Chapter 3
Arraying Concepts

The gain of an antenna divided by its system temperature, G/T, is one of the parameters that determine how much data can be sent over a communications link with a specified SNR. Our first goal in any study to understand arraying is to outline some of the practical aspects of arraying by treating the problem as adding individual G/T's. Next, we must recognize the bounds on performance achievable with current technology and attempt to parameterize both performance and cost in a way that can be related to antenna diameter. Then we must understand how the overall reliability and availability of an array are related to cost and how an array compares to a single aperture.

3.1 An Array as an Interferometer

Figure 3-1 shows two antennas located somewhere on the surface of a rotating Earth, viewing a distant radio source and forming a simple interferometer [1]. In vector notation, the difference in time of arrival, τ_g, of a radio wave from an infinitely distant source is simply

$$\tau_1 - \tau_2 = \tau_{\mathrm{g}} = \frac{\mathbf{B} \bullet \mathbf{i}}{c} = \frac{B \sin(\theta)}{c} \tag{3.1-1}$$

where B is the baseline vector extending from the intersection of axes on antenna number 1 to the intersection of axes on antenna number 2, i is a unit vector pointing to the radio source, and c is the speed of light (see Appendix A for how to determine the antenna intersection of axes). If the source is not at infinite distance, then the wave front is slightly curved and the vector expression is somewhat more complicated, but the process is essentially the same. We can write an expression for the difference in time of arrival in terms

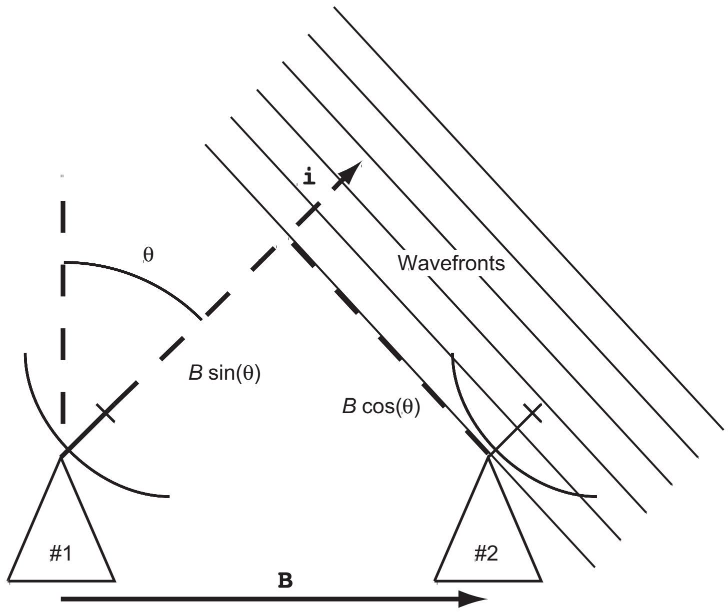

Fig. 3-1. A simple interferometer.

of the baseline and source directions. In effect, the accuracy with which we can calculate the delay is determined by the accuracy with which we can determine the baseline and source direction in a consistent reference frame.

Let us assume each antenna is observing a strong distant source at a radio frequency *f*, and the output of each antenna is connected to a multiplier by means of equal-length cables. The output of this multiplier, or correlator, at time *t*, then has the form

$$V_{\text{out}} \propto 2\sin(2\pi f t)\sin\left(2\pi f\left(t-\tau_g\right)\right) \tag{3.1-2}$$

If we expand this expression and run it through a low-pass filter, the result we are left with is

$$V_{\text{out}} \propto \cos\left(2\pi f \tau_g\right) \tag{3.1-3}$$

which is simply the coherent multiplication of the voltages from each element of the interferometer. Suppose the radio source being observed is a celestial source. Then τ_g will change by virtue of the Earth's rotation, and the output of the multiplier, or correlator, will exhibit the cosinusoidal variation described in Eq. (3.1-3) as the two signals go from in phase to out of phase.

If we know τ_g, or can somehow sense it, it is possible to build a compensating delay into one or both cables from the antennas such that the total cable delay and geometric delay is perfectly compensated. In this case, V_{out} for the multiplier is at maximum and the voltages are in phase. If we include an adding circuit in parallel with the multiplier, we can obtain the coherent sum of two antenna's voltages. It is just this kind of processing, using correlation to phase up the signals and then adding them, that constitutes a system that can perform antenna arraying.

