interference," since they interfere with our detection of the desired signal. In addition to artificial noise, there is natural noise, which is distinguished by its randomness. Were we to listen to any natural noise, we would hear a hissing sound, like that of the radio tuned between stations.



Where does this natural noise come from? There are many possible sources of noise; to avoid Earth chauvinism, we will consider first only noise sources that will affect any civilization in our Galaxy. First, we have noise from the Galaxy itself, caused by fast-moving electrons spiraling around magnetic fields. This noise is strongest at low frequencies, as shown on the plot above. Second, the Universe itself produces noise at all frequencies, but especially up to about 100 GHz. This noise is the radiation from the Big Bang itself, and its detection is the strongest evidence that the Big Bang actually happened. Third, radio receivers produce noise. Even if we have built the perfect receiver, the laws of quantum mechanics imply that it will produce a certain minimum noise, essentially the energy of one photon (hν). This minimum noise level, called the quantum limit, increases with the frequency being received, dominating the noise from the Universe at frequencies above about 30 GHz. Note that the noise is measured in kelvins by convention. For example, the line showing the quantum limit for receiver noise is $T=h\nu/k$.

Considering these three noise sources alone, we see that there is a broad valley of minimum noise between about 1 and 100 GHz, corresponding to wavelengths of 3 mm to 30 cm, a part of the radio region of the electromagnetic spectrum. Since broadcasting in this range of wavelengths would allow another civilization to have the best chance of picking up a weak signal, most scientists argue that 3 mm to 30 cm is the logical wavelength region for interstellar communications.

Not everyone accepts the argument that radio wavelengths are the only choice. Charles Townes has pointed out that sending pulses of light from lasers is an alternative method, which would allow the visible or infrared region to be competitive with the radio region. Indeed there are now a number of searches at the wavelengths of visible light, often referred to as optical SETI. One issue to consider is that the signal from the civilization must be distinguishable from the light from the star itself because we cannot spatially resolve planets from their stars. While stars emit radio waves, it is quite easy for transmitters to dominate the radio emission from stars. In contrast, stars radiate most of their energy at visible and near-infrared wavelengths. Nonetheless, powerful lasers that emit at very sharply defined wavelengths can be detected in the spectra of stars. A paper published in 2002 by Reines and Marcy reported a search for artificial signals in the stellar spectra that were used to search for planetary companions of 577 nearby stars (Chapter 3). They could have detected lasers with powers exceeding 60 kW that were focused by 10-m diameter telescopes toward the Earth. This scenario is very unlikely unless the other civilization already knows about our existence because the beam of such a signal is very small and few stars would be covered. No signals were found.

For the remainder of the chapter, we concentrate on radio techniques. Even the range from 1 to 100 GHz is a very big range of frequencies. Can we narrow it down still more? If we include the effects of the Earth's atmosphere, we see that the noise increases above about 10 GHz; this noise is caused by the presence of water and oxygen molecules in our atmosphere, so any planet like Earth would be likely to have these molecules in its atmosphere. On the other hand, broadcasting from space would remove this noise, so it is less certain that we can restrict searches to 1 to 10 GHz, even if we accept a little Earth chauvinism. Nonetheless, essentially all searches so far have concentrated on this range, often called the microwave region, essentially for historical reasons.

## B. Choice of Radio Frequency

When we examine the searches that have been made, we will find that most have concentrated on a very narrow range of frequencies. This fact reflects technological limitations, which are gradually being overcome. It is still of interest to examine the rationale for the choice of frequencies; this process is sometimes called the search for a "magic frequency."

Morrison and Cocconi suggested the first magic frequency in 1959. It is 1420 MHz, or 1.42 GHz, the frequency at which H atoms in space emit and absorb radiation. Before the discovery of interstellar molecules, this frequency seemed to be quite a unique choice. Since H is the most abundant atom in the Universe, one can still make a case for it, but it is less convincing now that over 100 molecules, some crucial for life, have been discovered to be emitting or absorbing radiation at many frequencies.

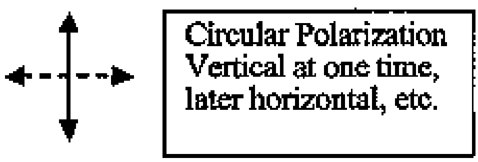
In fact, the first molecule discovered at radio frequencies led to another suggestion. In 1963, the molecular fragment, OH, was discovered in interstellar space by its absorption of radio waves at 4 frequencies between 1.612 and 1.720 GHz. Since H and OH combine to make $H_2O$, vital to life, the range of frequencies between 1.42 and 1.612 GHz was suggested as a range to be searched by Carl Sagan and Frank Drake, who dubbed this range the "water hole." They argued that species on Earth have always met around water holes, and interstellar species might do likewise. These suggestions were both made when research on H and OH was quite active; while they continue to be studied, attention has shifted to a considerable extent to higher frequencies where many more molecules can be studied. Thus these suggestions may be criticized as somewhat biased by the research interests of the time. I have never found them convincing.

A more universal frequency would be one that is constructed, in some unique way, from physical constants that would be well known to any technological civilization. To get a frequency, we would need to divide a constant with units of speed by a constant with units of distance. The only obvious choice for the speed is c, the speed of light. There are several choices for distance, but they are all related to each other by powers of a dimensionless quantity called the fine structure constant. Unfortunately, dividing the speed of light by any of these distances produces a frequency far above the radio region. Tom Kuiper and Mark Morris suggested that we simply multiply these ratios by the fine structure constant (which is approximately 1/137) enough times to get a frequency that is in the radio region. If we multiply by the fine structure constant 5 times, the resulting frequency is about 2.5568 GHz, which we will call the Kuiper-Morris frequency.

