JPL Publication 90-50

Scientific Applications of Frequency-Stabilized Laser Technology in Space

Bonny L. Schumaker

December 1, 1990

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

ABSTRACT

This report is a synoptic investigation of the uses of frequency-stabilized lasers for scientific applications in space. It begins by summarizing properties of lasers, characterizing their frequency stability, and describing limitations and techniques to achieve certain levels of frequency stability. Limits to precision set by laser frequency stability for various kinds of measurements are investigated and compared with other sources of error. These other sources include photon-counting statistics, scattered laser light, fluctuations in laser power and intensity distribution across the beam, propagation effects, mechanical and thermal noise, and radiation pressure. Methods are explored to improve the sensitivity of laser-based interferometric and range-rate measurements. Several specific types of science experiments that rely on highly precise measurements made with lasers are analyzed, and anticipated errors and overall performance are discussed. Qualitative descriptions are given of a number of other possible science applications involving frequency-stabilized lasers and related laser technology in space, applications that will warrant more careful analyses as technology develops.

ACKNOWLEDGMENTS

The large majority of the funding for this study was provided by the Communications and Information Systems Division (Code SC) of NASA's Office of Space Science and Applications (OSSA), for the purpose of understanding technology developments in the area of frequency-stabilized lasers needed to enable a broad range of scientific applications. Important additional funding was provided by the Solar System Exploration Division (Code SL) of OSSA as part of a study of solar-system science experiments that could be performed with first-generation laser communication links on planetary spacecraft.

Sincere thanks are extended to Lou Caudill of NASA Headquarters for initiating and funding the majority of this study and to William Rafferty of JPL for supporting the report's progress and completion. I am grateful to Peter Bender and John Anderson for sharing published and unpublished work, and to Bob Byer and Eric Gustafson for the loan of helpful reference material. Jim Ulvestad provided invaluable critical readings and technical discussions. Randy Bartman and JPL's Division 34 have provided a most enjoyable working environment. I thank Steve Synnott for reading the manuscript and giving needed encouragement. I also thank Jim Lesh for his support of this study in its early stages.

CONTENTS

Tables

Figures 221

EXECUTIVE SUMMARY

This report investigates aspects of frequency-stabilized lasers that are important for a variety of scientific applications in space, and describes in detail several kinds of scientific measurements that would use frequency-stabilized lasers and related technology. The much shorter wavelengths of lasers relative to radio or microwave emitters (microns *vs.* centimeters and longer) offer the potential for intrinsically finer resolution in measurements of angular position or distance. They also make laser links to spacecraft virtually immune to the plasma propagation effects that plague radio-science experiments in space and make dual-frequency radio links a necessity. Because of the much shorter wavelength (higher energy) associated with each photon, quantum effects are more noticeable than in centimeter-wavelength measurements. They are more noticeable even than thermal effects, since the energy $h\nu$ of a single optical photon is substantially greater than the analogous single quantum of thermal energy, $k_B T$, at room temperatures. Thus, whereas at radio wavelengths the error in measurements of phase or frequency typically is dominated by additive thermal noise, at optical wavelengths it is dominated by photon-counting statistics – the random, Poisson-distributed arrival times of the photons, an ultimate consequence of the quantum-mechanical nature of light and the "zero-point fluctuations" of the vacuum electromagnetic field.

The beam from a laser/telescope transmitter is narrow: $\lambda/D \approx 10$ microradians (μrad) $\approx$ 2 arcseconds for laser wavelength $\lambda = 1$ micron (μm) and telescope diameter $D = 10$ cm. Hence laser transmitters and receivers both can be substantially smaller than their microwave counterparts. The small size and lower mass offer practical advantages (such as eased launch constraints and more efficient fuel usage), and the higher frequencies offer wider communication bandwidths and higher data rates. These advantages are driving the present development of optical communication links for near- and deep-space applications. The small antenna size also suggests an important benefit to high-precision measurements involving range or range-rate measurements among orbiting spacecraft. In such experiments, the spacecraft house test masses that are maintained as nearly as possible in

inertial-free environments, surrounded by vacuum chambers and isolated from external accelerations of various origins, including mechanical, thermal, electrostatic, magnetic, and gravitational. This isolation requires DISturbance COmpensation Systems (DISCOS, or drag-free systems) to sense changing accelerations on the test masses and control mechanisms (*e.g.*, spacecraft thrusters or local electric fields) to compensate. In low Earth-orbit, bombardment by dust and gas molecules of varying velocities and densities causes a root-mean-square (rms) acceleration or drag that depends on the average density, the square of the spacecraft velocity, and the ratio of cross-sectional area to mass of the spacecraft or antenna. The reduced drag afforded by compact laser antennas also may make it possible to make measurements at lower orbit altitudes, where, for applications such as gravity-mapping or remote sensing, the signals are stronger.

The frequency stability of a laser is characterized most conveniently as a fractional quantity $\delta\nu/\nu$, the ratio of the frequency fluctuations (standard deviation) $\delta\nu$ to the nominal frequency ν. In the simplest measurement of a distance L made by monitoring the passage of cycles of a continuous-wave laser signal, the fractional error $\delta L/L$ in the measurement will be at least as large as the fractional fluctuations of the laser frequency:

$$\frac{\delta L}{L} \geq \frac{\delta\nu}{\nu} . \tag{1}$$

Many scientific applications using laser interferometers in space, such as the detection of gravitational waves, high-resolution measurements of anomalies in a planet's gravity field, or microarcsecond astrometry, require extremely precise measurement of distances ranging from several centimeters to millions of kilometers. Gravitational-wave detection imposes the most stringent demands on measurement precision and accuracy, requiring a sensitivity to fractional length changes of 10^{-20} or smaller in order to detect the gravitational waves expected from a variety of sources. Fortunately, eq. (1) is not the last word on the relation between the fractional frequency stability of a laser and obtainable accuracy for measurement of displacements. In practice, the intrinsic fractional frequency

stability can be considerably worse than the desired sensitivity to displacements, $\delta L/L$, provided that effects of the frequency fluctuations on the measurements can be calibrated adequately.

There are two general approaches to ease requirements on laser fractional frequency stability for a given displacement measurement sensitivity. The first is to measure the displacement relative to some other distance, *e.g.*, using a dual-arm interferometric approach. Since measurement is made of the relative change in length of two (or more) nominally equal arms of the interferometer, error sources that are correlated in the two arms, such as phase or frequency fluctuations in the common laser light that was split and sent down each arm, will cancel. For arm lengths held equal, say, to 0.1%, measurement error due to laser phase fluctuations also will cancel to 0.1%, and the requirement on laser fractional frequency stability given by eq. (1) will be eased by three orders of magnitude. If the desired measurement sensitivity still leaves the demands on laser frequency stability too high, it may also be possible, with a dual-arm interferometer, to calibrate the laser phase fluctuations and remove them during postprocessing of the data. Such a scheme has been proposed for the highly demanding space-based interferometric detection of low-frequency gravitational waves. It requires frequent sampling (multiple measurements over the round-trip light-travel time in each arm of the interferometer) in order to infer the Fourier components of the laser fluctuations from the spectra of the measured phase shifts. Given sufficient received coherent power, this approach can be used to reduce the measurement error caused by laser phase fluctuations sufficiently that it is no longer the dominant error.

One of the ultimate limiting sources of measurement error arises from photon-counting statistics, or the quantum-mechanical uncertainty principle that says that the accuracy with which the phase of a coherent light source can be measured scales inversely with the accuracy with which one can count the number of photons that have arrived at a detector in a given integration time. Lasers that are not amplitude-stabilized by feedback, no matter how stable in frequency, at best emit "coherent-state" light, consisting of photons whose arrival times at a detector are random and follow a Poisson

distribution; this randomness is a consequence of the quantum-mechanical zero-point fluctuations of the vacuum electromagnetic field. The resulting photocount statistics have the well-known $\sqrt{N}$ standard deviation, where N is the mean photocount. The photon-statistics limit to the fractional error $\delta L/L$ in a length measurement made with a laser wavelength λ is

$$\delta L/L = \frac{\lambda}{4\pi L\sqrt{N}} . \tag{2}$$

For example, for a laser interferometer with a 100-km baseline, a 60-watt laser operating at $\lambda = 1\ \mu$m, and an overall optics and detector efficiency of 2%, photon statistics would limit measurement sensitivity to 10^{-15} m in a 1-millisecond (ms) integration time. This would give a strain sensitivity of 10^{-20}, which still may be too poor for regular ground-based detection of gravitational waves. In astrometry, the limiting power comes not from a laser source, but from the brightness of the stars being observed. For a 10th-magnitude star observed at visible frequencies with 1-m apertures in a laser interferometer with a 10-m baseline and 2% overall efficiency (input starlight to detected photocount), photon statistics limit the angular precision to about 0.3 microarcsecond. Realization of such accuracy would require control of systematic errors and instrument geometry to about 10 picometers (pm). Recall from the discussion above that if a laser metrology system is used to accomplish this in an arrangement that provides common-mode cancellation of laser frequency fluctuations to 0.1%, the laser would have to be stable to about one part in 10^9, which is well within the reach of currently available technology.

When other sources of error (discussed below) are small enough that photon statistics dominate, there are techniques that can be used to increase the effective received coherent power and thus drive down even this error source. One of these is "power recycling," a technique likely to be important in ground-based laser-interferometric detectors of gravitational waves. Laser light that normally would exit the interferometer and be lost is recaptured and redirected coherently back into the interferometer. In this way the laser light circulating in the interferometer is built up to a level much

higher than the original input power, before a measurement of the relative phase shift is made. The success of this technique is limited ultimately by light losses (*e.g.*, imperfectly reflecting end mirrors). For very long baselines, the diffractive spreading of the light beam can produce an unacceptably weak signal at the measurement point. In this situation, the use of coherent laser transponders instead of retroreflectors is required; the corresponding diffractive losses scale as $1/L^2$ rather than $1/L^4$. Transponders would be required for space-based laser-interferometric detection of low-frequency (10^{-5} to 1 Hz) gravitational waves, and for high-resolution gravity-mapping (50 km spatial scales, $10^{-6}g$ gravity anomalies) of planets using laser interferometers aboard orbiting spacecraft.

Another kind of recycling technique, called "resonant recycling" or "resonating," can be used to increase the effective interferometer baseline and so increase the signal (phase shift) caused by a given relative displacement between two arms of an interferometer. Here the entire interferometer is made into a resonant cavity, and the laser light is allowed to resonate back and forth from arm to arm, thus building up a stronger phase shift between the two oppositely traveling light beams before they are recombined. Integration times may be limited ultimately by sign changes in the force being monitored, as is the case for detection of gravitational waves. In practice, more commonly it is light losses that limit the usefulness of these recycling techniques.

The presence of laser frequency fluctuations can make scattered light a difficult problem for precision interferometric measurements. The amount by which the phase of the main beam is changed due to the presence of scattered light scales with the square root of the ratio of power in the scattered light to power in the main beam and with the phase delay between the main and the scattered beam. This phase delay fluctuates as the laser frequency fluctuates. Furthermore, this error does not cancel in a dual-arm interferometer. One method of alleviating this involves deliberate phase modulation of the laser light, so that the phase delays tend to cancel when averaged over integral multiples of the modulation period. For best results, the modulation must be performed over a range of frequencies, tailored to suppress the dominant components of scattered light.

Fluctuations in laser power and in the spatial intensity distribution across the beam need not constitute a fundamental obstacle to measurement sensitivity. For most applications, they can be controlled adequately by feedback and mode-cleaning techniques. However, nanoradian (milliarcsecond) pointing accuracy is required for some applications. Radiation pressure from the randomly arriving laser photons typically is a negligible effect. But, unlike photon-statistics error, this effect produces a length-measurement error that grows, rather than decreases, with the square root of the integrated number of photons (laser power times integration time); hence it does preclude extending integration times arbitrarily in an effort to reduce the photon-statistics error. When laser power is high enough to make the radiation-pressure and photon-statistics errors comparable, the resulting limit to sensitivity is known as the "standard quantum limit" for detection of changes in the position of a free mass. Techniques for achieving sensitivities better than the standard quantum limit have been explored, but have not been implemented yet in practice.

Medium-induced phase noise (*e.g.*, refractive-index fluctuations) can be suppressed in some cases by using a dual-arm interferometer, although applications as demanding as gravitational-wave detection from the ground still require vacuums as good as 10^{-8} or 10^{-9} torr. The immunity of optical wavelengths to plasma-induced phase scintillation is an advantage for space applications in which ionospheric plasma effects otherwise would contribute unacceptable noise; lasers also can be used in conjunction with microwave transmitters to calibrate plasma effects and study plasma properties. Mechanical and thermal noise can constitute serious sources of systematic error and typically require the use of very stable materials, careful temperature control, and isolation from local accelerations. On Earth, local accelerations due to seismic noise and gravity gradients become increasingly significant at lower frequencies, so much so that Earth-based detection of gravitational waves at frequencies below about 10 Hz is deemed impossible. In space, active disturbance compensation systems are required to counteract fluctuating accelerations that arise from such effects as atmospheric drag, thermal and gravity gradients, and fluctuating radiation pressure from variations

in solar flux intensity.

The most demanding scientific application of frequency-stabilized laser technology being pursued currently is the detection of low-frequency gravitational waves (below 10 Hz), which by necessity must be done from space (Earth or solar orbit). Emitters of gravitational waves can be categorized as periodic sources, burst (or pulse) sources, and stochastic sources. Compact binary stars, supernovae, and individual compact objects are expected to be the most easily observed objects, producing dimensionless strains in an Earth-vicinity dual-arm interferometer that might be on the order of 10^{-20}. The appropriate measurement frequencies for these sources range from millihertz for periodic binaries to kilohertz for the millisecond bursts expected from collapsing objects. The advantages of space-based over ground-based laser interferometers in the low-frequency regime are (1) freedom from terrestrial gravity gradients and mechanical noise, thus enabling observations down to frequencies as low as 10^{-5} Hz or measurement time scales on the order of a day; (2) the potential for very long baselines — millions of kilometers with free-flying spacecraft; and (3) the ability to support test masses in a nearly inertial manner.

For the long baselines associated with space-based detection of low-frequency gravitational waves, laser transponders provide a critical advantage over retroreflectors. To achieve comparable measurement sensitivity limited by photon statistics, the laser power required with retroreflectors must be larger than that required with transponders by the ratio $(L\lambda/D_A D_B)^2$, where L is the baseline and D_A and D_B are the diameters of the apertures at each end of the link (transmit and receive aperture diameters at each station are assumed identical for this discussion). This ratio is $\sim 10^8$ for a 10-million-km baseline and 1-m apertures ($\lambda = 1$ μm)! Relative phase shifts corresponding to optical path-length changes as small as 10 pm over these long baselines must be monitored in order to reach strain sensitivities of 10^{-21}. Distances among free-flying spacecraft might be held equal to the 0.1% level, but active calibration of laser phase fluctuations still will be necessary. The drag-free technology needed to control or compensate for fluctuating accelerations

poses a significant challenge. It is estimated that the accelerations felt by the test masses in the spacecraft must be held to a level of $10^{-18}g$ or smaller. In addition to thermal and gravity gradients affecting the spacecraft, accelerations are caused by electrostatic forces (charge buildup from cosmic-ray impacts, for example) and random impacts of residual gas molecules on the test masses. The relative significances of these error sources depend on the measurement time scale. Photon statistics enforce a spectrally flat floor to measurement sensitivity on time scales of about 10 seconds to 10 minutes. Bombardment by residual gas molecules causes a fairly sharp decrease in sensitivity for time scales longer than about 10 minutes, and thermal gradients cause an even sharper decrease for time scales longer than about a day.

Another intriguing possible application of space-based laser interferometers is high-resolution mapping of planetary gravity fields. Short-wavelength spatial variations in the local gravity field and related undulations of the geoid could be mapped by monitoring changes in the range and range rate between spacecraft placed one behind another in a common polar orbit. A microwave version of this was proposed several years ago as the Geopotential Research Mission. Its best spatial resolution would have been a few hundred kilometers, and it would have been sensitive to gravity anomalies as small as $10^{-6}g$, or about 1 milligal (mgal). A laser version using coherent laser transponders rather than retroreflectors could provide spatial resolutions of 50 kilometers with similar sensitivity. The range-rate measurement accuracy σ_v required for sensing gravity anomalies depends on the desired spatial resolution $\lambda_l/2$ (λ_l is the spatial wavelength in a harmonic expansion of the gravity field), the orbit altitude h, the total number N_m of identical, independent measurements, and the along-track separation L between spacecraft in the following way:

$$\sigma_v \propto {N_m}^{1/2} \, {\lambda_l}^{3/2} \, e^{-2\pi h/\lambda_l} \, |\sin \pi L/\lambda_l|. \tag{3}$$

The maximum signal, hence the minimum requirement on range-rate accuracy, occurs for spacecraft separations L approximately equal to the desired spatial resolution $\lambda_l/2$. For measurement of a

1-mgal anomaly on Earth with 10^6 independent measurements at an orbit altitude of 160 km, range-rate measurement accuracies of approximately 50 nm/s would be required to achieve 50-km spatial resolution and 0.8 pm/s for 25-km resolution. The demand on range-rate measurement accuracy eases dramatically for spatial resolution poorer than about 75 km, or about half the orbit altitude. For example, in going from 50-km resolution to 100-km resolution, the range-rate accuracy requirement is relieved by a factor of about 500 — from 50 nm/s to 25 μm/sec.

To achieve a resolution better than about 100 km with an orbiting laser-interferometer gravity mapper, the dominant technology challenges are the drag-free system and the laser frequency stability. At an orbit altitude of about 160 km, the kinds of measurements described here would require disturbance compensation systems on the spacecraft capable of compensating for accelerations felt by measurement test masses down to about $10^{-13}g$. Technology for this has been proven, and development is in progress to improve performance by several orders of magnitude, motivated by the orders-of-magnitude more stringent drag-free requirements associated with space-based detection of low-frequency gravitational waves. For mapping gravity fields, a better drag-free system would permit a lower orbit altitude, hence stronger signals and improved sensitivity and spatial resolution.

Recall that in the detection of low-frequency gravitational waves, the requirements on laser frequency stability can be eased by making measurements frequently enough to calibrate the laser phase fluctuations. That approach cannot be used easily for this application because of the much smaller spacecraft separations ($\sim$ 100 km instead of 10^7 km). However, frequency-stability requirements could be eased by placing three, rather than two, spacecraft collinearly in the same orbit. If their separations were held equal to 0.1%, the resulting cancellation of laser frequency noise could ease the stability requirements by roughly three orders of magnitude. With two spacecraft whose separation is optimized for the desired spatial resolution, the laser fractional frequency stability (assuming one-second measurement integration times) would have to be at least as good as 5×10^{-12} to achieve 100-km spatial resolution, approximately 8×10^{-13} for 50-km resolution and 2.5×10^{-17} for

25-km resolution; the requirement continues to tighten sharply for better resolutions. Longer measurement integration times could ease these requirements somewhat, but not significantly. The use of three, rather than two, spacecraft for this application thus would provide considerable advantages and potential for high-resolution mapping.

Important areas of application that require coherent laser links but which do not make interferometric measurements of the sort described thus far are atmospheric lidar and scattering experiments. These experiments measure range and range rate from backscattered, Doppler-shifted laser signals. To be useful for weather forecasting, lidar measurements in Earth's atmosphere need range-rate measurement accuracies of 1 to 5 m/s and a vertical range resolution of about 1 km. Achievable range-rate accuracy σ_v is limited by the laser pulse spectral width to a value $\sigma_{v,\mathrm{bw}}$, by the spread of particle velocities in the scattering medium to $\sigma_{v,\mathrm{med}}$, and by photon statistics to $\sigma_{v,\mathrm{ph}}$. For a Gaussian pulse of spectral width $\delta\nu_p \simeq 0.3$ MHz [duration $\tau_p \equiv (2\pi\delta\nu_p)^{-1} \simeq 0.5$ μs at 1-μm wavelength], these accuracy limits are

$$\begin{aligned} \sigma_{v,\mathrm{bw}} &\simeq 0.15 \text{ m/s} ; \\ \sigma_{v,\mathrm{med}} &\simeq 0.25 \text{ m/s} \left(\frac{\sigma_{\mathrm{med}}}{0.4 \text{ m/s}}\right)^{1/2} ; \\ \sigma_{v,\mathrm{ph}} &\simeq 0.08 \text{ m/s} \left(\frac{1}{N_d}\right)^{1/2} . \end{aligned} \tag{4}$$

Here σ_{med} is the rms velocity spread in the scattering medium, and N_d is the number of detected photons. The range resolution σ_L for a 0.5-μs pulse (when the "range gate" is set equal to the pulse width τ_p) is on the order of 75 m ($= c\tau_p/2$) or larger. These values for range-rate accuracy and range resolution could be met with laser fractional frequency stabilities in the range 10^{-10} to 10^{-9}. Better stabilities for the reference laser used to generate the pulses could improve range-rate measurement accuracy by producing spectrally narrower pulses, but this improvement would be limited in practice by the maximum pulse duration set by the desired range resolution.

Another space-related scientific application that would use coherent laser links is occultation experiments in planetary (or satellite) atmospheres and in ring systems around giant planets. Atmosphere-occultation experiments attempt to measure refractivity profiles, which ultimately are used to estimate temperature and pressure as functions of height above the surface. Temperature measurements with errors smaller than 1% would require laser frequency stabilities ranging from 10^{-11} for Venus to 10^{-13} for Jupiter. Knowledge of the local vertical to a few microradians is also needed; this can be difficult to achieve for the giant planets because of differential motions in their atmospheres. Pointing the laser beam also may pose a challenge where atmospheric bending angles are large, as at Venus.

Ring-occultation experiments would use coherent laser links to measure the extinction and scattering from ring particles. Measurements of opacity and of the coherent phase shift caused by differential refraction can be used to solve for the column density and size distribution of ring particles whose sizes are comparable to or smaller than the laser wavelength. Information about the distribution of particles larger than the transmitter antenna (telescope) can be obtained from Doppler maps of the incoherently forward-scattered signal. To enable measurement of the coherent phase shift of laser light passing through ring structures, the laser may have to have a fractional frequency stability as good as 10^{-15}. But useful forward-scattering measurements of column density gradients could be made with frequency stabilities on the order of 10^{-11}. The Voyager radio-frequency experiments at Saturn gave information about particles ranging in size from 1 mm (wavelength-dependent extinction) to about 20 m (forward scattering). Similar experiments with lasers would provide complementary information about particles ranging from submicron size (extinction) to meter-size (forward scattering), and also would serve to verify information gained from radio measurements for these size ranges. Since past microwave experiments suggest that most ring particles range in size from centimeters to meters, the use of *both* laser and radio transmitters would be desirable for obtaining maximum information about ring particles.

Numerous other scientific applications of frequency-stabilized lasers and related technology exist, many of which use techniques similar or analogous to those discussed thus far. General mission enhancement will result from the use of coherent laser links for telemetry and navigation and tracking. Laser communication links from space are capable of much higher data rates than radio-frequency links, and coherent links would enable excellent performance even under conditions of high background light, *e.g.*, for missions near the Sun. Laser links to spacecraft can also be used for accurate, real-time remote optical tracking and navigation through astrometric and coherent Doppler measurements. Frequency-stabilized lasers also would enable a variety of improvements to onboard navigational capabilities and science instruments through the addition of sensitive tools such as coherent lidar and laser-gyro rotation sensors, as well as higher data-rate capabilities.

Coherent laser links to and among spacecraft will enable a wide variety of solar-system science experiments and astrophysical measurements. Ephemerides could be improved through remote optical tracking of laser-carrying spacecraft. The combination of remote optical tracking and *in situ* measurements among laser-carrying spacecraft offers the possibility of highly accurate measurements of planet and satellite masses and gravity fields. When combined with topographic data obtained with Earth-based photographic or radar data and perhaps augmented with lidar used on orbiters or during fly-by missions, this gravity data can provide information about the interiors of terrestrial planets, such as the degree of isostatic compensation, the radial temperature profile, and elastic properties. For the giant planets, the gravity data could be used in conjunction with independent information on rotation rate and radii to deduce the density distribution, put constraints on composition, and constrain possible models of the interiors. Cometary masses could be deduced by combining remote astrometric measurements made among the comet, spacecraft, and sun with precise measurements of range and range-rate made with coherent lidar from the spacecraft to the comet, and with onboard accelerometer measurements.

Studies of planetary atmospheres would benefit both from occultation experiments made with

long-range coherent laser links (described above) and from *in situ* coherent lidar measurements. Coherent laser scattering measurements would be particularly useful for cloud studies. Traditional "physical sampling" measurements involve capture of particles on substrates or aspiration of particles through insertion of instruments into the clouds. These provide highly local information only, they disturb the airstream, and they are tedious and time-consuming. In contrast, laser light-scattering measurements are efficient and nonintrusive. They would be applicable for particles in the 1 to 50 μm size range, which includes fog droplets (of order 1 μm) and cloud droplets (10 to 20 μm).

Coherent Doppler lidar also could be used to measure trajectories and other properties of cosmic and interplanetary dust particles and to provide insight into their origins. It is not known what fraction of the interplanetary dust particles comes from comets or asteroids, nor what fraction is actually interstellar dust that has penetrated the solar system. Chemical or isotopic information alone is inadequate to answer these questions because of the diversity of compositions among comets. Trajectory information is essential to correlate physical characteristics with specific sources. Conventional techniques to measure velocities of dust grains (such as those proposed for the space-station-based Cosmic Dust Collection Facility) involve capture and destruction of individual grains in order to measure their relative positions and times of flight between specified points. A coherent pulsed Doppler lidar system, operating with 10-μs pulses of relatively narrow spectral width, say 30 kHz (requiring fractional frequency stabilities on the order of 10^{-10} for the reference laser oscillator, assuming an operating wavelength of 1 μm), could measure velocities of particles in the 1 to 10 μm size range with a precision of about 35 m/s, or about 0.2 to 0.4% of the estimated average particle velocity. This presumes detector collecting areas on the order of 100 m^2 and integration times of 5 to 10 minutes, consistent with estimated impact rates of about 1,000/m^2-yr for particles of this size.

Studies of the Sun itself — its quadrupole moment and total angular momentum, its mass moments, density distribution, shape and dimensions, surface composition, and luminosity, and

properties of the solar wind and corona — traditionally have been carried out with dual radio-frequency links to spacecraft. Some of this information has been obtained through measurement of the differential effect of the solar plasma on the group delay of beams at two radio wavelengths, but much of it has relied on accurate tracking of the spacecraft as it passes within a few degrees of the Sun. A coherent laser link to such a spacecraft would be free of plasma-induced noise in the measurement of phase, and, when used in conjunction with precise astrometric tracking from Earth or Earth-orbit, could enable greatly improved measurements of the solar gravity field, including the magnitude and orientation of its quadrupole moment, J_2. Such measurements would provide information about rotation of the solar interior and lead to improved estimates of the solar mass distribution and total angular momentum J. This information would impact current theories of star and planet formation and also would enable sensitive solar-system tests of theories of gravitation. With laser fractional frequency stabilities on the order of 3×10^{-14} over several hours, or stable line widths of about 10 Hz at 1-μm wavelengths, Doppler velocity-measurement accuracies of about 0.1 mm/s could be achieved for laser-carrying spacecraft near the Sun, provided active disturbance-compensation systems were used to reduce stochastic accelerations on the spacecraft test mass to about $10^{-10}g$. (One-minute integrations were assumed for these estimates.) This tracking accuracy would enable inference of J_2 to an accuracy of about 2×10^{-8}, approximately 10% of its estimated value and about five times better than it is known now.

Finally, frequency-stabilized laser technology will play a vital role in solar-system tests of theories of gravitation, especially general relativity. The majority of such tests involve highly accurate tracking of a spacecraft under high-background conditions, generally near the Sun. This would be accomplished both with coherent laser links between Earth and laser-carrying spacecraft and with Earth-orbiting optical astrometric interferometers, whose accuracy is enabled by laser metrology systems employing frequency-stabilized lasers. Range-rate accuracies of 0.1 mm/s or better, ranging accuracies of a few centimeters of better, and/or submilliarcsecond angular accuracies are needed for

such tests to be significantly more conclusive than any made thus far. To achieve these performance levels, spacecraft must have active disturbance-compensation systems to reduce nongravitational accelerations to an effective level of $10^{-10}g$ or smaller over time scales on the order of a minute; this is roughly five orders of magnitude smaller than the buffeting that would be experienced by a solar-orbiting spacecraft due to fluctuations in the solar-wind intensity, for example. Alternatively, planetary orbiters or — still better — landers should be used. Candidate tests include measurement of the advance of the perihelion of Mercury (approximately 43 seconds of arc per century), which depends critically on our knowledge of J_2; measurement of the predicted time delay and deflection of light rays passing near a massive body; and measurement of the gravitational redshift of light, or, more generally, tests of the principle of equivalence. Measurements of the gravitational redshift would benefit from sensitive Doppler measurements made with a coherent laser link to a spacecraft near the Sun. Violations of the equivalence principle should also appear as measurable anomalies in orbits of planets and satellites; such tests would benefit significantly from the use of coherent laser transponders placed, for example, on the lunar surface.

The following table summarizes estimates of the fractional laser frequency stabilities $\delta\nu/\nu$ required for several types of scientific applications in space, as derived and discussed in this report. These estimates assume reasonable values for characteristics of the observing systems (as explained in the text) and the best information available about the phenomena to be observed (*e.g.*, expected signal strengths). In general, the listed values are on the conservative side of the allowable ranges. For interferometric astrometry, gravitational-wave detection, and mapping of Earth's gravity field, it is assumed that dual-arm interferometer configurations or their equivalent are used, so that laser frequency stability requirements are three orders of magnitude less severe than they would be in the absence of any common-mode cancellation of errors due to laser frequency fluctuations. The last column gives references to appropriate equations or discussions in the text.

Requirements on spacecraft laser frequency stability for a variety of scientific applications. Conservative estimates are given. For astrometry, gravitational-wave detection, and gravity-field mapping, it is assumed that dual-arm interferometers or their equivalent are used, to provide cancellation of errors due to laser frequency fluctuations to 0.1%. L denotes interferometer arm length.

Scientific application	Measurement technique	Observable	Information sought	$\delta\nu/\nu$	Reference
Microarcsecond astrometry	interferometry (space) $L \sim 2$ to 20 m	optical-path difference	star angular position and distance	5×10^{-9}	eq. (5.1.3)
Gravitational waves, low frequency	interferometry (space – solar orbit) $L \sim 10^7$ km	optical-path change	strain $\gtrsim 10^{-21}$	10^{-18*}	eq. (3.1.3)
Gravitational waves, pulses	interferometry (ground) $L \sim 5$ km	optical-path change	"	10^{-17}	eq. (3.4.6d)
High-resolution gravity-field mapping (1 mgal, resolution 40 to 50 km, 10^6 1-s measurements)	interferometry (in orbit)	change in range rate	gravity anomalies, geoid undulation	10^{-10} to 10^{-9}	eq. (5.3.10)
Atmospheric wind-sensing	pulsed Doppler lidar	velocity, range	wind speeds	10^{-10}	eq. (5.4.1)
Atmospheric occultation, Venus	one-way Doppler	Doppler shift, intensity	temperature, pressure profiles	10^{-11}	Table 5.5.1
Atmospheric occultation, Jupiter	one-way Doppler	"	"	10^{-13}	"
Ring occultation	one-way Doppler	"	phase shift complex extinction	10^{-15}	§5.5.2c
Ring occultation	one-way Doppler	"	forward-scattering cross section, column-density gradient	10^{-11}	eq. (5.5.16)

*With active laser phase calibration and appropriate tracking filters, this may be relaxed to 10^{-13} and perhaps made even less stringent.

1. INTRODUCTION

1.1 Purpose of report

The development of coherent microwave technology — stable radio-frequency sources such as the hydrogen maser, S- and X-band coherent transponders suitable for use on satellites and planetary spacecraft, and low-noise phase-sensitive receiver/amplifier/detector systems — has been the basis for many scientific and technical advances made during the last few decades. These advances have occurred in such areas as Earth science and remote sensing, planetary exploration and solar-system science, radio astronomy, and astrophysics. Coherent microwave technology also has contributed significantly to advances in atomic and molecular physics and has provided means to test a variety of predictions made by quantum physics and general relativity.

By comparison with microwave technology, the development of coherent laser technology at optical and infrared wavelengths is in its infancy. Since the first observation of a laser (fluorescent ruby) in 1960 (Maiman 1960), major advances have been made in understanding the lasing process and properties of lasing media and in manipulating lasing media to exhibit optimal gain as well as optimal spatial- and temporal-mode characteristics at desired wavelengths. Recent advances in solid-state laser technology, including the use of diode lasers to pump specially prepared crystals and techniques for controlling the temporal- and spatial-mode properties of the output light (Byer 1988; Fan and Byer 1988), have brought laser technology to the point where it is a viable tool for some sophisticated, hitherto impossible scientific measurements involving Earth-orbiting or interplanetary spacecraft. These include detection of long-period (one second to one day) gravitational waves from binary star systems and detection of gravitational pulses produced during the era of galaxy formation from the collapse of massive (10^4 to 10^6 solar masses) stars to form black holes; high-resolution mapping of planetary gravity fields with laser interferometers among coorbiting spacecraft (30 to 100 km resolution for Earth); light deflection and other sensitive tests of relativity and gravitation;

measurement of the solar quadrupole moment; and coherent-light-scattering experiments to probe the short-wavelength structure, composition, and dynamics of planetary rings and atmospheres.

Science measurements such as these put differing demands on laser sources for power, spatial or temporal coherence, reliability and lifetime, compactness, and ruggedness. Typically, they require low-noise infrared or optical detectors that can perform near the limits imposed by quantum mechanics. But a primary reason these measurements have not been practical at optical or infrared wavelengths is their stringent demand on laser frequency stability. Detection of low-frequency gravitational waves may be one of the most demanding in this regard, as it may require lasers with at least one watt of power and fractional frequency stabilities of 10^{-13} or better ($\leq$ 30 Hz) over several seconds to hours, operating autonomously on spacecraft that are coorbiting with Earth around the sun. Another demanding application is high-resolution mapping of the Earth's gravity field, down to scales of several tens of kilometers with sensitivities of 1 milligal ($10^{-6}g$) for gravity anomalies and 10 cm for geoid undulations. This could be accomplished with a laser interferometer between two spacecraft in orbit together (one behind the other), but would require fractional frequency stabilities as good as 10^{-14}; three spacecraft forming a collinear dual-arm interferometer might ease this requirement on laser frequency stability by several orders of magnitude. Other applications might require only a few hundred milliwatts of power and stabilities of 10^{-13} to 10^{-10} over shorter times. Compact, diode-pumped solid-state lasers that are excellent candidates for space-qualifiable lasers have been feedback-stabilized to a few hertz (approximately 10^{-14}) at output powers of a few milliwatts (Day, *et al.* 1990); at output powers of several hundred milliwatts they have exhibited short-term free-running stabilities of a few kilohertz (Kane, *et al.* 1987; Byer 1988; Fan and Byer 1988; Bush, *et al.* 1988). The prospects for improvements in both power and frequency stability are excellent.

Many science applications also require or would be enhanced by the addition of coherent laser transponders, devices that receive, amplify, and retransmit laser signals without losing informa-

tion about the phase of the received signal. The scientific applications of coherent optical and infrared phase measurements are myriad. Frequency-stabilized laser sources and phase-matching laser transponders will be key technologies for scientific use, both in space and in the laboratory.

This report describes several possible ways in which frequency-stabilized laser technology could be used in space for scientific measurements. The kinds of measurements that exploit high stability basically are of two types: interferometric measurements of angle and coherent Doppler measurements of relative velocity (range rate). Some applications actually use both kinds of measurements at once. The challenging applications just cited of low-frequency gravitational-wave detection and high-resolution mapping of planet gravity fields, for example, both involve coherent interferometric measurements of the changing separation — range rate — among laser-carrying spacecraft. The baseline lengths for these two applications are quite different — 10^7 km for the gravitational-wave interferometer and 10 to 100 km for the gravity-field mapper. Coherent laser systems are desirable for both because of the potential for high-accuracy range-rate measurements and quantum-limited signal-to-noise performance.

1.2 Outline of report

While scientific applications differ in their implementations of frequency-stabilized laser technology and in the objects of their measurements, many share features fundamental to measurements that exploit the temporal (and spatial) coherence of an electromagnetic wave. This report therefore begins (chapters 2 and 3) with a heuristic description of the temporal coherence of a laser and its relationship to the precision with which certain kinds of measurements can be made. General types of measurements possible with laser sources of high temporal coherence are described, and additional factors that limit measurement precision are pointed out.

Chapter 4 describes in detail two laser-based measurement techniques that would be used in various forms for all the applications discussed in this report. These techniques are interferometric

measurements and coherent measurements of relative velocity (range rate). Various limitations to these techniques are discussed, based on the descriptions of lasers and precision measurements given in chapters 2 and 3.

Chapter 5 examines in detail five areas of scientific application that use frequency-stabilized lasers and related technology. The first three — metrology systems (*e.g.*, for astrometric or imaging interferometers), gravitational-wave detection, and gravity-field mapping — involve interferometric measurements of an optical path difference. The last two — atmospheric wind-sensing and studies of planetary atmospheres and ring systems — involve precise Doppler measurements with coherent laser links. Discussion of the performance and limitations of these experiments is based on the analyses of the previous chapters. Chapter 6 describes qualitatively some additional scientific applications of frequency-stabilized lasers in space. Chapter 7 is a brief summary of the report, including a table summarizing the laser frequency stabilities needed for a variety of applications. (The Executive Summary provides a more detailed summary.) An appendix gives some supporting details for the analysis in section 5.3 of gravity-field mapping using coorbiting laser-carrying spacecraft.