For two identical antennas and receivers, this scheme for coherently adding the antenna signals doubles the SNR. However, it requires we implement a programmable delay line and calculate or derive, with some precision, the geometrical delay. The required precision of this delay is a function of the bandwidth of our receivers and can be determined as follows: Let us assume that our two antennas have identical receivers, centered at a frequency f_o, and have bandwidth Δf. If we make an error in the compensation of the geometric delay, we will in effect lose coherence, where the phase of the signal in the upper part of the band slips relative to the phase in the lower part. The requirement for coherence over the band becomes

$$\Delta f \, \Delta\tau << 1 \tag{3.1-4}$$

where Δf is in cycles and $\Delta\tau$ is in seconds. This requirement is simply stating that the phase shift across the bandpass due to an error in delay should be a small part of a cycle (less than or equal to 0.01 would work well). Therefore, for a bandwidth of 1 MHz, the error in delay compensation must be much less than a microsecond, or we will lose coherence in both the multiplication as well as the addition of the signals.

To see how errors in the length of the baseline (B) and errors in position of the source (θ, in radians) translate into errors in delay, we take the derivative of Eq. (3.1-1). Since these two errors are at right angles to each other, this derivative must take the form of a gradiant:

$$\nabla\tau_g = \Delta\tau_g = \left(\frac{\sin\theta}{c}\right)\Delta B\,\mathbf{u_B} + \left(\frac{B\cos\theta}{c}\right)\Delta\theta\,\mathbf{u}_{\boldsymbol{\theta}} \tag{3.1-5}$$

where vectors are indicated by boldface, the unit vectors are along the direction of B, and the direction of θ is at right angles to B.

The error in the calculation of geometric delay is simply the modulus of Eq. (3.1-5), or

$$\Delta\tau_g = \sqrt{\left(\frac{\sin\theta}{c}\right)^2 \Delta B^2 + \left(\frac{B\cos\theta}{c}\right)^2 \Delta\theta^2} \tag{3.1-6}$$

As an example, if our bandwidth were 10 MHz and we wished to keep our delay errors to 10^{-2} of the coherence function, then the above expressions indicate that the baseline error should be kept below 1 ns or 30 cm. A similar bound could be placed on the source position error $\Delta\theta$.

3.2 Detectability

The detectability of the signals that are discussed here will always relate to a sensitivity factor, known as *G/T*, where *G* is typically the gain of the antenna used to gather energy from the signal of interest and *T* is the total system temperature. Putting aside for the moment the question of how to coherently add apertures, the maximum possible sensitivity factor for an ideal array (i.e., no combining losses) is simply the sum of the sensitivity factors for each element, or

$$\left[\frac{G}{T}\right]_{\text{array}} = \sum_{i=1}^{N}\left(\frac{G}{T}\right)_i \tag{3.2-1}$$

In the case of a homogeneous array, having elements of equal collecting area and system temperature, the sensitivity factor is

$$\left[\frac{G}{T}\right]_{\text{array}} = \left(\frac{G}{T}\right)_0 \sum_{i=1}^{N} 1 = N\left(\frac{G}{T}\right)_0 \tag{3.2-2}$$

where the quantity in square brackets divided by $(G/T)_0$ is called the array gain and is usually expressed in decibels (dB). Figure 3-2 illustrates this by plotting the array gain versus the number of elements in the array (assumed to have equal G/T). It can be seen that, as the number of array elements increases, the incremental improvement in performance decreases. For instance (again assuming no combining loss), going from a single antenna to two antennas doubles the SNR and results in a 3-dB gain. However, going from two to three antennas results in a 4.8-dB overall gain, or an increase of 1.8 dB over the two-element array, and adding a tenth element to a nine-element array increases the SNR by only 0.46 dB.

For an inhomogeneous array, i.e., one having elements with different G_i's and T_i's, the arithmetic is more complicated but the reasoning is the same and can be evaluated easily. In this case, array gain typically is computed by adding *G/T* to the most sensitive element. If you array two antennas, the first having a *G/T* that is ten times the second, then the array gain will be about 0.4 dB. The cost of adding the second array element can be quantified, but only the customer can decide if the 0.4 dB is worth the cost.