While each of these suggestions has some merit, none has proven so compelling that all terrestrial astronomers agree. Thus, it seems like wishful thinking to believe that extraterrestrial astronomers should have all agreed on one or the other.

## C. Radio Techniques

Since almost all searches for extraterrestrial communication so far have been conducted at radio wavelengths, it is worth spending some time to understand the techniques that are used. The basic techniques of radio astronomy, radio communications, television, mobile telephones, etc., are the same. The information is transmitted by radio waves, which are just the low frequency end of the electromagnetic spectrum. When the radio waves reach an antenna (think of the long straight antenna on a car for example), the electric field, which is oscillating up and down at the frequency of the wave, causes electrons in the antenna to go up and down at the same frequency. This oscillating motion of the electrons is what we call an alternating current; it can be amplified by electronic devices. If the antenna is oriented vertically, it responds to vertically polarized radio waves, meaning that the electric field is oscillating up and down. If the oscillation is side to side, we call it horizontal polarization and the antenna must be oriented horizontally to pick it up. By combining vertical and horizontal polarization with the right phase shift between the peaks, one can also create circular polarization, which can either be right or left handed, depending on which way the electric field rotates around the direction of wave motion.



To detect weak signals, it helps to add the signals from more than one antenna, and some radio telescopes are just collections of antennas. However, at the frequencies from 1–10 GHz, it is more common to use a parabolic reflecting surface to collect waves from a large area and focus it onto a single antenna, after which it can be amplified as usual. Such devices are used as satellite dishes for TV reception, and larger versions are used for radio astronomy, where they are called radio telescopes.

The largest radio telescope in the world is near Arecibo, Puerto Rico. It is 300 m across, the size of three football fields laid end to end. It is built into a depression in the ground that happens to have nearly the right shape. This circumstance is what allowed such a large telescope to be built, but it also limits its flexibility, since the telescope cannot move, thereby restricting it to looking more or less straight up. Rotation of the Earth allows the telescope to sweep out a swath of the sky every 24 hours and some flexibility is provided by a movable secondary mirror, which allows the telescope to look in directions slightly different from straight up. This telescope was to have played a major role in the NASA search for signals, which we will describe below. After Arecibo, the next largest telescopes are in the 100 m category, but most of these can be pointed in any direction, so the part of the sky that they can see is limited only by their location on the Earth (a telescope located in northerly latitudes cannot see all of the southern skies and vice versa). A major new telescope of this size began operation in West Virginia in 2001.

After receiving and amplifying the radio waves, astronomers must decode whatever information they contain. For the naturally created signals studied by radio astronomers, the information is contained only in the amplitude of the signals and how the amplitude varies with different frequencies, directions, times, and polarizations. Indeed, most natural signals do not vary in time and are unpolarized. Artificial signals can have more complex information encoded into them. Suppose for example that the transmitter increases or decreases in amplitude. This is called amplitude modulation, since the transmitter is modulating, or changing slightly, the amplitude. This is the method used for AM (or Amplitude Modulation) radio. Another method is to use FM (Frequency Modulation) in which the frequency of the wave is altered slightly. The time scale of these changes is slow enough that there are many cycles of the basic frequency before any change occurs. The net result is that the signal is spread slightly in frequency around the basic frequency of the wave, called the carrier frequency.

For signals with information encoded into them, the electronics must contain devices that decode the information. So far most of the encoding and decoding have been done by analog electronics, but communication technology is switching to digital electronics. In this system, all the information is reduced to binary numbers, strings of 1s and 0s, or bits. For example, the amplitude and frequency of a particular sound can be represented by numbers. This method is already used in compact disk players and is spreading to other communications media. The primary advantage of digital techniques is that noise and distortion are much less of a problem. Each 1 or 0 is ultimately represented by a voltage level, which can be quite different; consequently, a small error in the signal will not cause the electronics to mistake a 1 from a 0, whereas any small error on an analog signal introduces noise or distortion.

A simple form of amplitude modulation, which is essentially digital, is to send pulses. The transmitter is turned on and off, or the amplitude is changed at regular intervals, with one amplitude representing 1 and another representing 0. This is the kind of signal that most people expect for extraterrestrial communication.