2. FREQUENCY-STABILIZED LASERS

Lasers are quantum-mechanical devices that generate intense, coherent, extremely monochromatic beams of light at optical or infrared wavelengths. This section considers, in turn, each of these features — coherence, high intensity, and monochromaticity — and how they are realized. Section 2.1 reviews some of the basic principles of lasers, focusing on achievement of coherence and high output intensity. Section 2.2 looks closely at the issues of monochromaticity — or how stable, single-frequency operation can be achieved.

2.1 Review of principles of lasers

2.1.1 Laser coherence

In the description of light as an electromagnetic wave or wave packet made up of certain frequencies, polarizations, and directions of motion (plain-wave modes), coherence is a measure of how well the amplitudes and phases of each component mode are defined. A familiar consequence of phase coherence, for example, is interference phenomena. The generation of intense coherent radiation at optical and higher frequencies would be as straightforward as it is at radio frequencies, if the classical, wavelike description of the electromagnetic field were adequate. According to that description, an oscillating electric dipole produces an electromagnetic field whose amplitude and phase at each point in space and time are related in a precise way to the amplitude and phase of the oscillating current of the dipole. But quantum theory maintains that the amplitude and phase of an electromagnetic wave are complementary observables and therefore cannot be measured simultaneously with arbitrary precision. Hence there is no realizable "state" or mode of the electromagnetic field with an exact amplitude and phase at every point in space and time. Of course, such idealizations — plane-wave modes — are used frequently as bases for mathematical descriptions of the electromagnetic field. But in nature (according to quantum theory), the state closest in character to the idealization of a plane

wave with perfectly defined amplitude and phase is a so-called "coherent state." Coherent states are plane-wave modes whose amplitudes and phases have uncertainties, or standard deviations, that are identical in magnitude (when suitably normalized, *e.g.*, to the square root of photon number, or the square root of energy divided by the energy $h\nu$ per photon), and equal to the minimum allowed by uncertainty principles. (For further discussion and definitions, see, *e.g.*, Louisell 1973 and references therein, or Loudon 1983.) These minimum "zero-point" or vacuum-field fluctuations contribute to the total electromagnetic field a combined energy density equivalent to one-half that of one photon, $h\nu/2$, for each possible frequency, polarization, and direction of motion (h is Planck's constant and ν is frequency). They are responsible, for example, for causing spontaneous transitions among atomic levels with consequent emission of photons. They are responsible for the "shot noise" associated with photon-counting detectors.

Zero-point fluctuations of the vacuum electromagnetic field also are responsible for the difficulty of generating coherent radiation at optical and higher frequencies compared to radio frequencies. To see why, consider what coherence means in terms of a quantum-mechanical photon description of the electromagnetic field, rather than the classical wave description. Roughly speaking, coherence is a measure of the relative contribution to the total energy flux in a beam of light arising from identical photons with specific energies, polarizations, and directions. The energy contribution from individual optical-frequency photons is several orders of magnitude larger than that from individual radio-frequency photons. Hence the relative energy contribution from vacuum fluctutations, which is not associated with photons of any preferred polarization or direction, causes negligible degradation of overall coherence at radio frequencies, but not at optical and higher frequencies. An X-band photon (wavelength $\lambda \simeq 3.6$ cm) has an energy of only 5.6×10^{-24} joules (J), roughly five orders of magnitude smaller than that of an optical photon at wavelength $\lambda = 0.5$ micron (μm), 4×10^{-19} J. At finite temperatures, an electromagnetic field also contains thermal photons, which have no preferred polarization or direction. The energy contribution to the field per mode from thermal photons is

almost two orders of magnitude greater than from X-band vacuum fluctuations even at cryogenic temperatures — $k_B T = 4 \times 10^{-22}$ J at only 30 K. [Here k_B is Boltzmann's constant, and T is temperature in kelvins (K).] Hence at radio frequencies, coherence is limited ultimately by the presence of thermal photons, not quantum-mechanical vacuum fluctuations. But at optical and higher frequencies, coherence is limited ultimately by vacuum fluctuations.

In traditional atomic lasers, optical or infrared photons are emitted as a result of certain radiative transitions among atomic states (energy levels associated with the valence electrons of atoms). The allowed transitions, which restrict the characteristics of the emitted radiation, are determined by internal symmetries and conservation laws. The probability that a particular transition will occur, resulting in emission of a photon with a particular frequency, polarization, and direction, is proportional to the relative contribution to the total radiation energy density in the vicinity of the atom from photons with those same characteristics. Because the vacuum fluctuations contribute an energy density equivalent to one-half that of one photon for every frequency, polarization, and direction, they will serve to stimulate the emission of photons of every such allowed characteristic. Such emission is called spontaneous because it occurs even if the mean number of photons with those characteristics is zero in the vicinity of the atom. The probability of such emission is determined by the nature of the transition, specifically by the energy difference between the atomic levels and the quantum-mechanical probability amplitude (matrix element) for a transition between the two atomic states. The emission of photons of certain allowed characteristics will be enhanced if additional photons with those characteristics are present — a process known as stimulated emission. In order to produce highly coherent light, consisting of many photons with identical characteristics, the probability for stimulated emission must be greater than that for spontaneous emission; *i.e.*, the mean number of photons present with the desired characteristics must be greater than one per mode.

To quantify the criteria for making stimulated emission dominate spontaneous emission, consider

a simple physical argument given by Einstein (1917) to explain empirical observations, an argument that does not appeal to quantum theory or refer to vacuum fluctuations. Denote two (infinitely narrow) energy levels of an atomic system by 1 and 2, between which radiative transitions occur and produce photons of energy $h\nu$. Denote the probability per unit time of a *spontaneous* transition from level 2 to 1 with emission of a photon of energy $h\nu$ by a. (Physically, $1/a$ is identified with the luminescence decay time.) Einstein postulated that the probabilities of *stimulated* transitions between the two levels, involving either absorption or emission of a photon of energy $h\nu$, were identical, and were proportional to the energy spectral density of the local radiation, ρ_ν (energy per unit volume and unit frequency interval). Denote that proportionality factor by b. (The coefficients a and b are calculated in most standard quantum mechanics texts, *e.g.*, Merzbacher 1970 or Schiff 1968; both are proportional to the square of the atomic-dipole matrix element.) The probabilities for absorption and emission of a photon therefore have the following forms:

$$P_{\rm abs} = \rho_\nu\, b\,; \qquad P_{\rm em} = P_{\rm spon} + P_{\rm stim} = a + \rho_\nu\, b\,. \tag{2.1.1}$$

The radiation spectral density ρ_ν is equal to the product of the number of plane-wave modes per unit spatial volume and unit frequency interval n_ν, the average number of photons in each mode $\overline{N_\nu}$, and the energy of each mode $h\nu$. The number of modes per unit volume in the frequency interval $d\nu$ is just the three-dimensional momentum-space volume element multiplied by 2 (for two polarizations):

$$n_\nu\, d\nu = 2\,\frac{d^3p}{h^3} = \frac{8\pi\nu^2}{c^3}\,d\nu\,. \tag{2.1.2}$$

The radiation energy spectral density ρ_ν is therefore

$$\rho_\nu = h\nu\, n_\nu\, \overline{N_\nu} = \frac{8\pi h\nu^3}{c^3}\,\overline{N_\nu}\,. \tag{2.1.3}$$

For a black body in thermal equilibrium at temperature T, the average number of photons per mode

is given by Planck's distribution (which follows from assuming a Boltzmann energy distribution):

$$\overline{N_\nu}^{\rm bb} \;=\; (e^{h\nu/k_{\rm B}T} - 1)^{-1} \;. \tag{2.1.4}$$

The average number of photons absorbed is $N_1\ P_{\rm abs}$, and the average number emitted is $N_2\ P_{\rm em}$, where N_1 and N_2 are the mean numbers of atoms in the lower and upper levels, respectively. In thermal equilibrium, the rate of absorption of photons must balance exactly the rate of emission of photons:

$$N_1\ \rho_\nu\ b \;=\; N_2\ (a + \rho_\nu\ b) \;. \tag{2.1.5}$$

With the equilibrium ratio of populations given by a Boltzmann distribution*, $N_2/N_1 \;= e^{-h\nu/k_{\rm B}T}$, the ratio of stimulated to spontaneous emission intensity is equal to the mean number of photons per mode, $\overline{N_\nu}^{\rm bb}$ [eq. (2.1.4)]:

$$\frac{I_{\rm stim}}{I_{\rm spon}} \;=\; \rho_\nu\ b/a \;=\; \overline{N_\nu}^{\rm bb} \;=\; (e^{h\nu/k_{\rm B}T} \;- 1)^{-1} \;. \tag{2.1.6}$$

This is exactly the result stated earlier: spontaneous emission is like stimulated emission, with the vacuum fluctuations serving as the stimulating radiation energy density, equivalent to an average of one photon in each mode of the field.

The condition for stimulated emission to dominate spontaneous emission is that the above ratio exceed unity, or that the product of temperature and wavelength satisfy

$$\left(\frac{T}{300\ {\rm K}}\right)\left(\frac{\lambda}{69\ \mu{\rm m}}\right) \geq 1 \;. \tag{2.1.7}$$

At submillimeter and longer wavelengths, this condition is met easily, even at cryogenic temperatures. But at optical and shorter wavelengths, the mean number of appropriate stimulating photons must

* Use of a Fermi distribution (appropriate for electron statistics) will lead to a similar result, since the energies of the atomic levels (relative to the chemical potential) are greater than $k_{\rm B}T$.

be boosted with an external pumping source in order to meet this condition. An external source of optical photons that produces a spectral flux density $I_\nu \equiv I_0/\Delta\nu$ (expressed here as a flux density I_0 evenly distributed over a bandwidth $\Delta\nu$, with units of W/m^2-Hz, for example) will produce a radiation energy spectral density $\rho_\nu = (8\pi/c)\, I_\nu$. Since the ratio of the Einstein coefficients is

$$\frac{a}{b} = h\nu\, n_\nu = \frac{8\pi h\nu^3}{c^3} \tag{2.1.8}$$

[eqs. (2.1.3) and (2.1.6)], the ratio of stimulated to spontaneous emission in the presence of an external pump is

$$\frac{I_{\rm stim}}{I_{\rm spon}} = I_\nu\, \frac{c^2}{h\nu^3} = \left(\frac{I_0}{100\ \text{mW/cm}^2}\right)\left(\frac{3\ \text{GHz}}{\Delta\nu}\right)\left(\frac{\lambda}{1\ \mu\text{m}}\right)^3 . \tag{2.1.9}$$

Thus, with a sufficiently strong pump source (*e.g.*, 100 mW/cm^2 of 1-μm wavelength radiation over a 3-GHz bandwidth), stimulated emission can be made to dominate spontaneous emission at room temperature and visible wavelengths.

2.1.2 Laser intensity

While coherence requires that stimulated emission dominate spontaneous emission, the other primary feature of laser light — high-intensity output — requires that stimulated emission also dominate absorption. Equations (2.1.1) and (2.1.5) above show that in thermal equilibrium the ratio of the light intensities of emitted (stimulated) and absorbed light is equal to the ratio of populations in the upper and lower levels:

$$I_{\rm stim}/I_{\rm abs} = N_2/N_1 = e^{-h\nu/k_{\rm B}T} . \tag{2.1.10}$$

Hence a laser requires a mechanism for maintaining a nonequilibrium population inversion between the upper and lower levels, with the ratio N_2/N_1 kept as large as possible. Note that external pumping with optical photons of the desired transition frequency will not produce the required population inversion. The pump light would induce transitions equally in both directions, leading

ultimately (in the limit that spontaneous emission is negligible relative to stimulated emission) to an equalization of the populations [eqs. (2.1.5–7)], but no inversion.

Physical characteristics of the electromagnetic radiation produced by a lasing system are controlled by placing the laser in a resonant cavity, whose natural vibration modes match and therefore amplify the waves from the laser, leading to the formation of steady, standing-wave oscillations. Because in practice the dimensions of cavities are much larger than a wavelength of light (*e.g.*, millimeters or longer), a given cavity may be compatible with many different modes. To produce highly monochromatic, coherent light, the cavity shape and characteristics must be optimized to allow the minimum possible of modes. Associated with each cavity-mode oscillation at frequency ν is a quality factor Q ($Q >> 1$), or time constant $\tau_c \equiv Q/2\pi\nu$ that can be viewed as the average photon lifetime in the cavity. The instantaneous energy density in the cavity mode at time t is $U(t) = U_0\ e^{-t/\tau_c}$. Thus $Q/2\pi$ is (approximately) equal to the ratio of the initial energy in the cavity mode to the energy lost per cycle:

$$\frac{U_0}{U_0 - U(t = 1/\nu)} = \frac{1}{1 - e^{-2\pi/Q}} \simeq \frac{Q}{2\pi} \equiv \nu\ \tau_c\ . \tag{2.1.11a}$$

If r is the cavity mirror reflectivity (assumed identical for both ends), t the transmissivity, and l the loss due to absorption, scattering, and diffraction, the overall decay time τ_c for a laser cavity of length L is

$$\begin{aligned} \tau_c &= \frac{L}{c(1-r)} \simeq 3.3\ \text{nsec} \left(\frac{L}{1\ \text{mm}}\right) \left(\frac{0.001}{1-r}\right), \\ (1-r) &\equiv t + l\ . \end{aligned} \tag{2.1.11b}$$

For a cavity length $L = 1$ mm and mirror reflectivity $r = 0.999$, the time constant is about 3.3 nsec, which corresponds to a cavity spectral full width at half maximum (FWHM), or "cold-cavity" line width, $\delta\nu_{lc} \equiv (2\pi\tau_c)^{-1} \simeq 50$ MHz. The average photon lifetime, or luminescence decay time a^{-1}, for an atomic transition sets an upper limit for the cavity decay time; for atomic dipole

transitions that produce optical or infrared photons, these times are typically 1–10 nsec. The loss coefficient associated with a laser cavity is

$$\alpha \equiv \frac{1-r}{L} = \frac{1}{c\tau_{lc}} \simeq 0.01 \text{ cm}^{-1} \left(\frac{1 \text{ mm}}{L}\right) \left(\frac{1-r}{0.001}\right) . \qquad (2.1.11c)$$

In atomic systems, creation of a population inversion typically requires three- or four-level systems, such as those depicted in Figure 1. Three-level systems (Figures 1a,b) suffer the disadvantage that initially all of the atoms are concentrated into the lowest level, and inversion must be reached against this strongly populated level. A four-level system (Figure 1c) need not have this disadvantage, and therefore can operate with a lower pump intensity. The chromium ion Cr^{3+} is a three-level quantum system commonly used, *e.g.*, in ruby lasers (Maiman 1960). The neodymium ion Nd^{3+} is a four-level system, commonly used in solid-state [glass or yttrium-aluminum-garnet (YAG)] lasers (Byer 1988).

Although lasers based on atomic systems must involve more than two energy levels, other kinds of two-level laser systems do exist. An example is a semiconductor laser (Figure 2), where an inversion is reached at a "P-N" junction between regions of different types of conductivity, N- type (electrons) and P-type (holes). At the junction, free electrons and holes recombine and emit light. Electrical pumping ensures a constant flow of electrons and holes toward the junction.

Gas lasers (Figure 3) are an example of another way to produce the requisite population inversion. Here two systems exchange energy. One, a two-level system, is pumped to excite its upper level, from which energy is transferred to the highest energy level of the second, three-level system. The second system undergoes a rapid, nonradiative transition to its second, metastable level, and inversion is reached between the first and second levels of the three-level system. A helium-neon laser operates in this way: helium atoms are pumped from the ground state to their 2^3s level by an electric discharge and this energy is then transferred by inelastic collisions to the neon atoms,

bringing the latter to their $2s$ state.

2.2 Laser frequency stability

Turn now to the monochromaticity of a laser. While other kinds of instabilities in lasers, such as fluctuations in power or beam geometry, can be important sources of error in some laser-based measurements, frequency stability is often one of the most severe error sources, and it can pose a significant technical challenge.

2.2.1 Line width and noise-power spectral density

A laser's output is centered about some frequency ν (the nominal radiative transition frequency), but its instantaneous frequency varies because of stochastic processes such as spontaneous emission and deterministic processes such as environmental or apparatus-related disturbances. Deterministic processes cause the center frequency to drift with time, and both stochastic and deterministic processes produce a finite spectral line width $\delta\nu$. This section provides a heuristic overview of the characterization and measurement of laser frequency stability. For a more rigorous discussion, the reader is referred elsewhere (*e.g.*, Allan 1966; Barnes, *et al.* 1971; Howe 1976).

The electric field $E(t)$ at the output of a laser operating at central frequency $2\pi\nu_0$ with amplitude E_0 and time-varying phase $\phi(t)$ has the form:

$$E(t) \;=\; E_0 \cos[2\pi\nu_0 t \;+\; \phi(t)] \;. \tag{2.2.1}$$

In general, the instantaneous phase $\phi(t)$ may result from some Gaussian random process. The instantaneous frequency excursion, or deviation from the central frequency ν_0, is

$$\Delta\nu(t) \;\equiv\; \frac{1}{2\pi}\,\frac{d\phi(t)}{dt} \;. \tag{2.2.2}$$

For time scales longer than a microsecond or so, frequency stability is often characterized by a time-averaged measurement of the observed frequency fluctuations. The conventional quantity is called an Allan variance (or its square root, the Allan standard deviation; Allan 1966). The Allan variance is calculated by making successive measurements of the fractional frequency excursions $\Delta\nu(t)/\nu_0$ at regular time intervals τ, squaring the difference between successive measurements and averaging those squared differences over all the measurements. Mathematically it is defined as

$$\sigma^2(\tau) \equiv \frac{1}{2N{\nu_0}^2} \sum_{j=1}^{N} [\overline{\Delta\nu}(t_{j+1}) - \overline{\Delta\nu}(t_j)]^2 , \tag{2.2.3a}$$

$$\overline{\Delta\nu}(t_j) \equiv \frac{1}{\tau} \int_{t_j}^{t_j+\tau} \Delta\nu(t)\, dt . \tag{2.2.3b}$$

Because it is an integral measure of the fluctuations and hence contains contributions from even the lowest frequencies, the Allan variance is appropriate for characterizing frequency stability even on very long time scales (minutes or hours). Throughout this report, a simple, heuristic description of frequency stability and its relation to precision science measurements is used, and Allan variances will not be referred to again explicitly. However, when required levels of fractional frequency stability are indicated for the various scientific measurements discussed, the Allan standard deviation is the quantity that must take on those indicated values.

On short time scales, frequency stability is characterized by a power spectral density and measured by direct spectral analysis. A power spectral density $S_\nu(f)$ of the frequency fluctuations is defined as the Fourier transform of the autocorrelation $C_\nu(\tau)$ of the instantaneous frequency excursions $\Delta\nu(t)$:

$$\begin{aligned} S_\nu(f) &= \int_0^\infty d\tau\, C_\nu(\tau)\, \cos 2\pi f\tau , \\ C_\nu(\tau) &\equiv \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} dt\, \langle \Delta\nu(t)\, \Delta\nu(t+\tau) \rangle , \end{aligned} \tag{2.2.4}$$

where the angle brackets $\langle \ldots \rangle$ denote an appropriate ensemble average of the measured values. When the spectral density $S_\nu(f)$ can be represented by power-law noise processes, it can be related in a simple way to the Allan variance (Barnes, *et al.* 1971). For example, "white" frequency noise corresponds to an Allan variance that increases for shorter measurement times as $1/\sqrt{\tau}$, while "flicker" frequency noise [$S_\nu(f) \propto 1/f$] corresponds to a constant Allan variance and white phase noise [$S_\nu(f) \propto f^2$] corresponds to an Allan variance that decreases for longer measurement times as $1/\tau^2$.

The spectral line width $\delta\nu$ of the laser is also an integral measure of the fluctuations and can be used to characterize long-term frequency stability; it will be used in a general way throughout the rest of this report to characterize laser frequency stability. Its relation to the power spectral density $S_\nu(f)$ of the fluctuations in laser frequency ν depends on the amplitudes and frequencies f of the fluctuations. For example, it can be shown (see, *e.g.*, Elliot, *et al.* 1982) that in the limit that the amplitudes of the fluctuations, or frequency excursions, are small and the fluctuations occur rapidly, the rms frequency excursion and the laser line width scale linearly with the power spectral density $S_\nu(f)$. In the other extreme of large frequency excursions that occur slowly, the rms frequency excursion and laser line width scale linearly with the *rms* power spectral density. These limits are given more explicitly below. Physical intuition into them can be gained with a simple model of the fluctuations as sinusoidal modulations of the laser phase, described below (after Salomon, *et al.* 1988).

Consider the laser frequency fluctuations to be a superposition of sinusoidal phase modulations with different modulation indices β_j and modulation frequencies f_j ($j = 1, 2, \ldots$) spread over some bandwidth B. These modulations cause the instantaneous phase $\phi(t)$ and frequency $\nu(t)$ of the laser output electric field to take on the forms

$$\phi(t) = \sum_j \beta_j \sin 2\pi f_j t \ , \qquad \nu(t) = \nu_0 + \sum_j \beta_j f_j \cos 2\pi f_j t \ . \tag{2.2.5}$$

Specialize now to a single modulation frequency f and associated modulation index β. The instantaneous frequency $\nu(t)$ fluctuates about ν_0 with a root-mean-square (rms) frequency excursion

$$\Delta\nu_{\rm rms} \simeq \frac{1}{\sqrt{2}} \beta f \ . \tag{2.2.6a}$$

The relative power contained in the modulation sidebands is given by the square of the first-order Bessel function ${J_1}^2(\beta)$, which is approximately equal to $\beta^2/4$ for $\beta << 1$. If the excursions $\Delta\nu_{\rm rms}$ are small and occur rapidly (high modulation frequency f), the sidebands will contain little power, because $\beta << 1$. Physically, this means that the high modulation frequency prevents any appreciable phase error from accumulating before the modulation direction reverses and the phase error begins to integrate back to zero. The laser line width $\delta\nu$ can be approximated by one-half the highest modulation frequency f_0 for which the sideband power becomes appreciable, which, for the sake of definition, can be said to occur when the modulation index β becomes as large as 0.5:

$$\delta\nu \simeq f_0/2 \ , \qquad \beta_0 \equiv 0.5 \ . \tag{2.2.6b}$$

If the frequency excursions have a white power spectral density $S_\nu(f) \equiv S_{\nu 0}$ (which is not an inappropriate model for feedback-stabilized lasers, discussed in section 2.2.3), the rms excursion $\Delta\nu_{\rm rms}$ will scale with the square root of an associated bandwidth B:

$$\Delta\nu_{\rm rms} = \sqrt{BS_{\nu 0}} \ . \tag{2.2.6c}$$

For excursions associated with the modulation frequency f_0, the appropriate bandwidth B is centered on f_0 and is of order f_0. Equating expressions (2.2.6a) and (2.2.6c) for modulation index $\beta_0 \equiv 0.5$ implies that the rms frequency excursion at modulation frequency f_0 is related to the power spectral density by

$$\Delta\nu_{\rm rms}(f_0) = 2\sqrt{2}\, S_{\nu 0} \ , \tag{2.2.7a}$$

and the laser line width $\delta\nu$, approximated by $f_0/2$, is

$$\delta\nu \simeq 4[\Delta\nu_{\rm rms}(f_0)]^2 / f_0 \simeq \sqrt{2}\,\Delta\nu_{\rm rms}(f_0) \simeq 4\,S_{\nu 0}\,. \qquad (2.2.7b)$$

This result agrees qualitatively with a rigorous analysis by Elliott, *et al.* (1982) of the effects of frequency modulation on laser line width and line shape. That analysis shows that in the limit of small, high-frequency excursions, the autocorrelation of the excursions is exponential, and the power spectrum of the frequency excursions is Lorentzian. With the assumption of Gaussian noise whose power spectrum is uniform up to a cutoff frequency $B \equiv f_0$, they find the FWHM of the power spectrum of the frequency fluctuations, denoted here by $\delta\nu_{\rm FWHM}$, to be related to the mean-square frequency excursion $\Delta\nu_{\rm rms}$ by

$$\delta\nu_{\rm FWHM} = \pi\,[\Delta\nu_{\rm rms}]^2 / B\,. \qquad (2.2.7c)$$

In the opposite limit of large frequency excursions that occur slowly ($\beta >> 1$), Elliot, *et al.* found that the power spectrum is Gaussian, and that the laser line width (also Gaussian in shape) scales with the *square root* of the power spectral density S_ν and is on the order of 2.35 times the rms frequency excursion.

Improvement of a laser's frequency stability, or reduction of its line width, requires reduction of the noise-power spectral density $S_\nu(f)$. This can be accomplished with feedback from a more stable frequency reference. Such techniques and their performance are considered in section 2.2.3. First, however, section 2.2.2 considers the limits to frequency stability without feedback, *i.e.*, for free-running lasers.

2.2.2 Free-running frequency stability

Figure 4 is an oscilloscope trace showing the free-running line width for a specially designed diode-pumped Nd:YAG laser ring oscillator, obtained by spectral analysis of the heterodyne beat signal between two such free-running oscillators (Kane, *et al.* 1987). The line width of any free-running atomic laser is prevented from being arbitrarily narrow by the process of spontaneous emission. The resulting minimum line width was shown by Schawlow and Townes (1958) to have the following form:

$$\delta\nu_{\text{spon}} = \frac{\pi h\nu}{P_0}(\delta\nu_{lc})^2 = \frac{h\nu}{4\pi P_0 {\tau_{lc}}^2} . \tag{2.2.8a}$$

Here P_0 is the average output power of the laser, and $\delta\nu_{lc} \equiv (2\pi\tau_{lc})^{-1}$ is the laser-cavity line width, inversely proportional to the cavity time constant τ_{lc}. This line width is approximately 1 Hz for a milliwatt of average power at 1-micron wavelength and a time constant of 4 nanoseconds (ns) or 40-MHz cavity line width:

$$\delta\nu_{\text{spon}} \simeq 1\ \text{Hz} \left(\frac{1\ \mu\text{m}}{\lambda}\ \frac{1\ \text{mW}}{P_0}\right)\left(\frac{4\ \text{ns}}{\tau_{lc}}\right)^2 . \tag{2.2.8b}$$

Comparison of expression (2.2.8a) with eq. (2.2.7c) and the surrounding discussion shows that the process of spontaneous emission can be viewed as causing rms frequency excursions that are smaller than the laser-cavity line width $\delta\nu_{lc}$ by the factor $\sqrt{Bh\nu/P_0}$, or $1/\sqrt{2\pi N}$, where $N \equiv P_0\tau/h\nu$ is the number of photons detected in an integration time $\tau = (2\pi B)^{-1}$.

It is common in practice to refer not to the time constant but to the loss coefficient α of the lasing medium, where $\alpha \equiv (c\tau_{lc})^{-1}$ [eqs. (2.1.11)], which is approximately 0.0083 cm^{-1} for a 4-nsec time constant. Semiconductor diode lasers tend to be extremely lossy, with $\alpha = 10\ \text{cm}^{-1}$ being not uncommon. [For a cavity length $L \simeq 0.3$ mm, $\alpha = 10\ \text{cm}^{-1}$ would correspond to end-mirror reflectivities of only 70%, according to eqs. (2.1.11.)] For a 1-mW average output power at 1-μm wavelength, such a diode laser would have a minimum free-running line width of about 1.5 MHz. In

contrast, a solid-state laser made from a Nd:YAG crystal has very low losses; α can be as small as 0.001 cm^{-1} (implying end-mirror reflectivities of order 0.9999 for a 1-mm crystal, or time constant $\tau_{lc} \simeq 33$ ns). The sponaneous-emission-limited line width for such a free-running laser would be as narrow as 15 mHz (millihertz, not megahertz!) for output powers of order 1 mW. In practice, free-running line widths tend to be several orders of magnitude broader than the Schawlow-Townes limit, because of thermal, acoustic, and other additional noise sources. Short-term free-running line widths of 3 to 10 kHz are among the narrowest that have been achieved for free-running diode-pumped Nd:YAG ring-laser oscillators (Bush, *et al.* 1988).

Expression (2.2.8a) can be derived to within a factor of two from purely classical arguments [see, *e.g.*, Jacobs (1979)]. Consider a single-mode laser cavity with time constant τ_{lc} and average number of photons per mode $\overline{N_\nu} >> 1$ [so that stimulated emission dominates spontaneous emission, per eq. (2.1.6)], in which there exists an appreciable population inversion between atomic levels 1 and 2 [$N_1 << N_2$, so that emission dominates absorption, per eq. (2.1.10)]. The average output power P_0 is the difference between the emitted and absorbed power:

$$P_0 = h\nu\, N_2\, a\, [1 + \overline{N_\nu}(1 - N_1/N_2)] \simeq h\nu\, N_2\, a\, \overline{N_\nu} \qquad (2.2.9a)$$

[eqs. (2.1.5) and (2.1.6)]. The average coherent radiation energy in the cavity is

$$U \equiv P_0\, \tau_{lc} \simeq h\nu\, \overline{N_\nu}\,, \qquad (2.2.9b)$$

where the cavity time constant τ_{lc}, or luminescence decay time, is approximately equal to the inverse of the rate $N_2 a$ at which photons are emitted spontaneously. The coherent power dissipated per mode by spontaneous emission is

$$\dot{U} \equiv h\nu\, N_2\, a = P_0\, /\overline{N_\nu}\,. \qquad (2.2.9c)$$

The classical cavity line width is given by

$$(\delta\nu)_{\rm cl} \equiv \frac{1}{2\pi}\frac{|\dot{U}|}{U} \simeq \frac{h\nu}{2\pi P_0\,\tau_c{}^2}\,. \tag{2.2.9d}$$

The reason this expression is larger than the Schawlow-Townes limit by a factor of two is that the argument makes no distinction between amplitude and phase fluctuations of the vacuum field. The vacuum fluctuations have no preferred phase, while the mean amplitude of the laser field does. On average, half of the spontaneously emitted photons are in phase with the laser field amplitude, and half are 90° out of phase. The former produce amplitude fluctuations in the laser output, while only the latter produce phase fluctuations, or spectral-line broadening.

Some recent quantum-mechanical analyses of laser frequency fluctuations suggest different explanations and predictions for ultimate limits on free-running frequency stabilities. (See, *e.g.*, Shapiro, *et al.* 1987; Gea-Banacloche 1987; Caves 1989). One analysis predicts an ultimate limit that is smaller than the Schawlow-Townes limit by a factor of two (Gea-Banacloche 1987). However, achievement of that limit requires illumination of one end of the lasing cavity with highly "squeezed" vacuum, which is difficult in practice. (For a review of nonclassical, squeezed states of light see, *e.g.*, Kimble and Walls 1987 or Schumaker 1986 and references therein.) Unlike ordinary vacuum, whose zero-point fluctuations (referred to simply as vacuum fluctuations elsewhere in this report) are distributed randomly in phase, the zero-point fluctuations in squeezed vacuum have a preferred phase; they are larger than ordinary zero-point fluctuations for some phases, and smaller for others. Hence when these zero-point fluctuations interfere at a photon detector with an electric field that has a specific amplitude and phase, the observed photon-counting statistics may be narrower or broader than the Poisson distribution (shot noise) that arises from ordinary vacuum fluctuations, depending on whether the reduced or enhanced zero-point fluctuations are in phase with the mean electric field (Schumaker 1984). The reader will ask why the use of squeezed vacuum doesn't bring the Schawlow-Townes limit down to zero, instead of one-half its nominal value given

by eq. (2.2.8a). The answer is that the quantum-limited free-running line width of a laser actually arises equally from two sources: measurement noise at a photon detector resulting from interference between the mean laser field and local vacuum fluctuations (which, in principle, can be reduced to zero by using squeezed vacuum), and spontaneous-emission noise resulting from (unsqueezed) vacuum fluctuations within the lasing medium.

2.2.3 Feedback-controlled frequency stability

Better frequency stability than is possible under free-running conditions can be achieved by locking a laser to a more stable reference source. This is accomplished by removing a fraction q of the laser light from the main beam and directing it to a frequency discriminator that is sensitive to small changes in the laser's instantaneous frequency, such as a resonant cavity. The output of the discriminator is used to generate an error signal, which then is used to drive a mechanism for making small corrections to the laser frequency. Even with a perfect correcting mechanism, the frequency fluctuations can be controlled only as well as they can be measured in the first place. The ultimate limit to frequency stability thus comes from measurement uncertainties within the controlling feedback loop. If this measurement noise is characterized by a spectral power density $N(\nu)$ and the servo loop has a spectral gain $G(\nu)$, the noise-power spectral density of the stabilized laser output can be reduced from its free-running value $S_\nu(f)$ to

$$\frac{S_\nu(f)}{G(\nu)} + \frac{(1-q)}{q} N_\nu(f) . \tag{2.2.10}$$

(See, *e.g.*, Barger, *et al.* 1973; Helmcke, *et al.* 1982; Hall 1986; Salomon, *et al.* 1988.)

In principle, all frequency noise associated with operation of the laser can be reduced arbitrarily by a servo loop with appropriate gain, and the ultimate frequency stability is limited by measurement noise within the feedback loop or the stability of the frequency reference. Essentially all techniques for measuring frequency fluctuations convert them to amplitude fluctuations, *e.g.*, by sending the

light through a resonant cavity with very sensitive spectral transmitting and reflecting properties. These amplitude fluctuations then are measured with photodetectors. The minimum measurement noise associated with conventional setups (excluding schemes that incorporate nonclassical light into the feedback loop) is due to shot noise at the photodetectors. Recall that shot noise refers to the Poisson-distributed intensity or photocount statistics that result from interference between the mean amplitude of the field impinging on the detector and the local vacuum fluctuations.

In practice, the requisite high gain and wide bandwidth [at least as wide as that of $S_\nu(f)$] of the feedback loop may be difficult to achieve, and the final laser frequency stability is more likely to be limited, for example, by technical noise in the laboratory environment. This technical noise may be deterministic in nature and arise from the environment (gravity, acoustic, temperature, and pressure effects) or from the apparatus (*e.g.*, electrical noise associated with the pumping source for the laser). For example, the strain induced in a 3-cm-long laser cavity made of Nd:glass by turning it 90° in the Earth's gravity field (equivalent to placing a 1-gm mass on one end of it) will shift the laser frequency by about 1 MHz. Temperature changes affect the frequency of a Nd:YAG laser oscillator by a few gigahertz per degree, and magnetic field effects change it at a rate of about 1.5 MHz per gauss (Byer, 1989). Stochastic noise sources associated with the laser source, such as excess quantum noise associated with the laser resonator design or jet-stream variations, also may be significant.

Figure 5b depicts a "fringe-side" frequency discriminator commonly used for laser frequency stabilization (Hall 1986; Drever, *et al.* 1983b). Part of the laser output is passed through a very narrow frequency filter (here shown as a resonant cavity), and the other part is merely attenuated. For maximum sensitivity, the cavity typically is tuned so that the laser frequency coincides with the half-maximum transmission point; frequency fluctuations in the laser field are reproduced as amplitude fluctuations in the detector output. The attenuated beam serves as a comparison to distinguish laser frequency fluctuations from actual laser amplitude fluctuations. The ratio of the

two detector outputs would reveal the frequency dependence of the filter independently of the input laser power; in practice, however, because of the short time scale of the frequency fluctuations (typically microseconds), it is simpler to subtract than to take the ratio of the two detector outputs. The attenuator is adjusted so that the two photocurrents are equal in magnitude and of opposite sign when the laser is tuned to the half-maximum transmission frequency. Their difference, amplified, becomes the error signal for adjusting the laser frequency. When the servo loop is closed, the differenced photodetector output averages to zero; thus, to first order, small intensity fluctuations in the laser will not affect the frequency stabilization.

The shot-noise-limited line width $\delta\nu_{\text{shot}}$ results from the photocurrent fluctuations in the two photodetectors, which unavoidably are fed into the frequency-feedback servo and treated as frequency fluctuations. It is related to the FWHM line width $\delta\nu_{\text{rc}}$ of the reference cavity in the frequency discriminator and the power P incident on the photodetector by

$$\delta\nu_{\text{shot}} = \frac{\pi h\nu(\delta\nu_{\text{rc}})^2}{\eta P} \simeq 1\ \text{mHz}\left(\frac{1}{\eta}\right)\left(\frac{1\ \text{mW}}{P}\right)\left(\frac{1\ \mu\text{m}}{\lambda}\right)\left(\frac{\delta\nu_{\text{rc}}}{1\ \text{MHz}}\right)^2 . \tag{2.2.11}$$

Here η is a subunity efficiency factor reflecting both photodetector quantum efficiency and efficiency factors associated with signal loss through the measurement process (*e.g.*, modulation and signal recovery). The incident power P is reduced from the laser output power P_0 by such factors as the on-resonance cavity transmission and the fraction f of the laser light that is diverted to the feedback loops; the transmission typically is adjusted to be close to 50% for maximum sensitivity to laser frequency fluctuations, and the fraction diverted for feedback is made as close to unity as requirements on the final frequency-stabilized laser output power will allow (*e.g.*, $f \simeq 0.9$). Comparison with the Schawlow-Townes limit [eq. (2.2.8)] shows that, with feedback, the limiting line width is set by the reference cavity, not the lasing cavity. In recent laboratory demonstrations, feedback has been used to achieve a relative line width of only 3 Hz on time scales of 100 ms between two diode-pumped solid-state ring oscillators ($\lambda = 1.062\ \mu$m) locked to the same high-

finesse interferometer (Day, *et al.* 1990); for the systems used, this line width was approximately 20 times broader than the shot-noise-limited line width.