Given these considerations, it seems reasonable that, for the case of large, costly elements, we not consider any element for addition to an array unless it

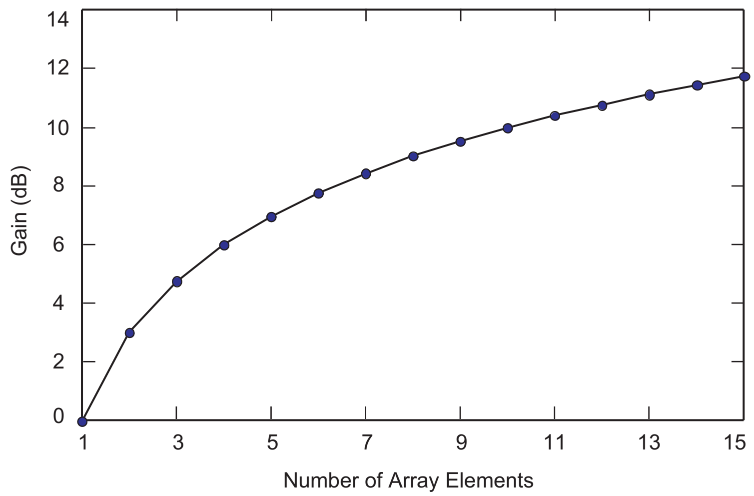

Fig. 3-2. Array gain as a function of the number of elements.

adds at least 10 percent to the aggregate G/T of the array. This suggests a rule of thumb that we not consider arrays larger than 10 elements. A particular example that might be of interest to the DSN is the arraying of, say, two 34-m elements with one 70-m element. If we assume all three have the same receiver temperature, then, since a 70-m antenna is about twice the diameter of a 34-m antenna, the G/T of the 70-m antenna is about four times that of the 34-m. Therefore, an additional 34 m will improve the G/T of an array of a 70-m antenna and a 34-m antenna by about one-fifth, or about 0.8 dB.

3.3 Gain Limits for an Antenna and Array

The gain, G, of an antenna is given in terms of its effective collecting area, A_e, at an operating wavelength, λ, as

$$G = \frac{4\pi}{\lambda^2} A_e \qquad (3.3\text{-}1)$$

The effective collecting area, as well, can be written as the product of the physical aperture area, A_p, times a factor, η, that is termed the aperture efficiency.

Ruze [2] has pointed out that various mechanisms cause deviations in the reflector surface that result in a systematic or random phase error. These errors can be mapped into the aperture plane and lead to a net loss of gain such that the relative gain is given by the expression

$$\frac{G}{G_0} = \exp\left[-\left(\frac{4\pi\sigma}{\lambda}\right)^2\right] \tag{3.3-2}$$

where σ^2 is the variance of the phase error in the aperture plane. While Eq. (3.3-1) predicts that the gain of an antenna should increase as the square of the frequency, Eq. (3.3-2) predicts that when $(\sigma/\lambda) > 1$, the gain drops rapidly. If we use Eq. (3.3-1) as the G_0 in Eq. (3.3-2) and then set its derivative with respect to λ equal to zero, we calculate that the gain will be a maximum at a wavelength λ_{min}, which is approximately equal to 13 times the root-mean-square (rms) surface error σ. This point is known as the gain limit of the antenna. Note that the concept of gain limit is equally valid for a synthesized aperture.

The phase error in the aperture plane of a single antenna is composed of several components: the surface roughness of the reflector (σ), mechanical distortions from a designed, specified parabolic shape, and the propagation medium, which could include the radome of the antenna if it has one, the atmosphere, and the ionosphere. Clearly, there are distortions in the effective aperture plane of an array that result in phase errors that are analogous to those of a single aperture. While most of these errors will be reduced with calibration by the arraying algorithm, any residuals will lead to a loss of gain for the array.

One of the potential disadvantages of an array is due to the fact that its physical extent is always larger than the equivalent single-antenna aperture that it synthesizes. As a result, phase errors due to atmospheric fluctuations, which increase as the distance between individual elements increases, can limit the gain of the array. A typical example of this phenomenon is in the case of the troposphere, where over short distances (<1 km) the phase fluctuations are coherent because they come from the same atmospheric cell. Therefore, for antennas close together, the phase variations between the two antennas cancel each other out. As the distance between the antennas increases, the phase variations are coming from different atmospheric cells and are no longer coherent. Therefore, cancellation no longer takes place.