## D. Recognizing and Decoding the Message

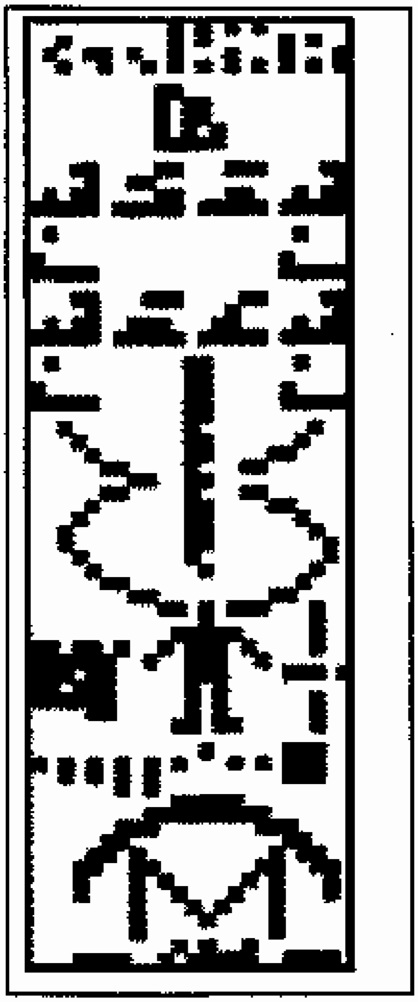
If a pulsed signal arrived at a radio telescope used only for astronomy, it might not be recognized, since most observations are not set up to detect variations on short time scales. An exception is the study of interstellar scintillation, which observes variations in the signals from very distant galaxies caused by fluctuations in the density of electrons in the interstellar gas in our Galaxy. These changes are random, but occur on fairly short time scales. In the early 1960s, scientists in the United Kingdom were studying these variations, and one of the graduate students, Jocelyn Bell, noticed a repeated pattern of rapid, regular variations. Upon closer examination, these turned out be extremely regular pulses. Since no natural phenomenon was known to produce pulses, these were jokingly called LGMs, or Little Green Men. However, it soon became clear that no information was encoded in the pulses, and a natural explanation was found for these LGMs, or pulsars, as they are now known. They are rapidly spinning neutron stars, the ultradense remnants of supernova explosions, whose intense magnetic fields create beamed radio waves. As the star spins, the beam of radio waves sweeps past the direction of the Earth, and we see a pulse, analogous to the regular flash of light from the rotating beam of a lighthouse.

This example indicates how a signal might be recognized even if we were not looking for it, as long as the equipment is suitable. Another example is provided by the case of the cosmic masers. A maser is a device that amplifies microwaves, in the same way that a laser can amplify light. Invented in the 1950s, masers were used to build extremely sensitive radio receivers, although they have now been largely superseded by other devices. The discovery of cosmic masers came in 1965, when radio astronomers were studying absorption from interstellar OH molecules, which had been discovered a few years earlier. When the telescopes were pointed toward certain regions, emission, rather than absorption, was seen. Other peculiarities attracted considerable attention, and it was soon discovered that the source of the emission was too compact for the emission to be produced by normal processes. While no one suggested extraterrestrial signals, it is not totally implausible that an advanced civilization might attract attention by doing something peculiar to a naturally occurring phenomenon, ensuring that scientists from other civilizations would study it. The emission later proved to be the result of a naturally occurring maser. For quite different reasons, I studied the nature of the emission in a way that would have revealed whether any information was encoded into it. None was found.

These two examples indicate that it is unlikely that we have already received a signal and not recognized it; in both cases, follow-up studies that would have revealed information were performed. Nonetheless, many signals may be present that have not been examined at all, because they do not fall into the range of any radio astronomical investigation. To find these signals will require a dedicated search strategy.

Even after a signal is found, there is the problem of decoding it. We usually think of codes as trying to hide information, and the science of constructing codes that are hard to break is called cryptography. For interstellar communication, we have the opposite goal: to construct a signal that is easy to decode, an effort that we might call anticryptography. A little thought will convince you that this is no easy task when we know nothing about the language or psychology of the recipient. Most suggestions revolve around mathematics, in which the meaning of symbols can be decoded by the relationships between them.

The most popular suggestion for an opening message is a picture, which could convey an immediate image of certain facts about the sender. The picture shown below was broadcast in the only known deliberate attempt to send a signal into space. In the 1970s, Frank Drake and Carl Sagan used the Arecibo telescope to broadcast a series of pulses toward the globular cluster, M13. Since this cluster is very far away and composed of stars with few heavy elements, this was not a serious attempt to communicate, but rather a demonstration, which we will use as an example of how a message might be encoded.



How do we get a picture from a series of 1s and 0s? The simplest way is to use the 1s and 0s to represent light and dark, with the pattern of 1s and 0s building up a black and white picture. But how do we know how to arrange a string of 1s and 0s to make a picture? We need to tell the receiving civilization where a new line begins or, equivalently, how many rows and columns are in the picture. Drake and Sagan did this by making the total number of bits the product of two prime numbers. They sent 1679 bits in their message and then repeated it over and over. Any receiving civilization could recognize that the pattern repeated and identify the 1679 bits as the message. Since 1679 is the product only of two prime numbers, 23 and 73, one could deduce that the message was a two dimensional grid with 23 rows and 73 columns, or vice versa. It would then be a simple matter to try both ways to see which version made a sensible picture. In this case, the pattern with 23 columns and 73 rows makes the picture to the left. A simple variation would be to make the total number of bits be the product of three prime numbers; then the third number could be the frame number for a movie made of successive two-dimensional images.

While the picture method seems plausible, we cannot be sure that another civilization will choose it. Even if we receive and reconstruct a picture, understanding what the other civilization intended by the picture may be difficult. It is interesting to see what parts of the picture you can understand. Drake and Sagan asked experts in various fields to try to interpret the picture that they sent. They found that experts in biology could understand the part of the picture portraying terrestrial biology, for example. It is not as clear that an alien would have the same success. In addition, further messages must establish a code for communicating other information. If we accidentally intercepted a signal that was not designed to be easy to decode, we might well not be able to decipher it.