Recent quantum-mechanical analyses of the limiting line widths possible with feedback stabilization suggest that the shot-noise limit of eq. (2.2.11) might be surpassed by a factor of two if squeezed light is used in the feedback loop, but improvement beyond that is not likely (Caves 1989). In one of these analyses, the expression (2.2.11) is reduced by the factor

$$\frac{1}{2}\left[1-\eta_{\rm d}\,(1-\zeta)\;+\;\epsilon_{\rm loss}\right]\,. \tag{2.2.12}$$

Here ζ represents phase fluctuations in the vacuum incident on the far end of the reference cavity, with $\zeta = 1$ for ordinary (unsqueezed) vacuum; $\eta_{\rm d}$ is the photodetector quantum efficiency; and $\epsilon_{\rm loss}$ is a small but positive loss term designed to be small in fringe-side locking stabilization schemes. If the loss term can be ignored, and maximally squeezed vacuum were available to use in the feedback loop ($\zeta = 0$), the line width given in eq. (2.2.11) could be reduced only by the factor $(1-\eta_{\rm d})/2$. While the use of nonclassical light may become a routine refinement to precision laser measurement techniques in the future, at present more straightforward approaches are being pursued, such as increasing laser output powers and finding more stable reference cavities (or alternatives to cavities, such as atomic or molecular spectral lines).

3. FREQUENCY-STABILIZED LASERS FOR PRECISION MEASUREMENTS

3.1 How does frequency stability limit measurement precision?

The previous chapter distinguished among different characterizations of laser frequency stability. In this chapter and for most of the rest of the report, such care is abandoned in favor of a simpler approach, which will suffice to provide a general understanding of the relationship between frequency stability and measurement precision for most applications. Denote the typical magnitude of a laser's frequency fluctuations over a time scale of interest by $\delta\nu$. Define a corresponding "coherence time" τ_{coh} as that time over which the fluctuations produce a phase uncertainty of 2π, or one cycle:

$$\tau_{\text{coh}} \equiv (\delta\nu)^{-1}. \tag{3.1.1}$$

If such a laser is used as a clock, with time measured by the number of wave periods counted, then, for times longer than τ_{coh}, the measurements will have an uncertainty of at least one wave period. A corresponding coherence length is defined by $c\tau_{\text{coh}}$. Measurements of free-space electromagnetic propagation distance time-tagged to this oscillator would have uncertainties of at least one wavelength for distances longer than $c\tau_{\text{coh}}$.

Fractional frequency stability is defined as $\delta\nu/\nu$, where ν is the nominal center frequency of the laser. Short-term free-running line widths as narrow as 6 to 10 kilohertz (kHz) have been achieved with diode-pumped solid-state (Nd:YAG) ring-laser oscillators operating near $\lambda = 1$ μm with several milliwatts of average output power, corresponding to fractional frequency stabilities of a few parts in 10^{-11} (Bush, *et al.* 1988; Kane, *et al.* 1987). The discussions in section 2.2 [eqs. (2.2.8)] indicate that these stabilities are several orders of magnitude worse than the limits set by losses in the gain medium alone. The poorer performance is caused primarily by thermal and other environment-related fluctuations, as well as fluctuations in the power of the diode laser used for pumping. With

feedback and appropriate control of environmental effects and pump-power fluctuations, millihertz line widths are possible [eq. (2.2.11)]; line widths of a few hertz have been demonstrated recently (Day, *et al.* 1990).

The effect of frequency stability on measurement precision or sensitivity can be seen by considering the simplest "single-arm" interferometer: a beam splitter, a stable reference clock, and a fringe counter (Figure 6). Fluctuations δL in the distance L can be monitored by observing the phase shifts $\delta\phi_L = (2\pi/\lambda)\,\delta L$. However, frequency fluctuations $\delta\nu$ in the laser light also will produce phase shifts $\delta\phi_\nu = 2\pi(L/c)\,\delta\nu$; uncertainty in these must be smaller than the signal phase shifts $\delta\phi_L$ to avoid ambiguity in the measurement. The minimum discernible length change δL_ν due to such ambiguity depends on the certainty with which the laser frequency fluctuations are known, and it scales with the path length L being measured:

$$\delta L_\nu = L\,\frac{\delta\nu}{\nu}\,. \tag{3.1.2}$$

In a dual-arm interferometer (Figure 7), in which only changes in the *difference* between the optical path lengths in the two arms are sought, fluctuations in the measured phase shifts that arise from fluctuations in the laser phase or frequency are correlated in the two arms. Hence they cancel on subtraction to the extent that the two arm lengths are equal or their difference known. If the two arm lengths (L_1, L_2) are held equal to a fraction f (0.1%, say), then the minimum perceptible change in path difference $L_1 - L_2$ can be of order $1/f$ (1,000) times better than would be possible in a single-arm measurement, if laser frequency fluctuations are the dominant error:

$$\delta(L_1 - L_2) \simeq L\,f\,\frac{\delta\nu}{\nu} \tag{3.1.3}$$

[$L \equiv (L_1 + L_2)/2$; see §4.1.1.]

Many applications of interferometry measure an optical path L in order to infer some other

quantity such as the arrival angle of light or the strength of a perturbing force, for which the measure of sensitivity is $\delta L/L$, the smallest measurable *fractional* change in path length. For example, astrometric (dual-arm) interferometers must be sensitive to fractional changes in the relative optical path of at least 5×10^{-12} in order to measure angular separations of stars to microarcsecond (5-picoradian) accuracy. This requires knowledge of the relative optical path lengths to 50 picometers (pm) or better for baselines of 10 meters or shorter (see section 5.1). A measurement sensitivity of 10^{-20} or better is needed to have a reasonable hope of detecting gravitational waves (see section 5.2). Planetary gravity fields can be mapped by forming laser interferometers among two or more orbiting spacecraft and measuring changes in their separations induced by local gravity anomalies (see section 5.3). The required accuracies for range-rate measurements among the spacecraft depend on the desired sensitivity and resolution (minimum measurable strength and spatial extent of anomalies) and on the orbit altitude, since lower orbits result in stronger signals. Accuracies of order 50 nm/s could enable sensitivity to gravity anomalies as weak as one milligal (1 mgal $\simeq 10^{-6}g$) with spatial resolutions of 50 km. These range-rate accuracies would require control or calibration of laser frequency fluctuations to approximately 10^{-12} over time scales of several minutes.

3.2 Why optical-frequency stable oscillators?

Why is it desired to have stable oscillators at optical frequencies? One reason is the potential for improved temporal and spatial resolution due to the shorter wavelengths. Another reason is immunity to externally induced phase fluctuations that fall off inversely with frequency, such as those associated with propagation through regions containing free charged particles (*e.g.*, planetary ionospheres or the solar corona; see section 3.4.4 below). Important practical reasons also include such factors as the need in space-based applications for compact hardware (*e.g.*, to reduce drag and weight) and the need for small antenna beam width (*e.g.*, to avoid multipath reflections from

surrounding objects and to minimize propagation losses). Finally, a wide variety of scientific applications require frequency-stable lasers in order to probe directly short-wavelength phenomena or the short-wavelength properties of media or fields.

3.3 What kinds of measurements are enabled?

The heuristic picture given in section 3.1 of the relationship between frequency stability and measurement sensitivity suggests three different quantities that are measurable to high resolution with stable laser oscillators: time, distance, and frequency. All of these capabilities support important scientific applications in space.

As measures of time, frequency-stable laser oscillators offer superior resolution for producing, controlling, or monitoring sampling-time intervals. The combination of high temporal resolution and long coherence times enables extremely precise measurement of time delays on scales of minutes or more, perhaps even interplanetary light-travel times.

As measures of distance, stable laser oscillators already have proved themselves both practical and, in many cases, essential. They are commonly used for instrument or system metrology. Their uses as ranging devices include such applications as remote sensing, altimetry, navigation and tracking, and maneuvering of vehicles. Finally, they are valuable tools in the measurement of many kinds of forces — gravitational, rotational, frictional, and vibrational.

As measures of frequency, stable laser oscillators are useful in two general ways. The first is in measurements of relative velocity (range rate) *via* Doppler shifts. Velocity measurements can be made relative to "deterministic" targets such as spacecraft or large natural bodies, or stochastic targets such as dust, aerosols, or particulates in planetary atmospheres or the interplanetary medium. For stochastic targets, measurements of Doppler-shifted back- or forward-scattered light can provide information about velocities, sizes, composition, and spatial distribution of the targets. The second

general way in which stable laser oscillators are used as a means of measuring frequency precisely is spectroscopy. In space, a major application of this occurs in atmospheric remote sensing (Hinkley 1976). In the laboratory, important applications exist in atomic, molecular, and quantum physics.

3.4 What other factors limit measurement precision?

Although uncontrolled or uncalibrated frequency fluctuations of an oscillator present a fundamental limit to achievable precision, in many applications they do not constitute the dominant error. Other error sources that can be significant include photon statistics (uncertainty in measured photocount at the output of a photodetector), scattered light, fluctuations in laser power and intensity distribution across the beam, medium-induced phase noise (*e.g.*, refractive-index fluctuations), and mechanical and thermal noise. These are discussed in the following sections, as is the influence of radiation-pressure fluctuations and the "standard quantum limit."

3.4.1 Photon statistics

A potentially important limitation to measurement precision at optical frequencies is photon statistics, a quantum-mechanical contribution to fluctuations in the output of a photon detector. Although quantum theory permits the existence of light for which this error could be arbitrarily small, sophisticated laboratory techniques involving feedback or nonlinear optical processes are required to generate it. (See, *e.g.*, Björk and Yamamoto 1988; Kimble and Walls 1987; Machida and Yamamoto 1988; Wu, *et al.* 1986.) Aside from these "nonclassical" states, all light describable by semiclassical theory produces a standard deviation at least as large as $\sqrt{N}$ in the photocount at the output of a photon detector, if N is the mean photocount over a particular integration period. These minimum fluctuations, characterized by Poisson statistics, are known as shot noise. Recall from the discussion in section 2.1.1 that shot noise actually arises from the superposition of vacuum fluctuations onto the mean amplitude of the electric field impinging on the photodetector.

The uncertainty principles at the heart of quantum theory impose an inverse relation between the precisions with which one can count photons, the photocount standard deviation σ_N, and the precision $\delta\phi_{\rm ph}$ with which one can measure the phase (or phase shift) of a light beam (Carruthers and Nieto 1968):

$$\delta\phi_{\rm ph} \geq \frac{1}{2\sigma_N}, \qquad \sigma_N \equiv \langle N^2 \rangle - \langle N \rangle^2 . \tag{3.4.1a}$$

(Here $\langle N \rangle$ denotes the mean photocount, used in place of N here for clarity.) This implies a minimum uncertainty in a measurement of length or length change

$$\delta L_{\rm ph} = \frac{\lambda}{2\pi}\,\delta\phi_{\rm ph} \geq \frac{\lambda}{4\pi\sigma_N} . \tag{3.4.1b}$$

In a dual-arm interferometric measurement such as that depicted in Figure 7, the input laser light is divided in two at a beam splitter, and half sent down each arm of the interferometer. A beam splitter is a four-port device: two inputs, two outputs. Typically, the second input port of this beam splitter is unused; *i.e.*, only laser light (first input port) and the vacuum field (second input port) enter the interferometer arms. The light exiting the arms is recombined at a second beam splitter, whose outputs are sent to two photodetectors. The fields at the output of a beam splitter are additive superpositions of the complex amplitudes of the input fields, with opposite signs (*i.e.*, superpositions $E_1 \pm E_2$ for input fields E_1, E_2). When these two superposed fields are directed onto a photodetector and the photodetector outputs are differenced, only contributions due to interference between the two input fields remain (*e.g.*, $E_1{}^* E_2$, where $E_1{}^*$ is the complex conjugate of the complex field amplitude E_1). Hence the rms fluctuations σ_{N_d} of the differenced photodetector output $N_d \equiv N_1 - N_2$ typically are dominated by the product of the mean amplitude of the input laser field (approximately equal to $\sqrt{N_0}$, the square root of the mean photocount $N_0 \equiv P_0\tau/h\nu$ that would be obtained if the laser power P_0 were incident directly on the photodetector for the same integration time τ) and the rms amplitude fluctuations σ_0 of the vacuum field. For ordinary vacuum, $\sigma_0 = 1$, and the photocount statistics of the differenced outputs mimic exactly those of the

shot noise ($\sigma_{N_0} = \sqrt{N_0}$) that would be observed with the laser alone. In a "squeezed" vacuum field, σ_0 can be made to be smaller than 1, so that fluctuations in the differenced detector outputs can be made smaller than the laser shot-noise level: $\sigma_{N_d} = \sqrt{N_0}\,\sigma_0 < \sqrt{N_0}$. (For more discussion see, *e.g.*, Caves 1981; Schumaker 1984; Grangier, *et al.* 1987; Kimble and Walls 1987; Schumaker, *et al.* 1987; Wu, *et al.* 1986.)

The minimum discernible change in the difference in arm lengths $\ell \equiv L_1 - L_2$ is found by applying expression (3.4.1b) to the differenced measurement:

$$\delta\ell_{\rm ph} \simeq \frac{\lambda}{4\pi}\,\frac{1}{\sqrt{\eta N_0}}\,\frac{1}{\sigma_0} = \frac{1}{4\pi}\left(\frac{hc\lambda}{\eta P_0 \tau}\right)^{1/2}\frac{1}{\sigma_0}\,, \qquad N_0 \equiv P_0\tau/(h\nu)\,. \tag{3.4.2a}$$

Here P_0 is the input laser power, τ is the measurement integration time, λ is the laser wavelength, and η is a subunity efficiency factor arising from imperfect detector quantum efficiencies, imperfectly reflecting mirrors, and propagation losses. Note that the photon-statistics error is independent of arm length for a given number of detected photons. Of course, the efficiency factor η may depend on arm length because of propagation losses. For example, $\eta \propto L^{-2}$ for one-way propagation losses, or $\eta \propto L^{-2n}$ if the light is reflected back and forth for n round trips before being detected. (In the latter case, mirror losses typically are more severe than propagation losses; multireflection interferometers are discussed in section 4.1.2.) For a 600-mW laser operating at a wavelength of 1 μm, a 1-ms integration time, and an overall receiver and detection efficiency $\eta \simeq 0.02$, the photon-statistics error (without resorting to squeezed vacuum) for measurements of the difference ℓ in arm lengths would be about 10^{-14} m, and would scale with other parameter values as follows:

$$\delta\ell_{\rm ph} \simeq 10^{-14}\ {\rm m}\ \left(\frac{\lambda}{1\ \mu{\rm m}}\,\frac{0.6\ {\rm W}}{P_0}\,\frac{1\ {\rm ms}}{\tau}\,\frac{0.02}{\eta}\right)^{1/2}. \tag{3.4.2b}$$

While 10^{-14} m is an impressively small error in distance, it is unacceptably large for ground-

based laser-interferometer detectors of gravitational waves, which require measurement sensitivities $\delta\ell/L$ of 10^{-20} or smaller (see section 5.2). Practical baselines for ground-based interferometers cannot exceed about 100 km, and at integration times longer than a few milliseconds, measurement error is dominated not by photon statistics but by ground noise. To reach the necessary sensitivities, techniques must be used to increase the effective laser power and the effective arm lengths, in addition to the possibility of using squeezed vacuum. Some of these techniques are discussed in section 4.1 and in the section on gravitational-wave detection (section 5.2).

The photon-statistics limit to measurement precision applies not only to laser light but also to broadband light, such as starlight. It presents a fundamental limit to the measurement sensitivity of space-based astrometric and imaging interferometers in astronomy (see section 5.1). Ten minutes of integration on a star of apparent visual magnitude $m_v = 10$ (see Allen 1973 for a definition of the visual magnitude system) with a 1-meter aperture D_r and 2% overall receiving and detection efficiency η will produce a mean photocount N of approximately 10^7 over the wavelength range 0.5 to 0.6 μm:

$$N \simeq 10^{10} \times 10^{-0.4m_v} \frac{\pi {D_r}^2}{4} \tau\eta \simeq 10^7 \times 10^{-0.4(m_v-10)} \left(\frac{D_r}{1\text{ m}}\right)^2 \left(\frac{\tau}{10\text{ min}}\right)\left(\frac{\eta}{0.02}\right) . \quad (3.4.3)$$

The corresponding photon-statistics error for measurement of the relative optical paths in an interferometer would be $\delta\ell_{\text{ph}} \simeq 10$ pm. Astrometric interferometers are being developed for use in Earth orbit that will perform close to this limit. For an interferometer with arm lengths L of 2 to 20 meters, knowledge or control of the relative optical paths to 10 to 100 pm would enable microarcsecond angular measurement precision. Such interferometers will enable studies of astrophysical phenomena, extend our knowledge of the distance scale of the universe, search for other planetary systems, and enable previously impossible, stringent tests of general relativity that are sensitive to effects of second order in the gravitational potential. (See, *e.g.*, Stachnik 1989; Misner, *et al.* 1973; Reasenberg 1988; Reasenberg, *et al.* 1988; Shao, *et al.* 1988; Vessot 1984; Mozurkewich, *et al.* 1988.)

A final reminder may be in order about why, at optical frequencies, it is reasonable to quote a limiting precision based on photon statistics alone and to neglect additive noise sources that are thermal in nature. Such additive noise dominates in radio interferometry, where equivalent system or antenna "noise temperatures" are several tens of kelvins. A more complete expression than (3.4.1a) for the minimum measurable phase shift $\delta\phi$ due to photon statistics *and* the presence of additive thermal noise sources characterized by a temperature T_s is

$$\delta\phi = \frac{1}{2}\left[\frac{1}{N} + \frac{k_B\hat{T}_s}{N} + \frac{(k_B\hat{T}_s)^2}{N^2} + \frac{(k_B\hat{T}_s)}{N^2}\right]^{1/2}, \tag{3.4.4a}$$

where N is the mean number of detected photons with energy $h\nu$, and $\hat{T}_s \equiv T_s/h\nu$. At X-band wavelengths ($\lambda \simeq 3$ cm), radio photons have energies $h\nu$ equivalent to those of thermal photons associated with $T = 0.15$ K; thus, even at cryogenic system temperatures, $T_s \simeq 60$ K, $k_B\hat{T}_s \simeq 400$ at X-band. At visible wavelengths ($\lambda \simeq 1$ μm), optical photons have energies equivalent to thermal photons with $T \simeq 14,500$ K, so $k_B\hat{T}_s \leq 0.02$ even at room temperature. Since the mean number of detected photons $N >> 1$ in both the optical and radio regimes, this leads to expression (3.4.1a) for the limiting phase error at optical wavelengths and to the following expressions at radio wavelengths for the cases of high signal-to-noise ratio (SNR), ($N >> k_B\hat{T}_s$) and low SNR ($N << k_B\hat{T}_s$), respectively:

$$\delta\phi \simeq \frac{1}{2}\frac{k_B\hat{T}_s}{N} \qquad (N >> k_B\hat{T}_s >> 1), \tag{3.4.4b}$$

$$\delta\phi \simeq \frac{1}{2}\left(\frac{k_B\hat{T}_s}{N}\right)^{1/2} \qquad (N << k_B\hat{T}_s) \tag{3.4.4c}$$

(*cf.* discussion in Crane and Napier 1986).

3.4.2 Scattered light

If an appreciable fraction of the laser light arriving at a detector has not come directly from the source, but instead has been scattered into the detector after deflection by various optical surfaces between the source and the detector, confusion can ensue in measurements of the laser phase. At the detector, the phase of the scattered laser light is delayed by an amount Φ_{sc} relative to the main beam, because the light has traveled an additional distance ΔL_{sc}:

$$\Phi_{sc} = 2\pi\nu\Delta L_{sc}/c \, . \tag{3.4.5}$$

Since the scattered beams are coherent with the main beam, their fields add coherently at the detector. The phase and amplitude of the combined beam can be determined by adding the fields vectorially on a phasor diagram, where the length of each vector denotes the field's amplitude and the direction the field's phase (see Figure 8). If the ratio of the field amplitudes of the scattered and main beams is ϵ_{sc}, implying a relative intensity ${\epsilon_{sc}}^2$ in the scattered beam, the phase of the combined beam will differ from that of the main beam by

$$\phi = \epsilon_{sc} \sin\Phi_{sc} \, . \tag{3.4.6a}$$

A realistic value for ϵ_{sc} might be of order 10^{-4}. The phase shift will fluctuate in response to motions of the scattering elements (changes in ΔL_{sc}) or fluctuations $\delta\nu$ in the laser frequency. The latter produce proportional fluctuations $\delta\phi$ in the phase of the combined beam:

$$\delta\phi = \epsilon_{sc} \cos\Phi_{sc} \, \delta\Phi_{sc} = \epsilon_{sc} \cos\Phi_{sc} \, 2\pi \, \Delta L_{sc} \, \delta\nu/c \, . \tag{3.4.6b}$$

To the extent that the laser frequency fluctuations are uncalibrated, they cause an error in the inferred optical path length

$$\delta L = \epsilon_{sc} \cos\Phi_{sc} \, \Delta L_{sc} \, \delta\nu/\nu \, . \tag{3.4.6c}$$

In practice, the path delay ΔL_{sc} cannot be controlled well enough to maintain the condition $\cos \Phi_{sc} = 0$, and the path delays easily can be as large as L. Hence the error in measured optical path can be of the order

$$\delta L \simeq \epsilon_{sc} L \, \delta\nu/\nu \, . \tag{3.4.6d}$$

If there were no scattered light, fluctuations in the phase of the main beam due to laser frequency fluctuations could be made to cancel in a dual-arm interferometer [eq. (3.1.3)]. This will not work for fluctuations in the combined phase ϕ because the path delays ΔL_{sc} cannot be controlled well enough. However, it is possible to calibrate the effect of the laser frequency fluctuations and thus suppress the error caused by scattered light by phase modulating the laser light at a modulation frequency higher than that associated with fluctuations in the laser frequency and higher than the measurement frequency range (see, *e.g.*, Man, *et al.* 1978; Schilling, *et al.* 1981). For a modulation signal $\psi(t)$ with modulation period T_m, the phase delay between the combined and the main beams at the detector will have the form [eq. (3.4.6a)]

$$\phi(t) = \epsilon_{sc} \sin[\Phi_{sc} + \Delta\psi_{sc}(t)] = \epsilon_{sc} (\sin \Phi_{sc} \cos \Delta\psi_{sc} + \cos \Phi_{sc} \sin \Delta\psi_{sc}) ,$$

$$\Delta\psi_{sc}(t) \equiv \psi(t) - \psi(t - \Delta L_{sc}/c) \, . \tag{3.4.6e}$$

High-frequency fluctuations in the phase delay $\phi(t)$ arise almost entirely from the modulation signal $\psi(t)$. Hence the average over one modulation period T_m of the phase delay between the combined and scattered beams is

$$\bar{\phi} = \frac{\epsilon_{sc}}{T_m} \left[\sin \Phi_{sc} \int_0^{T_m} \cos \Delta\psi_{sc} \, dt + \cos \Phi_{sc} \int_0^{T_m} \sin \Delta\psi_{sc} \, dt \right] . \tag{3.4.6f}$$

The scattered light will add coherently to the main beam (*i.e.*, the average phase delay $\bar{\phi} = 0$) on time scales of order T_m only if a modulation signal $\psi(t)$ is applied for which *both* of the above

integrals vanish. For such phase modulation, the phase delay of the scattered light relative to the main beam will average to zero for integration times equal to integer multiples of the modulation period, and sharp measurements of the phase of the main beam are possible. The challenge of finding and generating an appropriate modulation signal is increased by the fact that typically there are many scattered light components of appreciable strength ϵ_n, which experience different path delays ΔL_n. The sine integral will vanish if the time delays for the strongest scattered components, $\Delta L_n/c$ ($n = 1, 2, \ldots$), are integer multiples of the modulation period T_m. Conditions for the cosine integral to vanish are less straightforward. In general, there is no simple analytical modulation function $\psi(t)$ that will work. Note, for example, that the modulation amplitude (maximum value of $|\Delta\psi_{sc}(t)|$) must be larger than $\pi/2$ in order for the cosine function to take on both positive and negative values in the integration interval. While it is possible to modulate the laser light with broadband white noise, greatest success requires modulation functions designed specifically to suppress the strongest components of the scattered light.

Phase modulation of the laser light has been used successfully to suppress optical feedback effects from reflected light (Man, *et al.* 1978), and it has been applied to ground-based laser-interferometric gravitational-wave detectors (see, *e.g.*, Schilling, *et al.* 1981; Schnupp, *et al.* 1985). For interferometers in which the arms are operated as optical delay lines (as in Figures 7 or 11a), it can be shown that the strongest components of scattered light are those that have traversed the delay lines an integer number of times (see, *e.g.*, Billing, *et al.* 1983). In this case, the appropriate modulation functions are special combinations of orthogonal elementary square-wave functions, or Walsh functions (Harmuth 1977). In particular, to calibrate the scattered-light components that have traversed the delay lines up to n times, an appropriate modulation function would be a repeated series of n elementary square-wave functions. The modulation frequency ν_m might be chosen to be on the order of 10 times higher than the measurement frequency of interest — *e.g.*, 10 kHz for detection of 1-ms gravitational-wave pulses, and the periods of individual square-wave functions

making up the modulation range from $1/n^2$ to $1/n$ times the modulation period ${\nu_m}^{-1}$ (Schnupp, *et al.* 1985). There is a practical limit to this technique, however: as the number n of scattered-light components to be calibrated becomes very large, required switching times may become faster than are realizable.

3.4.3 Other laser instabilities: power, beam geometry

Fluctuations in laser output power or intensity distribution across the beam contribute to phase or path-length measurement error when they interact with asymmetries or imperfections in the measuring apparatus, such as misalignments that cause imperfect interference between wave fronts. Mechanical stabilization typically is used to control path lengths to a fringe or better. For further control, an optical feedback loop is used to generate an error signal that controls some mechanism such as a Pockels cell for adjusting the physical path length. Typically, such a loop works by nulling the average phase difference Φ between wave fronts of two interfering beams at a detector, where Φ is defined as the spatial average over the beam cross section of the local, static phase differences $\phi_0(x,y)$ weighted by the ratio of the local intensities $I_0(x,y)$ to the total average power P_0:

$$\Phi = \frac{1}{P_0} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx\, dy\, I_0(x,y)\, \phi_0(x,y) ,$$

$$P_0 \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx\, dy\, I_0(x,y) . \qquad (3.4.7)$$

Suppose first that there is no significant spatial variation in intensity across the beam cross section [$I_0(x,y) \equiv I_0$]. Power fluctuations in the laser output will cause fluctuations in the error signal Φ used to control the path length and hence fluctuations in the path length around its desired operating length. However, fluctuations in a laser's output power can be reduced to the shot-noise level with active feedback, similar to the feedback-stabilization of laser frequency described in section 2.2.3. With this technique, fluctuations in laser output power need not be a dominant error in measurements of phase or of path length.

Time-varying fluctuations $\delta I(x,y,t)$ in the intensity distribution across a laser beam produce fluctuations $\delta\Phi(t)$ in the error signal (average phase difference) used to control path length:

$$\delta\Phi(t) \;=\; \frac{1}{P_0}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx\; dy\; \delta I(x,y,t)\; \phi_0(x,y)\;. \tag{3.4.8}$$

Local intensity fluctuations can be caused, for example, by lateral jitter of the laser beam or pulsations of the laser beamwidth. A simple lateral jitter, in which the beam preserves its shape but "walks" laterally (in the x-direction, say) by a small, variable distance $\xi(t)$, will produce local intensity fluctuations

$$\delta I(x,y,t) \;=\; \xi(t)\; \partial_x I_0(x,y)\;, \tag{3.4.9a}$$

where $I_0(x,y)$ is the mean local intensity. In the presence of a small tilt α between the interfering wavefronts, which would cause a local phase delay $\phi_0 = 2\pi x\alpha/\lambda$, these intensity fluctuations will produce fluctuations $\delta\Phi(t)$ in the error signal that are of order

$$\delta\Phi(t) \;\simeq\; \frac{2\pi}{\lambda}\,\alpha\,\xi(t)\;. \tag{3.4.9b}$$

These fluctuations enforce a minimum detectable change in path length $\delta L \simeq \alpha\,\xi$, or a few picometers for milliradian ($\simeq 0.05$ degree) tilts. Fortunately, this error can be reduced with straightforward "mode-cleaning" techniques for minimizing nonuniformities in the intensity distribution across the beam. Before entering the interferometer, the laser beam is sent through a resonant (Fabry-Perot) cavity consisting of two spherical mirrors. With proper choices for the mirror curvatures and separation, the transverse modes of the laser (which are responsible for the fluctuations in intensity distribution) can be suppressed and only the lowest longitudinal mode of the laser transmitted. (This rejection of the transverse modes introduces small fluctuations in the output power, which usually are negligible.)

3.4.4 Medium-induced phase noise

Refractive-index fluctuations, which produce fluctuations in the phase and group velocities of a propagating wave, can be a significant source of error in optical measurements of phase shifts or path lengths. Scattering or absorption of laser light by the intervening medium also will randomize the phase, in addition to attenuating the amplitude. Some of these effects are described more quantitatively in chapter 4 (see, for example, section 4.2.2 on coherent pulsed lidar with stochastic targets). Fluctuations δn in the refractive index will produce fluctuations of the phase of a laser beam that scale with the path length L:

$$\delta\phi_n = \frac{2\pi}{\lambda} L\, \delta n \,, \tag{3.4.10a}$$

and cause error in measurement of the path length

$$\delta L_n = \frac{\lambda}{2\pi} \delta\phi_n = L\, \delta n \,. \tag{3.4.10b}$$

Under some conditions, this error can be suppressed by using a dual-arm interferometer.

If the path is very long (perhaps interplanetary distance), the path-length measurement will depend primarily on the path-averaged refractive index and will be insensitive to small-scale random fluctuations. In this case, deleterious effects of the medium can be alleviated by the use of two frequencies, provided the medium is dispersive and the frequency dependence of the refractive index at the two frequencies is known. Measurement resolution still is limited by the noise level of the uncorrelated random phase fluctuations at each frequency, however. Because the use of multiple frequencies can isolate spatial and temporal refractive-index variations, it also can serve to probe properties of the medium. In planetary atmospheres, for example, refractivity profiles at several different frequencies can provide information about pressure, temperature, density, and composition (see section 5.5).

While many sources of medium-induced phase noise are strongly frequency-dependent, on the whole there is no prevailing pattern that makes their effects significantly less harmful at optical frequencies than at radio frequencies. One exception to this statement is the effect of charged particles. A local density of charged particles (plasma) modifies both the phase and group velocities ($v_{\rm ph}$ and $v_{\rm g}$, respectively) of a wave, but the effect decreases approximately with the inverse square of the frequency (see *e.g.*, Thompson, *et al.* 1986):

$$v_{\substack{\rm ph \\ \rm g}} = c\left(1 \pm An_e/f^2\right); \qquad A = \frac{e^2}{8\pi^2 m_e} \simeq 40 \text{ m}^3\text{s}^{-2} \quad . \tag{3.4.11}$$

At optical frequencies, neither the overall delay nor fluctuations in the delay are significant error sources. (The former is on the order of 10^{-16} seconds or less for signals passing the Sun at 20 solar radii, or 5° Sun-Earth-probe angle.) At radio frequencies, multiple frequencies can be used to calibrate the overall delay, but cancellation of the fluctuations can be imperfect because of noncommon paths. This immunity of optical signals to plasma-induced phase fluctuations might be exploited to achieve more accurate single-frequency coherent measurements than possible with radio oscillators of comparable stability. Alternatively, it could be used in conjunction with multiple-frequency radio measurements to improve the calibration of charged-particle effects and to enhance measurements on solar and planetary ionospheres (*e.g.*, measurements of electron densities and magnetic-field intensities and distributions).

3.4.5 Mechanical and thermal noise

Sensitive measurements on Earth are limited at low frequencies by ground noise — seismic and acoustic motions, whose effects on path-length measurements grow roughly as the inverse square of frequency, and gravity gradients arising from naturally-ocurring density variations in the ground and atmosphere, whose effects increase as the inverse fourth power of frequency (Saulson 1984). At

kilohertz frequencies, for example, ground noise can cause rms motions on the order of picometers. At frequencies above about 10 Hz, good isolation from seismic and acoustic noise (to about six orders of magnitude) can be achieved with appropriate materials (Drever, *et al.* 1983a). But at measurement frequencies below about 10 Hz, the contribution of local gravity gradients to ground noise becomes significant enough that sensitive measurements of forces as small as those due to the passage of a gravitational wave (producing a strain of 10^{-20}, say) will never be achievable on Earth.

Although the effects of ground noise can be alleviated at measurement frequencies above about 10 Hz, thermal-induced motions still threaten high-precision distance measurements. In the most sensitive of such measurements, laser-interferometric detection of gravitational waves, test masses are carefully isolated from ground noise and suspended freely in vacuum chambers that approximate an inertial environment. Still, each test mass has internal modes of vibration and exhibits low-frequency "pendulum oscillations" due to the noninertial environment (Drever, *et al.* 1983a; Hough, *et al.* 1983). A mass m at temperature T with a resonant frequency ν_0 and a quality factor Q (damping time $\tau \equiv Q/2\pi\nu_0$), when measured continuously over a bandwidth B, will exhibit an rms motion (position fluctuation) due to internal thermal vibration of

$$\delta L_{\rm vib} \simeq \left(\frac{B k_{\rm B} T}{2\pi^3 Q m {\nu_0}^3}\right)^{1/2} \simeq 2.6\times 10^{-18}\,{\rm m}\left(\frac{1\ {\rm kHz}}{\nu_0}\right)\left(\frac{B}{\nu_0}\right)^{1/2}\Pi\ ,$$
$$\Pi \equiv \left(\frac{10^6}{Q}\cdot\frac{T}{300^\circ{\rm K}}\cdot\frac{10\ {\rm kg}}{m}\right)^{1/2}\ . \qquad (3.4.12a)$$

The same mass, suspended in a near-inertial environment, will exhibit pendulum-type oscillations at frequencies ν_p that typically are smaller than about 1 Hz. [For an equivalent pendulum length L_p, $\nu_p \simeq (2\pi)^{-1}(g/L_p)^{1/2} \simeq 1$ Hz $(25\ {\rm cm}/L_p)^{1/2}$.] The resulting error in length measurements made at frequency ν_m is

$$\delta L_{\rm pend} = \left(\frac{4 B k_{\rm B} T \nu_p}{(2\pi)^3 Q m {\nu_m}^4}\right)^{1/2} \simeq 8\times 10^{-20}\,{\rm m}\left(\frac{1\ {\rm kHz}}{\nu_m}\right)^{3/2}\left(\frac{B}{\nu_m}\right)^{1/2}\left(\frac{\nu_p}{1\ {\rm Hz}}\right)^{1/2}\Pi\ . \qquad (3.4.12b)$$

Like ground noise, these thermal noise sources preclude Earth-based detection of gravitational waves at frequencies much lower than about 10 Hz. But they are a challenge at higher frequencies, too. For example, the high resonant frequencies ($\gtrsim$ 1 kHz) needed to minimize vibrational noise [eq. (3.4.12a)] require rigid construction of test masses and optical elements, which may conflict with practical requirements of adjustability for some optical elements (*e.g.*, beam splitters).

Although it is free from ground noise, a space environment poses its own challenges to high-precision measurements of phase shifts or path lengths. Baselines for interferometric measurements can be quite long — perhaps tens of millions of kilometers — which can offer greater sensitivity in the measurement of forces, as in the detection of gravitational waves or mapping of planetary gravity fields, for example. But a variety of external perturbing noise sources exist to produce measurement errors. Solar-intensity fluctuations produce thermal distortions and fluctuating thermal gradients, which produce anisotropic radiation-pressure fluctuations and resulting motions of test masses. This error source can be controlled with thermal shielding and calibration followed by data correction, but its effect increases roughly as the inverse fifth power of measurement frequency and can become appreciable for measurements made on time scales of a day or longer (measurement frequencies $\nu_m <$ 10^{-5} Hz) (Bender, *et al.* 1988). Random impacts of gas molecules that exist in the imperfect vacuum surrounding test masses also produce random motions. For space-based detection of low-frequency gravitational waves (periods of hours and longer), it is estimated that accelerations produced by these impacts must be controlled or corrected for to about $10^{-18}g$, which would require vacuums of approximately 10^{-11} torr for 10-kg test masses (Stebbins, *et al.* 1989). Cosmic-ray impacts cause test masses to gain net electrical charges, and these must be sensed and cancelled by a driving field. Impacts from massive particles such as protons with energies greater than about 100 MeV would transfer significant momentum to the test masses, although these occurrences would be rare. Other possible effects that could produce significant accelerations and motions to test masses include outgassing from the spacecraft (*e.g.*, from attitude-control thrusters), stochastic buffeting due to the

solar wind, and changing gravitational accelerations (*e.g.*, from fuel motion and depletion).

To compensate for some of these external perturbations in space, sensitive "disturbance compensation systems" (DISCOS), also called "drag-free" systems, are being developed. An early version of such a system, flown on the TRIAD satellite in 1972, was able to reduce the accelerations experienced by a test mass to about $5 \times 10^{-12} g$, down to subhertz frequencies (Staff of the Space Dept., *et al.* 1974). These systems use an error signal generated by the motion of an isolated test mass to control spacecraft thrusters and cancel the perturbing acceleration. A DISCOS system similar to that used on TRIAD would be needed in order to map Earth's gravity field with 50-km resolution and 1-mgal ($\simeq 10^{-6} g$) sensitivity (see section 5.3). Much more sophisticated DISCOS systems, capable of creating an inertial environment down to about $10^{-18} g$, are required for space-based detection of low-frequency gravitational waves (see section 5.2).