3.4 System Temperature

In characterizing the performance of antenna and receiver systems, it is common practice to specify the noise power of a receiving system in terms of the temperature of a matched resistive load that would produce an equal power level in an equivalent noise-free receiver. This temperature is usually called the "system temperature" and consists of two components: the temperature corresponding to the receiver itself due to internal noise in its front-end amplifier, and the temperature corresponding to antenna losses or spurious signals coming from ground radiation, atmospheric attenuation, cosmic background, and other sources. The term "antenna temperature" usually is used to express the power received from an external radio source and is related to the intensity of the source as well as to the collecting area and efficiency of the antenna. In what follows, we will use this terminology to characterize various receiver systems that have been used in the DSN [3]. Clearly, any improvement that can be made in the area of system temperature on a specific antenna should be considered before taking the steps to array several such antennas.

There is a new generation of transistor amplifiers called high electron mobility transistors (HEMTs). Figure 3-3 illustrates the state of this technology in 1989. In this figure, the effective noise temperature of an 8.4-GHz (X-band) HEMT amplifier is plotted against the physical temperature of the device. It can be seen that the noise temperature of the amplifier varies almost linearly with the physical temperature. The data were fitted with a straight line (shown as the solid line) that indicates the amplifier noise improves at the rate of 0.44 kelvin per kelvin, or 0.44 K/K, in the region where the physical temperature is >150 K.

Figure 3-4 shows HEMT amplifier noise performance versus frequency for three common cooling configurations. The first is at room temperature, the second is cooled to approximately –50 deg C with a Peltier-effect cooler, and the third uses a closed-cycle helium refrigerator capable of lowering the device temperature to 15 K. Note that cooling has the most benefit at the higher frequencies. It is also important to remember that this technology has been highly dynamic for the past several years. As in most areas of microelectronics, there have been rapid improvements in performance, accompanied by reduced costs.

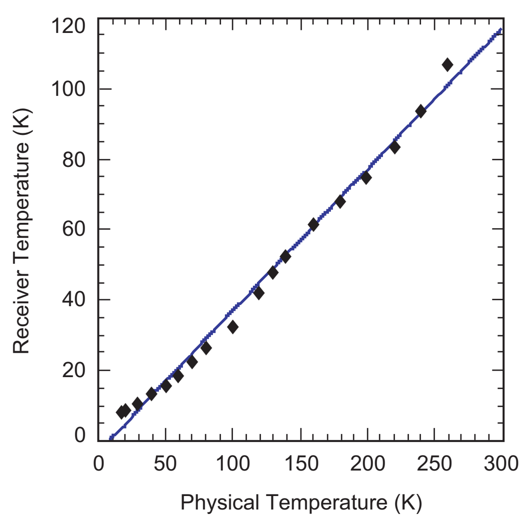

Fig. 3-3. Amplifier performance versus temperature.

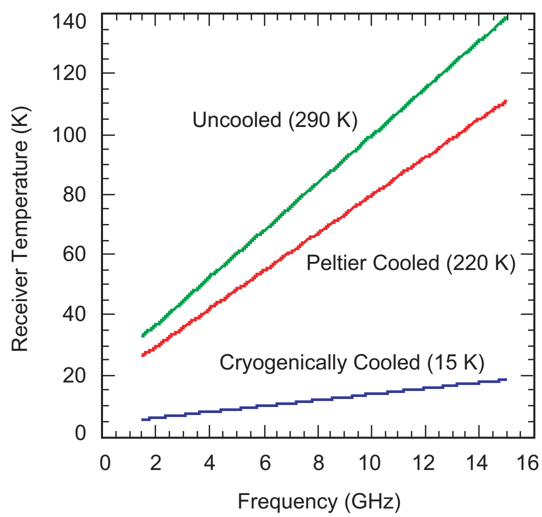

Fig. 3-4. Amplifier performance versus frequency.

Table 3-1 lists the various noise contributions to the total system temperature we might expect for a HEMT RF package at both 4 GHz (C-band) and 13 GHz (Ku-band). The atmospheric contribution comes from thermal noise generated by atmospheric gases and varies as the amount of atmosphere along the line of sight, i.e., as the secant of the zenith angle Z. The cosmic blackbody background is a constant 2.7 K. Spillover and scattering will depend on antenna [e.g., prime focus, Cassegrain, or beam waveguide (BWG)], feed, and support structure design.