## E. Past and Future Searches

Before describing the history and future prospects for SETI, let us consider some possible situations. If an advanced civilization is trying to contact emergent civilizations like ours, they may have set up beacons to attract attention. In the absence of knowledge of where and when new civilizations are arising, these beacons would have to broadcast in all directions at all times. The advanced civilization might also choose to broadcast at several frequencies to increase the chances that a rather primitive civilization might pick up the signal. To keep a powerful, omnidirectional beacon operating for extended periods would require enormous energy sources, so we can be sure that this civilization would be very advanced.

The problem would be greatly simplified if the advanced civilization knew where to look. Large radio telescopes collect the energy into small beams, so that much more power can be concentrated on the intended target. Existing transmitters on the Arecibo telescope are powerful enough to communicate with a similar telescope across much of the Galaxy. At the same time, the small beam selects only stars lying within that beam as possible recipients. The problem is knowing where to look or to transmit.

One attractive way to solve the problem is to detect unintentional signals (called leakage radiation because the radiation leaks through the Earth's ionosphere) from the other civilization. We are probably not yet capable of doing this, but another civilization might be. What kind of leakage signals has the Earth produced? The first leakage radiation from the Earth was produced in the late 1930s when frequencies high enough to escape the ionosphere began to be used, but the real increase in leakage radiation occurred after World War II, with the rapid expansion of television and defense radars.

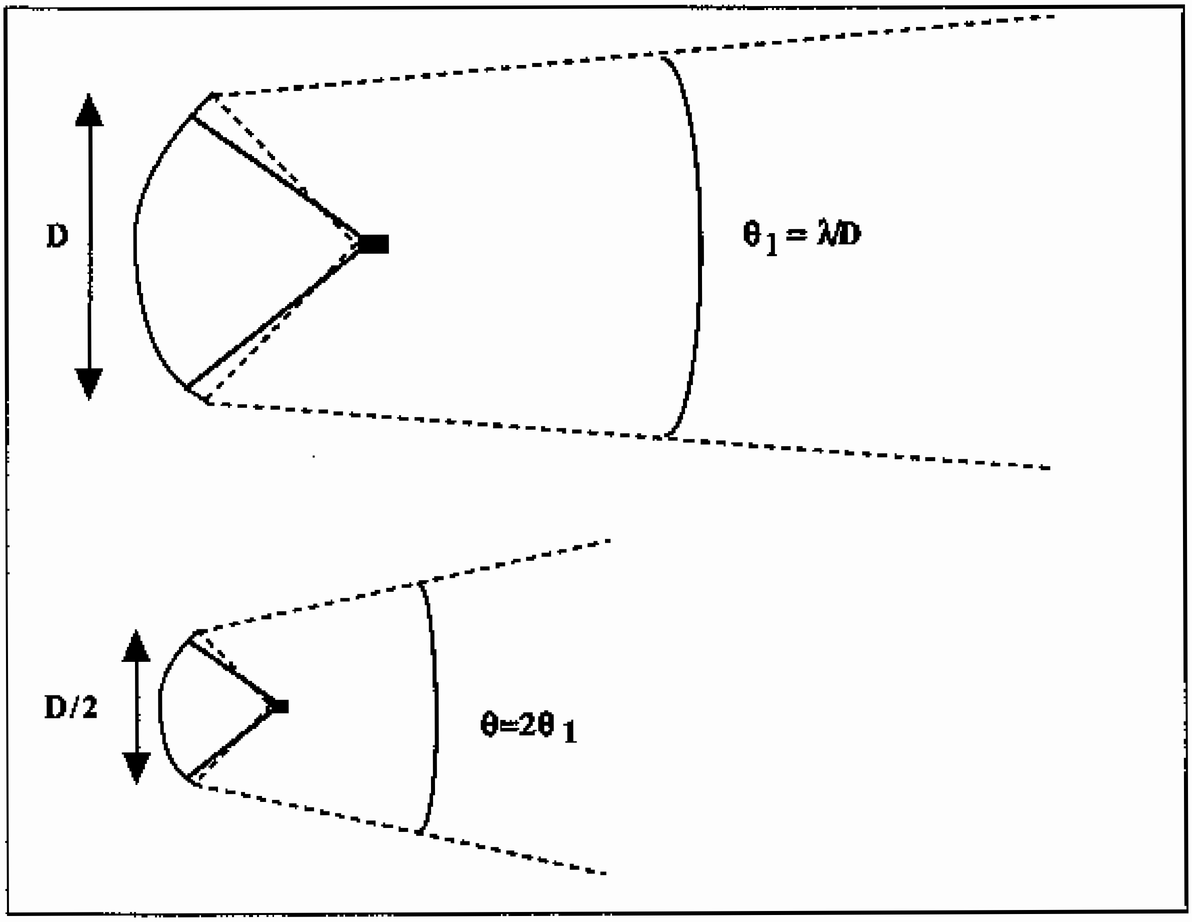
Television transmission has grown steadily to the point where it exceeds $10^7$ watts (10 Megawatts). While a distant civilization would be unable to distinguish individual stations, the rotation of the Earth, combined with the concentration of transmitters in the U.S., Europe, Japan, and Australia, would produce a regular, 24-hour cycle in the strength of this signal. An unusual, but regular pattern like this is something that scientists are likely to try to study.

Defense radars are much more powerful than individual radio stations, but there are fewer of them, and they broadcast for only a fraction of the time. While they would be the most intense of our leakage radiation, they would be hard to study because they would be unlikely to repeat in a regular fashion. In addition, the frequencies are changed often for security reasons. Finally, the end of the cold war may mean that they will eventually become obsolete.