3.4.6 Radiation pressure and the "standard quantum limit"

Radiation pressure refers to momentum transfer to a target from incident photons. Each time a photon of energy $h\nu$ bounces off a mirror, it transfers a momentum $p = 2h\nu/c = 2h/\lambda$ to the mirror, which causes the mirror (of mass m) to move a distance $\simeq p\tau/m$ in a time τ. Consider a dual-arm interferometer. If the mean numbers of photons in each arm during an averaging time τ are N_1 and N_2, respectively, and the light is allowed to make n round trips in each arm before being recombined — *i.e.*, the light bounces off the end mirrors n times — the mean difference in momentum imparted to the two ends of the interferometer is

$$p \equiv p_1 - p_2 \simeq N_d \cdot 2nh/\lambda, \qquad N_d \equiv N_1 - N_2 \,. \tag{3.4.13a}$$

If this did not vary, it would produce a constant offset in the phase difference between the light beams exiting the interferometer arms and would not be a source of measurement error. However, fluctuations in the net transferred momentum p (characterized by a standard deviation σ_p) will cause

an error in measurement of the difference in optical paths for the two arms

$$\delta\ell_{\rm rp} \simeq \frac{\tau}{m}\,\sigma_p = \frac{2n\tau h}{m\lambda}\,\sigma_{N_d} . \tag{3.4.13b}$$

Here m is the mass of the (identical) end mirrors or test masses suspended thereon. Recall from section 3.4.1 how the rms fluctuations in the differenced output photocount N_d of a dual-arm interferometer (standard deviation σ_{N_d}) depend on the input laser power P_0 and wavelength λ, the averaging time τ, and the in-phase rms amplitude fluctuations σ_0 of the light entering the interferometer with the laser light (usually vacuum, in which case $\sigma_0 = 1$):

$$\sigma_{N_d} \simeq \left(\frac{\eta P_0 \tau\lambda}{hc}\right)^{1/2} \sigma_0 . \tag{3.4.14}$$

Here η is an overall efficiency factor arising from photodetector inefficiencies, propagation losses, and power losses on reflection. Hence the error in measurements of the difference in arm lengths caused by radiation-pressure fluctuations is

$$\begin{aligned}\delta\ell_{\rm rp} &\simeq \frac{2nh\tau}{m\lambda}\left(\frac{\eta P_o \tau\lambda}{hc}\right)^{1/2}\sigma_0 = \left(\frac{4h\eta P_0\tau^3}{m^2 c\lambda}\right)^{1/2} n\sigma_0 \\ &\simeq 3\times 10^{-20}\ \text{m}\ (n\sigma_0)\left[\frac{P_0}{10\ \text{kW}}\,\frac{\eta}{0.1}\,\frac{1\ \mu\text{m}}{\lambda}\left(\frac{10\ \text{kg}}{m}\right)^2\left(\frac{\tau}{10^5\ \text{s}}\right)^3\right]^{1/2} .\end{aligned} \tag{3.4.15a}$$

Contrast this with the measurement error due to photon-counting statistics [eqs. (3.4.2)]:

$$\begin{aligned}\delta\ell_{\rm ph} &\simeq \frac{1}{4\pi}\left(\frac{hc\lambda}{\eta P_0\tau}\right)^{1/2}\frac{1}{n\sigma_0} \\ &\simeq 4\times 10^{-21}\ \text{m}\left(\frac{1}{n\sigma_0}\right)\left[\frac{10\ \text{kW}}{P_0}\,\frac{0.1}{\eta}\,\frac{\lambda}{1\ \mu\text{m}}\,\frac{10^5\ \text{s}}{\tau}\right]^{1/2} ,\end{aligned} \tag{3.4.15b}$$

where the factor n has been added to account for multiple reflections (see eqs. 4.1.1). These errors become comparable when

$$\eta P_0 \tau^2 = \frac{\lambda m c}{8\pi n^2 {\sigma_0}^2} \simeq 120\ \text{W s}^2\,\frac{1}{(n\sigma_0)^2}\left(\frac{m}{10\ \text{kg}}\,\frac{\lambda}{1\ \mu\text{m}}\right) , \tag{3.4.16a}$$

under which condition their magnitudes can be expressed solely in terms of the mass m of the end mirrors and the averaging time τ:

$$\delta\ell_{\rm ph} \simeq \delta\ell_{\rm rp} \simeq \left(\frac{\hbar\tau}{m}\right)^{1/2} \simeq 10^{-19}\ {\rm m}\left(\frac{\tau}{1\ {\rm ms}}\ \frac{10\ {\rm kg}}{m}\right)^{1/2} \qquad (3.4.16b)$$

($\hbar$ is Planck's constant divided by 2π). For an overall efficiency $\eta \simeq 0.02$, a mass $m = 10$ kg, laser wavelength $\lambda = 1\ \mu$m, and ordinary vacuum injected with the laser light ($\sigma_0 = 1$), radiation-pressure and photon-statistics errors become comparable when $P_0\tau^2 = 6$ kW-s^2 — *e.g.*, for $P_0 = 60$ W with 10-s averaging times, or $P_0 = 1.7$ W with 1-minute averaging times. If the product $\eta P_0\tau^2$ is smaller than the value defined by eq. (3.4.16a), photon-statistics error will dominate radiation-pressure error.

The limiting error described by eq. (3.4.16b) is known as the "standard quantum limit" (SQL) and is viewed by some scientists as a fundamental limit to the precision with which the position of a free mass m (or separation between two identical free masses) can be determined by successive measurements made at time intervals τ.* Its derivation follows directly from a basic tenet of quantum mechanics — that the position and momentum of a free mass m are conjugate observables (*i.e.*, non-commuting Hermitian operators in a Hilbert space) and as such obey an uncertainty principle. The product of their standard deviations for a single (simultaneous) measurement of both must be greater than $\hbar/2$:

$$[\hat{x},\hat{p}] \equiv \hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar \implies \sigma_x\,\sigma_p \geq \hbar/2\ . \qquad (3.4.17a)$$

Here $\hat{x}$ and $\hat{p}$ denote operators in the Heisenberg representation, whose time dependence describes that of the measured values for position and momentum. The Heisenberg position operator for an undisturbed free mass m evolves over a time τ as

$$\hat{x}(\tau) = \hat{x}(0) + \hat{p}\,\tau/m\ . \qquad (3.4.17b)$$

* For discussions of the SQL see, *e.g.*, Braginsky, *et al.* 1980; Caves, *et al.* 1980; Yuen 1983; Bondurant and Shapiro 1984; Caves 1985; Schumaker 1985.

Because $\hat{x}$ and $\hat{p}$ do not commute, $\hat{x}(0)$ and $\hat{x}(\tau)$ also do not commute and, therefore, obey an uncertainty principle:

$$[\hat{x}(0),\ \hat{x}(\tau)] \ = \ i\hbar\tau/m \ \Longrightarrow \ \sigma_x(0)\,\sigma_x(\tau) \ \geq \ \hbar\tau/2m \ . \tag{3.4.17c}$$

Because of this minimum value for their product, the sum of the mean-square position uncertainties at time $t = 0^+$ (just after one measurement, for example) and at time $t = \tau^-$ (just before a second measurement) also has a minimum-allowed value*:

$$\sigma_x{}^2(0^+) \ + \ \sigma_x{}^2(\tau^-) \ \gtrsim \ \hbar\tau/m \ . \tag{3.4.17d}$$

If an instantaneous measurement of position was made at time $t = 0$ and a result obtained with an uncertainty $\sigma_x(0^+)$, and the object is untouched until a time $t = \tau$ later, how accurate can a second position measurement made at time $t = \tau$ be? The minimum uncertainty $\sigma_x(\tau^+)$ that can be associated with the result of the second measurement will be the rss (square root of the sum of the squares) of the uncertainty $\sigma_x(\tau^-)$ in the object's position just prior to the second measurement and the intrinsic resolution σ_M of the measurement apparatus:

$$\sigma_x{}^2(\tau^+) \ = \ \sigma_x{}^2(\tau^-) \ + \ \sigma_M{}^2 \ . \tag{3.4.17e}$$

Clearly it would not be efficient to use an apparatus whose measurement resolution was finer than one's knowledge of the object's position after the first measurement. Hence it can be assumed that $\sigma_M \geq \sigma_x(0^+)$, in which case eq. (3.4.17d) implies that the uncertainty in the second measurement

* The reader may ask why uncertainties are characterized by standard deviations. The implicit assumption of Gaussian wave functions for the object(s) and a Gaussian-preserving measurement process are key to this argument, but whether they are essential to a proof that the SQL cannot be surpassed has not been proven.

must be at least as large as $(\hbar\tau/m)^{1/2}$:

$$\sigma_x(\tau^+) \geq [\sigma_x{}^2(0^+) + \sigma_x{}^2(\tau^-)]^{1/2} \gtrsim (\hbar\tau/m)^{1/2} . \tag{3.4.17f}$$

This is the SQL for discrete measurements spaced by time intervals τ. Applied to ground-based detection of millisecond (kilohertz) gravitational-wave pulses, for example, it would limit the precision of interferometric measurements of the change in relative optical paths to

$$\delta\ell_{\rm sql} \equiv \left(\frac{\hbar\tau}{m}\right)^{1/2} \simeq 10^{-19}\ {\rm m} \left(\frac{\tau}{1\ {\rm ms}}\ \frac{10\ {\rm kg}}{m}\right)^{1/2} . \tag{3.4.18a}$$

An analogous expression is obtained for continuous measurements made over a bandwidth B centered on measurement frequency ν_m:

$$\delta\ell_{\rm sql} \simeq \left(\frac{\hbar B}{\pi m {\nu_m}^2}\right)^{1/2} \simeq 6\times 10^{-16}\ {\rm m} \left(\frac{B}{\nu_m}\ \frac{10^{-5}\ {\rm Hz}}{\nu_m}\ \frac{10\ {\rm kg}}{m}\right) , \tag{3.4.18b}$$

(Bondurant 1986), where the measurement frequency $\nu_m \simeq 10^{-5}$ Hz would be appropriate for space-based measurements of periodic gravitational waves from rotating binary stars (see section 5.2).

In almost all practical applications, other sources of measurement error will be more significant than the SQL or radiation-pressure fluctuations. However, it is of intellectual interest at least to consider the consequences of recent analyses that suggest that the effects of radiation-pressure fluctuations might be cancelled by inserting specially designed nonlinear optical elements such as Kerr cells (Bondurant 1986). Such elements produce a phase shift proportional to the incident intensity, just as radiation pressure causes changes in path length (hence phase shifts) that are proportional to photon intensity. Feedback would be used in conjunction with the nonlinear elements to control mirror motions. The technique would work only over a narrow bandwidth, but it appears feasible in principle. Its existence causes one to question whether the SQL is really a fundamental limit. Any approach to reducing photon-statistics measurement error (increased laser power, longer

integration times, use of squeezed vacuum) necessarily increases radiation-pressure measurement error, and this trade-off was invoked historically to support the existence of the SQL as a limit to measurement precision that could not be surpassed. But if one can reduce photon-statistics error arbitrarily and then deliberately suppress the resulting radiation-pressure fluctuations, perhaps there is no ultimate limit, in principle, to the precision with which one can measure changes in the position of an object.

4. LASER-BASED MEASUREMENT TECHNIQUES FOR SCIENTIFIC APPLICATIONS IN SPACE

This chapter describes two generic measurement techniques involving frequency-stabilized lasers and associated technology that are fundamental to a wide variety of scientific experiments. The first technique is laser interferometry, and the second is coherent measurement of relative velocity (range rate) using Doppler-shifted continuous-wave or pulsed optical signals. General features of each technique are discussed, applicable to short and long baselines (laboratory to planetary distance scales). Emphasis is placed on long-baseline experiments, however, since they are more likely to be performed in space or use space-related technology and because they place greater demands on laser frequency stability and laser-related technology in general. Methods are described for reducing dominant errors and otherwise enhancing overall measurement sensitivity. Two types of coherent velocity measurements are considered: those made on deterministic targets such as spacecraft (individual targets moving with well-defined relative velocities) and those made on stochastic targets such as aerosols or dust particles (multiple targets moving with random relative velocities). Several specific types of science experiments based on these techniques are analyzed in chapter 5; this chapter lays the groundwork for those analyses, and should aid the reader in extending them to other possible applications, such as those described in chapter 6.

4.1 Interferometric measurements

This section extends the discussion in chapter 3 of interferometric measurements and their error sources to examine the following issues: dual-arm *vs.* single-arm configurations, optimum arm lengths and techniques for increasing effective arm lengths, calibration of the effects of laser frequency fluctuations in interferometric measurements (which may alleviate demands on intrinsic laser frequency stability), and techniques to reduce photon-statistics error.

4.1.1 Dual-arm and single-arm configuations

Dual-arm interferometers, which measure the relative change in path length between two arms of approximately equal length, offer several advantages over single-arm interferometers. Errors that are correlated in the two arms, such as laser frequency fluctuations or (in some cases) medium-induced phase fluctuations, will cancel in relative measurements to the extent that the arm lengths L_1 and L_2 can be held equal. In practice, the degree of cancellation, or the fraction $f \equiv (L_1 - L_2)/L$, might be 0.1% or smaller ($L \simeq L_1 \simeq L_2$). For some kinds of measurements, it is possible to orient the two arms so that the effect being measured occurs with opposite sign in each arm, in which case differencing the signals not only cancels some errors but also doubles the signal. This is true, for example, for the gravitational-wave detectors described in section 5.2 and the gravity-field mapping technique described in section 5.3.

4.1.2 Optimizing arm length

Interferometers infer changes in arm length or the difference in two arm lengths by measuring the phase shift of laser light that has traversed the arm(s). In some cases, *e.g.*, where the physical length change is a result of a strain induced by an external force, the length change ΔL will scale with the length L itself; hence longer arms provide stronger signals. The maximum desirable physical path length may be set by a combination of different factors: The signal may cease to increase beyond a certain maximum arm length, *e.g.*, because the force causing the change is of limited duration; losses or other sources of measurement error may become unacceptably large for path lengths longer than a certain maximum; or a maximum length may be dictated by practical constraints such as available real estate, cost, difficulties of maintaining a sufficiently noise-free environment, or horizon-limited line of sight.

When physical lengths of interferometer arms cannot be increased further, sensitivity to length changes or differences ΔL still can be enhanced by effectively amplifying the signal before measuring

it — increasing the phase shift in the laser light that results from a given change in the interferometer arm lengths ΔL. This can be accomplished by making the laser light traverse each arm many times. The following subsections consider three techniques by which the optical paths, and hence the observed phase shifts, can be increased and interferometer sensitivity improved.

4.1.2a Optical delay lines

A common way to increase the optical path lengths in an interferometer is to operate each arm as an optical delay line, reflecting the light back and forth many times before allowing it to exit the arms and recombine with light from the other arm. Such multireflection Michelson-type interferometers have been investigated for use with ground-based gravitational-wave detectors (Weiss 1982; Drever, *et al.* 1983a). They are depicted in Figures 7, 10a, and 11a. If the light makes n round trips before exiting either arm, the measured phase shift $\Delta\phi_n$ will be n times larger than it would be for one round trip, implying the potential for an n-fold improvement in sensitivity for measurement of a given physical length change ΔL. The maximum desirable number of bounces is limited ultimately by losses or, in some cases, by the characteristic time for the effect being measured. If, for example, the length change to be measured is produced by a periodic force, then after one-half period the force will change sign and begin producing a length change in the opposite direction, thus reducing the signal to be measured.

Multireflection techniques suffer eventually from power losses due to imperfectly reflecting mirrors in the delay lines, as well as the effects of scattered light (which are aggravated by laser frequency fluctuations and motions of or asymmetries in the optical elements; see section 3.4.2). Power losses lead to increased measurement error due to photon-counting statistics. Suppose each end mirror in the delay line has a power reflectivity R (nominally very close to unity), so that the fractional power loss upon reflection from either end mirror is $1 - R$. If the light makes n round trips before being recombined at the beam splitter ($2n - 1$ reflections), the power in the recombined beam is smaller

than the original input power by the factor R^{2n-1}. With no losses, the minimum detectable length change would be n times smaller for light that makes n round trips instead of 1. Photon-statistics error, which scales with the inverse square root of power arriving at the detector, will be R^{n-1} times worse because of power losses associated with the additional reflections. Hence the minimum detectable length change (due to photon statistics) with a multireflection system in which the light makes n round trips in each arm is smaller than achievable with a single-reflection system by the factor

$$\frac{\delta L_{\mathrm{ph},n}}{\delta L_{\mathrm{ph},1}} = \frac{R^{1-n}}{n} \simeq \frac{e^{(n-1)(1-R)}}{n} , \tag{4.1.1}$$

since for $R \simeq 1$, $\ln R \simeq R - 1$ ($e \simeq 2.718$ is the base of the natural logarithms). Figure 9 shows the resulting ratio of sensitivities $\delta L/L$ for n round trips *vs.* a single round trip, for various values of mirror reflectivity R. Maximum sensitivity improvement occurs for an optimum number of round trips

$$n_0 = |\ln R|^{-1} \simeq (1-R)^{-1} , \tag{4.1.2a}$$

which produces an improvement, or reduction in minimum detectable length change, of

$$\frac{\delta L_{\mathrm{ph},n_0}}{\delta L_{\mathrm{ph},1}} \simeq (1-R)e \simeq 2.7 \times 10^{-4} \left(\frac{1-R}{10^{-4}}\right) . \tag{4.1.2b}$$

Mirrors are available that exhibit a fractional power loss $1 - R$ smaller than 10^{-4}, implying potential improvement by a factor of several thousand. Depending on a particular apparatus, other error sources such as scattered light or mechanical or thermal stability of the interferometer may defeat full realization of this improvement.

4.1.2b Fabry-Perot cavities

An alternative to a multireflection Michelson-type interferometer, which provides the desired longer light-storage times but avoids problems of scattered light, is to operate each arm as a Fabry-Perot

cavity, set initially to be in precise resonance with the laser frequency (Drever, *et al.* 1983a; Spero 1986). Fabry-Perot-type interferometers are depicted in Figures 10b and 11b. If the optical path lengths in the two arms are equal, the light from each arm will be in phase and give a null when differenced and photodetected. Relative displacements between the two arms thus show up as intensity changes at the photodetector. While this technique is not plagued with deleterious scattering effects or the need for large delay-line mirrors, it does put a more stringent requirement on laser frequency stability in that the laser must be kept in close resonance with the cavities. If absorption losses are ignored, then, for mirror reflectivities R, this system provides a potential improvement in measurement sensitivity that is greater by a factor e than that provided by a multireflection Michelson. In practice, the performance of the Fabry-Perot system is likely to be comparable to a Michelson because of the effects of absorption losses and/or because mirror reflectivities in the Fabry-Perot may have to be limited in order to prevent the light-storage time from *exceeding* the characteristic time of the phenomenon producing the length change being measured (*e.g.*, the period or duration of a gravitational wave). Both multireflection Michelson and Fabry-Perot interferometers are being developed for use in ground-based detection of gravitational waves.

4.1.2c Resonant recycling

Recall that the motive behind making interferometer arms into multipass optical delay lines or optical cavities was to increase the phase shift experienced by the laser light for a given change or difference in arm lengths. These schemes have been carried still further (in theory, not yet in practice) by gravitational-wave experimenters, who plan to turn the entire interferometer into a resonator (Drever 1983; Vinet, *et al.* 1988; Meers 1988). The basic idea is illustrated in Figures 10a and 10b for multireflection Michelson and Fabry-Perot-type interferometers, respectively. The technique is especially appropriate for narrowband measurements of a periodic force that produces a length

change or strain of opposite sign in each interferometer arm, so that the phase shifts experienced by light circulating in the two arms are also of opposite sign.

Consider a multireflection Michelson interferometer for illustration. The delay lines in each arm are adjusted so as to store the light for one-half cycle ($\tau_{GW}/2$) of the expected gravitational wave, or in general the characteristic time $\tau_F/2$ over which a length change of constant sign is produced (*i.e.*, for $n = c\tau_F/4L$ round trips). But instead of being extracted and recombined coherently with light from the other arm, the light exiting each arm is sent into the other arm. This exchange continues for as long as possible before mirror losses cause the photon-statistics error to increase enough to cancel the improvement in strain sensitivity offered by the increased signal. Since the sign of the external force being measured changes with each half-cycle, the light circulating in one direction around the interferometer will experience a monotonically increasing phase shift, and the light circulating in the other direction will experience a decreasing phase shift. The increase in signal is directly proportional to the number of "exchanges" made.

When the interferometer arms are Fabry-Perot cavities instead of optical delay lines, a single resonating mirror suffices to produce the exchanges. The cavity lengths are adjusted to produce cavity resonances near the laser frequency. The path lengths between the corner mirrors and the resonating mirror are adjusted to produce two resonant modes of the coupled cavities displaced symmetrically about the laser frequency at half the measurement frequency — *i.e.*, at frequencies $\nu \pm 1/\tau_F$. This requires that the corner-mirror reflectivity R_c satisfy

$$1 - R_c = \frac{4\pi L}{c\tau_F} \tag{4.1.3a}$$

(Thorne 1987). The laser frequency then is adjusted so that it drives the lower-frequency mode. Force-induced displacements of the end mirrors occur with frequency $1/\tau_F$ and thereby upconvert

photons into the upper mode. The resonating mirror reflectivity R_R also must be made to satisfy

$$1 - R_R \;=\; 2\,\frac{1-R}{1-R_c}\,, \tag{4.1.3b}$$

where R is the reflectivity of the end mirrors. This technique is most useful if mirror reflectivities are high enough and arm lengths long enough to permit storage times on the order of the characteristic times of the signals being measured. In the case of gravitational waves, ground-based detectors will focus on measurement frequencies around a kilohertz (see section 5.2); hence storage times on the order of a millisecond are desirable. In an interferometer with 1-km arms, that would require ~75 round trips to be made in each arm. For space-based interferometers studying lower frequency (millihertz and lower) gravitational waves, storage times of minutes to hours would be required. With 10^7-km arms, a 1-hour storage time would require ~ 30 round trips in each arm.

4.1.3 Frequency stability: calibration *vs.* control

Several aspects of the role of frequency stability in precision measurements have been discussed, such as the use of dual arms or differenced measurements to suppress frequency fluctuations and the increased demands on frequency stability resulting from the presence of scattered light. Here a technique is described for calibrating the effects of laser frequency fluctuations on the interferometric measurements. Subsequent removal of the effects from the measurement data enables detection of more subtle effects, such as those caused by the passage of a gravitational wave. If enough measurements can be made to calibrate the spectrum of laser phase fluctuations, this technique can ease considerably the demands on intrinsic frequency stability of the laser output.

Consider a dual-arm configuration, which itself eases requirements on laser frequency stability by the fraction to which the arm lengths can be held equal (section 4.1.1). Next, suppose the length L_1 of one of the arms is known and maintained well enough that fluctuations ΔL_1 in it (inferred from fluctuations in the measured phase shift of the laser light) arise primarily from fluctuations in

the laser frequency $\delta\nu$ and from a force (such as a passing gravitational wave) whose effect is to be measured:

$$\Delta L_1 = \Delta L_{1F} + \Delta L_{1\nu} \ , \qquad \Delta L_{1\nu} = L_1 \, \frac{\delta\nu}{\nu} \ . \tag{4.1.4a}$$

Observed fluctuations $\Delta\ell$ in the difference of the two arm lengths $\ell \equiv L_1 - L_2$ also arise from the force and from laser frequency fluctuations:

$$\Delta\ell = \Delta\ell_F + \Delta\ell_\nu \ , \tag{4.1.4b}$$

$$\Delta\ell_\nu \ = \ \ell \, \frac{\delta\nu}{\nu} \ = \ \frac{\ell}{L_1} \, \Delta L_1 \ - \ \frac{\ell}{L_1} \, \Delta L_{1F} \ . \tag{4.1.4c}$$

By assumption, changes ΔL_{1F} induced by the force in the measured length L_1 are much smaller than those ($\Delta L_{1\nu}$) due to laser frequency fluctuations, else the latter would not pose an obstacle to measurement sensitivity. This is true also of the changes in the difference of the two arm lengths ℓ, even though $\Delta\ell_F$ might be as large as $2\Delta L_{1F}$ and $\Delta\ell_\nu$ might be smaller than $\Delta L_{1\nu}$ by the factor ℓ/L_1. Hence the laser-induced changes $\Delta\ell_\nu$ in ℓ can be calibrated from the observed changes ΔL_1 in L_1:

$$\Delta\ell_\nu \ \simeq \ \frac{\ell}{L_1} \, \Delta L_1 \ . \tag{4.1.4d}$$

The observed changes ΔL_1 can be measured very precisely, perhaps limited only by photon statistics — *i.e.*, with an error of the order

$$\delta(\Delta L_1) \ \geq \ \frac{\lambda}{4\pi\sqrt{N_1}} \ , \tag{4.1.4e}$$

if N_1 is the mean number of detected photons having exited arm 1 [eqs. (3.4.1)]. Recall from eqs. (3.4.2) that this error is as small as 10^{-14} m for a laser power $P_0 = 600$ mW, an overall efficiency factor $\eta = 2\%$, laser wavelength $\lambda = 1$ μm, and integration time of 1 ms. Hence the contribution of laser frequency fluctuations to the observed fluctuations in the differenced arm lengths can be known to a precision even better than this, by the factor ℓ/L_1, which typically would be on the order of 10^{-3} or smaller. This calibration is more than adequate to be able to distinguish the changes

caused by a force from those caused by laser frequency fluctuations. Even a force as weak as a passing gravitational wave might produce a strain of 10^{-20}, giving a change in measured arm length on the order of $10^{-20}L$, or 10^{-10} m for 10^7-km arms. More rigorous Fourier-domain analyses of this method of calibrating laser phase noise have been performed for studies of space-based detection of low-frequency gravitational waves, with conclusions similar to these (Faller, *et al.* 1984; Bender 1989; Schumaker 1990). Use of this calibration technique could ease the demands on laser frequency stability so much as to bring it within current state-of-the-art technology. Suitable tracking filters must be implemented to permit an adequate spectral sampling of the laser phase fluctuations, of course.

4.1.4 Beating photon statistics

Whether through control at the laser source or calibration during the measurement process, the error induced by laser frequency fluctuations can be reduced dramatically. Another, more fundamental source of error in all measurements involving light is photon statistics. Section 3.4.1 discussed the origin and magnitude of this error. The most straightforward way to reduce this error is to increase the received laser power P_0, since the error scales as $1/\sqrt{P_0}$. Steps could be taken to increase the laser's average output power, such as coherently combining multiple lasers (see, *e.g.*, Byer 1988; Nabors, *et al.* 1990). The use of squeezed light to drive down the photon-statistics error was discussed briefly in section 3.4.1 and is promising in principle, though nontrivial to implement. In this section, two additional techniques are described for increasing the amount of power in the laser signals exiting the arms of an interferometer, given a fixed amount of input laser power. The first technique is to use active laser transponders instead of retroreflectors in two-way measurements. It is appropriate for long-baseline measurements of the sort envisioned in space, where propagation losses cause significant degradation of received laser power. The second technique is to "recycle"

the laser light by enclosing the entire interferometer in a resonant optical cavity and allowing the circulating laser light to build up to an intensity many times that of the input laser light. This technique is being pursued in conjunction with ground-based laser interferometers for detection of gravitational waves.

4.1.4a Transponders *vs.* reflectors

Consider a laser transmitter ("A") and a target ("B") separated by a distance L. Light from the transmitter is sent to the target and then returned, in one case after reflection and in another case after being received, amplified, and retransmitted coherently (*i.e.*, without loss of phase integrity). The photon-statistics limit to the precision with which changes in the distance L can be measured was shown in section 3.4.1 to be

$$\delta L_{\rm ph} = \frac{\lambda}{4\pi\sqrt{N_d}} \ , \tag{4.1.5}$$

where N_d is the number of photons detected back at the transmitter (assuming squeezed light is not used to lower the detected photocount variance below $\sqrt{N_d}$). The detected photocount N_d depends on whether the laser light was reflected or transponded from the target. Suppose the transmitter antenna (used both for transmitting and receiving, say) has diameter D_A and the target antenna diameter D_B. The beam divergences associated with them are $\theta_A \equiv s_A\ \lambda/D_A$ and $\theta_B \equiv s_B\ \lambda/D_B$, respectively. It may be necessary or desirable to make one or both of the antennas nondiffraction-limited, *e.g.*, to ease pointing requirements on the target antenna; thus, practical values for s_A and s_B might be 5 and 15, say, corresponding to 50- and 150-microradian beamwidths, respectively, for 10-cm apertures and 1-μm wavelength laser light.

The ratios of power received at B to power transmitted from A and power received at A to

power transmitted from B are given by the respective"Fresnel factors" F_{AB} and F_{BA}:

$$F_{AB} \equiv \frac{P_r(B)}{P_A} = \left(\frac{D_B}{L\theta_A}\right)^2 = \left(\frac{D_A D_B}{s_A \lambda L}\right)^2 ;$$
$$F_{BA} \equiv \frac{P_r(A)}{P_B} = \left(\frac{D_A}{L\theta_B}\right)^2 = \left(\frac{D_A D_B}{s_B \lambda L}\right)^2 . \tag{4.1.6}$$

If the target B is a retroreflector, and A transmits a laser power P_0 to B which B simply reflects back to A, the detected photocount back at A over an integration time τ is

$$N_d{}^{\rm rr} = N_0 \, \eta^{\rm rr} \, F_{AB} \, F_{BA} \qquad \text{(retroreflector)} ,$$
$$N_0 \equiv \frac{P_0}{h\nu} \tau , \tag{4.1.7a}$$

where $\eta^{\rm rr}$ is an overall efficiency factor for the case of a retroreflector. If, instead, the target B has a transponder that retransmits a power P_t back to A, the detected photocount at A will be

$$N_d{}^{\rm tr} = N_t \, \eta^{\rm tr} \, F_{BA} \qquad \text{(transponder)} ,$$
$$N_t \equiv \frac{P_t}{h\nu} \tau , \tag{4.1.7b}$$

where $\eta^{\rm tr}$ is an overall efficiency factor for the case of a transponder. The photon-statistics measurement errors for these cases are

$$\delta L_{\rm ph}{}^{\rm rr} = \frac{s_A s_B \lambda^3}{4\pi\sqrt{N_0 \eta^{\rm rr}}} \left(\frac{L}{D_A D_B}\right)^2 = \frac{s_A s_B}{4\pi} \left(\frac{hc\lambda^5}{P_0 \eta^{\rm rr} \tau}\right)^{1/2} \left(\frac{L}{D_A D_B}\right)^2 , \tag{4.1.8a}$$

$$\delta L_{\rm ph}{}^{\rm tr} = \frac{s_B \lambda^2}{4\pi\sqrt{N_t \eta^{\rm tr}}} \left(\frac{L}{D_A D_B}\right) = \frac{s_B}{4\pi} \left(\frac{hc\lambda^3}{P_t \eta^{\rm tr} \tau}\right)^{1/2} \left(\frac{L}{D_A D_B}\right) . \tag{4.1.8b}$$

For either retroreflectors or transponders, the photon-statistics error $\delta L_{\rm ph}$ can be reduced with larger antennas, shorter wavelengths, higher power lasers, longer integration times, and improved overall efficiencies. With transponders, the photon-statistics error scales with interferometer arm

length L, hence the minimum detectable strain $\delta L_{\rm ph}/L$ is independent of arm length. With retroreflectors, propagation losses cause the photon-statistics error to scale with L^2, hence the minimum detectable strain increases for longer arm length L.

An example of an application for which the use of transponders instead of retroreflectors would enhance science return appreciably is the mapping of planetary gravity fields with laser interferometers formed from two or more coorbiting spacecraft separated by distances L (see section 5.3). The photon-statistics error [eq. (4.1.8a)] in measurement of the spacecraft separations is

$$\delta L_{\rm ph}{}^{\rm rr} \simeq 2\ {\rm pm} \left(\frac{\lambda}{1\ \mu{\rm m}}\right)^{5/2} \left(\frac{2\ {\rm s}}{\tau}\ \frac{10\ {\rm mW}}{P_0}\ \frac{0.05}{\eta^{\rm rr}}\right)^{1/2} \left(\frac{s_A s_B}{75}\right) \left(\frac{L}{50\ {\rm km}}\ \frac{10\ {\rm cm}}{D_A}\ \frac{10\ {\rm cm}}{D_B}\right)^2 , \quad (4.1.9a)$$

and scales as indicated for different parameter values. If transponders are used instead of retroreflectors, and the transponders retransmit a power P_t equal to the power P_0 transmitted originally, the photon-statistics measurement error is about 1.5 orders of magnitude smaller for the same parameter values:

$$\delta L_{\rm ph}{}^{\rm tr} \simeq 0.08\ {\rm pm} \left(\frac{\lambda}{1\ \mu{\rm m}}\right)^{3/2} \left(\frac{2\ {\rm s}}{\tau}\ \frac{10\ {\rm mW}}{P_0}\ \frac{0.05}{\eta^{\rm tr}}\right)^{1/2} \left(\frac{s_B}{15}\right) \left(\frac{L}{50\ {\rm km}}\right) \left(\frac{10\ {\rm cm}}{D_A}\right) \left(\frac{10\ {\rm cm}}{D_B}\right) . \quad (4.1.9b)$$

4.1.4b Power recycling

When one of the output ports of a laser interferometer (a Michelson-type, for illustration) is tuned to a dark fringe, most of the laser power exits the other output port. This power can be recycled, or fed coherently back into the interferometer, leading eventually to a circulating light flux within the interferometer that is much higher than the input laser power P_0. The technique is depicted in Figures 11a and 11b for multireflection Michelson and Fabry-Perot types of interferometers, respectively. It has been analyzed for application to ground-based laser-interferometer detectors of

gravitational waves (Drever, *et al.* 1983a; Vinet, *et al.* 1988; Meers 1988). Ideally, the light-storage time in the arms should be comparable to the characteristic time of the effect being measured (*e.g.*, the force, such as a passing gravitational wave, producing the change in arm lengths). Losses occur at the beam splitters and at imperfectly reflecting mirrors (the end mirrors in the interferometer arms as well as the recycling mirror), and also as a result of imperfect interference between wavefronts at the beam splitter due to misalignments or asymmetry between the interferometer arms. In practice, the light-storage time τ_s may have to be shorter than the characteristic time τ_F of the force being measured, because losses associated with the longer storage times become great enough to degrade measurement sensitivity (through photon-statistics error) despite the increased signal resulting from the longer optical path. If the dominant losses come from delay-line mirrors, for example, and the recycling mirror is adjusted to provide maximum light buildup, overall measurement sensitivity can be improved with recycling by a factor $[2\tau_s/\tau_F(1-R_e)]^{1/2}$. Here $R_e \simeq 1 - R^{2n-1}$ is the effective overall power reflectance for a delay line involving $2n-1$ reflections (n round trips) from mirrors with power reflectivities R. Under optimum conditions, τ_s would be of order $\tau_F/2$, and the improvement factor would be approximately $R^{1/2-n}$. This factor is approximately $eR^{1/2}$ when $n = n_0 \equiv -(\ln R)^{-1}$, the optimum number of bounces derived in eqs. (4.1.1) and (4.1.2). Without power recycling, the use of multiple reflections gave a maximum improvement factor of $(1-R)e$ [eq. (4.1.2b)]. Thus, power recycling can compensate effectively for the increase in photon-statistics error that arises from power losses associated with multiple reflections in a delay line.

4.2 Coherent measurements of relative velocity

The range rate (line-of-sight component of relative velocity) between two oscillators in relative motion can be determined by measuring the Doppler shift in the frequency of one relative to the other. Measurement precision in many circumstances is limited by phase or frequency fluctuations

in the oscillator(s) used for reference. Two-way measurements offer the advantage that the effects of low-frequency fluctuations in the oscillator's phase (fluctuations that occur over periods longer than the round-trip light-travel time) will tend to cancel. Range-rate measurement techniques for discrete objects such as spacecraft moving with well-defined velocities (deterministic targets) differ from those for multiple targets exhibiting a spread of velocities, such as aerosol or dust particles entrained in wind (stochastic targets). For deterministic targets, two-way measurements generally are made by transmitting signals to a target on spectrally pure, continuous-wave (CW) carriers, which are either reflected back or transponded coherently (received, amplified, and retransmitted without loss of frequency or phase integrity). For stochastic targets, measurements typically are made by transmitting short pulses and measuring the time delay of the returned (scattered) signals, by comparison with a stable reference oscillator. The measurements are described as coherent because the phase of the return signal is compared with the phase of a reference oscillator, whose frequency stability therefore must be very good. This section describes these two types of measurements, and discusses briefly some scientific applications and the motivations for performing them with lasers — *i.e.*, at optical instead of longer wavelengths.

4.2.1 Deterministic targets

The ability to make accurate measurements of range rate over long distances — tens of kilometers to interplanetary distances — is crucial to several kinds of science experiments, some of which are described briefly here and in more detail in chapter 5. It also will enhance planetary missions by enabling accurate tracking of laser-carrying spacecraft under conditions of high background light, *i.e.*, with the sun or bright planets in the field of view, since the heterodyne detection process associated with coherent laser links and measurements of Doppler shifts provides superb spectral filtering (see section 6.1). Under low-background conditions, coherent optical Doppler tracking provides a data

type complementary to those types obtained with incoherent optical-tracking techniques such as ranging (by direct detection of pulsed laser light) or angular tracking (with filled-aperture astrometric techniques or with interferometers).

The use of optical frequencies (lasers) instead of microwave frequencies for range rate measurements on deterministic targets offers several potential benefits, especially for Earth-orbiting or planetary missions. The small size and mass of optical components may ease launch costs or logistics and can be important for minimizing range-perturbing forces among spacecraft. The narrow antenna beamwidths offer greater power efficiency (at the expense of possible pointing challenges) and decreased multipath reflections from the spacecraft body and appendages. Measurement precision and sensitivity also are potentially better at optical frequencies because of the finer resolution offered by shorter wavelengths and because of decreased susceptibility to some noise sources, such as plasma-induced phase scintillation. At optical frequencies, the primary limitations to the precision of velocity measurements on deterministic targets are photon statistics and laser frequency fluctuations. Coherent laser links, in which both intensity and phase of the laser signal are monitored, offer potentially better measurement sensitivity and accuracy than do incoherent links, which monitor only laser intensity. To drive down errors arising from photon statistics, which may be especially significant over long propagation distances, coherent laser transponders should be used instead of retroreflectors for two-way measurements.