3.5 Reliability and Availability

In the following discussion, we will compare results for communication links made up of arrays of various sizes. As we will see, there are certain advantages for availability that occur when using a large number of smaller elements verses a small number of large elements to achieve a given level of performance.

The specification of a communications link requires knowledge of the availability of the link components, one of which is the ground aperture, or array element. If we were to operate an array with no link margin (by margin, we mean extra capacity over what is necessary to meet requirements), we would find that increasing the array size beyond some number N_{max} leads to the interesting conclusion that the total data return is decreased!

Table 3-1. Range of total system temperature.

Noise Source	4.0 GHz	13.0 GHz
Atmosphere (K)	5.0 s(Z)	7.8 s(Z)
Cosmic background	2.7	2.7
Spillover, scattering	4–8	4–8
Microwave losses	4–12	4–16
Subtotal	16–28	19–31
Receiver temperature		
Room temperature (290 K)	40	110
Peltier (220 K)	30	90
Cryogenic (15 K)	8	17
Total (zenith)		
Room temperature (290 K)	56–68	129–141
Peltier (210 K)	46–58	109–121
Cryogenic (15 K)	24–36	36–48

In order to clarify this assertion, consider the following simplified argument. Define the availability, A_T, of a system to be the percentage of time that the system is operable for scheduled support. Thus, the down time required for maintenance is not counted. We should keep in mind that the overall availability is a product of all subsystem availabilities, although, for the remainder of this discussion, we will focus on the antenna availability. The total data return, D_T, can be written in terms of the system availability, A_T, and the integral of the data rate:

$$D_T = A_T \int D_R(t)dt \tag{3.5-1}$$

where the integral is taken over the interesting portion of the mission. Suppose the data rate, $D_R(t)$, is adjusted to the highest level that can be supported by the total ground aperture used to receive the signal. If we use an array on the ground of N elements, each having availability p, and the total signal from the array is near the detection threshold, then the total data return can be written in the form

$$D_T = Np^N f(t) \tag{3.5-2}$$

where $f(t)$ is some function of time and includes all of the factors that enter into link performance (e.g., distance, antenna gain, duration of a pass, etc.), and p^N is the availability of the entire array. Very often $f(t)$ cannot be increased, and the total data return can be increased only by increasing the ground array (e.g., a signal of interest transmits only for a finite duration and does not repeat). Since $p<1$, we see that D_T has a maximum value at the value of N given by

$$N_{\max} = \frac{-1}{\ln(p)} \tag{3.5-3}$$

A graph of $N_{\max}$ as a function of the individual array-element availability p is shown in Fig. 3-5. Using Eq. (3.5-2), we see for an array whose size is greater than $N_{\max}$ that the data return drops precipitously. This result stems directly from our assumption that the data rate would be increased to take advantage of *all* the ground aperture—that is how it is done with a single antenna. In fact, use of an array requires that we consider antenna availability in a different way than we do for a single antenna. In a link with a single antenna, the antenna is a single point of failure. In an array, the concept of availability must be merged with that of link margin.

In Appendix B, we derive relations that give the array availability as a function of the number of antenna elements (spare elements) over and above the minimum number needed to achieve the required G/T. In order to make a comparative assessment of the performance of various arrays, Fig. 3-6 shows the array availability plotted as a function of the fraction of extra elements that are devoted to sparing for three array sizes (designated in the figure by N_e for the number of required elements) and for a fixed-element availability of $p = 0.9$. The following interesting observation can be made: *The availability of*

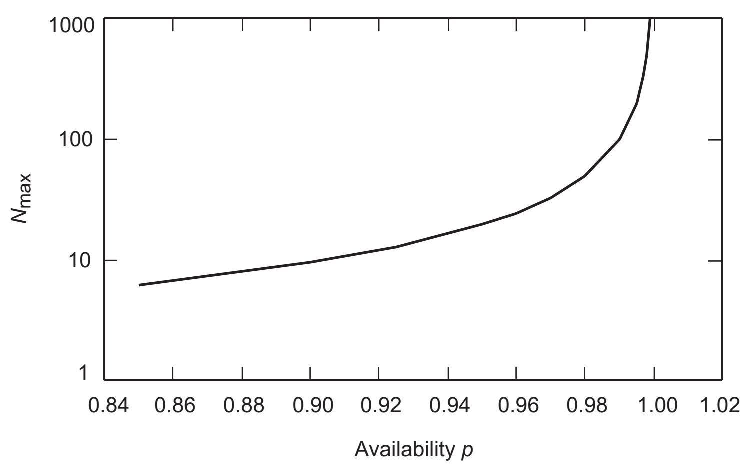

Fig. 3-5. N_{max} versus availability.