Since the leakage radiation is spreading in all directions from the Earth at the speed of light, leakage radiation would have reached all stars within a distance given by the current year minus the year in which it was emitted. In 2002, the earliest transmissions would have traveled over 61 light years, reaching over 2000 of the nearest stars. However, these transmissions were very weak; if we consider only transmissions after about 1960, when television transmissions exceeded about 3 Megawatts, these would have covered about 42 ly, or about 800 stars. Thus, we have announced our presence to a substantial number of stars. If the nearest civilization, based on your Drake equation, lies within this range, it may well know about us. In this case, the civilization could have broadcast a signal intended for us, rather than an omnidirectional beacon. If the civilization is only half as far away as our leakage radiation has extended, then there has been time for a return signal from them to reach us. This is the scenario explored in the novel, **Contact**, by Carl Sagan and the movie based on the novel.

Even if we can assume that a much more advanced civilization is broadcasting, either in all directions, or specifically in our direction, we have a formidable problem in knowing where and when to listen, what frequency and what polarization to use, and how the signal will be coded. This problem has been likened to searching for a needle in a haystack. Earthly haystacks are three dimensional, but the cosmic haystack has even more dimensions. The frequency is one of these dimensions, and the direction is another. If we assume that the other civilization is broadcasting continuously, we can look at any time, removing that dimension. Without this assumption, the problem becomes nearly impossible to solve. We can observe in two polarization modes simultaneously, and we can design our detection devices to recognize various kinds of coding.

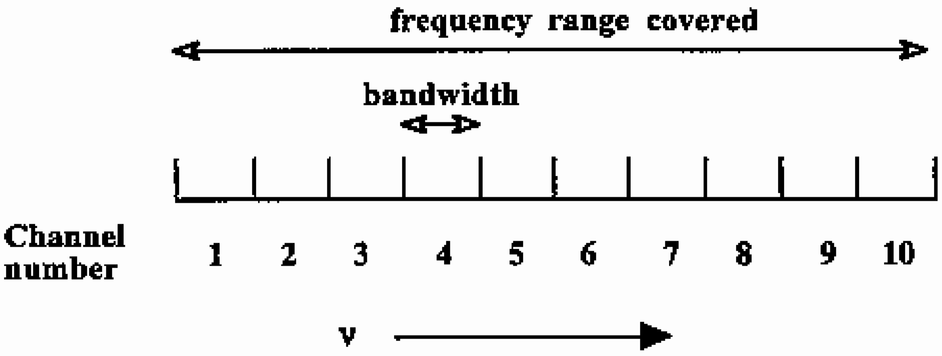
We are left with the questions of where to look and what frequency to use. In both cases, we run into a tradeoff. Let us consider first the issue of what direction to look. We do not know which direction to look, so we have to consider looking in many directions. The number of directions searched is one dimension of the cosmic haystack. On the other hand, we have to be able to detect a weak signal in the presence of noise. Our ability to detect a weak signal is called our sensitivity, often measured in W $m^{-2}$ (Watts per square meter); the smaller the sensitivity, the better. Thus, the sensitivity is another dimension of the cosmic haystack. The sensitivity improves (gets smaller) in proportion to the area (or diameter squared) of our telescope, and in proportion to the square root of the time we spend collecting signals. Mathematically, $S \propto D^{-2} t^{-1/2}$. To detect a weak signal, we would want to collect data in a given direction for a long time, using a large telescope.



Now we come to the tradeoff: clearly, the longer we spend looking in one direction, the fewer the number of directions we can examine. There is another, more subtle factor in this tradeoff. While we have described the telescope as looking in a certain direction, in fact, it receives power from a range of directions; the range of directions over which it receives signals well is called the beam of the telescope. Basic optics requires that $\theta$, the angular size of the beam (in radians) is approximately the wavelength of the radio waves divided by the diameter of the telescope ($\theta=\lambda/D$). Thus, for a fixed wavelength, the beam angle decreases as the diameter of the telescope increases, and the number of directions we need to point the telescope to cover a certain region of the sky increases in proportion to the square of the diameter. If we know nothing about the direction to look, we want to cover as much area on the sky as possible; in a limited time, we would then want to use a small diameter telescope. This choice would limit our sensitivity, so we could detect only very strong signals.