One example of an experiment requiring highly accurate range rate measurements is the detection of low-frequency gravitational waves (see section 5.2). Recall from discussions in previous chapters that detection of low-frequency gravitational waves can be accomplished only in space since the required isolation from seismic and other vibrational noise cannot be achieved on Earth and because maximum sensitivity to waves whose characteristic periods range from minutes to days (10^2 to 10^6 seconds) requires a separation between inertial test masses of at least 10^7 km. Current estimates indicate a need for lasers with average output powers of a few watts and fractional fre-

quency stabilities of about 10^{-15} over short (subsecond) time scales. Frequency fluctuations which occur more slowly can be calibrated during the measurement process, as described in section 4.1.3.

High-resolution mapping of planetary gravity fields is another important science application involving coherent laser links between spacecraft (see section 5.3). This could be accomplished by forming a laser interferometer between two or more orbiting spacecraft (one behind the other, separated by a distance L) and measuring the changes induced in their range and range rate as they pass over an anomaly in the gravity field. The sensitivity of the measurement (strength of the minimum discernible gravity anomaly) is proportional to the precision with which the range rate can be measured. The magnitude of the change induced in the range rate depends on the strength and spatial extent of the gravity anomaly, as well as the orbit altitude of the spacecraft. For an orbit altitude of 160 km, range-rate accuracies on the order of 20 μm/s may be just adequate to sense 1-mgal variations in the geopotential with 100-km resolution (1 mgal $\equiv 10^{-3}$ cm/s^2 $\simeq 10^{-6}g$), or to sense 5-cm rms undulations of the geoid. Steep improvements in range rate measurement accuracy are required to gain moderate improvements in spatial resolution with comparable sensitivity. Sensitivity to 1-mgal anomalies with 50-km spatial resolution (at 160-km orbit altitude), for example, requires range rate accuracies on the order of 50 nm/s; the two-fold improvement in spatial resolution requires more than a 400-fold improvement in range rate accuracy (see Table 5.3.2). This level of range rate accuracy could be accomplished with coherent laser links, but it would require laser fractional frequency stabilities of 10^{-12} or better over time scales of a few seconds. The decreased drag associated with the smaller "antennas" used for laser links in such experiments could be an important additional advantage over radio links, in that it could permit lower orbit altitudes. A change from 160 km to 140 km in altitude would reduce the required range rate accuracies by a factor of ~ 3.5 for 100-km spatial resolution and a factor of ~ 12 for 50-km spatial resolution [eqs. (5.3.7)and (5.3.8)].

Figure 12 depicts two different schemes for making two-way measurements of Doppler shifts on

optical carriers. In the first (Figure 12a), the target spacecraft transponds the received laser signal coherently, *i.e.*, it phase-locks an onboard laser to the received laser signal and uses this laser to retransmit a signal. In practice, it may be adequate for some experiments to lock to the (short-term average) frequency, rather than the phase of the incoming signal. In the second scheme (Figure 12b), the target spacecraft has a free-running reference laser that has good short-term frequency stability, but need not be as stable over long times as the reference laser on the other spacecraft. The beat frequency between the laser on the target spacecraft and the incoming signal is measured and encoded as data and transmitted back on the free-running laser carrier, where it is compared with a second beat-frequency measurement made between the original reference oscillator and the received signal from the target spacecraft. The second scheme is easier to implement and for many applications may provide adequate measurement precision. Its performance is explored in more detail below. If the frequency-stabilized reference laser at the originating station has a coherence time that is considerably longer than the round-trip light-travel time, the slow fluctuations in its frequency will cancel when the two beat frequencies are differenced, and performance of this second scheme can approach the performance achievable with a fully phase-coherent technique. Even in the presence of higher-frequency laser phase fluctuations, good performance might be maintained if the laser phase noise can be calibrated by the method described in section 4.1.3.

Suppose two spacecraft (SC) labelled "1" and "2" (or one spacecraft and one base station, for example) are separated by a nominal distance L, and SC 2 is moving away from SC 1 with a line-of-sight velocity (range rate) $v > 0$. At time $t = 0$ SC 1 transmits a signal on a spectrally pure carrier of frequency $\nu_1(0) \equiv \nu_1$. When SC 2 receives the signal a time $\tau = L/c$ later, it measures the carrier frequency to be $\gamma\nu_1$, where γ is the nonrelativistic Doppler factor

$$\gamma \equiv \left(\frac{1-\beta}{1+\beta}\right)^{1/2} \simeq 1-\beta, \qquad 0 < \beta \equiv v/c << 1 \; . \tag{4.2.1}$$

(The symbol β will be used for v/c to make the equations simpler.) If SC 2 has a free-running laser

whose average operating frequency is ν_2, it will measure a beat frequency between this laser and the laser light arriving from SC 1 equal to

$$\epsilon_2(\tau) = |\nu_2(\tau) - \gamma\,\nu_1(0)| \,. \tag{4.2.2}$$

In practice, the range rate v will have been estimated in advance and the average frequencies of the lasers on SC 1 and 2 will be separated so as to avoid the need for broad detection bandwidths (at least at SC 2). If the frequency difference $\nu_1(0) - \nu_2(\tau) \equiv \nu_1 - \nu_2$ is chosen such that

$$\nu_1 - \nu_2 \;\approx\; \frac{\beta}{(2-\beta)}\,(\nu_1 + \nu_2)\,, \tag{4.2.3}$$

the beat frequency $\epsilon_2(\tau)$ measured at SC 2 will be very small. SC 2 then retransmits a signal on its carrier of ν_2, which has encoded in it the value of the beat frequency $\epsilon_2(\tau)$ that it measured. When SC 1 receives this signal, it measures the beat frequency between the incoming carrier and its own reference laser, whose frequency at this time is $\nu_1(2\tau)$:

$$\epsilon_1(2\tau) = \nu_1(2\tau) - \gamma\,\nu_2(\tau) \,. \tag{4.2.4}$$

If condition (4.2.3) is satisfied and $\nu_1(2\tau) \approx \nu_1(0)$, this beat frequency will be on the order of $\beta\,(\nu_1 + \nu_2)$. To make this smaller, it might be desirable to use a down-shifted laser at SC 1 for this beat-frequency measurement, *i.e.*, to make $\nu_1(2\tau)$ smaller than $\nu_1(0)$ by an amount of order $\beta\,[\nu_1(0) + \nu_2(\tau)]$. A comparison is made between the beat frequency $\epsilon_1(2\tau)$ measured at SC 1 and the value of the beat frequency $\epsilon_2(\tau)$ measured at SC 2, and the difference used to infer the range rate v:

$$\begin{aligned} \epsilon \equiv \epsilon_1(2\tau) - \epsilon_2(\tau) &= \nu_1(2\tau) + \gamma\nu_1(0) - (1+\gamma)\nu_2(\tau) \\ &\simeq (2-\beta)\,[\nu_1(0) - \nu_2(\tau)] + \nu_1(2\tau) - \nu_1(0)\,. \end{aligned} \tag{4.2.5}$$

The variance of the measured beat-frequency difference ϵ is defined by

$$\langle[\Delta\epsilon(2\tau)]^2\rangle \;\equiv\; \langle[\epsilon(2\tau)]^2\rangle \;-\; \langle\epsilon(2\tau)\rangle^2 \,, \tag{4.2.6a}$$

where angle brackets denote an appropriate ensemble average. It depends on the variances of the measured laser frequency ν_1 at times $t = 0$ and $t = 2\tau$, the range rate $v = \beta c$, and the degree of correlation between the fluctuations of SC 1's laser frequency at times $t = 0$ and $t = 2\tau$:

$$\langle [\Delta\epsilon(2\tau)]^2 \rangle \simeq \langle [\Delta\nu_1(2\tau)]^2 \rangle + (1 - 2\beta) \langle [\Delta\nu_1(0)]^2 \rangle - 2(1 - \beta) \langle \Delta\nu_1(0)\Delta\nu_1(2\tau) \rangle . \quad (4.2.6b)$$

There is no dependence on frequency fluctuations of SC 2's free-running laser because factors of β^2 or smaller have been neglected relative to those of order unity. If the coherence time of the laser at SC 1 is much longer than the round-trip light-travel time 2τ — *i.e.*, appreciable fluctuations in laser 1's frequency (or phase) occur only over time scales longer than 2τ — the frequency fluctuations at times $t = 0$ and $t = 2\tau$ will be highly correlated. Hence the last term can be large and negative, *i.e.*, these low-frequency fluctuations will tend to cancel and not contribute significant error to the measurement of the difference in beat frequencies ϵ. If laser frequency fluctuations on shorter time scales are appreciable, this cancellation is imperfect or absent, the latter occurring if the coherence time of the laser at SC 1 is shorter than 2τ, causing the correlated noise term to be zero.

The time averages of these variances are autocorrelation functions [eq. (2.2.4)], which are related to the (one-sided) noise-power spectral density $S_{\nu_1}(f)$ of laser 1 by

$$\lim_{t_m \to \infty} \frac{1}{t_m} \int_{-t_m/2}^{t_m/2} dt \, \langle \Delta\nu_1(t)\Delta\nu_1(t + 2\tau) \rangle \equiv C_{\nu_1}(2\tau) = \int_0^\infty df S_{\nu_1}(f) \cos(4\pi f\tau) . \quad (4.2.7)$$

Estimates of range rate $v = c\beta$ are made by averaging the differenced beat-frequency measurements $\epsilon(t)$ over times t_m with some weight function $h(t)$ [eq. (4.2.5)]. The mean-square error ${\sigma_v}^2$ in such estimates is

$$\begin{aligned} {\sigma_v}^2 &= \left(\frac{c}{2\nu_0}\right)^2 \lim_{t_m \to \infty} \frac{1}{t_m} \int_{-t_m/2}^{t_m/2} dt \, |h(t)|^2 \langle [\Delta\epsilon(t)]^2 \rangle \\ &\simeq \left(\frac{c}{\nu_0}\right)^2 \int_0^{f_c} df \, |H(f)|^2 \, S_{\nu_1}(f) \, \sin^2(2\pi f\tau) , \end{aligned} \quad (4.2.8a)$$

$$2\nu_0 \equiv \nu_1(0) + \nu_2(\tau) . \qquad (4.2.8b)$$

Here $H(f)$, the filter response function, is the Fourier transform of the weight function $h(t)$, f_c is the receiver cut-off frequency, and eq. (4.2.7) has been used.

The noise-power spectral densities of most oscillators typically exhibit "flicker" frequency noise at lower frequencies [$S(f) \propto 1/f$], white frequency noise at higher frequencies, and white phase noise [$S(f) \propto f^2$] at still higher frequencies. Diode-pumped solid-state lasers, among the most promising candidate frequency-stabilized lasers for space applications, typically exhibit white frequency noise above about 100 kHz and flicker frequency noise, which is thought to be caused primarily by power fluctuations in the pump laser, down to about 100 Hz. Below about 100 Hz, the noise curves for some free-running lasers have been observed to roll over to a $1/\sqrt{f}$ dependence, rather than changing to the expected $1/f^2$ random-walk spectrum (Day, *et al.* 1990).

The filter response function $H(f)$ should be chosen so as to attenuate the noise spectrum at least as fast as the laser noise-power spectral density makes it rise. If, for example, the laser exhibits flicker or white frequency noise in the detection bandwidth, a uniform weighting scheme, equivalent to cycle-counting or "start-stop" phase measurements, might be adequate. Its frequency response function is a sinc function,

$$H(f) = \frac{\sin \pi f t_m}{\pi f t_m} \qquad \text{(uniform weighting)} . \qquad (4.2.9a)$$

If the laser exhibits flicker or white *phase* noise [$S_\nu(f) \propto f$ or f^2, respectively] in the detection bandwidth, a triangular weighting scheme might be more beneficial, equivalent to differencing contiguous phase measurements averaged over periods t_m. Its filter response function is the square of that for uniforming weighting (MacArthur and Posner 1985).

Diode-pumped Nd:YAG lasers have been operated with free-running line widths as narrow

as 100 Hz on time scales of a few tens of microseconds, and 0.5 to 1 kHz on time scales of 1 to 10 ms, corresponding to short-term fractional frequency stabilities better than a few parts in 10^{12} (Kane, *et al.* 1987; Bush, *et al.* 1988). With feedback stabilization, similar lasers have been operated with line widths of only a few hertz on time scales of 100 ms (Day, *et al.* 1990), and in principle this can be improved by orders of magnitude (see discussion in section 2.2.3). Suppose then, for illustration, that at frequencies above a few hundred hertz (on time scales shorter than a few milliseconds), fluctuations in the laser frequency are small enough that they contribute negligibly to the error in a measurement. Then, for two-way Doppler measurements over distances smaller than about 100 km, the sine function in expression (4.2.8a) can be replaced by its argument. If the measurements are weighted uniformly [filter response function given by eq. (4.2.9a)], the resulting range rate measurement error is

$$\sigma_v{}^2 \simeq 2\left(\frac{c}{\nu_0}\right)^2\left(\frac{\tau}{t_m}\right)^2 \int_0^{f_c} df\; S_{\nu_1}(f)\,[1-\cos 2\pi f t_m]\;. \tag{4.2.9b}$$

The maximum error can be obtained by ignoring the part of the integrand proportional to $\cos 2\pi f t_m$. Using the language of chapter 3 with $(\delta\nu_1)^2 \equiv \int d\nu S_{\nu_1}(f)$, the range rate measurement error therefore is related to the fractional frequency stability of laser 1, the separation between SC ($L \equiv c\tau$), and the measurement averaging time t_m by

$$\sigma_v \;\simeq\; \sqrt{2}\left(\frac{c\tau}{t_m}\right)\left(\frac{\delta\nu_1}{\nu_1}\right) \;\simeq\; 1.4\ \text{nm/s}\left(\frac{1\ \text{s}}{t_m}\,\frac{c\tau}{100\ \text{km}}\right)\left(10^{14}\,\frac{\delta\nu_1}{\nu_1}\right)\;. \tag{4.2.9c}$$

Here ν_0 in the denominator has been replaced with ν_1, a reasonable approximation since $v << c$. Note that if the reference laser were not stable in frequency over the round-trip light-travel time, the factor (τ/t_m) would be absent, and measurement precision could be considerably worse.

4.2.2 Stochastic targets

Techniques for measuring the velocities of stochastic targets, such as dust particles (aerosols) entrained in wind, differ from those used on deterministic targets. In the most common technique, short pulses are transmitted and ranges and range rates are inferred from the time delays of the received (one-way or forward-scattered) or returned (two-way or back-scattered) signals. This technique is referred to generally as Doppler radar, or Doppler lidar when performed at optical frequencies. Figure 13 shows a schematic diagram of a coherent Doppler lidar system.

Doppler lidar measurements can be accomplished incoherently or coherently, *i.e.*, with direct-detection receivers or heterodyne receivers. Incoherent systems typically use high-resolution Fabry-Perot etalons to convert intensity measurements into accurate Doppler-frequency measurements. Coherent schemes compare the phase of the scattered signal with the phase of a reference oscillator in a heterodyne measurement. Thus, in coherent schemes, measurement precision depends crucially on the frequency stability of the oscillator. In principle, coherent schemes can measure Doppler shifts more accurately, and under high-background conditions (*e.g.*, scattered sunlight) they offer superior spectral filtering. Some incoherent receivers that use combinations of interference filters and scanning etalons to block background light are currently competitive with coherent systems. Comparisons among the different systems abound in the literature (see, *e.g.*, Abreu 1979; Hays, *et al.* 1984; Kane, *et al.* 1984; McDermid, *et al.* 1985; Menzies 1985, 1986; Rye 1989).

Coherent Doppler lidar systems can operate in a continuous-wave (CW) mode, rather than a pulsed mode. Obtaining good range resolution is less straightforward than with pulsed sytems, in which range resolution is determined by the pulse duration and can be shorter than a few hundred meters. Range resolution can be achieved using an optical depth-of-focus effect (Huffaker 1978), or the carrier can be frequency-modulated and range information extracted from the modulation frequency in the scattered wave (Hinkley 1976). CW systems have the disadvantage of being restricted

to ranges smaller than half the coherence length of the reference oscillator; for fractional frequency stabilities of 10^{-10}, the coherence length $c\tau_{coh} \equiv \lambda(\delta\nu/\nu)^{-1}$ for a 1-μm laser is only 10 km. The discussion here will be restricted to *coherent pulsed* systems.

Coherent Doppler radar or lidar on stochastic targets is used for a variety of scientific measurements, many in Earth's atmosphere. For example, coherent microwave radar has been used to study the structure and development of storms by analyzing the scattering from raindrops (Ray, *et al.* 1975; Meneghini, *et al.* 1983), to monitor turbulence in clear air by sensing reflections from density fluctuations (Doviak and Jobson 1979), and to measure wind velocity by studying the scattering from dust particles (aerosols), which exist even in the "clearest" air (Post, *et al.* 1981; Weinman 1988). Coherent infrared lidar systems using carbon-dioxide (CO_2) lasers operating at 10.6-μm wavelength are being developed for wind-sensing applications, possibly operated from an Earth-orbiting satellite or a space shuttle (Huffaker 1978; Harney 1983). Section 5.4 of this report compares the performance of coherent Doppler lidars at 1-μm and 10-μm wavelengths for remote (satellite-based) wind-sensing in Earth's atmosphere. Similar techniques could be used from spacecraft to probe regions around planets and satellites or the interplanetary medium. Solid-state lasers operating in the 1-μm range have size, reliability, power output, and frequency stability that make them superior candidates for lidar systems to be used in space.

The choice of wavelength for a Doppler lidar system can depend on many factors: achievable velocity-measurement precision and range resolution; efficiency (including pulse energies required to achieve a given performance and overall electrical-to-optical efficiency of the laser); practical considerations such as the reliability, size, and lifetime of the laser and other components; and eye-safety considerations. Some of the latter issues are addressed in section 4.2.2d and 5.4, but the discussion here (subsection 4.2.2a–c) will focus primarily on achievable measurement precision.

The precision with which velocity measurements can be made with a pulsed coherent lidar

system depends ultimately on the strength of the backscattered signal. For a given pulse energy, the strength of the backscattered signal depends on wavelength λ in several ways. First, it scales linearly with wavelength (because signal strength is defined in units of photon flux, not energy, hence pulse energies must be divided by $h\nu \equiv hc/\lambda$). Second, it depends on the backscatter cross section β, which is determined by the size, distribution, and possibly shape of the scattering particles and which can have a complicated wavelength dependence (see subsection 4.2.2c). Third, it can be degraded because of interference from background scattered light, an effect which becomes important through the presence of turbulence. For aperture diameters D larger than the transverse atmospheric coherence length ρ_0, the degradation is approximately proportional to $(2\rho_o/D)^2$ [see eq. (4.2.25) below; Clifford and Wandzura 1981; Clifford and Lading 1983]. In Earth's atmosphere, $\rho_0 \propto \lambda^{6/5}$ (approximately); thus, for a given aperture size $D > 2\rho_0$ (at both wavelengths λ_1 and λ_2), the degradation at $\lambda_1 = 1$ μm is 250 times worse than at $\lambda_2 = 10$ μm. This effect is important for ground-based systems, but can be minimized with Earth-orbiting systems.

The strength of the backscattered signal often is not the limiting factor for measurement precision. In wind-sensing performed in Earth's atmosphere with coherent lidar, for example, the primary error sources are (1) the spread in velocities of the scattering particles due to turbulent eddies and bulk wind-shear effects and (2) the spectral width of the transmitted pulses. (Wind-sensing is discussed in more detail in section 5.4.)

Subsection 4.2.2a below describes the dominant contributions to measurement error and resulting limits to performance for velocity measurements made with coherent pulsed Doppler lidar systems. Subsection 4.2.2b derives expressions for the strength of the return signal, neglecting effects of turbulence or wind shear. These expressions are combined with others in subsection 4.2.2a to quantify achievable measurement precision for reasonable parameter values. Subsection 4.2.2c discusses the backscatter cross section and its wavelength dependence, as well as effects of turbulence. Subsection 4.2.2d summarizes requirements on laser technology for applications using coherent

pulsed Doppler lidar.

4.2.2a Velocity- and range-measurement precision

Consider a stochastic target made up of randomly moving particles of various sizes, which exhibit a range of velocities (standard deviation) σ_{med} due to wind shear, turbulence, or similar effects. The precision with which the average (line-of-sight) velocity of the particles can be measured by a coherent pulsed Doppler lidar system is limited by the strength of the received signal, the bandwidth of the transmitted pulse, and the velocity spread of the target medium. In the heuristic language of chapter 3, the first of these sources of measurement error corresponds to photon statistics, and the second to laser frequency spread associated with the finite pulse duration. The third error source, the uncertainty or spread in velocities of the target particles, is analogous to stochastic fluctuations in an interferometer's arm length due to thermal or mechanical noise.

Conventional expressions for lidar velocity precision are based on radar analysis methods. The treatment of measurement noise in the radar analyses is not optimum for optical frequencies and photon-counting detection techniques, but in appropriate limits the methods lead to correct numerical predictions. These discrepancies are pointed out briefly toward the end of this subsection and in subsection 4.2.2b, and more appropriate expressions are indicated. Since the conventional, complicated radar expressions are suspect, and since it is always satisfying to try to understand complicated things in simple ways, this section continues in the spirit of previous chapters by deriving limits for velocity-measurement precision based on simple, heuristic arguments. (The results will be surprisingly accurate, in the sense of agreement with the appropriate limits of the complicated radar equation.)

In general, the rms error (standard deviation) σ_v in a measurement of relative velocity (range

C-2

rate) v made at wavelength λ is related to the uncertainty $\delta\nu$ in Doppler frequency by

$$\sigma_v \sim \frac{\lambda}{2}\,\delta\nu\,. \tag{4.2.10a}$$

This would imply that to achieve a range-rate accuracy σ_v, the fractional frequency stability of the reference laser used to generate the pulses must be known to a level

$$\frac{\delta\nu}{\nu} < \frac{2\sigma_v}{c}\,. \tag{4.2.10b}$$

In Doppler measurements made with pulses, there is a minimum spread or uncertainty $\delta\nu_{\rm bw}$ in the measured Doppler frequencies due to the finite spectral width associated with a pulse of finite duration. The spectral widths associated with Gaussian pulses of duration (FWHM) τ_p or rectangular pulses of duration τ_r are

$$\delta\nu_{\rm bw}^{\rm G} \gtrsim (2\pi\tau_p)^{-1}\,; \qquad \delta\nu_{\rm bw}^{\rm rect} \gtrsim (\pi\tau_r)^{-1}\,. \tag{4.2.11a}$$

Assume Gaussian pulses for the remaining discussion. The uncertainty in estimates of range rate due to the spectral width of the transmitted pulse is therefore on the order of

$$\begin{aligned}\sigma_{v,\rm bw} &\simeq \frac{\lambda}{2}\,\delta\nu_{\rm bw} = 0.15\ \text{m/s}\left(\frac{\lambda}{1\ \mu\text{m}}\,\frac{\delta\nu_{\rm bw}}{0.3\ \text{MHz}}\right)\\ &\gtrsim \frac{\lambda}{4\pi\tau_p} \simeq 0.15\ \text{m/s}\ \left(\frac{\lambda}{1\ \mu\text{m}}\,\frac{0.5\ \mu\text{s}}{\tau_p}\right).\end{aligned} \tag{4.2.11b}$$

Note that this level of measurement error can be identified with an equivalent uncertainty in laser frequency or fractional frequency stability:

$$\frac{\delta\nu_{\rm bw}}{\nu} \sim \frac{2\sigma_{v,\rm bw}}{c} \sim 10^{-9}\left(\frac{\sigma_{v,\rm bw}}{0.15\ \text{m/s}}\right). \tag{4.2.11c}$$

If this is the dominant source of measurement error, it can be reduced by narrowing the spectral width of the transmitted pulses, *i.e.*, by improving the frequency stability of the reference laser.

Increasing the pulse duration also could reduce the pulse spectral widths, but such reduction must be weighed against a loss in range resolution [eq. (4.2.14) below].

The measurement error caused by the velocity spread (standard deviation) σ_{med} of the target particles takes a little more thought — it is not simply proportional to σ_{med}. When Doppler measurements are made with pulses, there is no getting away from error due to the finite spectral width of the pulses. The result is that when the spread in velocities of the target particles is the dominant source of error, the uncertainty in estimates of range rate is "shared" between σ_{med} and $\sigma_{v,\text{bw}}$:

$$\sigma_{v,\text{med}} \sim (\sigma_{v,\text{bw}}\, \sigma_{\text{med}})^{1/2} . \tag{4.2.12a}$$

The spread in velocities of the target particles arises largely from the presence of turbulent eddies and bulk wind-shear effects. In Earth's atmosphere, aerosol particles entrained in wind typically exhibit a spread in velocities $\sigma_{\text{med}} \approx 0.4$ m/s (Huffaker 1978; Kane, *et al.* 1984; Menzies 1986). Combining this with expression (4.2.11b) for $\sigma_{v,\text{bw}}$ implies an error in range-range estimates of

$$\sigma_{v,\text{med}} \simeq 0.25 \text{ m/s} \left(\frac{\lambda}{1\ \mu\text{m}} \frac{\delta\nu_{\text{bw}}}{0.3 \text{ MHz}} \frac{\sigma_{\text{med}}}{0.4 \text{ m/s}} \right)^{1/2} . \tag{4.2.12b}$$

Note that the velocity spread of the particles implies a maximum time τ_{corr} over which successive signal samples are correlated:

$$\tau_{\text{corr}} \equiv \frac{\lambda}{4\pi\sigma_{\text{med}}} \simeq 0.2\ \mu\text{s} \left(\frac{\lambda}{1\ \mu\text{m}} \frac{0.4 \text{ m/s}}{\sigma_{\text{med}}} \right) . \tag{4.2.12c}$$

This defines a coherent integration time or maximum appropriate detector integration time τ_d for a single measurement.

Consider now the contribution to measurement error from photon statistics. If a pulse returns with a Doppler shift ν_D, and heterodyne detection is used to measure the phase relative to a stable reference oscillator, the uncertainty in that phase measurement will be related to the uncertainty $\delta\nu_D$

in the Doppler frequency by $\delta\phi \sim 2\pi\tau_p\delta\nu_D$. Recall that the minimum uncertainty $\delta\phi$ is related to the standard deviation of the detected photocount N by $\delta\phi \geq 1/(2\sigma_N) \approx 1/(2\sqrt{N})$ [see eq. (3.4.1a) and surrounding discussion]. The approximation $\sigma_N \approx \sqrt{N}$, where N is the mean integrated photocount (Poisson statistics), is reasonable for most of the applications described here. This is because signal strengths typically are large compared to the noise powers associated with intensity fluctuations *not* of shot-noise origin, such as detector dark current, additive thermal noise, or background scattered photons. The resulting error in range-rate estimates is

$$\begin{aligned} \sigma_{v,\mathrm{ph}} \sim \frac{\lambda}{2}\,\delta\nu_{\mathrm{ph}} \equiv \frac{\lambda}{2}\,\frac{\delta\phi_{\mathrm{ph}}}{2\pi\tau_p} \gtrsim \frac{\lambda}{8\pi\tau_p\sqrt{N}} \simeq \frac{\sigma_{v,\mathrm{bw}}}{2\sqrt{N}} \\ \simeq 0.08\ \mathrm{m/s}\left(\frac{\lambda}{1\ \mu\mathrm{m}}\,\frac{\delta\nu_{\mathrm{bw}}}{0.3\ \mathrm{MHz}}\right)\left(\frac{1}{N}\right)^{1/2}. \end{aligned} \tag{4.2.13}$$

A derivation of the received signal strength N is given in subsection 4.2.2b; for ranges, pulse energies, and receiver/detector sizes and efficiencies typical of many experiments in Earth's atmosphere, and a backscatter coefficient appropriate for 1-μm wavelengths in Earth's middle troposphere, N can be as large as 20 to 100.

Two further quantities are important for defining the performance of a lidar system. First is the range resolution σ_L, defined by

$$\sigma_L \equiv \frac{cM\tau_p}{2} \simeq 75\ \mathrm{m}\left(M\,\frac{\tau_p}{0.5\ \mu\mathrm{s}}\right). \tag{4.2.14}$$

Here τ_p is the pulse duration, and the multiplicative factor M accounts for additional sources of uncertainty in measurements of pulse arrival times, including the response time τ_{det} of the detector and the characteristic time τ_{sc} for optical interaction with the scattering particles. The combination of these effects determines an optimum range "gate" $\tau_g \equiv M\tau_p$ or duration of the sampling window for comparing returning pulses with pulses from the reference oscillator. The second important quantity is the maximum range rate or maximum Doppler shift expected. If the maximum Doppler

shift to be detected is f, then one must sample the backscattered light at this frequency, and the detector bandwidth should be at least as wide as f. This will permit unambiguous measurements of range rate up to a maximum range rate v_{Ny} (the "Nyquist" velocity) given by

$$v_{Ny} \equiv f\lambda/2 = 50 \text{ m/s} \left(\frac{f}{100 \text{ MHz}}\right) \left(\frac{\lambda}{1 \ \mu\text{m}}\right) . \tag{4.2.15}$$

The expressions derived above for range-rate error or velocity-measurement precision are correct qualitatively and agree well with those obtained from the Doppler *radar* equation in the limits where each of these error sources dominates. For target velocities that exhibit a Gaussian distribution, the Doppler radar expression for the uncertainty in a single measurement of velocity is

$$\sigma_v = \frac{\lambda}{4\pi^{1/4}} \left(\frac{f}{M\tau_p}\right)^{1/2} \left[W + \frac{8\sqrt{\pi}\, W^2}{\text{SNR}_w} + \frac{g}{2\pi^{3/2}} \left(\frac{1}{\text{SNR}_w}\right)^2\right]^{1/2} \tag{4.2.16a}$$

(see, *e.g.*, Huffaker 1978; Kane, *et al.* 1984; Menzies 1986; Zrnic 1979). When care is taken to define signal-to-noise ratios in ways appropriate for optical rather than radio frequencies (as described in subsection 4.2.2b), this expression can be used to describe the performance of coherent Doppler lidar systems under general conditions. Here SNR_w is the "wideband" signal-to-noise ratio, shown in subsection 4.2.2b to be related to the number of detected photons N by

$$\text{SNR}_w{}^2 \simeq (2f\tau_p)\, N \tag{4.2.16b}$$

[see eq. 4.2.24c)], where f is the sampling frequency ($f \simeq \pi b_d$, where b_d is the spectral interval over which detected photons are integrated), and τ_p is the pulse width. The quantity W in eq. (4.2.16a) is the fractional Doppler frequency spread of the return signal due to the spread in target velocities and the finite pulse bandwidth:

$$W \equiv \frac{1}{v_{Ny}} \left(\sigma_{med}{}^2 + \sigma_{bw}{}^2\right)^{1/2} \equiv \frac{\sigma_0}{v_{Ny}} , \tag{4.2.16c}$$

where the symbol σ_0 has been introduced to stand for the root-sum-square (rss) of the scattering particles' spread in velocities and the velocity-measurement error due to the pulse spectral width [eq. (4.2.11b)]. The numerical value of the factor g in the third term of eq. (4.2.16a) depends on the method of numerical analysis used; $g = 1$ for a complex-covariance method of estimation (based on maximum-entropy spectral moments), and $g = 2\pi^2/3$ for a Fourier method of estimation (based on variance of the Fourier power spectral moments) (see Zrnic 1979). It will be assumed here that $g \equiv 1$. The quantity $M\tau_p$ is the range gate defined in eq. (4.2.14) above. Note that improvement in accuracy by a factor of order $N^{-1/2}$ can be achieved by averaging N independent measurements.

To make the complicated radar expression (4.2.16a) more tractable, rewrite it in terms of the three contributions to measurement error derived earlier. Measurement precision σ_v thus is found to be the rss of three quantities,

$$\sigma_v \;=\; (A^2 \;+\; B^2 \;+\; C^2)^{1/2} \;, \tag{4.2.17a}$$

with A, B, and C defined as follows:

$$\begin{aligned} A \;&\equiv\; \alpha \left[{\sigma_{v,\mathrm{bw}}}^4 \;+\; {\sigma_{v,\mathrm{med}}}^4\right]^{1/4} \;=\; \alpha\,(\sigma_0\,\sigma_{v,\mathrm{bw}})^{1/2} \;=\; \alpha\,\sigma_{v,\mathrm{bw}}\,[1 \;+\; (\tau_p/\tau_{\mathrm{corr}})^2]^{1/4} \;, \\ B \;&\equiv\; \frac{2\sigma_0}{\xi^{3/4}\,N^{1/4}} \;, \\ C \;&\equiv\; \sigma_{v,\mathrm{ph}} \;=\; \frac{\sigma_{v,\mathrm{bw}}}{2N^{1/2}} \;, \end{aligned} \tag{4.2.17b}$$

where

$$\alpha \;\equiv (\pi/4)^{1/4} \;\simeq 0.94, \qquad\qquad \xi \;\equiv\; 2f\tau_p \;, \tag{4.2.17c}$$

and M has been set equal to unity (*i.e.*, $\tau_g \equiv \tau_p$). Except for the (puzzling) factor B, the heuristic expressions derived earlier for the separate contributions to measurement error from pulse bandwidth ($\sigma_{v,\mathrm{bw}}$), spread in scattering-particle velocities ($\sigma_{v,\mathrm{med}}$), and photon statistics ($\sigma_{v,\mathrm{ph}}$) agree extremely well with this expression. The terms A, B, and C are evaluated quantitatively in section 5.4

for comparable lidar systems operating at 1-μm and 10-μm wavelengths, with system parameters appropriate for the application of global wind-sensing. For typical system parameters, the B term will make a negligible contribution relative to the A or C terms, and the simple expressions derived at the start of this section will be good indicators of performance. Note, too, that B is the only term in this expression that has an explicit dependence on sampling frequency, and it can be suppressed by use of an appropriately high sampling frequency f.

Note the following interesting feature about the expressions (4.2.17) for velocity-measurement precision: a longer pulse duration τ_p would appear to reduce the error contribution from all three terms. Of course, any such improvement would come at the expense of poorer range resolution ($\sigma_L \sim c\tau_p/2$), so in practice the optimum pulse width will represent an acceptable compromise between range resolution and velocity-measurement precision. This observation does not support the "conventional wisdom" that says no benefit comes from having pulses τ_p longer than the correlation time τ_{corr} set by the spread of velocities in the scattering medium (*i.e.*, from reducing $\sigma_{v,\text{bw}}$ below σ_{med} (see, *e.g.*, Menzies 1986). The correlation time τ_{corr} *does* define a coherent integration time, which defines a maximum detector integration time τ_d and thus limits the detected photocount N.

A general question of interest is, for what ranges of signal strength (N) will the terms A, B, or C contribute appreciably to measurement error? Examination of eq. (4.2.17b) shows that A is always at least as large as $\sigma_{v,\text{bw}}$, and C can be larger than $\sigma_{v,\text{bw}}$ only for $N \lesssim 1$. Thus, A *always* dominates C except in the extreme photon-statistics-limited case in which there is less than one detected photon per integration time (over the entire detection bandwidth). In terms of the quantities $Y \equiv 4\pi W \equiv 4\pi\sigma_0/v_{\text{Ny}}$ and ξ defined above, the relative magnitudes of A, B, and C, will

be determined by the following conditions:

$$
\begin{aligned}
A \geq B &\Leftrightarrow \xi N \geq \frac{4}{\pi} Y^2 \simeq Y^2 \, ; \\
A \geq C &\Leftrightarrow \xi N \geq \left(\frac{4}{\pi}\right)^{1/2} Y^{-1} \simeq Y^{-1} \, ; \\
B \geq C &\Leftrightarrow \xi N \geq Y^{-4} \, .
\end{aligned}
\tag{4.2.18}
$$

For a lidar system operating at a given wavelength, only three parameters need be assumed in order to evaluate the quantities A, B, and C as functions of the number of detected photons N: the spread in velocities of the scattering particles (σ_{med}); the maximum relative velocity (range rate) to be measured (v_{Ny}); and the pulse duration (τ_p). From these one can infer the quantities $\sigma_{v,\text{bw}}$, σ_0, and the sampling rate f. In turn, this establishes values for the above-defined quantities Y and ξ. Table 4.2.1 lists operational parameters for 1-μm and 10-μm lidar systems, assuming pulse widths $\tau_p = 2$ μs, a Nyquist velocity $v_{\text{Ny}} = 50$ m/s, and a spread in particle velocities $\sigma_{\text{med}} = 0.4$ m/s. The terms A, B, and C are evaluated as functions of the number of detected photons N. With these assumed parameters, neither tabulated case is photon-statistics limited (provided, of course, that the integration time is long enough that $N \gtrsim 1$), and measurement precision for each is set by the term A.

Table 4.2.1 Operational parameters for $\lambda = 1$ μm and $\lambda = 10$ μm lidar systems, based on assumptions of pulse width $\tau_p = 2$ μs, Nyquist velocity $v_{\mathrm{Ny}} = 50$ m/s, and spread in velocities of scattering particles $\sigma_{\mathrm{med}} = 0.4$ m/s. The range resolution for both systems, assuming a range gate τ_g equal to the pulse width τ_p, is $c\tau_p/2 \simeq 300$ m. See eqs. (4.2.17) for definitions of A, B, and C. Fourier-transform-limited Gaussian pulses are assumed $[\sigma_{\mathrm{bw}} \equiv (\lambda/2)\delta\nu_{\mathrm{bw}} = \lambda/(4\pi\tau_p)]$.The detector integration time τ_d has been set equal to the correlation time $\tau_{\mathrm{corr}} \equiv \lambda/(4\pi\sigma_{\mathrm{med}})$, and is therefore 10 times larger for the 10-μm system.