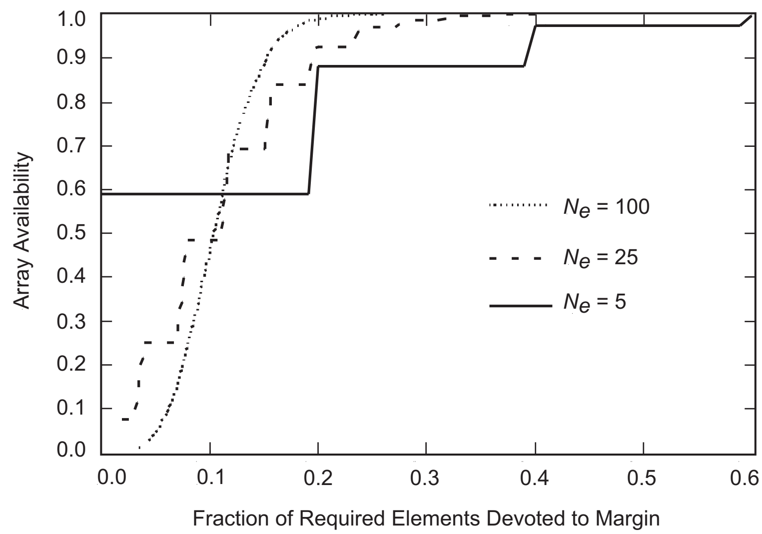

Fig. 3-6. Array availability versus element margin.

the array can be increased by increasing the number of spare elements. The array availability starts with a value much below the element availability, but increases rapidly and surpasses the element availability for a margin of less than about 30 percent, or 1 dB. The rate of increase of array availability is faster for arrays with a larger number of elements, even though it starts with a much smaller value. At some point as the sparing level increases, all the arrays with different numbers of elements reach approximately the same availability, beyond which a given sparing results in higher availability for larger arrays than for smaller arrays.

For larger arrays, sparing can be increased more gradually, since each additional element constitutes a smaller fraction of the total array. For an element availability of 0.9 for example, the minimum availability of a two-element array is 0.81, which increases to 0.972 by the addition of one element. This is the smallest increment possible and constitutes a 50 percent increase in the collecting area, or a 1.76-dB margin. In contrast, for a 10-element array with the same element availability, the minimum array availability is 0.349, but by the addition of three elements (a 30 percent increase, or a 1.1-dB margin), an array availability of 0.966 is achieved. Typically, for a given level of sparing or percentage of increase in the collecting aperture, a higher array availability is achieved in arrays with larger numbers of elements.

This discussion demonstrates some of the advantages of a large array of smaller apertures in comparison with a small array (few elements) of larger apertures, in terms of providing a more gradual way of increasing the performance margin or, conversely, a more gradual degradation in case of element failure. Furthermore, since a higher array availability is achieved in arrays with larger numbers of elements (for a given margin or percentage of increase in the collecting aperture), the designer of a large array can trade off element reliability for cost, while still maintaining the same overall reliability as that of an array with a smaller number of elements with higher individual reliability. Interestingly enough, the smaller elements used in larger arrays typically have a much higher reliability than do their larger counterparts to begin with, since they are less complex and easier to maintain.

References

[1] A. R. Thompson, J. M. Moran, and G. W. Swenson, Jr., *Interferometry and Synthesis in Radio Astronomy*, New York: Wiley, 1986.

[2] J. Ruze, "Antenna Tolerance Theory—A Review," *Proceedings of the IEEE*, vol. 54, no. 4, pp. 633–640, April 1966.

[3] *DSMS Telecommunications Link Design Handbook*, JPL 810-5, D-10379, Rev. E., Jet Propulsion Laboratory, Pasadena, California, January 2001.