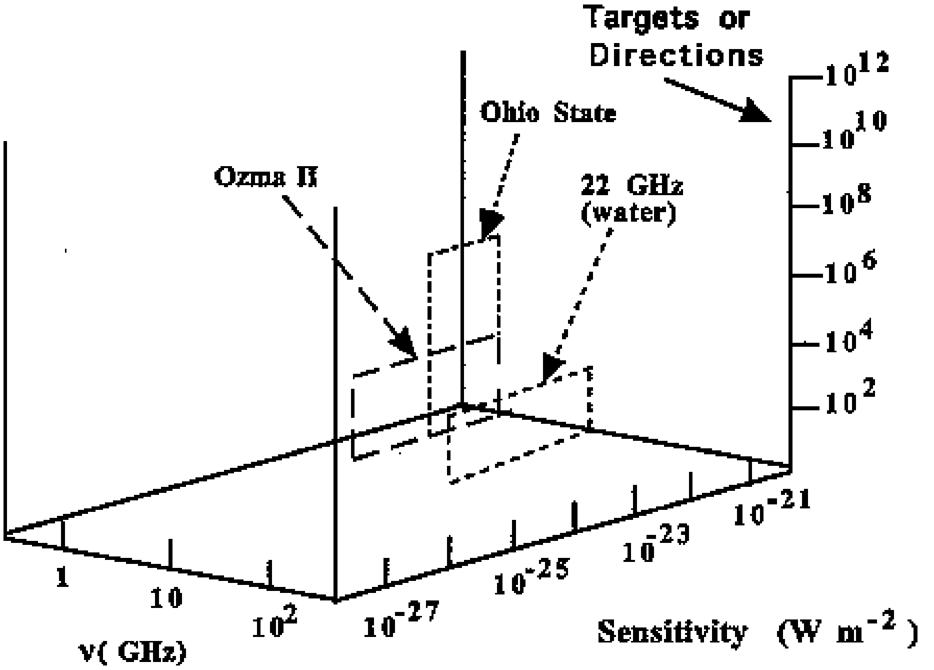
If the signal is strong, we can use a small telescope and only listen for a short time in any one direction. The small telescope in the figure has half the diameter of the large one, so its beam is twice as large; it can then cover a large region of the sky in a shorter time. On the other hand, if the signal is weak, we need a large telescope, and/or we need to observe a given direction for a longer time to collect more signal; this strategy limits the number of directions we can look. The best strategy depends on the nature of the signals: if there are only a few signals, but they are very strong, the first strategy is preferred; if there are many weak signals, the second strategy is preferred.

A similar tradeoff exists in the third dimension of our cosmic haystack, the frequency. If the signal is concentrated in a small range of frequencies, our ability to detect it is maximized by using detectors that select only that small range of frequencies. However, if we do not know the frequency of the signal, we have to search a large number of frequencies, favoring a detector that responds over a large frequency range. To some extent, technology can help us get around this problem. With modern electronics, we can observe many frequency channels, each selecting a narrow range of frequencies, simultaneously; such a device is sometimes called a spectrum analyzer. A diagram of a spectrum analyzer with ten channels is shown in the next figure.



Most of the early searches used conventional equipment designed for radio astronomy, which typically has 100s to 1000s of channels. Given a limited number of these channels, there is a tradeoff between the bandwidth of a channel and the frequency range covered. If a searcher believes in a magic frequency, the searcher would select a narrow bandwidth; if not, a larger bandwidth might be chosen. As with the tradeoff between sensitivity and number of directions searched, we shall see how past and future searchers have chosen between these strategies as we consider the portions of the cosmic haystack that they have searched or that they plan to search.

The first effort at SETI was conducted in 1960 by Frank Drake, who used a telescope of modest size (26 m) in West Virginia to examine two nearby stars at 1.42 GHz, the frequency suggested by Morrison and Cocconi the previous year. A narrow range of frequencies (0.4 MHz) was searched. This project could have succeeded only if there are many civilizations broadcasting quite strong signals very precisely at the 1.42 GHz frequency. Drake called this attempt Project Ozma, after a princess in the stories of the land of Oz. This choice symbolizes well the quixotic nature of the search; indeed Drake saw it primarily as a demonstration of the *idea* of searching. Project Ozma has also been used as a benchmark; later searches can be characterized by the probability of success relative to that of Ozma (effectively by the ratio of the volumes of the cosmic haystack searched).



Several searches illustrated in the above picture of the cosmic haystack were carried out in the late 1960s and early 1970s in the former Soviet Union, but the next one we will mention in detail was conducted in 1972–1976 by Ben Zuckerman and Pat Palmer. They used a larger (91 m diameter) telescope in West Virginia to survey the nearest 670 stars of "suitable" type. Stars were included if their masses were low enough to allow a life on the main sequence of about 4 billion years; very low mass stars were included, as well as some binary stars with separations that would allow stable planetary orbits. Calling their project Ozma II, Zuckerman and Palmer also observed at 1.42 GHz. Two different bandwidth strategies were adopted with the 384 available channels: with 192 channels, they covered 10 MHz (channel width was then about 52 kHz); with the other 192 channels, they covered 625 kHz, with a channel width of about 4 kHz. They could have detected a signal as weak as $10^{-23}$ W $m^{-2}$; with this sensitivity (much better than that of Ozma), they could have detected a 40 Megawatt transmitter on a 100 m diameter telescope aimed in our direction.

Both Ozma and Ozma II assumed that 1.42 GHz was the magic frequency, and they were looking for signals from a relatively small number of nearby stars, thus implicitly assuming that N is very large. Both were carried out in a limited time, using telescopes that were normally used for radio astronomy. A different approach was begun in 1973 by F. Dixon and D. Cole, who used the Ohio State radio telescope to begin a continuous survey that was not biased toward nearby stars. Like the Arecibo telescope, the Ohio State telescope could not point to arbitrary directions, but observed a swath of directions determined by the Earth's rotation. This search had modest sensitivity (about $10^{-21}$ W $m^{-2}$, nearly 100 times worse than Ozma II), so it could only detect extremely strong transmissions from the distant stars that it was optimized for. Like the previous searches, they searched at 1.42 GHz, covering a modest bandwidth (500 kHz) in a small number of 10 kHz channels. The main feature of this program was its continuous operation and ability to check on interesting signals. Some interesting signals were detected, but none repeated. Mostly by virtue of covering many directions in the sky, this project is estimated to be about $10^6$ times as likely as Ozma to have found a signal. The Ohio State telescope has now been closed.