Parameter	$\lambda = 1$ μm	$\lambda = 10$ μm
$\sigma_{v,\mathrm{bw}}$	0.04 m/s	0.4 m/s
σ_0	0.4 m/s	0.57 m/s
τ_d	0.2 μs	2.0 μs
f	100 MHz	10 MHz
$Y \equiv 4\pi\sigma_0/v_{\mathrm{Ny}}$	0.1	0.13
$\xi \equiv 2f\tau_p$	800	80
A	0.12 m/s	0.45 m/s
B	$0.005/N^{1/4}$ m/s	$0.04/N^{1/4}$ m/s
C	$0.02/N^{1/2}$ m/s	$0.2/N^{1/2}$ m/s

To achieve a velocity error σ_v smaller than 0.5 m/s with the 10-μm system or better than 0.1 m/s with the 1-μm system, the term A must be reduced. This requires reducing either σ_{med} or $\sigma_{v,\mathrm{bw}}$ [eq. (4.2.17b)]. Typically, one has no control over the velocity spread σ_{med} of the scattering medium; hence, one must reduce $\sigma_{v,\mathrm{bw}}$. If the reference laser used to generate the pulses is essentially perfectly stable in frequency and the (Gaussian) pulses are Fourier-transform-limited — *i.e.*, the pulse spectral width is due entirely to the finite duration of the pulse — then $\sigma_{v,\mathrm{bw}} = \lambda/(4\pi\tau_p)$ [eq. (4.2.11b)], and improvement is possible only by increasing the pulse duration τ_p. Of course, this will degrade range resolution [eq. (4.2.14)]. In general, the pulse spectral width $\delta\nu_{\mathrm{bw}}$ may be considerably wider than $(2\pi\tau_p)^{-1}$, which is due in part to frequency instabilities in the reference laser used to generate the pulses; thus, $\sigma_{v,\mathrm{bw}} \equiv (\lambda/2)\delta\nu_{\mathrm{bw}} > \lambda/(4\pi\tau_p)$. In the current example, what pulse spectral width is required to reduce term A by 50%, *i.e.*, to give an error of about 6 cm/s for the 1-μm system and about 25 cm/s for the 10-μm system? Note that reduction beyond these values could make term C start to contribute (the number N of detected photons in an integration time is assumed

to be on the order of 1 or more). Such a reduction requires that $\sigma_{v,\mathrm{bw}}$ be reduced by a factor of 4, to $\sim$ 1 cm/s for the 1-μm system and to $\sim$ 10 cm/s for the 10-μm system. These correspond to pulse spectral widths of about 20 kHz for each system, requiring pulse durations of at least 8 μs, and reference-laser frequency stabilities of $\sim 7 \times 10^{-11}$ and $\sim 7 \times 10^{-10}$ for the 1-μm and 10-μm sources, respectively. Note that for the 10-μm system to achieve an accuracy comparable to the 6-cm/s accuracy of the 1-μm system (assuming $N >> 10$, so that the A term still dominates the B and C terms), its pulse spectral width would have to be on the order of 2 kHz, requiring a pulse duration on the order of 100 μs and a reference-laser frequency stability of 7×10^{-11}.

In some cases, the figure of merit in terms B and C may not be the required number of detected photons, but rather the pulse energy, for example. To compare systems at different wavelengths in these more general ways, one must know how N scales with wavelength. As shown in subsection 4.2.2b, it depends linearly on wavelength through the detector integration time τ_d [eq. (4.2.12c)] and the number of photons per pulse $E_p/h\nu$; it depends on wavelength in a more complicated way through the backscatter coefficient β and the overall efficiency η. The detector integration time τ_d is limited by the correlation time $\tau_{\mathrm{corr}} \simeq \lambda/4\pi\sigma_{\mathrm{med}}$. The efficiency η includes effects of propagation losses (*e.g.*, absorption) as well as receiver and detector efficiencies; for simplicity here, assume η is approximately constant for the wavelengths being compared. Hence for a given minimum required number of detected photons, the minimum required product of pulse energy E_p and backscatter coefficient β will be 100 times *higher* for the 1-μm system than for the 10-μm system. However, if the scattering particles are smaller than 1 μm, and Rayleigh scattering applies, the backscatter coefficient will scale as λ^{-4}; hence the minimum required pulse energies will be 100 times *lower* for the 1-μm system than for the 10-μm system. Thus, there is always an advantage (in terms of improved velocity-measurement precision for given pulse energies) to using shorter wavelengths, as long as Rayleigh scattering still applies, *i.e.*, as long as the wavelength is still longer than the characteristic size of the scattering particles. Conversely, for particles large relative to the laser wavelength (the

geometric-optics scattering limit), the backscatter coefficient is relatively insensitive to wavelength, and better measurement precision can be achieved for given pulse energies with longer wavelengths. Further performance trade-offs between 1-μm (Nd:YAG) lidar systems and 10-μm (CO_2) lidar systems are discussed in subsection 4.2.2d, and in more detail in section 5.4 for the application of global wind-sensing.

The above discussion has quantified to some extent the statement made in the introduction to this section that the use of shorter wavelengths (*e.g.*, 1-μm instead of 10-μm lasers) offered the potential for better measurement precision. Now that their validity has been assured, the simple expressions [eqs. (4.2.11–13)] can be used to justify this claim in a general way by considering the trade-offs between velocity-measurement precision and range resolution for different wavelengths and comparable signal strengths N. When the uncertainty in measured velocity stems primarily from the transmitted pulse bandwidth [eq. (4.2.11b)], the product of velocity uncertainty and range resolution scales linearly with wavelength:

$$\sigma_{v,\mathrm{bw}}\, \sigma_L \;\propto\; \lambda \;. \tag{4.2.19a}$$

When the velocity uncertainty is due primarily to the spread in target velocities [eq. (4.2.12b)], measurement precision improves only as the square root of the wavelength for a given range resolution:

$$\sigma_{v,\mathrm{med}}\, \sigma_L{}^{1/2} \;\propto\; (\lambda\, \sigma_{\mathrm{med}})^{1/2} \;. \tag{4.2.19b}$$

This $\lambda^{1/2}$ dependence can be understood as improvement in precision of the measured Doppler shift by the inverse square root of the number of cycles counted; for a particular range gate or measurement integration time, more full cycles of return signal are received at shorter wavelengths. Finally, if photon statistics are the dominant error source [eq. (4.2.13)], then, for a given number of detected photons N, the product of velocity precision and range resolution scales linearly with

wavelength:

$$\sigma_{v,\mathrm{ph}}\,\sigma_L \propto \frac{\lambda}{N^{1/2}}\,. \tag{4.2.19c}$$

4.2.2b Received signal strength

The previous discussion indicated the role played by photon statistics (or signal-to-noise ratio, or received signal strength) in the performance of coherent lidar systems. For quantitative examples, it assumed that the signal was strong enough to ensure that at least one signal photon would be detected per integration time ($N \geq 1$). This subsection examines signal strength more closely, and shows under what conditions the assumption $N \geq 1$ is reasonable.

The flux s of backscattered "signal" photons arriving at the detector (units of s^{-1}) is given approximately by

$$s = \frac{c}{2}\,\frac{E_\mathrm{p}}{h\nu}\,\beta\,\frac{A_\mathrm{r}}{L^2}\,\eta\,. \tag{4.2.20a}$$

Here E_p is the pulse energy, β is the volume backscatter coefficient (in units of $\mathrm{m}^{-1}\ \mathrm{sr}^{-1}$ throughout this section), A_r is the receiver area, L is the range, and η is an overall efficiency that includes such factors as system optical efficiencies, detector quantum efficiencies, and the two-way transmission (at wavelength λ) along the path between the lidar and the medium. The backscatter coefficient is discussed in subsection 4.2.2c. In general, this would be an integral equation with β and η functions of the range L and the pulse energy E_p replaced by the product of the pulse width τ_p and the temporal power profile of the pulse evaluated at a time $2L/c$ prior to the measurement. Equation (4.2.20a) is valid if these dependences on range can be ignored, and the pulse width is small compared to the round-trip light-travel time. The factor of 1/2 reflects the fact that the signal arises from a slab of target medium of thickness $c\tau_\mathrm{p}/2$. A typical value for β in Earth's atmosphere might be $10^{-8}\ \mathrm{m}^{-1}\mathrm{sr}^{-1}$ (see section 5.4). With this value for β, pulse energies of 100 mJ, an overall efficiency $\eta \simeq 0.1$, a range $L \simeq 10$ km, a 0.4-m diameter circular receiver aperture ($A_r \equiv \pi {D_r}^2/4$), and a 1-μm

laser wavelength, the signal photon flux would be $s \approx 10^8\ \mathrm{s}^{-1}$; it scales for different parameters as

$$s \simeq 10^8\ \mathrm{s}^{-1} \left(\frac{E_\mathrm{p}}{0.1\ \mathrm{J}}\ \frac{\lambda}{1\ \mu\mathrm{m}}\ \frac{\beta}{10^{-8}}\ \frac{\eta}{0.1}\right) \left(\frac{D_r}{0.4\ \mathrm{m}}\ \frac{10\ \mathrm{km}}{L}\right)^2 . \tag{4.2.20b}$$

It is useful to decompose the signal flux into its spectral components $\tilde{s}(\nu) \equiv \tilde{s}$, which have units of $\mathrm{s}^{-1}\ \mathrm{Hz}^{-1}$. If the backscattered signal photons arrive with a uniform spread in frequency over a bandwidth B_sig, the spectral flux of arriving signal photons is

$$\tilde{s} \equiv s/B_\mathrm{sig} . \tag{4.2.21a}$$

[A "tilde" ($\sim$) above a quantity indicates a Fourier component.] The signal spectral width B_sig is related to the velocity-measurement error σ_v by

$$B_\mathrm{sig} \simeq \left(\frac{2}{\lambda}\right) \sigma_v . \tag{4.2.21b}$$

If velocity error is not dominated by photon statistics, but rather by a combination of the pulse spectral width and the spread in particle velocities, a good estimate for B_sig can be made by setting σ_v equal to term A in eq. (4.2.17a). Provided the pulse width τ_p is not very much longer than the correlation time τ_corr [eqs. (4.2.12c) and (4.2.17b)], one finds that

$$B_\mathrm{sig} \simeq (2\pi\tau_p)^{-1} . \tag{4.2.21c}$$

"Noise" photons, consisting of detector thermal and dark-current photons and background backscattered photons, are also counted by the detector, and arrive with a spectral flux denoted here by $\tilde{n}$. A useful quantity known as the "narrowband SNR" can be defined without reference to a particular detector. It is equal to the ratio of the mean spectral flux of photons at the detector, $\tilde{n}_d$, to its standard deviation, $\sigma_{\tilde{n}_d}$. Under the assumptions that (1) the signal photons (backscattered laser

light) obey Poisson statistics and (2) the noise photons behave statistically like additive thermal noise, it can be shown that the mean spectral flux of photons at the detector and its standard deviation are given by

$$\tilde{n}_d \;=\; \tilde{s} + \tilde{n}\;, \tag{4.2.22a}$$

$$\sigma_{\tilde{n}_d} \;=\; \left[\tilde{s}(1+2\tilde{n}) + \tilde{n}(\tilde{n}+1)\right]^{1/2} \tag{4.2.22b}$$

(Schumaker 1984, 1989). If the spectral flux of noise photons is much smaller than that of signal photons ($\tilde{n} << \tilde{s}$, with $\tilde{s} > 1$), the standard deviation of the detected spectral flux can be simplified to

$$\sigma_{\tilde{n}_d} \;\simeq\; \left[\tilde{s}(1+2\tilde{n})\right]^{1/2}\;. \tag{4.2.22c}$$

In this limit, the narrowband SNR takes on its conventional "Poisson" character:

$$\mathrm{SNR_n} \;\equiv\; \frac{\tilde{n}_d}{\sigma_{\tilde{n}_d}} \;\simeq\; \tilde{s}^{1/2}\,[1-(\tilde{n}/\tilde{s})^2] \;\simeq\; \tilde{s}^{1/2}\;. \tag{4.2.22d}$$

If the detector has a rectangular response function (corresponding to a uniform weighting of the measurements, or cycle counting), the total number of signal photons detected in an integration time τ_d ($\tau_d \lesssim \tau_p$) is

$$N \equiv \tau_d\, s \;=\; B_{\mathrm{sig}}\, \tau_d\, \tilde{s}\;. \tag{4.2.23a}$$

The narrowband SNR thus is related to the square root of the total number of detected signal photons by

$$\mathrm{SNR_n} \;\simeq\; \left(\frac{1}{B_{\mathrm{sig}}\tau_d}\right)^{1/2} N^{1/2}\;. \tag{4.2.23b}$$

The narrowband SNR gives the ratio of signal power density to noise power density at the signal (laser) frequency. In practice, because of turbulence and shear effects (σ_{med}), the signal energy is spread over a range of frequencies. Because of this, it is conventional to define a "wideband" SNR, obtained from statistics of the detected photon flux averaged over a spectral interval b_d related to the sampling frequency f by $b_d \simeq f/\pi$. The mean flux n_d of detected photons in each spectral bin

(units of s^{-1}) and its standard deviation are related to the spectral quantities defined in eqs. (4.2.22) by

$$\begin{aligned} n_d &= b_d\, \tilde{n}_d \ , \\ \sigma_{n_d} &= \left(\frac{b_d}{\tau_d}\right)^{1/2} \sigma_{\tilde{n}_d} \ . \end{aligned} \tag{4.2.24a}$$

The wideband SNR is given by

$$\mathrm{SNR_w} \equiv \frac{n_d}{\sigma_{n_d}} = (b_d \tau_d)^{1/2}\, \mathrm{SNR_n} \simeq \left(\frac{b_d}{B_{\mathrm{sig}}}\right)^{1/2} N^{1/2} \ . \tag{4.2.24b}$$

Setting $b_d = f/\pi$ and $B_{\mathrm{sig}} = (2\pi\tau_p)^{-1}$ [eq. (4.2.20c)], one finds the following relation between wideband SNR and number of detected photons:

$$\mathrm{SNR_w} \simeq (2f\tau_p)^{1/2}\, N^{1/2} \ . \tag{4.2.24c}$$

This relation was used in subsection 4.2.2a [eq. (4.2.16b)] to rewrite the complicated radar equation for velocity-measurement precision in terms of the quantities σ_{med}, $\sigma_{v,\mathrm{bw}}$, and N. For the parameters indicated in eq. (4.2.20b) and in Table 4.2.1, appropriate for a 1-μm coherent lidar system, the number of detected photons per integration time and the wideband signal-to-noise ratio are

$$N \simeq 20 \ , \qquad \mathrm{SNR_w} \simeq 90 \ . \tag{4.2.24d}$$

4.2.2c Backscatter cross section and effects of turbulence

The strength of the returned lidar signal depends linearly on the cross section for backscattering at the lidar wavelength. The backscatter volume cross section β (m^{-1} sr^{-1}) depends on the size of the scattering particles. For Rayleigh scattering, where particle sizes are smaller than the wavelength, the cross section scales as λ^{-4}. For particles considerably larger than the wavelength (the "geometric

optics" regime), the scattering has a very weak dependence on wavelength. For particles on the order of the wavelength ("Mie scattering"), the wavelength dependence is less pronounced than in Rayleigh scattering and does not have a simple analytic form, except when the distribution of sizes follows a simple power-law (the "Junge" model) (Hinkley 1976). Typically, in this case, the cross section still grows with shorter wavelength. For example, in Earth's atmosphere, the scattering particles comprising most of the aerosol mass are approximately 1 μm in size. Because Rayleigh scattering would not apply for a 1-μm laser, the ratio of backscatter at 1 μm to backscatter at 10 μm is far less than 10,000—typically in the range 40 to 200 (Kane, *et al.* 1984). Some discussion of the backscatter cross section at 1- and 10-μm wavelength for Earth's atmosphere appears in section 5.4.

Signal strength is maximized by using high pulse energies and choosing wavelengths that maximize the backscatter cross section. It is equally important that interference from background scattered light be minimized. The amount of background light received depends on the number of spatial modes seen by the receiver (*i.e.*, the receiver field-of-view). In the absence of turbulence, if the receiver is diffraction-limited (one spatial mode), the amount of background light detected is independent of aperture size; thus, one should use a large, diffraction-limited aperture, so that maximum signal is collected from a single diffraction-limited spot on the focal plane. However, when turbulence (refractive-index fluctuations) is present, the signal power collected in a single spatial mode stops increasing with receiver area when the aperture size D approaches the transverse coherence length ρ_0 of the turbulent medium in the focused beam, because more and more background light interferes. (Note that it is turbulence immediately in front of the telescope that is important; turbulence in the scattering medium has no effect, since the scattering process is incoherent.) To a good approximation, the reduction in signal strength due to turbulence in the focused beam can be described by the factor (Clifford and Wandzura 1981)

$$\eta_{\text{turb}} \equiv \frac{1}{1 + D^2/4{\rho_0}^2}. \tag{4.2.25}$$

In Earth's atmosphere, $\rho_0 \propto \lambda^{6/5}$. Thus, for example, the coherence length at 10 μm is $\sim$ 15.8 times longer than at 1 μm, and, as noted earlier, for a given aperture $D > 2\rho_0$, the degradation at 1 μm is $\sim$ 250 times worse than at 10 μm. Signal loss due to turbulence is an important problem for ground-based systems and may result in the need to collect signal in several different spatial modes. The effect is more important at shorter wavelengths, both because of the shorter transverse coherence length of the atmosphere, and because the diffraction-limited spot size is smaller (the area of the latter scales as λ^2). The negative effects of turbulence can be minimized by using small apertures and essentially eliminated by using a system in Earth orbit.

4.2.2d Laser technology requirements

The application of coherent pulsed Doppler lidar poses several challenges to laser technology. (1) A frequency-stabilized, low-power reference laser is needed to ensure that the frequency at which the pulses are transmitted does not fluctuate or drift appreciably over the round-trip light-travel time. This laser serves both as a regulator for the transmitted pulses and as a local oscillator for heterodyne detection of the returned signal. (2) Pulses from the low-frequency reference laser must be amplified to produce high pulse energies, and these must be produced at high repetition rates. (3) The reference-laser frequency must be tunable over a wide range to compensate for Doppler shifts due not only to relative motion of the scattering particles but, more importantly, to the potentially large relative velocities encountered with an orbiting spacecraft.

A system designed to have a velocity-measurement error under $\sim$ 0.5 m/s must have a reference laser whose fractional frequency stability is on the order of 10^{-9} or better, *i.e.*, well under 1 MHz at 1-μm wavelengths. For a range of 15 km, this stability must hold for at least 100 μsec. Pulse energies of at least a few hundred millijoule are desirable, at repetition rates up to several tens of hertz. The tuning requirement imposed by the velocities of the scattering particles alone is mild:

at 1-μm wavelength, scattering particles with velocities of order 1 m/s produce Doppler shifts of order 2 MHz. However, for systems in Earth orbit, the relative velocities between spacecraft and scatterers can be large and the tuning requirement severe: the Doppler shift at 1-μm due to an orbiting spacecraft with a 7.2-km/s ground velocity and an observation angle 10° from the vertical is approximately 2.5 GHz. Use on satellites, *e.g.*, for Earth remote sensing or wind-sensing, also requires that the lasers be efficient, reliable, and long-lived, such that they can operate unattended for several years.

Solid-state lasers in general are excellent candidates for use with Doppler lidar. They provide spectral coverage from 0.8 μm to about 3.5 μm and can be frequency-doubled to serve at shorter wavelengths, in contrast to dye lasers, which do not operate in CW mode at wavelengths longer than about 1 μm. Better frequency stabilities can be obtained with solid-state lasers than with dye lasers. In solid-state lasers, the primary cause of frequency instabilities is refractive-index fluctuations due to variations in pump power or ambient temperature; these are lower in magnitude and typically occur at lower frequencies than the frequency jitter caused by flowing dye streams. Solid-state lasers are more reliable and longer-lived than dye lasers (infrared laser dyes also tend to degrade quickly because of thermal decomposition); when pumped with diode lasers (as opposed to flash lamps), their overall electrical-to-optical efficiencies have been made as high as 10% (Byer 1988). Eye-safety considerations point to lasers that operate at wavelengths longer than about 1.4 μm, a regime accessible by a variety of solid-state lasers currently under development (Moulton 1987).

Among the strongest candidate solid-state lasers for applications such as remote sensing and laser backscattering measurements in Earth's atmosphere are alexandrite and titanium sapphire lasers ((Moulton 1987). Alexandrite lasers operate in both CW and pulsed modes between 0.7 and 0.8 μm (roughly); titanium sapphire lasers can operate over a broader wavelength range, roughly 0.65 through 1.2 μm for pulsed operation and a slighly narrower range for CW operation. For longer-wavelength operation, from about 1.5 to 2.5 μm, cobalt magnesium-fluoride (Co:MgF_2) lasers are

good candidates. Alexandrite lasers have been used to measure water vapor and oxygen concentrations in the atmosphere (0.726 and 0.760 μm, respectively). They typically produce pulse energies of 1 to 200 mJ at 10-Hz repetition rates, although they have been made to produce pulses of 0.5 J at 250-Hz repetition rates for more demanding applications such as laser isotope separation (Moulton 1987). When pumped by frequency-doubled neodymium lasers ($\sim$ 0.5-μm wavelength), which in turn are pumped by laser diode arrays, titanium sapphire lasers are among the most efficient, reliable, and broadly tunable of all solid-state laser systems. Their tuning range is broad enough to serve a variety of remote sensing applications, *e.g.*, they can detect both the 0.726- and 0.938-μm water bands, as well as the 0.760-μm oxygen band (Hinkley 1976). These lasers can produce 1-J pulses at 10-Hz repetition rates.

For specific applications in which such broad tunability is not required, the preferred solid-state lasers are typically neodymium- or holmium-doped crystals or glasses. Neodymium-doped YAG lasers, for example, which operate at 1.06 μm, have been developed extensively in recent years because of their potential for a wide variety of scientific uses, especially for space-related applications. Low-power ($<$ 400 mW) Nd:YAG ring lasers operated in special nonplanar configurations have exhibited single-mode operation and free-running line widths as narrow as 100 Hz on time scales of a few tens of microseconds, 0.5 to 1 kHz over times of 1 to 10 ms, and several tens of kHz for times up to 1 second (Kane, *et al.* 1987; Bush, *et al.* 1988; Nilsson, *et al.* 1989). High peak powers can be obtained using multipass Nd:YAG or Nd:glass slabs as amplifiers. These can produce pulse energies of 1 to 10 J at repetition rates up to 100 Hz (Byer 1988; Fan and Byer 1988; Reed, *et al.* 1988). Acousto-optic modulators are used to chop the amplified beam into pulses a few microseconds in duration and to produce the desired frequency offset to compensate for the Doppler shift of the return signal. These lasers can be tuned easily over ranges of several tens of GHz (Byer 1988; Fan and Byer 1988). They are small and light (low-power diode-pumped versions, tunable over 20 GHz, weigh much less than 1 kg), rugged, and reliable. When pumped with small diode lasers instead

of flash lamps, the Nd:YAG lasers are more efficient than CO_2 lasers (approximately 10% *vs.* 5%) and offer longer operating lifetimes (Kane, *et al.* 1987; Byer 1988; Fan and Byer 1988). However, a potentially important practical drawback to the use of 1-μm Nd:YAG lasers for some applications is the eye-safety problem: ANSI recommendations would limit pulse energies to about 40 mJ with a 1-m diffraction-limited telescope (Menzies 1986; Sliney and Wolbarsht 1982). Development is under way of solid-state lasers with the desirable characteristics of Nd:YAG but at slightly longer wavelengths. Holmium-doped solid-state lasers, which operate at 2 μm, are a possible alternative candidate currently being explored. Further discussion and comparison of different lidar systems for the application of wind-sensing by Earth-orbiting spacecraft can be found in section 5.4, and references therein.

5. SCIENCE EXPERIMENTS REQUIRING STABILIZED LASERS: EXAMPLES

5.1 Metrology for astrometric and imaging interferometers

Laser interferometers have obvious important practical application for instrument metrology. Use of a sensitive laser metrology system to calibrate or control the relative motions and distortions of elements within an instrument or measuring apparatus can enable much better measurement accuracy than could be achieved by depending only on mechanical rigidity and thermal isolation for instrument stability. One kind of scientific instrument that demands very precise metrology is a phase-coherent optical interferometer for astrometry or astronomical imaging purposes. In contrast to amplitude or intensity interferometers, which make visibility measurements of the interference fringes, phase-coherent interferometers actively control the relative path lengths of the interfering beams in order to maintain high fringe visibility. On the ground, atmosphere-induced phase fluctuations limit angular resolutions to a few milliarcseconds, even for very bright astronomical point sources (Shao, *et al.* 1987, 1988). In space, microarcsecond astrometry and 100-microarcsecond imaging resolution on fairly dim objects are thought to be achievable with modest-size instruments (2- to 20-meter baselines), but they require that sytematic errors, particularly those associated with knowledge or control of the optical path lengths in the interferometer, be reduced to the level of photon-statistics error (Reasenberg 1986; Reasenberg, *et al.* 1988; Shao, *et al.* 1988).

Figure 14 is a simplified diagram of the way in which such an interferometer and its laser metrology system might work. Starlight arriving at an angle θ relative to the interferometer axis arrives at the two telescopes with a relative delay. From the telescope primary mirrors, the received light signals are guided to and combined at a beam splitter. Fringes, corresponding to constructive or destructive interference between the beams, are formed as the delay changes by amounts comparable to a wave period. The delay depends on the position of the star relative to the interferometer baseline; once the baseline orientation is calibrated, the angular positions of stars can be measured.

The wider the optical bandpass of the system, the stronger is the detected signal, and hence the more precisely the delay can be measured. However, since the fringe intensity pattern is proportional to the Fourier transform of the optical bandpass, a wider optical bandpass means a narrower interference pattern (fewer fringes). In order to track these fringes and maintain high visibility, the relative path lengths from mirror to beam splitter must be controlled or monitored to a small fraction of a wavelength. To achieve angular precisions comparable to that determined by photon statistics, this "small fraction" should be as small as 10^{-4}, *i.e.*, tens of picometers.

For a nominal baseline length L and equal paths from mirrors to the beam splitter, the path-length difference for starlight arriving at an angle θ is $L \sin\theta$. Adjacent fringes correspond to changes in this difference by one wavelength; hence the angular spacing between fringes is

$$\Phi = \frac{\lambda}{L\cos\theta} \simeq 10 \text{ mas} \left(\frac{\lambda}{0.5\ \mu\text{m}}\right) \left(\frac{10 \text{ m}}{L\cos\theta}\right) \tag{5.1.1}$$

("mas" is milliarcseconds). Achievement of the microarcsecond angular precision desired of these instruments clearly requires some serious fringe-splitting. If the relative path lengths from the primary mirrors to the beam splitter could be controlled perfectly, and other sources of phase fluctuations at the detectors (*e.g.*, scattered light) were negligible, photon statistics would limit angular precision to $\sigma_{\text{ph}} \sim \Phi/(4\pi\sqrt{N})$, where N is the number of detected signal photons [eqs. (3.4.1)]. Recall that a 10-minute observation of a star of apparent visual magnitude 10 with a 1-meter telescope and 2% overall optics and detection efficiency over a detection bandwidth (wavelength) from 0.5 to 0.6 μm will produce a mean detected photocount N of about 10 million photons [eq. (3.4.3)], hence a photon-statistics path-length error of about 13 pm, or about 0.3 microarcsecond over a 10-m baseline. To achieve 1-microarcsecond measurement precision with a nominal 10-m baseline would require real-time control of (uncorrelated) fluctuations in the relative path lengths to better than 50 pm. Interferometers this ambitious are being studied seriously for eventual use in Earth orbit (Reasenberg 1986; Reasenberg, *et al.* 1988; Shao 1988).

How realistic are these requirements of picometer-level control of relative path lengths? It is shown below that the demands on laser frequency stability to achieve this level of metrology, in principle, are not extreme, provided scattered light can be minimized. However, thermal and mechanical noise, which produce distortions and fluctuations in the shapes and positions of the optical elements, are likely to pose serious challenges.

Refer again to Figure 14 and its accompanying description, which shows in a simplified way the general function of a laser metrology system for an optical interferometer. The requirement on laser fractional frequency stability is minimized by the use of a dual-arm configuration for the metrology interferometers used to measure and control the relative path lengths. If the path lengths are held equal to a fraction f, the fractional frequency fluctuations of the laser must be controlled or calibrated at least to the level σ/f, where σ is the desired angular measurement precision:

$$\frac{\delta\nu}{\nu} = \frac{\sigma}{f} \simeq 5 \times 10^{-9} \left(\frac{\sigma}{1\ \mu\text{as}}\right) \left(\frac{10^{-3}}{f}\right) . \tag{5.1.2}$$

As explained in section 3.4.2, a dual-arm configuration doesn't help overcome the effects of scattered laser light, since the statistics of scattered light are not correlated in the two arms. If the scattered light arriving at the beam splitter carries an intensity equal to a fraction $\epsilon_{sc}{}^2$ of the main beams, the requirement on laser fractional frequency stability (or knowledge) to achieve an angular precision σ is

$$\frac{\delta\nu}{\nu} = \frac{\sigma}{\epsilon_{sc}} \tag{5.1.3}$$

[see eqs. (3.4.6)]. In practice, the factor $\epsilon_{sc}{}^2$ typically can be kept to 0.1% or smaller with modest effort, implying a requirement on laser frequency stability no worse than $\delta\nu/\nu \simeq 2 \times 10^{-10}$ for 1-μas angular measurements.

The laser metrology systems in these interferometers are designed to monitor the average changes in path lengths due to motions and distortions of the optical elements over time scales

that range typically from 10 minutes to several hours. Hence fluctuations occurring on shorter time scales, such as mechanical vibrations, do not need to be tracked, and the metrology servos can have narrow ($\sim$ millihertz) bandwidths. (Contrast this with the much more stringent requirements for ground-based laser-interferometer gravitational-wave detectors, in which fluctuations on time scales of milliseconds are important; see section 5.2.) Of course, the metrology measurements, while very precise, still provide biased measurements of the angle θ. Each optical element introduces its own bias, and the overall measurement bias will vary with time and must be calibrated and corrected for in the data analysis. The time scales over which the metrology bias and related errors can be tolerated will depend on the throughput of the instrument – *i.e.*, the rate at which measurements can be performed as well as the precision of each measurement. To provide significant science return in mission lifetimes under 10 years, it is reasonable to expect that a space-based astrometric or imaging interferometer should be able to make enough measurements to estimate instrument bias parameters and star positions to the desired precision several times per day (Reasenberg, *et al.* 1988). In this case, linear, slowly changing instrument biases (time scales of several hours) could be accepted.

Although long-term linear changes in measurement bias can be tolerated, requirements on thermal and mechanical stability of the interferometer elements (including both starlight and metrology mirrors, retroreflectors, and beam splitters) are still severe on time scales of 10 minutes to several hours. Thermal-induced distortions of the optical elements can arise from local thermal gradients or nonuniform coefficients of thermal expansion. Commonly used glasses such as zerodur and ULE typically have linear coefficients of thermal expansion of about 3×10^{-8} K^{-1} in temperature ranges of 10 to 50 K, which may be uniform to about 25%, or 10^{-8} K^{-1} (Shao, private communication 1988). The thicknesses of these elements generally are between 1 and 5 cm. Hence, even with these stable materials, temperatures would have to be controlled to a few hundredths of a Kelvin to keep the magnitudes of distortions below 10 pm, as required for microarcsecond precision from short-baseline ($L < 10$ m) interferometers.

5.2 Gravitational-wave detection

What are gravitational waves*? Where do they come from? What are our chances of detecting them? Why should we try?

This section will attempt to address each of these questions in turn. The reader who has persevered this far in the report is reasonably prepared to address the third question already, because, in a sense, most of this report has addressed it. The detection of gravitational waves by using a laser interferometer to monitor the difference in displacements of two "inertial" test masses at the ends of the arms of a dual-arm interferometer (as in Figures 7 and 9 through 11) requires extremely good measurement sensitivity ($\delta\ell/L$), better than any other scientific application being pursued at this time. All of the error sources described in this report can pose serious challenges to the gravitational-wave experimenter. On the ground, seismic and gravity-gradient noise restricts measurements to short time scales (100 msec or less, typically 1 msec; see section 3.4.5). Since baseline lengths are limited on the ground, sensitivity must be enhanced by suppressing photon-statistics noise (*e.g.*, by power "recycling" techniques that increase effective laser power); by increasing the effective arm lengths and resulting optical signal (phase shift) for a given strain induced by a gravitational wave (by operating the interferometer arms as multireflection delay lines or resonant optical cavities); and by suppressing scattered light; (see section 4.1). Knowledge or control of laser frequency fluctuations is essential in order to distinguish their effects on the observed optical phase shifts from the effects of a passing gravitational wave, especially in the presence of appreciable scattered light. In space, baselines can be very long (millions of kilometers, if spacecraft in solar orbit are used), so one can make

* The reader is reminded that "gravitational" waves should be distinguished from "gravity" waves, although general usage has blurred the distinction. Gravitational waves are transmissions of signals caused by changes in the structure of space-time that give rise to changes in the gravitational force felt by an object. Gravity waves are any waves, such as ocean waves, in which gravity is the restoring force.

measurements of gravitational waves whose characteristic time scales are quite long, from seconds to days. Since the long baselines make propagation losses significant, the ends of the interferometer arms must have coherent laser transponders instead of retroreflectors, and laser phase (frequency) fluctuations must be controlled or calibrated very accurately. At measurement frequencies between about 1 mHz and 1 Hz, sensitivity is likely to be dominated by photon statistics (shot noise). At very low frequencies, below 1 mHz or so, sensitivity is degraded by low-frequency external perturbations to the positions of the test masses inside the spacecraft. Such perturbations include residual gas pressure inside the chambers containing the test masses; thermal gradients due, for example, to fluctuations in solar intensity; and changing gravitational attraction between spacecraft caused by fuel usage and spacecraft motion (Stebbins, *et al.* 1989).

5.2.1 What are gravitational waves?

Gravitational waves are "ripples in the curvature of space-time," caused by coherent bulk motions of matter (such as the collapse of stellar cores) or by coherent vibrations of space-time curvature (a subtle concept, but an accurate description of what is expected to happen around a black hole).* General relativity predicts their existence and gives them two fundamental characteristics: first, they travel with the speed of light; second, they are "transverse and traceless." The latter characteristic means that they produce a quadrupolar, divergence-free force field, with two polarizations (see Figure 15). In the language of quantum field theory, which says that the energy in a radiation field must be quantized, this means that the quanta of the gravitational field ("gravitons") have zero rest mass and charge and are bosons with spin 2; because of their zero rest mass, their spins must be parallel to their motion (as for photons), giving them only two independent spin states, or polarizations.

Gravitational waves are not precisely defined quantities, since the nonlinear character of the

* For a more complete tutorial on this subject, the reader is referred to a number of excellent references: Thorne 1987; Deruelle and Piran 1983; Misner, *et al.* 1973.

gravitational field means that the contributions of gravitational waves to space-time curvature cannot be separated from the "background" contributions of everything else (the solar system, Galaxy, and beyond). However, in all astrophysical scenarios for their production, the length scales over which the waves vary in amplitude are much shorter than the length scales over which background contributions to curvature vary; gravitational waves truly can be thought of as "ripples." This famous "shortwave approximation" (Misner, *et al.* 1973) makes it possible to decouple the gravitational waves from the background and treat them as independent entities.

The force exerted by a gravitational wave on a mass m is related to its position x (displacement from a specified origin) by the fourth-rank Riemann curvature tensor $R_{\alpha\beta\gamma\delta}$ (Misner, *et al.* 1973). In general, each of the subscripts $(\alpha, \beta, \gamma, \delta)$ can take on any of four values (0,1,2,3), representing the three components of space (x,y,z) and one of time (t). The transverse traceless nature of gravitational waves means that the only nonzero components of the Riemann curvature tensor are $R_{\alpha 0 \gamma 0}$; hence a second-rank tensor is adequate to describe the force. Since the Riemann tensor has dimensions of (1/time)^2 or (1/length)^2, it is appropriate to define a dimensionless second-rank tensor h_{jk} by

$$\frac{\partial^2 h_{jk}}{\partial t^2} \equiv -2R_{j0k0}{}^{\rm GW} \ . \tag{5.2.1a}$$

The jth spatial component of the force $F_j{}^{\rm GW}$ induced on a mass m by a passing gravitational wave is related to the kth component of the mass's position by

$$F_j{}^{\rm GW} = -m\, R_{j0k0}{}^{\rm GW}\, x^k = \frac{m}{2}\, \frac{\partial^2 h_{jk}}{\partial t^2}\, x^k \ . \tag{5.2.1b}$$

Since the resulting changes δx^j in the mass's position are so tiny compared to the mass's distance from the origin, the latter can be regarded as constant. Hence the displacement due to a gravitational wave is related linearly to the components of its original position by the "strain" tensor h_{jk}:

$$\delta x^j = \frac{1}{2}\, h_{jk}\, x^k \ , \tag{5.2.1c}$$

and the force can be written as

$$F_j^{\rm GW} = m \frac{\partial^2(\delta x^j)}{\partial t^2} . \qquad (5.2.1d)$$

If a gravitational wave is travelling in the z-direction, its *transverse* nature guarantees that the only nonzero components of the wave field, or strain tensor, are the xx, yy, xy, and yx components. The *traceless* nature of the wave reduces these four components to two independent components (the two polarizations referred to above, and shown in Figure 15): $h_{xy} = h_{yx}$, and $h_{xx} = -h_{yy}$. These components of the strain tensor are sometimes referred to as the gravitational-wave amplitudes, since the energy carried by the waves is proportional to their square. If the arms of a dual-arm interferometer are oriented at right angles in a plane perpendicular to the direction of travel of a gravitational wave, and the arms are parallel to the lines of force (the x and y axes in Figure 15a), the gravitational wave will induce displacements of equal magnitude but opposite sign in the two arms. The difference in fractional displacement of the two interferometer arms, $\delta\ell/L$ in the language used in previous chapters of this report (ℓ is the difference in the arm lengths), is thus a direct measurement of the magnitude of the gravitational-wave amplitude h; *i.e.*, $\delta\ell/L \simeq h$ (in this example, $h = h_{xx}$).

5.2.2. Where do they come from?

All expected sources of gravitational waves strong enough to be detected by an observer in the solar system are of astrophysical origin. Summaries of the variety of sources are given in many papers (*e.g.*, Press and Thorne 1972; Thorne 1987). Current understanding of the strength of the waves produced by various events and the frequency with which they might occur (hence the distance to the nearest one) is very limited, and predictions typically carry uncertainties of several orders of magnitude. Nevertheless, semiquantitative predictions can be made, and these are crucial to the effort to detect gravitational waves. The waves produced can be classified according to their temporal characteristics: *Periodic* waves, which are superpositions of sinusoids whose frequencies are well defined and essentially constant over an observing time; *Bursts*, or pulses, which last for

very short times, on the order of milliseconds or less; and *Stochastic* waves, which last a long time and fluctuate stochastically. These types and their possible sources are discussed below.

5.2.2a Periodic sources of gravitational waves

Periodic gravitational waves are characterized by a well-defined set of discrete frequencies and a single polarization. For example, the amplitude might be

$$h(t) = h_0 \cos 2\pi\nu t \ , \tag{5.2.2}$$

for frequency ν. The expected strengths h for these sources are summarized in Figure 16a, together with the kinds of detection schemes that might be used to measure them. Figure 16b is a more detailed summary of low-frequency periodic sources (binary systems) and their strengths; these waves would be detectable with sufficiently sensitive space-based long-baseline (10^7-km) interferometers.

Pulsars, or rotating neutron stars, are a possible source of periodic waves with frequencies in the 10- to 10^4-Hz regime accessible to ground-based laser-interferometric detectors. Gravitational waves will be emitted if there are deviations from symmetry around the rotation axis. Such deviations might exist because of deformations in the star's solid crust (or core, if that is solid) produced in "star quakes," or in young neutron stars (younger, say, than a few hundred thousand years) where accretion has not yet smoothed the stars in the process of causing them to spin up to rapid rotation rates (Pandharipande, *et al.* 1976; Schumaker and Thorne 1983). Deviations from symmetry also could arise from magnetic pressure, if the internal magnetic fields are strong enough. However, estimates indicate necessary field strengths to be stronger than observed. For example, for the Crab pulsar, the internal field would have to be about 10 times as strong as the measured surface field (Thorne 1987; Zimmermann 1978). Pulsars rotating more rapidly than a critical rate (which corresponds typically to 1- to 2-msec periods) might develop a "CFS" instability (Chandrasekhar 1970; Friedman and Schutz 1978) that would produce strong hydrodynamic waves in the star's

surface layers, which would radiate gravitational waves. Pulsars that are strong X-ray emitters are good candidates for exhibiting this kind of instability; the expected gravitational-wave amplitude scales with the square root of the X-ray flux. All estimates of their amplitudes give a value of h smaller than about 10^{-24}, which puts them beyond the capability of existing detectors. But they may be observable with future laser-interferometric detectors — perhaps on the moon, where seismic noise is minimal, excellent vacuums can be maintained, and baselines longer than 100 kilometers may be possible. These sources are in the bottom right-hand corner of Figure 16a.

Binary stars are potentially much stronger sources of periodic gravitational waves than are pulsars. Characteristics of the waves that would be emitted from binaries are well understood, for they depend only on the mass and orbital parameters of the system. The orbital periods typically are several hours or longer, and the frequency associated with the gravitational waves emitted is twice the orbital frequency. Such millihertz-frequency waves require long-baseline space-based laser interferometers for their detection. For a binary system with reduced mass $\mu \equiv M_1 M_2/M$ and total mass $M \equiv M_1 + M_2$, rotating with a frequency ν_b, and at a distance r from Earth, the expected gravitational-wave amplitude is (Eardley 1983)

$$h \simeq 8.7 \times 10^{-21} \left(\frac{\mu}{M_\odot}\right) \left(\frac{M}{M_\odot}\right)^{2/3} \left(\frac{100\ \mathrm{pc}}{r}\right) \left(\frac{\nu_b}{1\ \mathrm{mHz}}\right)^{2/3} \quad \text{(binaries)}\ , \tag{5.2.3}$$

where $M_\odot$ is the solar mass and 1 pc $\simeq 3\times 10^{13}$ km. (Note that the *energy* carried by the gravitational waves would have the expected $1/r^2$ dependence.) The shortest-period binary system known consists of a white dwarf and a neutron star, and has an orbital period of 11 minutes, hence a gravitational-wave frequency of about 3 mHz (Priedhorsky, *et al.* 1986). Double white-dwarf and double neutron-star binaries also are good candidates (see Figure 16b and Hils, *et al.* 1990).

5.2.2b Burst sources of gravitational waves

Bursts, or pulses, of gravitational waves are superpositions of sinusoids (with complicated time dependences) that last only for a few cycles, or a total time shorter than typical observation times, which are on the order of milliseconds for ground-based detectors and several minutes to hours for space-based detectors. (See Figure 16c for a summary of expected strengths and frequencies of burst sources, and Figure 16d for a more detailed look at expected burst sources at frequencies below 1 Hz, as would be appropriate for space-based detectors.) The amplitude at Earth of a gravitational wave emitted in a burst is related to the total energy radiated in the form of gravitational waves, $E_{\rm GW}$, their characteristic frequency $\nu_{\rm p}$ [pulse duration $\tau_{\rm p} \equiv (2\pi\nu_{\rm p})^{-1}$], and the distance r from Earth by (Thorne 1987)

$$h \simeq 2.7 \times 10^{-20} \left(\frac{E_{\rm GW}}{M_\odot c^2} \frac{1\ {\rm kHz}}{\nu_p}\right)^{1/2} \left(\frac{10\ {\rm Mpc}}{r}\right) \quad \text{(bursts)} . \tag{5.2.4}$$

The detection of bursts is not as straightforward as for periodic waves. If the time dependence of the wave amplitude $h(t)$ can be anticipated, an appropriate Wiener filter is used with the detector. The Fourier transform of a Wiener filter is the ratio of the Fourier transform of the expected signal to the detector noise spectral density; a Wiener filter thus suppresses those frequencies in the signal that are associated with high levels of detector noise. The observed output signal is then deconvolved with the Wiener filter and the expected noise contribution subtracted off. It has been shown that with this kind of detection scheme, the gravitational-wave amplitude h required to conclude at the 90%-confidence level that a gravitational wave has been seen in three different measurements (3 times per year) by two identical detectors is 3 to 5 times larger than the effective strain due to noise (Thorne 1987).

Type II supernovae, produced by the collapse of the cores of massive, old stars to form neutron stars, are excellent candidates for strong bursts of gravitational waves. Type I supernovae also may

be good candidates; their generation is not well understood, but is thought to be associated with the nuclear explosion of mass-accreting white dwarfs or the collapse of white dwarfs to form neutron stars. In the latter case, little optical radiation would result, but gravitational waves should be emitted (Thorne 1987). The gravitational-wave strength from a supernova depends on the degree of nonsphericity involved in the collapse; no gravitational radiation would be emitted in a perfectly spherical collapse. Type II supernovae may be highly spherically symmetric, although very little factual evidence exists about them. Type I supernovae, if produced as currently thought, are more likely to deviate considerably from spherical symmetry. In the Milky Way Galaxy, both types of supernovae occur *approximately* once every 40 years; out to distances of 10 Mpc (the Virgo cluster of galaxies), they may occur as often as several times per year (Tammann 1981). The collapse of white dwarfs to neutron stars — if such events really occur — could be as frequent as four times per year in our Galaxy. Most calculations of the collapse to neutron stars predict strengths on the order of $h \sim 10^{-23}(10 \text{ Mpc}/r)$ in two primary frequency regimes: $\sim$ 1 kHz, caused by the initial collapse and "bounce," and $\sim$ 10 kHz, caused by pulsations of the newly formed neutron star (see, *e.g.*, Eardley 1983 or Müller 1984). If the neutron star rotates in an end-over-end mode, the strength of emitted gravitational waves might be as much as 50 to 500 times stronger (Ipser and Managan 1984; Eardley 1983).

Stars or star clusters also might collapse to form black holes, emitting significant amounts of gravitational radiation. The characteristic frequency of the waves scales inversely with the mass M of the black hole:

$$\nu_{\text{BH}} \simeq \frac{c^3}{5\pi GM} \simeq 1.3 \text{ kHz} \left(\frac{10\ M_\odot}{M}\right) \tag{5.2.5a}$$

(Thorne 1987). Masses of black holes are expected to range from 2 $M_\odot$ to 10^{10} $M_\odot$. Collapse of stars to form black holes of a few solar masses is expected to occur at the rate of about one per year out to distances on the order of 10 Mpc. Holes as massive as $\sim 10^6$ $M_\odot$, which probably form only in galactic nuclei, may form at a rate of only a few per year throughout the entire observable universe,

i.e., out to the Hubble distance (several thousand megaparsecs) (Rees 1983). If gravitational waves carry away an energy equal to a fraction ϵ of the final black hole mass M, their amplitude at a distance r from Earth will be on the order of

$$h_{\rm BH} \simeq \left(\frac{15}{2\pi}\epsilon\right)^{1/2} \frac{M}{r} \simeq 7 \times 10^{-21} \left(\frac{\epsilon}{0.01}\right)^{1/2} \left(\frac{M}{10\ M_\odot}\right) \left(\frac{10\ {\rm Mpc}}{r}\right) \qquad (5.2.5b)$$

Potentially very strong (relatively speaking!) sources of gravitational-wave bursts are compact binary systems composed of neutron stars or black holes, which are close enough that they may spiral in and coalesce because of "radiation reaction," the loss of energy removed from the system in the form of gravitational waves. The binary pulsar PSR 1913+16 is such a system (Taylor and Weisberg 1982); it should coalesce $\sim 3.5 \times 10^8$ years from now. The frequency of the gravitational wave, twice the orbital frequency, increases with time until reaching its maximum during the final coalescence. This maximum frequency is expected to be on the order of 1 kHz for neutron stars (masses on the order of one solar mass); for black holes, the frequency may be lower, scaling inversely with the mass of the primary (more massive) black hole (Thorne 1987). Generally, one would observe at lower frequencies, *e.g.*, 100 Hz, since the binaries spend more time at the lower orbital frequencies and because Earth-based receivers tend to be less noisy at 100 Hz than at kilohertz frequencies. The characteristic amplitude of the gravitational waves from such compact binary systems located a distance r from Earth is

$$h \sim 4 \times 10^{-22} \left(\frac{\mu}{M_\odot}\right)^{1/2} \left(\frac{M}{M_\odot}\right)^{1/3} \left(\frac{100\ {\rm Mpc}}{r}\right) \left(\frac{100\ {\rm Hz}}{\nu}\right)^{1/6} . \qquad (5.2.6)$$

Low-frequency gravitational-wave bursts could be emitted when a star or a small black hole passes near a supermassive black hole ($\sim 10^8 M_\odot$). The characteristic frequencies and amplitudes of

the waves for non-head-on collisions are (Kojima and Nakamura 1984)

$$\nu \sim \frac{1}{20M_1} = 10^{-4}\ \mathrm{Hz} \left(\frac{10^8 M_\odot}{M_1}\right) \tag{5.2.7a}$$

and

$$h \sim \frac{M_2}{2r} = 2 \times 10^{-21} \left(\frac{M_2}{M_\odot}\right) \left(\frac{10\ \mathrm{Mpc}}{r}\right), \tag{5.2.7b}$$

where M_1 is the mass of the massive black hole, and M_2 is the mass of the infalling star or small black hole. Within 10 Mpc there are of order 100 galaxies of comparable or greater mass than our own ($\sim 10^{11}\ M_\odot$), including M87, which may contain a central black hole of $\sim 4 \times 10^9 M_\odot$ (Sargent, *et al.* 1978; Young, *et al.* 1978; Richstone 1988). It is thought that these events could happen as frequently as once per year (Thorne 1987).

5.2.2c Stochastic background of gravitational waves

Figure 16e (from Thorne 1987) shows expected strain amplitudes and characteristic frequencies for various stochastic sources of gravitational waves. The most likely source of a stochastic background of gravitational waves is the superposition of the waves from *galactic* binaries, and, to a lesser extent (by a factor of about 6) from *extragalactic* binary systems. This background would exist predominantly at frequencies below a few tens of millihertz (Hils, *et al.* 1990), and it may be an obstacle to observation of low-frequency gravitational waves from other sources. A broadband periodic source that produces a strain h will be observable over the background over an integration time τ only if $h > (2\pi\nu\tau)^{-1/2}\ h_{\mathrm{background}}$. At 0.1 mHz, the combined contribution from binaries to the stochastic background might be on the order of $h_{\mathrm{background}} \sim 10^{-18}$ (Figure 16e). This means, for example, that integration times of order several months ($\tau \gtrsim 10^7$ sec) will be required to detect (unambiguously) gravitational waves from the white-dwarf binaries i Boo and SS Cyg ($h \sim 10^{-20}$; see Figure 16a).

Other possible sources of stochastic background include: (1) leftover gravitational radiation from the demise of postulated pregalactic massive stars ("Population III" stars), either as supernovae or during collapse to black holes or black-hole binaries (Rees 1983; Carr 1986); (2) primordial gravitational waves, produced during the first $\sim 10^{-43}$ (!) seconds [the Planck time, $(\hbar G/c^5)^{1/2}$] after the big bang when space-time became quantized and gravitons last scattered off matter (Thorne 1987); (3) very-low-frequency gravitational waves (10^{-7}–10^{-5} Hz) produced during phase transitions in the early expansion of the universe (Hogan 1986); and (4) vibrations of "cosmic strings" — one-dimensional "defects" in the vacuum produced during a very early phase transition associated with postulated grand-unified interactions (Thorne 1987).

5.2.3 What are our chances of detecting them?

The above description of gravitational-wave sources indicates that sensitivities $\delta\ell/L$ should be at least as good as 10^{-21} to have reasonable prospects of definitively observing several gravitational-wave events per year. The primary technique for gravitational-wave detection being developed at this time is laser interferometry, with baselines up to a few kilometers on the ground and up to 10 million kilometers in space. The discussions in chapter 3 and section 4.1 of interferometric measurement techniques and associated errors are relevant to gravitational-wave detection. Dual-arm interferometers would be used to suppress (by cancellation) errors that are correlated in the two arms, such as laser frequency fluctuations and, to a limited extent, refractive-index fluctuations. To improve sensitivity in the face of errors that are uncorrelated in the two arms and are independent of arm length ("fixed-end effects"), the optical paths in the interferometer arms would be increased in order to increase the optical phase shift produced by a given gravitational-wave-induced strain. This could be accomplished by building physically larger interferometers, by operating the interferometer arms as optical delay lines or optical cavities, and by circulating the light back and forth

between arms before recombining the beams (resonant recycling). Mechanical and thermal noise leads to stringent requirements on seismic and acoustic isolation of test masses and to the need for sophisticated disturbance-compensation systems in space. Radiation-pressure fluctuations from the laser light itself (section 3.4.6) are not a significant source of measurement error at typical laser power levels, and fluctuations in laser power or nonuniformities in the intensity distribution across the beam could be made insignificant with intensity-feedback stabilization and mode-cleaning techniques (section 3.4.3). Photon statistics, one of the most important and fundamental error sources on the ground and in space, will be reduced as laser sources of higher power (and requisite frequency stability) become available. In the meantime, or in conjunction with more powerful lasers, techniques such as power recycling of the laser light are valuable (section 4.1.4), as may be the use of squeezed vacuum light (section 3.4.1). For the very long baselines used in space, coherent laser transponders must be used instead of retroreflectors to drive down the photon-statistics error. These error sources and others have been discussed in the previous chapters, along with ways to minimize them. This section summarizes the limiting sensitivities for laser-interferometric gravitational-wave detection as determined by these error sources.

It is useful to ask what the magnitudes of various system parameters must be in order to reach measurement sensitivities $\delta\ell/L$ of order 10^{-21} or 10^{-22} with laser interferometers, since the discussion in section 5.2.2 indicated that such sensitivities would be required in order to have a reasonable expectation of detecting gravitational waves from the kinds of sources believed to exist. Begin by considering ground-based interferometric detectors designed for detection of gravitational waves whose characteristic times (periods or pulse durations) are on the order of 2 milliseconds, say. For maximum signal (strain produced by a passing gravitational wave), the light-storage time in each arm of the interferometer should be on the order of 1 millisecond, or half the period of the gravitational wave. For arms 1 kilometer in length operated as optical delay lines, for example, this would require that the light make $n = 150$ round trips in each arm before being recombined with light

from the other arm ($2nL/c = 1$ ms). The photon-statistics error, described by equations (3.4.2) or (3.4.15b), indicates that for a laser operating at 1-μm wavelength, to reach 10^{-21} strain sensitivity would require an effective power $\eta P_0 \geq 60$ W, equal to the laser output power P_0 multiplied by all appropriate efficiency factors that account for imperfect optics, mirror losses, detector quantum efficiencies, and other losses. The sensitivity would scale for other values of system parameters in the following way:

$$\frac{\delta\ell_{\rm ph}}{L} \simeq 10^{-21} \left(\frac{150}{n}\right) \left(\frac{1\ {\rm km}}{L}\right) \left(\frac{\lambda}{1\ \mu{\rm m}} \frac{1\ {\rm ms}}{\tau} \frac{60\ {\rm W}}{\eta P_0}\right)^{1/2} . \tag{5.2.8a}$$

(The use of classical, unsqueezed light at the input ports of the beam splitter has been assumed.) This sensitivity is very close to the standard quantum limit for position measurements on 10-kg test masses in 1-ms integration times [eq. (3.4.18a)]:

$$\frac{\delta\ell_{\rm SQL}}{L} \simeq \sqrt{2}\,\frac{\delta L_{\rm SQL}}{L} = \left(\frac{2\hbar\tau}{m}\right)^{1/2} \frac{1}{L} \simeq 1.4\times 10^{-22} \left(\frac{1\ {\rm km}}{L}\right) \left(\frac{\tau}{1\ {\rm ms}} \frac{10\ {\rm kg}}{m}\right)^{1/2} . \tag{5.2.8b}$$

It is also comparable to the limiting sensitivity due to pendulum oscillation modes of the test masses [eq. (3.4.12b)] at measurement frequencies ν_m on the order of 1 kHz:

$$\begin{aligned} \frac{\delta\ell_{\rm pend}}{L} \simeq \sqrt{2}\frac{\delta L_{\rm pend}}{L} &\simeq 10^{-22} \left(\frac{1\ {\rm km}}{L}\right) \left(\frac{1\ {\rm kHz}}{\nu_m}\right) \left(\frac{B}{\nu_m} \frac{\nu_0}{1\ {\rm Hz}}\right)^{1/2} \Pi\ , \\ \Pi &\equiv \left(\frac{10^6}{Q} \frac{T}{300\ {}^\circ{\rm K}} \frac{10\ {\rm kg}}{m}\right)^{1/2} . \end{aligned} \tag{5.2.8c}$$

Distinguishability of gravitational waves from vibrational noise would require test masses to have resonant frequencies as high as several kilohertz or more, and quality factors (Q) as high as 10^6 [eq. (3.4.12a)]:

$$\frac{\delta\ell_{\rm vib}}{L} \simeq \sqrt{2}\,\frac{\delta L_{\rm vib}}{L} \simeq 7.4\times 10^{-22} \left(\frac{1\ {\rm km}}{L}\right) \left(\frac{5\ {\rm kHz}}{\nu_0}\right) \left(\frac{B}{\nu_0}\right)^{1/2} \Pi\ . \tag{5.2.8d}$$

Techniques such as the laser phase modulation decribed in section 3.4.2 might be used to suppress the scattered light until its amplitude is down to 10^{-5} that of the main beam. Presuming that relative lengths of the two interferometer arms also could be monitored to this level (1 cm for a 1-km baseline), calibration or control of laser fractional frequency fluctuations to a level of 10^{-16} still would be required to reach a photon-statistics-limited sensitivity of 10^{-21}; at 1-μm laser wavelength, this means $\delta\nu \lesssim 20$ mHz:

$$\frac{\delta\nu}{\nu} \lesssim \frac{L}{\sqrt{2}\epsilon_{sc}\Delta L_{sc}}\left(\frac{\delta\ell_{ph}}{L}\right) \sim 10^{-16}\left(\frac{10^{-5}}{\epsilon_{sc}}\right)\left(\frac{L}{\sqrt{2}\Delta L_{sc}}\right)\left(\frac{\delta\ell_{ph}}{L}\cdot 10^{21}\right). \qquad (5.2.8e)$$

Even with the best available materials and techniques for isolation, ground noise will be an appreciable source of measurement error at 10^{-21} sensitivity for measurement frequencies on the order of 1 kHz with a 1-km baseline (see section 3.4.5). Refractive-index fluctuations due to residual gas in the interferometer arms also might begin to be troublesome at this level of sensitivity. Perhaps the only potential error source that would *not* pose a challenge to achieving this sensitivity, given the parameters necessary to meet the photon-statistics limit of eq. (5.2.8a), is radiation-pressure fluctuations [eq. (3.4.15a)]:

$$\frac{\delta\ell_{rp}}{L} \simeq 5\times 10^{-26}\left(\frac{1\text{ km}}{L}\right)\left(\frac{10\text{ kg}}{m}\right)\left(\frac{1\ \mu\text{m}}{\lambda}\,\frac{\eta P_0}{60\text{ W}}\right)^{1/2}\left(\frac{\tau}{1\text{ ms}}\right)^{3/2}. \qquad (5.2.8f)$$

Next, consider a long-baseline ($\sim 10^7$ km) space-based laser interferometer formed among three (or more) spacecraft in solar orbit, of the sort currently under study for the detection of low-frequency (< 1 Hz) gravitational waves (Bender 1980; Faller and Bender 1984; Faller, *et al.* 1984; Stebbins, *et al.* 1989). Such an interferometer is depicted in Figure 17; Figures 16b and 16d show the expected spectral sensitivity together with the expected strengths of periodic and burst sources of gravitational waves, respectively. Such a system is designed to measure gravitational waves with periods ranging from 1 second to 10^5 seconds ($\sim$ 1 day) by measuring relative optical phase shifts between laser light propagating down and back in each arm. Diode-pumped solid-state

lasers, such as Nd:YAG, are excellent candidates for use in such systems because they can operate under frequency-stabilized conditions with high output powers (on the order of 1 watt), and exhibit desirable characteristics for space qualification (robust, compact, and long-lived). A space-based detector has three primary advantages over a ground-based detector: (1) it is free from Earth's gravity gradients and mechanical noise and so can operate with high sensitivity at frequencies below 1–10 Hz, *i.e.*, all the way down to about 10^{-5} Hz; (2) it can have a very long baseline; and (3) the test masses at the ends of the interferometer arms can be supported in a more nearly inertial manner. However, over such great distances, propagation losses can become important, thus increasing the photon-statistics error unless laser powers are increased accordingly. Also, at frequencies below about 10 mHz, the space environment is far from quiet; significant spurious accelerations between the test masses arise from such causes as residual gas pressure in the vacuum chambers housing the test masses, cosmic-ray impacts, and thermal gradients arising from fluctuations in the solar radiation. (These are discussed further below.)

Consider first the effects of photon statistics on measurement sensitivity. Because propagation losses can be so severe over the long baselines associated with a space-based instrument, two cases must be distinguished: one in which coherent laser transponders are used at the ends of the interferometer to receive, amplify, and retransmit the received laser signal, and the other in which retroreflectors are used (see subsection 4.1.4a). The limiting strain sensitivity $\delta\ell_{\rm ph}/L$ due to photon statistics for these two cases can be found from eqs. (4.1.8). Assume for illustration here that the arm lengths L are on the order of 10^7 km, the laser wavelength is 1 μm, and that measurement integration times τ of $\sim 10^4$ s ($\sim$ 3 hours) are used. Also assume that the telescopes are diffraction-limited, *i.e.*, that $s_A = s_B = 1$ in eqs. (4.1.6) (this was not assumed for the example of planet gravity-mapping because the latter application is by nature less demanding technologically and more feasible in the near future). With retroreflectors, just to reach a sensitivity of 10^{-20} would require 0.5-m telescopes

and 1.6 kW of average laser output power (assuming an overall efficiency factor $\eta^{rr} \simeq 0.02$):

$$\frac{\delta\ell_{ph}{}^{rr}}{L} \simeq 10^{-20}\left(\frac{s_a s_b}{1}\right)\left(\frac{\lambda}{1\ \mu m}\right)^{5/2}\left(\frac{10^4\ s}{\tau}\right)^{1/2}\left(\frac{L}{10^7\ km}\right)\left(\frac{0.5\ m}{D_A}\ \frac{0.5\ m}{D_B}\right)^2\left(\frac{1.6\ kW}{P_t}\cdot\frac{0.02}{\eta^{rr}}\right)^{1/2}. \tag{5.2.9a}$$

In contrast, smaller telescopes ($D_A \simeq D_B \simeq 25$ cm) and much smaller transponded power P_t (on the order of 16 W for overall efficiency $\eta^{tr} \simeq 0.02$) are adequate to give a photon-statistics-limited strain sensitivity of 10^{-23}:

$$\frac{\delta\ell_{ph}{}^{tr}}{L} \simeq 10^{-23}\left(\frac{s_B}{1}\right)\left(\frac{\lambda}{1\ \mu m}\right)^{3/2}\left(\frac{10^4\ s}{\tau}\right)^{1/2}\left(\frac{0.25\ m}{D_A}\ \frac{0.25\ m}{D_B}\right)\left(\frac{16\ W}{P_0}\cdot\frac{0.02}{\eta^{tr}}\right)^{1/2}. \tag{5.2.9b}$$

Note that because of L^{-2} propagation losses in each direction, the photon-statistics-limited strain sensitivity with transponders is independent of baseline length, whereas the sensitivity with retroreflectors actually decreases with increasing baseline.

Uncontrolled or uncalibrated laser frequency fluctuations will also introduce error into the measurements. To achieve a measurement sensitivity of 10^{-21} with a 10^7-km baseline, displacements must be measured at the 10-picometer level, which requires measuring the phase of the laser light to approximately 10^{-5} cycles. This probably would be accomplished by down-shifting the Doppler signals, narrow-band filtering, and using optimized digital filters for sampling and measuring the phase (more sophisticated techniques than those described in section 4.2.1). Even if the arm lengths can be held equal to 10^{-4}, enabling some suppression of laser frequency fluctuations correlated in the two arms, calibration of the laser frequency fluctuations to 10^{-17}, or about 3 mHz, still would be necessary over times as long as hours. A technique such as that described in section 4.1.3 should be used, in which the length of one arm is held constant or monitored very precisely. This length is used as a reference to calibrate the effects of laser phase (or frequency) fluctuations and so distinguish them from the effects produced by a passing gravitational wave. Successful use of such a calibration technique could ease the requirement on intrinsic (controlled) fractional frequency stability of the

laser, possibly bringing it down to to 10^{-12} or so, *i.e.*, to well within the reach of current technology for a space-qualifiable, long-lived ($\geq$ 10 years) laser.

Some other sources of measurement error or causes of change in the relative optical paths that would be difficult to distinguish from changes induced by a gravitational wave include the following: thermal expansion of the telescopes; "beam walk" across irregularities in optical surfaces; "aperture walk" across nonuniformities in the laser beam (instabilities in laser power and beam geometry such as those discussed in section 3.4.3); mispointing of the telescopes or misalignments of optical elements; and refractive-index changes in the interplanetary medium. The contribution to the last source from fluctuations in electron density could be calibrated with radio links between the spacecraft (see section 3.4.4). These error sources must also be addressed in precision astrometric and imaging instruments, although the tolerances for a gravitational-wave laser interferometer are more demanding.

At the sensitivity levels required for gravitational-wave detection, particularly at measurement frequencies below about 1 mHz, some additional error sources become significant, whose suppression is very difficult. These are fluctuating displacements of the test masses due to stochastic external accelerations. They arise from many different phenomena: residual gas pressure around the test masses; cosmic-ray impacts; thermal gradients and fluctuating thermal radiation pressure due to fluctuations in the solar flux; fluctuations in the solar-wind pressure; fluctuating gravitational attraction between the test masses and the surrounding spacecraft (caused, for example, by fuel usage and motion or distortions induced in the spacecraft); electrical charging-up of the masses due to cosmic-ray impacts with resulting acceleration from electric fields in the cavities; nonuniform outgassing from the spacecraft; and fluctuations in the interplanetary magnetic field (diamagnetism of the test masses). For quantitative discussion of these see Stebbins, *et al.* (1989); a brief discussion is given below.

These spurious accelerations must be suppressed to less than about $10^{-18}g$ ($g \simeq 9.8$ m/s^2) in a 10^{-4}–10^{-3} Hz bandwidth from nominal magnitudes that may be on the order of $10^{-13}g$, in order to allow detection of gravitational waves. Such suppression will require state-of-the-art materials technology and design for thermal insulation and mechanical stability and the use of active disturbance-compensation systems (DISCOS, or drag-free systems) that outperform by several orders of magnitude any such systems built to date (Stebbins, *et al.* 1989; see also section 3.4.5). Although studies to date indicate that the desired drag-free performance is achievable, engineering design and demonstration of the required drag-free system are major parts of the challenge of developing a space-based interferometric gravitational-wave detector.

Figures 16b and 16d show the expected spectral density of gravitational-wave amplitudes from low-frequency periodic sources (binaries) and burst sources, as might be observed with a space-based interferometer like the one considered here. Also shown in those figures is the expected sensitivity of such an interferometer. (These figures assume 10^6-km arm lengths rather than the 10^7-km values assumed here and in recent studies of such a system, but the spectral sensitivity curves would be similar qualitatively.) Sensitivity decreases above about 0.01 to 0.1 Hz because the arm length becomes longer than half the wavelength of the expected gravitational waves. At frequencies between 1 mHz and 0.01 to 0.1 Hz, photon statistics are the limiting error source. At lower frequencies, 10^{-5} to 10^{-3} Hz, the dominant error source is the random impacts on the test masses of gas molecules, due to imperfect vacuums in the cavities surrounding the masses. It has been estimated that if the residual gas density can be reduced to an internal pressure of a few times 10^{-11} torr, and arm lengths are 10^7 km, the sensitivity of the interferometer as limited by this noise source would be 10^{-22} at $\nu = 10^{-4}$ Hz, and would scale roughly as $\nu^{-3/2}$ (Bender, private communication).

At still lower frequencies, around 10^{-5} Hz and lower, the dominant error source is expected to be fluctuations in the net thermal radiation pressure on the test masses due to fluctuations in the solar intensity (Stebbins, *et al.* 1989). From observational data on solar intensity fluctuations

together with reasonable assumptions about the degree of thermal shielding achievable for the cavities containing the test masses, it is estimated that this noise source will limit sensitivity to about 10^{-19} at frequencies of 10^{-5} Hz, and should rise steeply with decreasing frequency, as $\nu^{-16/3}$ (Bender, *et al.* 1988; Stebbins, *et al.* 1989).

A potentially dominant white-noise source of spurious accelerations of the test masses is impacts from cosmic rays (protons) with energies up to a few GeV. The effect of impacts from *galactic* cosmic rays is estimated to be roughly 30 times smaller than the effect of residual gas pressure described above. However, during solar flares, the flux of *solar* cosmic rays could rise to several orders of magnitude higher than the flux of galactic cosmic rays; while the fraction of time when this might occur is likely to be very small, the antenna performance would be seriously compromised during that short time.

5.2.4 Why should we try to detect them?

(After reading the preceding subsections, how can the reader ask this? Or perhaps, after reading about all the difficulties involved, how could he or she possibly not ask this in a serious way?)

The answer to this question could be long and detailed, but instead let it be short: the predictions and implications of general relativity and other theories of gravitation impact our understanding of the universe in profound ways. On cosmological scales, they have ramifications for the structure, evolution, and destiny of the universe. On microscropic scales, they are tied intimately to the fundamental laws of physics and to the (unified?) field theories that explain them. Suffice it to describe here just a few of the significant impacts that would come from the regular (several times per year, say) detection of gravitational waves.

General relativity predicts that gravitational waves, like electromagnetic waves, should travel at

the speed of light, c. Alternative theories predict a different speed for gravitational waves, typically a higher speed by a factor of 10^{-6} or more (Thorne 1974). This could be verified with simultaneous optical and gravitational-wave observations of supernovae. The difference in propagation speed $\Delta c/c$ would lead to a delay Δt in the arrival of the initial light outburst relative to the initial gravitational-wave burst. For a supernova occurring at a distance r from Earth, this delay would be

$$\Delta t \approx \frac{r}{c}\frac{\Delta c}{c} \approx 2 \text{ weeks} \left(\frac{r}{10 \text{ Mpc}}\right)\left(\frac{\Delta c}{c}\, 10^{9}\right) . \tag{5.2.10}$$

Thus, a deviation of 10^{-9} from c for the propagation speed of gravitational waves would show up as a time delay of approximately two weeks between the optical and gravitational-wave bursts from supernova explosions in the Virgo cluster of galaxies (roughly 10 Mpc from Earth). If $\Delta c/c$ were several orders of magnitude larger, say on the order of 10^{-6}, the time delay for supernovae in the Virgo cluster would be $\sim$ 30 years, which would make it difficult to relate detected gravitational waves unambiguously to a particular optical supernova. In that case, the test would be better made with a supernova occurring near the center of the Milky Way Galaxy ($r \simeq 10$ kpc); there a deviation $\Delta c/c$ as large as 10^{-6} would produce a time delay of roughly two weeks, while a deviation $\Delta c/c \simeq 10^{-9}$ would produce a delay of about 20 minutes.

In any metric theory of gravity*, a gravitational wave can have at most six independent polarization states (Eardley 1983). Some metric theories contain additional constraints that reduce this number: in general relativity, the number is two; in the Dicke-Brans-Jordan theory (Brans and Dicke 1961), it is three. Both of these metric theories predict that the polarization state of a gravitational wave is an invariant: different observers undergoing different motions while observing a gravitational wave will measure the same polarizations. This is not the case in general for other metric theories.

* Metric theories are based on a mathematical representation built around a four-dimensional space-time metric tensor and obey all the laws of special relativity in their local Lorentz frames (the Einstein equivalence principle). Virtually all viable theories of gravitation are metric.

With regular observations of gravitational waves, enough statistical information on the polarizations of gravitational waves could be gathered to support or disprove the prediction of general relativity that only two possible polarizations exist.

Observation of the gravitational waves from known astrophysical sources could provide a great deal of information about the sources. One well-studied example is rotating neutron stars, or pulsars. Observational data on the amplitudes and relative phases of the spectral components of the gravitational waves would provide definitive information about such a star's structure and dynamics. The combination of data gained by measuring gravitational waves emitted after a star quake with timing data gained by measurements of the electromagnetic energy radiated would reveal information about the coupling of the solid crust (and possibly solid core) to the fluid mantle.

5.3 Planet gravity-field mapping

For decades, scientists have known that the best way to achieve a global characterization of the Earth's gravity field with high sensitivity (1-mgal gravity anomalies and 1 to 10-cm geoid undulations) and high spatial resolution (50 to 150 km) would be to monitor changes in the accelerations experienced by orbiting spacecraft (Wolff 1969; Douglas, *et al.* 1980; Kaula 1983). Conventional satellite-tracking data are adequate only for determining long-wavelength variations in the geopotential, *i.e.*, for spatial resolutions of several thousand kilometers (Koch and Witte 1975). Satellite altimetry has been used successfully to measure features as small as 100 to 200 km in extent with sensitivities of a few mgal (Rapp 1979), but altimetry has the disadvantages that it doesn't distinguish the geoid from sea level, and it doesn't give coverage of the continents. There are two basic techniques for measuring the gravity field from orbit: the use of onboard gravity gradiometers or *in situ* monitoring of changes in the relative velocities of two or more satellites, one in a low orbit (equipped with appropriate drag-free systems) and the others either in the same low orbit or in higher orbits. Although the recent development of superconducing gravity gradiometers has begun

to make them competitive with the other techniques, theoretical calculations still suggest that the best sensitivity and resolution could be achieved with two or more drag-free satellites in the *same* low-altitude polar orbit, one behind the other (Breakwell 1979). Global coverage would be achieved by using several such combinations of spacecraft distributed in longitude. The net change in range rate between the satellites as they pass over an anomaly in the geopotential scales linearly with the strength of the anomaly and the spacecraft separation, provided the separation is on the order of or smaller than the desired measurement resolution (defined conventionally as one-half the horizontal spatial extent or wavelength of the variations being measured). It falls off exponentially with the ratio of orbit altitude to measurement resolution.

A mission of this sort was proposed first nearly a decade ago as GRAVSAT, which later came to be known as the Geopotential Research Mission (GRM; *e.g.*, Keating, *et al.* 1986). It would have involved a pair of spacecraft in a single polar orbit at 160-km altitude, the lowest possible before atmospheric drag would lead to unacceptable measurement degradation. The spacecraft separation was to vary between about 150 and 550 km, and the changing range rate between them was to be monitored with two-way, dual-radio-frequency Doppler tracking between the spacecraft. With 1-μm/s range-rate precision and a year's worth of measurements taken at 10-second intervals (with 4-second averaging times), this mission was expected to map the Earth's gravity field with a sensitivity of 2.5 milligal for gravity anomalies and 10 cm for geoid undulations, down to spatial resolutions on the order of 110 km, or approximately $1° \times 1°$. Higher resolution was precluded because of the greater demand on range-rate measurement precision.