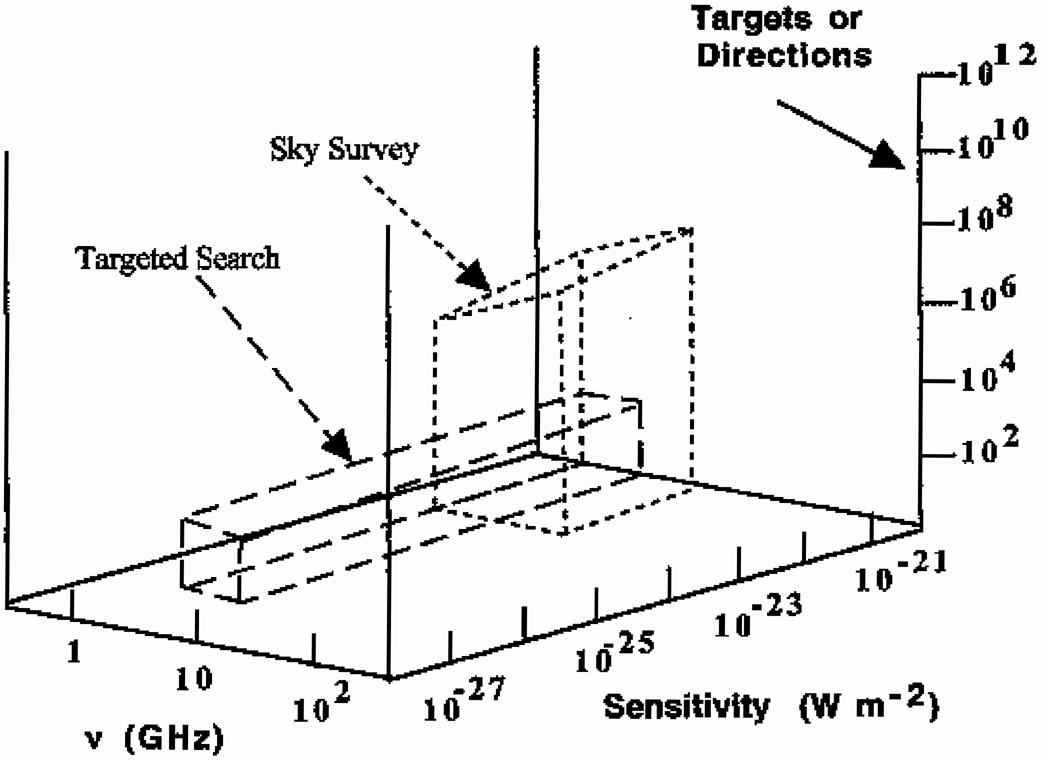
All the projects described so far have used only a small number of frequency channels, as typically used for radio astronomy. With support from the Planetary Society, Paul Horowitz took the next step, building a spectrum analyzer that was specialized for SETI. He chose to build a device with a very large number of very narrow channels. Originally, Horowitz took his spectrum analyzer to different telescopes for short periods, but in 1983, it was permanently installed on a 26 m diameter telescope at Harvard, and project Sentinel began. The telescope was fixed in position to survey the sky as the Earth rotated, rather than searching toward nearby stars. The search frequency was again 1.42 GHz. The spectrum analyzer for Sentinel had 128,000 channels, many more than previously used, but each was extremely narrow, so that the total coverage was only 6 kHz. Use of such narrow channels assumes that the transmitting civilization continuously adjusts the frequency to correct for Doppler shifts; Horowitz had thus made a very different assumption about the nature of the signal than had previous searchers. In 1985, the spectrum analyzer was expanded to 8 million channels, but it still had a very narrow bandwidth (400 kHz). The project was renamed META (Megachannel Extraterrestrial Assay). Because of the large number of channels and continuous operation, META expanded the volume of the cosmic haystack over that searched by Ozma by a factor of over $10^9$.

In 1993, Horowitz and Sagan reported on five years of observations, in which they covered the sky three times. There were 8 signals that were truly hard to explain as noise. They had some tendency to lie close to the plane of the Galaxy, as would be expected if they were signals from distant civilizations. The problem is that none of the signals repeated, making it impossible to prove that they came from extraterrestrial civilizations. Horowitz and Sagan concluded that they were probably glitches in the electronics. Project Phoenix (described below) re-observed the directions of the META signals in 1997 to 1998 with better sensitivity and found no signals. In October 1995, Horowitz began project BETA (for Billion channel Extraterrestrial Assay). He did not quite achieve a billion channels, but he came close. BETA has $2.5 \times 10^8$ channels; by using 8 different settings, he covers $2 \times 10^9$ Hz. This bandwidth allows him to cover the water hole. Using chips donated by an electronics firm, he was able to construct the spectrum analyzer for about $250,000. The search was suspended in the late 1990s when a severe windstorm blew the telescope off its mount.

So far, all the searches we have examined in detail have been conducted at 1.420 GHz. Some searches at other frequencies, notably the OH frequencies, the Kuiper-Morris frequency, and the frequency used to detect interstellar $H_2O$, have been conducted; but they have not been nearly as complete. Consequently, the dimension of the cosmic haystack that has been least explored is the frequency dimension. The coverage of this dimension can be greatly expanded by using many channels, each much wider than used by Horowitz.

The most ambitious search was one planned by NASA. With great fanfare, it began in October 12, 1992, the 500th anniversary of the discovery of America by Columbus. Within a year, it fell prey to a budget-cutting move in Congress and was stopped. It is worth discussing the plans to see what it would have accomplished. Two separate strategies were adopted: the sky survey would have covered as much of the sky as possible with modest sensitivity (adopting the strategy of the Ohio State and META projects); the targeted search would have achieved higher sensitivity toward nearby stars (adopting the strategy of Ozma II). Both planned to use spectrum analyzers with over a million channels. We will consider both in some detail.

The sky survey would have used NASA telescopes with 34 m diameters, located in California and Australia, which have been used to track spacecraft. The spectrum analyzer would have had a channel width of 20 Hz; the initial system would have had $2 \times 10^6$ channels, thereby covering 40 MHz simultaneously. By 1996, the number of channels was to have increased to $16 \times 10^6$, allowing 320 MHz to be covered at once. By changing the frequency received by the telescope, they planned to cover the entire range of 1–10 GHz. Even with the expanded spectrum analyzer, this would have required about 30 different settings. To cover the entire sky with 30 different frequency settings in the 6 years planned for this project, they could have spent only a few seconds looking in any given direction at any given frequency. The sensitivity would have been modest, consistent with the goal of detecting very strong transmitters.



The targeted search was to cover about 800 of the nearest stars that are similar to the Sun, thus covering the suitable stars within about 75 ly. Of these, 244 can be seen with the Arecibo telescope, while the remainder would have required other, smaller telescopes. The spectrum analyzer had narrower channels (1 Hz wide) than the one used in the sky survey, but it would also have had $16 \times 10^6$ channels. All frequencies between 1 and 3 GHz were to be observed, a range that covers both the water hole and the Kuiper-Morris frequency, but not the entire 1–10 GHz region. The plan was to observe each star for 1000 seconds, yielding a sensitivity of about $10^{-27}$ W $m^{-2}$. The figure shows the parts of the cosmic haystack that the two searches would have filled.

The researchers involved with the targeted search have continued a more modest effort with private funding. Named after the mythical bird that rises from its own ashes, Project Phoenix is using non-NASA telescopes to achieve many of the original goals of the targeted search. They conducted a search of 200 stars like the Sun within 150 ly, using a 64-meter telescope in Australia. They avoided binary stars and stars younger than $3 \times 10^9$ yr. They observed each for 5 minutes, using a spectrum analyzer with $28 \times 10^6$ channels, each 1 Hz wide. With multiple frequency settings, they covered 1.2 to 3.0 GHz. They could have detected a 1 Megawatt transmitter if it were beamed toward us with a similar telescope. One interesting feature of this search was an immediate follow-up by a second telescope, located 150 miles away. This technique allowed them to discriminate against terrestrial interference very effectively. They found no convincing extraterrestrial signals. Similar observations were obtained with telescopes in West Virginia. Updates can be found on their home page on the World Wide Web (http://www.seti.org), as can the plans for the Allen Telescope Array. This joint project of the SETI Institute and the University of California will be an array of about 350 telescopes, each 6-m in diameter located at Hat Creek Observatory in northern California. It will be usable for radio astronomy, but its main goal will be SETI. Using advances in electronics, it will cover a wide frequency range (about 1-10 GHz) with many channels. It will be able to direct all the telescopes toward a single star or use some telescopes to look at different stars. It can extend the targeted search to over 100,000 stars.

Even the full NASA SETI program would not have been the ultimate search. Since the main advance in the NASA program was a vast expansion of the frequency coverage, the next logical advance would be to improve the sensitivity and sky coverage. We noted earlier the trade off between these, which is caused by a large telescope having a small beam. It is possible to build systems that have more than one beam (essentially they look in more than one "direction" at a time). These systems are just beginning to be built, mostly at higher frequencies, but so far only a few (about 30) beams are used at once. It is more costly to do this than to observe many frequencies at once, since the entire receiver must be copied, rather than just the last stages.

Another way to observe many directions simultaneously is to use more than one telescope. This is clearly the most expensive option, but it has some other capabilities. Barney Oliver proposed an advanced SETI system some time ago. It would have consisted of 1000 telescopes, each with a diameter of 100 m (the largest fully steerable telescope now operating). Arranged in a circular pattern, the telescopes would have resembled a giant eye, so the project was called Cyclops. With this system, the nearest 1000 stars could be continuously monitored. In another mode of operation, all the telescopes could look in the same direction, adding their signals to improve sensitivity. In this mode, Cyclops could detect a 1000 Megawatt beacon at a distance of 1000 ly. The most exciting capability of Cyclops would be the detection of leakage radiation similar to that of the Earth from any planet within 100 ly, allowing us to detect civilizations that are not actively trying to contact us. As originally conceived, Cyclops would have covered only the water hole, but its frequency range could easily be expanded. Of course, the drawback of Cyclops was cost. In current dollars, it would cost about $50 billion and take 10–20 years to build.

While there is little immediate prospect for a full-fledged Cyclops, the idea of using many telescopes working together has already been used for radio astronomy. For example, the Very Large Array (VLA) in New Mexico contains 27 telescopes spread over about 20 miles. By combining the signals from widely separated telescopes, a much smaller beam can be "synthesized"; while this would not be an advantage in a sky survey mode, it would be helpful in pinning down the star from which a signal originated. In his novel, **Contact**, Carl Sagan imagines a VLA that has been expanded to 131 telescopes and renamed ARGUS, after a mythical creature with 100 eyes. The Allen Telescope Array will be similar, except that it uses more, but smaller telescopes.