Significantly better performance might be achieved with a laser version of the GRM, that is, using spacecraft equipped with frequency-stabilized lasers whose separations are monitored with coherent laser links. The analysis described in this section suggests that such a mission could map Earth's gravity field to spatial resolutions on the order of 50 km or better. The improvement provided by the use of frequency-stabilized lasers results both from more accurate Doppler measurements and

from the possibility of using lower-altitude orbits.

This section begins by deriving expressions that relate measurement sensitivity in terms of gravity-anomaly strengths or geoid undulations to the accuracy of range-rate measurements among spacecraft. Simplifying but realistic approximations are made where possible to keep this presentation tractable, but comparison with results of more rigorous analyses is provided. Relations among measurement sensitivity, spatial resolution, range-rate measurement precision, and laser fractional frequency stability are summarized in tables. Error sources that impact measurement sensitivity by limiting range-rate measurement accuracy are discussed, including photon statistics, fluctuations in atmospheric refractive index, and "drag" effects such as fluctuations in solar intensity.

5.3.1 Requirements on Doppler accuracy and laser stability

Consider two or three spacecraft in identical circular polar orbits around Earth (mass M_e, radius R_e) or any planet of mass M and mean radius R, at an orbit altitude h (Figure 18). Suppose the spacecraft are separated by nominal distances L, and that they form single-arm or collinear dual-arm interferometers of the sort depicted in Figures 6b and 7a. The dependences on altitude h of orbital periods τ and velocities v follow from Kepler's law:

$$\begin{aligned} \tau(h) &\simeq 91.5 \text{ min} \left(1 + \frac{h}{R}\right)^{3/2} \left(\frac{R}{R_e}\right)^{3/2} \left(\frac{M_e}{M}\right)^{1/2} , \\ v(h) &\simeq 7.2 \text{ km/s} \left(1 + \frac{h}{R}\right)^{-1/2} \left(\frac{M}{M_e}\right)^{1/2} \left(\frac{R_e}{R}\right)^{1/2} . \end{aligned} \tag{5.3.1}$$

Throughout this section, numerical values given apply to Earth, but the extension to an arbitrary planet of radius R, mass M, and surface gravitational acceleration g is straightforward. Henceforth, assume an orbital velocity $v_0 \equiv 7.2$ km/s and mean Earth radius $R \equiv R_e \equiv 6370$ km. If the gravitational field were spherically symmetric, the orbital velocities of the spacecraft would remain constant, as would the separation L (or L_1 and L_2) between spacecraft. An anomaly in the

gravitational field, such as that caused by a local concentration of mass, will cause each spacecraft to speed up as it approaches the anomaly and slow down as it recedes. (The opposite occurs for an anomaly opposite in sign, *i.e.*, a local sparsity of mass.) The relative velocity (range rate) produced between two spacecraft by this perturbing acceleration will change sign as the spacecraft pass over the anomaly, reaching its maximum absolute value when the two spacecraft are at an angular distance of approximately h/R on either side of the anomaly. The gravitational potential can be represented as an expansion in spherical harmonics, Y_{lm},

$$\begin{aligned} U(r > R, \Omega) &= \frac{GM}{R} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left(\frac{R}{r}\right)^{l+1} J_{lm}(R)\, Y_{lm}(\Omega) \,, \\ J_{lm}(R) &\equiv \frac{4\pi}{2l+1} \left(\frac{1}{R}\right)^{l+3} \frac{1}{M} \int d^3r' \, (r')^l \, \rho(r', \Omega') \, Y_{lm}(\Omega') \,, \end{aligned} \tag{5.3.2}$$

where G is the gravitational constant, Ω denotes solid angle, and the integration is over the volume of the planet defined by the region of nonzero mass density ρ (Heiskanen and Moritz 1967; Kaula 1966). A similar expression can be used for a description of the field in terms of surface layers and surface mass densities. The spatial wavelength λ_l and wavenumber k corresponding to the harmonic of degree l are, respectively,

$$\lambda_l \equiv 2\pi R/l, \qquad k \equiv l/R = 2\pi/\lambda_l \,. \tag{5.3.3}$$

The goal here is to estimate the magnitude of short-wavelength variations in the gravity field ($\lambda_l \leq 400$ km, or $l \geq 100$) from measurements of the changing separation, or relative velocity v, between satellites. This is a standard "inversion" problem in which measurements of v are to be transformed into a uniform representation of the surface gravitational potential. The "resolution" associated with harmonics complete through degree l is conventionally taken to be $\lambda_l/2$, corresponding to an angular resolution $\alpha_l \equiv \lambda_l/2R = \pi/l$.

The potential at an altitude h also can be written in terms of the potential on the surface by

using Poisson's formula (Heiskanen and Moritz 1967):

$$U(r = R + h, \Omega) \;=\; R\,(r^2 \;-\; h^2) \int \frac{d\Omega'}{4\pi} \; \frac{U(R, \Omega')}{r^2 \;+\; R^2 \;-\; 2rR\cos\psi} \; . \tag{5.3.4a}$$

Here ψ is the angle between $\vec{r}$ and $\vec{R}$, where $\vec{r}$ is the line from the center of the Earth to the point at an altitude h above the gravity anomaly, and $\vec{R}$ is the line from the center of the Earth to the point on the Earth's surface at which the surface potential is measured. (The latter point may be the "subspacecraft" point; see Figure 18.) Because the orbits are polar, coverage will be essentially uniform at low latitudes; hence, for analysis, the spherical Earth can be replaced with an infinite flat plane on which coordinates are specified by x, y values. In this "flat-Earth" approximation (Breakwell 1979), Poisson's formula becomes a two-dimensional convolution involving the potential on the surface:

$$\begin{aligned} U(x, y, h) \;&=\; \frac{h}{2\pi(x^2 \;+\; y^2 \;+\; h^2)^{3/2}} \;\star\; U_0 \; , \\ U_0 \;&\equiv\; U(x, y, h = 0) \; . \end{aligned} \tag{5.3.4b}$$

Here the symbol $\star$ denotes (two-dimensional) convolution. Both for ease in calculations and because the ultimate goal is to derive measurement uncertainties that correspond to a particular resolution or wavelength, it is appropriate to consider the Fourier components of all quantities of interest. The two-dimensional Fourier transform of the potential at a height h is

$$\begin{aligned} \tilde{U}(k_x, k_y, h) \;&=\; e^{-kh}\; \tilde{U}(k_x, k_y, h = 0) \;\equiv\; e^{-kh}\; \tilde{U}_0(\vec{k}) \; , \\ k \;&\equiv\; ({k_x}^2 \;+\; {k_y}^2)^{1/2} \; . \end{aligned} \tag{5.3.4c}$$

Representative values for wavelength, harmonic order, angular resolution, and the important factor in the exponent of this expression, $kh = lh/R = 2\pi l/\lambda_l$, are given in Table 5.3.1. Note that the flat-Earth approximation is not valid at low frequencies, *i.e.*, wavelengths $\lambda_l \gtrsim R$.

Table 5.3.1 Relations among wavelength, spherical harmonic order, and angular size of gravity anomaly. The last column, when exponentiated, describes the dependence of measurement performance on orbit altitude h, here assumed to be 160 km. R is the Earth's mean radius, $\sim$ 6370 km.

λ_l (km)	$l = 2\pi R/\lambda_l$	$\alpha = \lambda_l/2R$	lh/R
400	100	1.80°	2.5
200	200	0.90°	5.0
100	400	0.45°	10.0
80	500	0.35°	12.6
50	800	0.23°	20.1
40	1000	0.18°	25.1

The equilibrium condition ${v_0}^2/2 = (U_0 + \text{constant})$ implies that the maximum change in range rate δv_s between the spacecraft (from its equilibrium value of zero) is related to a perturbation in the potential $\delta U(x, y, h)$ at the point located midway (in the x-direction) between the two spacecraft by

$$\begin{aligned} v_0\ \delta v_s &= \delta U(x + L/2, y, h) - \delta U(x - L/2, y, h) \\ &= \delta U(x, y, h) \star \delta_D(y) \left[\delta_D(x + L/2) - \delta_D(x - L/2)\right] , \end{aligned} \tag{5.3.5a}$$

where δ_D denotes a Dirac delta function. In the Fourier domain, this relation is

$$\begin{aligned} \delta\tilde{v}_s(\vec{k}) &= \frac{2i}{v_0}\, e^{-kh}\, \delta\tilde{U}_0(\vec{k})\, \sin\left(\frac{kL\cos\phi}{2}\right) , \\ k\cos\phi &\equiv k_x . \end{aligned} \tag{5.3.5b}$$

where the angle ϕ denotes the direction of the two-dimensional vector $\vec{k}$. Perturbations δv_s in relative velocity (range rate) due to local variations in the potential are the signals of interest, as opposed to perturbations in relative velocity due to other causes, including measurement error; hence the subscript "s" on δv_s. Most statistical models of the Earth ascribe fluctuations in the surface potential to white-noise fluctuations in thin layers below the surface (Morrison 1976; Jordan 1978). The perturbations $\delta\tilde{U}_0$ therefore are isotropic:

$$\delta\tilde{U}_0(\vec{k}) = \delta U_0(k) . \tag{5.3.6a}$$

Perturbations in the local gravitational acceleration g and the local geoid height n also are isotropic, and are related to perturbations in the potential as follows (in Fourier space):

$$\delta\tilde{g}(k) \;=\; k\delta\tilde{U}_0(k)\;; \tag{5.3.6b}$$

$$\delta\tilde{n}(k) \;=\; \frac{1}{g}\,\delta\tilde{U}_0(k), \qquad g \;\equiv\; \frac{GM}{R^2} \;\simeq\; 9.8\ \mathrm{m/s^2}\;. \tag{5.3.6c}$$

Signal-induced velocity perturbations are not isotropic:

$$|\delta\tilde{v}_s(\vec{k})| \;\equiv\; |\delta\tilde{v}_{s,\max}(k)|\,\sin\!\left(\frac{kL\cos\phi}{2}\right)\;. \tag{5.3.7a}$$

In the "linear signal regime," where $kL\cos\phi << 2$ or $L\cos\phi << \lambda_l/\pi$, the sine function can be approximated by its argument, and the signal grows linearly with spacecraft separation L. To simplify this presentation, ignore the nonisotropic nature of the signal and assume it has its maximum value.* This maximum change in range rate is related linearly to the (Fourier transforms of the) gravity anomaly strength and the geoid undulation:

$$|\delta\tilde{v}_{g,\max}| \;=\; \frac{2R}{v_0}\,\frac{1}{l}\,e^{-lh/R}\,|\delta\tilde{g}| \;=\; \frac{\lambda_l}{\pi v_0}\,e^{-2\pi h/\lambda_l}\,|\delta\tilde{g}|\;; \tag{5.3.7b}$$

$$|\delta\tilde{v}_{n,\max}| \;=\; \frac{2g}{v_0}\,e^{-lh/R}\,|\delta\tilde{n}| \;=\; \frac{2g}{v_0}\,e^{-2\pi h/\lambda_l}\,|\delta\tilde{n}|\;. \tag{5.3.7c}$$

Values of the maximum changes in range rate corresponding to a Fourier component l, or wavelength λ_l, are given in Table 5.3.2.

* A rigorous calculation including angular dependences and effects of Earth's rotation, but valid only in the flat-Earth approximation and in the linear signal regime, appears in Breakwell (1979). That calculation gives a result that agrees closely with the result of the simpler analysis done by this author and presented here.

Table 5.3.2 Maximum changes in range rate $|\delta\tilde{v}_{g,\max}|$ and $|\delta\tilde{v}_{n,\max}|$ induced by a 1-mgal gravity anomaly and a 5-cm geoid undulation, respectively, occurring over spatial extents λ_l; velocity-measurement accuracies σ_v required to detect 1-mgal gravity anomalies and 5-cm geoid undulations of wavelength λ_l; and required frequency stability or knowledge $\delta\nu/\nu$ to achieve the range-rate accuracy shown for 1-mgal anomalies, assuming a dual-arm interferometer (three coorbiting spacecraft) with spacecraft separations nominally equal to 0.1%. Spacecraft separation is assumed to satisfy the condition for maximum signal, $L = \lambda_l/2$. An orbit altitude of 160 km is assumed, with $v_0 \equiv 7.2$ km/s and $R \equiv 6370$ km. It is assumed that 10^6 independent measurements are made, each with a 1-s integration time. For a given wavelength λ_l, spacecraft separation, and orbit altitude the uncertainties in anomaly strength and geoid undulation scale linearly with range-rate measurement precision.

λ_l (km)	$\lvert\delta\tilde{v}_{g,\max}\rvert$ (1 mgal)	$\lvert\delta\tilde{v}_{n,\max}\rvert$ (5 cm)	σ_v (1 mgal)	σ_v (5 cm)	$\delta\nu/\nu$ (1 mgal)
400	14.3 μm/s	11.0 μm/s	810. μm/s	620. μm/s	$4. \times 10^{-6}$
200	580. nm/s	895 nm/s	23. μm/s	36. μm/s	2.3×10^{-7}
100	1.9 nm/s	5.8 nm/s	54. nm/s	165. nm/s	$1. \times 10^{-9}$
80	120. pm/s	500. pm/s	3.1 nm/s	12. nm/s	7.8×10^{-11}
50	41. fm/s	250. fm/s	820 fm/s	5. pm/s	3.3×10^{-14}
40	0.2 fm/s	1.7 fm/s	3.8 fm/s	30. fm/s	$2. \times 10^{-16}$

What is the connection between measurements of change in range rate and variations in the gravitational potential? The answer requires use of some simple aspects of estimation theory, which are described in the Appendix (section 8). Results are that, for gravity anomalies and geoid undulations, respectively, the rms errors at harmonic l or wavelength $\lambda_l \equiv 2\pi R/l$ are

$$\begin{aligned} \sigma_{\tilde{g}} &= \sigma_v \left(\frac{\pi}{N_m}\right)^{1/2} \left(\frac{v_0}{2R}\right) l^{3/2}\, e^{lh/R} \\ &= \sigma_v v_0 \left(\frac{2\pi^4 R}{N_m \lambda_l^3}\right)^{1/2} e^{2\pi h/\lambda_l} \; ; \end{aligned} \tag{5.3.8}$$

$$\begin{aligned} \sigma_{\tilde{n}} &= \sigma_v \frac{v_0}{2g} \left(\frac{\pi}{N_m}\right)^{1/2} l^{1/2}\, e^{lh/r} \\ &= \sigma_v \frac{v_0}{g} \left(\frac{\pi^2 R}{2N_m \lambda_l}\right)^{1/2} e^{2\pi h/\lambda_l} \; . \end{aligned} \tag{5.3.9}$$

The velocity-measurement accuracies σ_v required to sense 1-mgal gravity anomalies or 5-cm geoid undulations over various spatial extents are given in Table 5.3.2. These values assume a 160-

km orbit altitude, an average over $N_m = 10^6$ identical, independent measurements, and that the spacecraft separation L is adjusted to match the desired spatial resolution in order to maximize signal ($L = \lambda_l/2$). Figures 19a and 19b are plots of the maximum change in range rate and the required measurement accuracy for a 1-mgal gravity anomaly, as a function of wavelength. Note the faint signal (small velocity change) and high range-rate accuracy needed to sense variations in the gravity field over spatial scales as short as 25 to 50 km.

The required range-rate accuracies shown in Table 5.3.2 can be converted to requirements on control or calibration of laser fractional frequency fluctuations over the measurement integration times t_m by using eq. (4.2.9c) (which assumes a cycle-counting technique for Doppler measurements — a suboptimal technique, but one which will suffice to give a pessimistic rather than optimistic requirement on laser frequency stability). Assume, too, that three spacecraft are used, and that they form a collinear dual-arm interferometer with arms held equal to about 0.1%. This will ease the requirements on control or calibration of the laser frequency fluctuations by a factor of about 1,000. (If only two satellites are used, use of a "smarter" Doppler measurement technique as well as active calibration of laser frequency fluctuations may be able to compensate.) Under these assumptions, the required laser fractional frequency stability (controlled or calibrated) scales with the required range-rate accuracy σ_v, the measurement integration time t_m, and the spacecraft separation $L \equiv c\tau$ as follows:

$$\frac{\delta\nu}{\nu} \simeq 10^{-11} \left(\frac{100\ \text{km}}{L}\right) \left(\frac{t_m}{1\ \text{s}}\right) \left(\frac{\sigma_v}{1\ \text{nm/s}}\right) . \tag{5.3.10}$$

For illustration, consider the situation in which 10^6 independent 1-second measurements are made with spacecraft separations of roughly 50 km ($\lambda_l \simeq 100$ km). Using $\sigma_v \simeq 54$ nm/s (Table 5.3.2), eq. (5.3.10) says that the laser fractional frequency fluctuations must be controlled or calibrated to roughly 10^{-9} in order to sense 1-mgal variations in the gravity field with spatial resolutions on the order of 50 km. For similar sensitivity with better spatial resolution (shorter wavelengths), the frequency stability requirements become much more stringent. For example, to achieve 25-km spatial

resolution (λ_l = 50 km) with the spacecraft separated by 25 km (to maximize signal), fractional fluctuations in the laser frequency must be controlled or calibrated to a few parts in 10^{-14}.

The single-measurement requirements for range-rate accuracy and corresponding laser frequency stability derived above were eased by a factor of 1,000 by the assumption that information was to be gained not from individual measurements but by an average over roughly one million ($N_m \sim 10^6$) independent measurements. Generally speaking, the separations among spacecraft are sensitive to a specific gravity anomaly for only a small fraction of the orbit that passes directly over that anomaly ($\approx$ 10% for anomalies of spatial extent $\lesssim$100 km; see, for example, Fig. 5 of Douglas, *et al.* 1980). Further, only 1 of every $\sim$ 16 orbits will pass directly over that anomaly. Therefore, a given pair of spacecraft may be sensitive to a specific anomaly only $\sim$ 0.6% of the time. With measurements made once per second, the total number of measurements per day that are sensitive to a specific ground area is $\sim$ 500, so it would require 5 to 6 years to get 10^6 measurements. However, initial ideas for GRM envisioned eight pair of spacecraft, which would reduce the time period for getting 10^6 measurements to about 8 months. The actual appropriate value for the number of measurements N_m made over specific locations will depend on many factors and must be determined in the context of the specific parameters and data-analysis procedures envisioned for the experiment.

5.3.2 Photon statistics

Measurement of the changing separations and range rates among spacecraft could be limited by poor signal-to-noise ratio caused by an inadequate photon flux with which to make the Doppler measurements. An approximate idea of the minimum laser power needed to keep photon-statistics errors from precluding the necessary range-rate accuracies can be obtained by comparing the range-rate accuracies listed in Table 5.3.2 with the photon-statistics error $\delta L_{\rm ph}$ in a length measurement for a given integration time τ. The latter was calculated in subsection 4.1.4a for the cases of retrore-

flectors and transponders [eqs. (4.1.9)]. For illustration, consider the requirement of 820 fm/s for single-measurement range-rate accuracy σ_v in order to map the gravity field with 1-mgal sensitivity and 25-km spatial resolution. Most of the parameters in eqs. (4.1.9) can be assumed to have their indicated values — 10-cm nondiffraction-limited transmit and receive apertures for all spacecraft (deliberate beam-spreading by a factor of 5, say, for the central spacecraft and 15 for the outer spacecraft), 5% overall optics and detection efficiencies for either transponders or retroreflectors, and 1-μm laser wavelength. In a 1-second integration time, photon statistics will cause a measurement error of about 710 fm if retroreflectors are used and the initial laser output power is 10 mW, but an error of only about 60 fm if the outer spacecraft have coherent laser transponders that retransmit with an average output power of 10 mW. This factor-of-12 advantage with transponders over retroreflectors depends on the ratio of telescope apertures to spacecraft separation [$D_A D_B / s_A \lambda L$, from eqs. (4.1.8)]; the laser output power would have to be increased by a factor on the order of 150 with retroreflectors to make up this advantage, bringing the requirement on average laser output power (with retroreflectors) to about 1.5 W. As the range-rate accuracy requirements become more stringent for higher-resolution gravity mapping, the error contribution from photon statistics becomes a greater threat that can be subdued only by using higher laser powers, and the use of coherent laser transponders instead of retroreflectors becomes essential.

5.3.3 Medium-induced phase fluctuations

Since ionospheric charged-particle effects are negligible at optical frequencies, medium-induced phase errors would arise primarily from fluctuations in the atmospheric index of refraction. In general, the fluctuations or perturbing accelerations that have the greatest effect on measurement accuracy are those whose characteristic periods are commensurate with the time required to produce a signature from the shortest wavelength components of the force being measured. At an orbit velocity of 7.2 km/s, the spacecraft travel distances on the order of 100 km in approximately 15 seconds. Hence the maximum deviation from zero of the relative velocity between the two coorbiting spacecraft typically will be caused by external perturbing forces with 15-second periods.

Fluctuations in the refractive index due to atmospheric density fluctuations tend to occur over periods considerably longer than 10 to 15 seconds at an altitude of 160 km (Bender, private communication). The exception to this is short-period, large fluctuations due, for example, to the use of spacecraft thrusters. Other short-period variations in density caused by processes such as ionospheric disturbances could constitute an important limit to measurement sensitivity. The average path-delay in each arm due to a nonunity-average index of refraction does not pose a fundamental limit to measurement sensitivity, provided it can be adequately calibrated (in a single-arm interferometer) or cancelled (in a dual-arm interferometer). At 160-km altitude above Earth, the average index of refraction differs from unity by no more than a few parts in 10^{13}; this would produce an average path correction of about 10 nm for a 50-km arm length. If the difference in arm lengths is 50 m or less, the corresponding difference in path delay between the two arms of a dual-arm interferometer would be 10 pm or less.

5.3.4 External accelerations

Fluctuating relative accelerations on each spacecraft will result in fluctuating motions of the isolated test masses used to define the precise change in position of each spacecraft, and could pose a serious problem. Active disturbance compensation systems, such as were planned in the GRM proposal, are required. Fluctuating accelerations could arise, for example, from buffeting by the solar wind or radiation pressure, varying gravitational forces on the test masses as fuel moves and is depleted, and outgassing, especially from the spacecraft attitude-control systems (see also discussions in sections 3.4.5 and 5.2.3). For a rough idea of the magnitude of such forces, note that the gravitational acceleration felt by a test mass due to a 10^4-kg point mass 3 meters away is on the order of 7×10^{-3} mgal. Fluctuations in the gravitational acceleration experienced by the test mass as fuel moves within the spacecraft might be 2 to 3 orders of magnitude smaller than this, or about 10^{-5} mgal $\simeq 10^{-11} g$ (*e.g.*, movement of 10 to 100 kg of fuel over distances of a meter or two). If these changes occur over time scales of 10 to 15 seconds, they could produce fluctuations in the measured range rate between spacecraft on the order of 1 nm/s, unacceptably large for high-resolution mapping (see Table 5.3.2). Stated a bit more rigorously, the Fourier components of an external acceleration corresponding to periods T of 10 to 15 seconds, which have the form $a \equiv a_0 \sin(2\pi t/T)$, will produce periodic fluctuations δv in the measured range rate v between spacecraft:

$$\delta v = a_0 \cdot \frac{T}{2\pi}\left[1 - \cos(2\pi L/v_0 T)\right] \quad . \tag{5.3.11}$$

Thus, a sinusoidal perturbing acceleration of strength $a_0 \simeq 10^{-11} g$ with a 10- to 15-second period could produce changes in the range rate on the order of 0.3 nm/s over time scales of 10 to 15 seconds, for spacecraft separated by 25 to 50 km. All perturbing accelerations to the test masses with fluctuations of this magnitude or stronger must be calibrated or cancelled with compensating accelerations applied to the test masses.

One of the most severe sources of perturbing accelerations in this application is the solar wind, whose intensity varies unpredictably. The average deceleration due to the solar wind experienced by a spacecraft in orbit depends on the time of the solar cycle, the projected frontal area of the spacecraft, the local atmospheric density (a function of orbit altitude, latitude, and solar flux), and the square of the orbital velocity. At an altitude of 175 km in orbit around Earth, a spacecraft experiences an average deceleration of about 5 mgal $\simeq$ 50 μm/s^2. Over time scales of 10 to 15 seconds, this may fluctuate by as much as 0.1% (Bender, private communication), producing fluctuations in the relative velocity between spacecraft as large as 150 nm/sec.

The only solution to fluctuating external accelerations is an active compensation system. A successful "DISCOS" (DISturbance COmpensation System) unit was flown in 1972 on the U.S. Navy TRIAD-1 satellite ("Staff of the Space Dept.," *et al.* 1974) (see section 3.4.5). It was able to compensate for perturbing accelerations down to $10^{-11}g$ along all three axes. The test mass in this system was allowed to move within a dead zone of about 1 cm, and in the region surrounding the mass, gravitational gradients were held to less than $10^{-11}g$/mm (Keating, *et al.* 1986). These and much more advanced drag-free systems currently are under intense study for use with space-based laser-interferometric detectors of gravitational waves (Bender, *et al.* 1988), and for a relativity experiment intended to measure the Earth's "frame-dragging" effect to 1%, or 0.5 milliarcsecond per year (Everitt, *et al.* 1988; see section 6.2.6). Studies indicate that the required performance of approximately $10^{-14}g$ at 10-second periods for high-resolution mapping of Earth's gravity field can be achieved and exceeded. The requirements on the DISCOS systems for the gravitational-wave detectors are much more severe, by 4 to 5 orders of magnitude.

5.4 Atmospheric wind-sensing

Programs have been under way for the last decade to develop CO_2 (wavelength $\lambda = 9.11$ to 10.06 μm) pulsed coherent lidar systems for ground-based, airborne, and satellite measurement of atmospheric wind velocity. Such systems would be used to measure horizontal components of the wind and wind shear, as well as the vertical distribution of the horizontal wind field. The primary science driver for Earth-orbiting systems is improved meteorological forecasting. This application requires single-measurement accuracies of 1 to 5 m/s for each component of the horizontal wind field (whose magnitudes may reach on the order of 100 m/s or 230 mph), and a vertical range resolution of approximately 1 km (Huffaker 1978). For twice-daily global coverage with horizontal averaging over 300×300-km cells, pulse repetition frequencies of at least 1 to 10 Hz are required, as well as a scanning capability. The scanning results in large Doppler shifts of the backscattered signals due to the high relative velocity of the orbiting spacecraft. At 10-μm wavelengths, these Doppler shifts could be as large as 1 GHz (for a conical scan technique, a spacecraft at 800-km altitude, and a 55° nadir angle; Menzies 1986). In contrast, the signal of interest — the Doppler shift due to backscatter from aerosols — would be only 200 kHz at 10-μm wavelengths, for aerosols with a relative velocity (line-of-sight component) of approximately 1 m/s. A pulse repetition frequency f must be used that is high enough to measure maximum velocities of 50 to 100 m/s; this would be $\sim$ 10 to 20 MHz at 10-μm wavelengths. However, as discussed in subsection 4.2.2b, such high sampling frequencies can be avoided with bandwidth reduction techniques in which the signal is detected in several narrow overlapping filters of width b_d ($b_d \sim f/10$, for example), each of which is followed by an analog-to-digital filter; the Doppler information is then processed in parallel. (See Figure 13 and discussion in subsection 4.2.2b.) In general, pulse energies in the range of 1 to 10 J are desired, although this requirement depends strongly on laser wavelength for a given velocity-measurement precision. In addition, for satellite-borne operation, lidar systems also must satisfy strict size and weight conditions, and have lifetimes of several years.

Theoretical studies made over the last 5 to 10 years have suggested that remote wind-sensing lidar systems at optical frequencies could provide several advantages over infrared and microwave systems (Kane, *et al.* 1984; Menzies 1985, 1986). As shown in section 4.2.2, for comparable received signal strengths (detected number of photons), the shorter-wavelength system could offer a smaller velocity error and better range resolution [eqs. (4.2.19)]. In Earth's atmosphere, if turbulence effects are negligible, the received signal strengths for 1-μm and 10-μm pulsed coherent lidar systems with similar pulse energies are roughly comparable, because the backscatter cross section is 10 to 50 times larger at 1 μm than at 10 μm (Kent, *et al.* 1983; Patterson, *et al.* 1980). Turbulence effects are worse at 1 μm than at 10 μm, because the transverse atmospheric coherence length scales roughly as $\lambda^{6/5}$ (see subsection 4.2.2c); however, these effects can be minimized by using small apertures, and eliminated by going to Earth-orbit. The Doppler shifts at 1-μm wavelengths are 10 times higher than at wavelengths of 10 μm. Thus an effective sampling frequency of 100 to 200 MHz would be required to measure relative velocities of 50 to 100 m/s. The bandwidth reduction techniques noted above could be used to reduce this high sampling frequency.

A performance comparison based on relative photon efficiencies (required number of transmitted photons per pulse for a given velocity-measurement precision) was carried out by Menzies (1985, 1986) for four different candidate lidar systems. These included two coherent (heterodyne) systems based on 10-μm CO_2 lasers and 1-μm Nd:YAG lasers, and two incoherent (direct-detection with Fabry-Perot filters) systems using frequency-doubled (0.5-μm) Nd:YAG lasers and Raman-shifted XeCl excimer lasers (0.35 μm). The performance of the coherent systems was found to be superior to that of the incoherent systems. The comparison in this reference was restricted to conventional, flash-lamp-pumped Nd:YAG lasers, which (in a pulsed mode) operate at $\sim$ 1% electrical-to-optical efficiency, while $C0_2$ lasers operate at $\sim$ 5% efficiency. Even so, the Nd:YAG sytem was found to be competitive with the $C0_2$ system, but the latter was preferred because of the eye-safety problem (see subsection 4.2.2d). A change from conventional to diode-pumped Nd:YAG lasers, which operate

with electrical-to-optical efficiencies near 10%, makes the latter perform better and more efficiently than the CO_2 systems. Because of the eye-safety problem, it is possible that for the time being CO_2 systems will remain the favored candidate for near-Earth applications. (See subsection 4.2.2d for more details.)

Armed with the groundwork laid in section 4.2.2, the reader can make a quick but accurate performance comparison of different kinds of coherent pulsed lidar systems that might be used to provide a global wind-sensing capability, without having to resort to numerical computations of complicated equations such as the radar equation (4.2.16a). Such a comparison is outlined below, for systems operating at 1.06 μm and 10.6 μm. Similar performance evaluations could be made for a variety of science applications, such as measurements in planetary atmospheres and rings made from spacecraft or landers, or measurements on interplanetary dust, as well as other meteorological and commercial applications mentioned in section 4.2.2 for Earth's atmosphere. One need only use different expressions or numerical values for the backscatter coefficient, efficiency η (dependent on the receiver/detector system as well as propagation losses between spacecraft and target medium), and desired range resolution and then optimize system parameters such as receiver area, pulse energy, pulse width, and sampling frequency to give the desired velocity-measurement precision and range resolution. Results and equations from section 4.2.2 will be used freely here, with minimal explanation and no derivation.

The system parameters to be assumed, and potentially optimized, include the laser wavelength λ, pulse duration τ_p, pulse energy E_p, receiver diameter D_r (receiver area $A_r \equiv \pi {D_r}^2/4$), and overall receiver/detector efficiency η_r. For wind-sensing, the range L typically will not exceed 10 km. As stated above, a range resolution of 1 km is desired; the range-gate duration $\tau_g \equiv M\tau_p$ therefore must be no longer than 6.7 μs [eq. (4.2.14a)], regardless of wavelength. For the sake of comparison, assume here that a pulse width $\tau_p = 3.2$ μs is used for both the 1-μm and the 10-μm systems. Assumed values for other system parameters (E_p, D_r, η_r) are listed in Table 5.4.1. If the system is to be able

to measure a maximum relative velocity (line-of-sight component) of 50 m/s (the "Nyquist velocity" v_{Ny}), the sampling frequencies f must be 100 MHz and 10 MHz for the 1-μm and 10-μm systems, respectively ($v_{Ny} \equiv f\lambda/2$).

Other parameters relevant to evaluating performance depend on the nature of the scattering particles, including their size and distribution as well as the transmission efficiency of the laser light through through the medium. For this application, the aerosol volume-backscatter coefficient β (units of $m^{-1}s^{-1}$) will be approximated for $\lambda = 1$ μm from model calculations by Kent, *et al.* (1983) for midtropospheric altitudes, and will be taken from JPL (Menzies 1986) and NOAA (Post 1984) measurements for $\lambda = 10$ μm. At 5-km altitude, the backscatter coefficient at $\lambda = 1$ μm is approximately 10^{-8} m^{-1} sr^{-1}, compared to 2×10^{-10} at $\lambda = 10$ μm. At low altitudes ($\sim$ 2.5 km), losses due to water vapor absorption can be significant for the 10-μm laser; however, the aerosol backscatter coefficient rises by so much at the lower altitudes that it is a reasonable approximation to ignore these losses. At 1-μm, the extinction is far less, and again negligible relative to the backscattering. This factor of 50 difference between the backscatter coefficients at 1 μm and 10 μm approximately cancels the $\sim \lambda^2$ dependence of the other factors contributing to the total number of signal photons detected (see Tables 5.4.1 and 5.4.2), making the signal strengths comparable for the two systems, if equal pulse energies are used. (Possible degradation in signal strength due to turbulence in front of the receiver is neglected here; as discussed in subsection 4.2.2c, it could be a greater problem at 1 μm than at 10 μm.)

Another parameter to be assumed, which in practice must be measured or taken from models, is the expected spread (standard deviation) in velocities of the aerosols, due to turbulence and wind-shear effects. A value of $\sigma_{med} = 0.4$ m/s will be used here, consistent with values used by Huffaker (1978) and Menzies (1986) describing conditions of light turbulence and an average wind shear appropriate for midlatitude locations. This value corresponds to a correlation time τ_{corr} (time over which successive signals are correlated) of approximately 2 μs at $\lambda = 10$ μm and 0.2 μs at

$\lambda = 1$ μm. Detector integration times τ_d are taken to be equal to these values for the respective lidar systems. These quantities are summarized in Table 5.4.1 below.

Table 5.4.1 Assumed system parameters and backscatter cross section for performance comparison between 1-μm and 10-μm coherent pulsed Doppler lidar systems for global wind-sensing. The maximum expected relative velocity $v_{Ny} = f\lambda/2$ is 50 m/s. The standard deviation σ_{med} of the velocity distribution for the aerosols is 0.4 m/s. Detector integration times τ_d are set equal to the resulting correlation times $\tau_{corr} \sim \lambda/(4\pi\sigma_{med})$. The range L is 10 km. A pulse duration $\tau_p = 3.2$ μs is used for both systems, adequate to provide 1-km range resolution. See section 4.2.2 for further discussion and definition of parameters.

Parameter	$\lambda = 1$ μm	$\lambda = 10$ μm
E_p	0.1 J	0.1 J
D_r	0.4 m	0.4 m
η	0.1	0.1
τ_p	3.2 μs	3.2 μs
f $(= 2v_{Ny}/\lambda)$	100 MHz	10 MHz
τ_d	0.2 μs	2.0 μs
β	10^{-8}	2×10^{-10}

The derived quantities which are used with equations (4.2.17) to calculate velocity-measurement precision are summarized in Table 5.4.2 (see section 4.2.2 for discussion and definitions). Clearly, with the parameters chosen here, the 1-μm and 10-μm systems both can provide precisions in the 1 to 5 m/s range. The precision of the 1-μm system is about a factor of 3 better than that of the 10-μm system. Even if the pulse energy of the 1-μm system were dropped from its nominal value here of 0.1 J to 40 mJ (the ANSI recommendation for maximum pulse energy because of eye-safety considerations), the velocity precision for the 1-μm system would remain about 0.1 m/s. This is so because with the other parameter values assumed here, the 1-μm system is operating with such a high SNR (the quantity A is the dominant contributor to measurement error) that a cut in pulse energy by 40% is barely noticeable.

Table 5.4.2 Derived quantities for comparison of achievable velocity-measurement precision for the 1-μm and 10-μm lidar systems described in the text and in Table 5.4.1. See section 4.2.2 for discussion and definitions.

Parameter	$\lambda = 1$ μm	$\lambda = 10$ μm
$\sigma_{v,\mathrm{bw}} \equiv \lambda/4\pi\tau_p$	0.025 m/s	0.25 m/s
$\sigma_0 \equiv (\sigma_{\mathrm{med}}{}^2 + \sigma_{\mathrm{bw}}{}^2)^{1/2}$	0.4 m/s	0.47 m/s
$Y \equiv 4\pi\sigma_0/v_{\mathrm{Ny}}$	0.10	0.12
s	10^8 s^{-1}	2×10^7 s^{-1}
$N_d \equiv s\tau_d$	20	40
$\xi \equiv 2f\tau_p$	640	64
$\mathrm{SNR_w} \equiv (\xi N_d)^{1/2}$	113	51
A	0.10 m/s	0.32 m/s
B	0.003 m/s	0.02 m/s
C	0.003 m/s	0.02 m/s
$\sigma_v \equiv (A^2 + B^2 + C^2)^{1/2}$	0.10 m/s	0.32 m/s

Where does laser frequency stability enter in? The spectral width (full width at half maximum) of the pulse is $\delta\nu_{\mathrm{bw}} = 2\sigma_{v,\mathrm{bw}}/\lambda \geq (2\pi\tau_p)^{-1}$, where equality in the second relation holds only for a Fourier-transform-limited Gaussian pulse. In practice, the pulse spectral widths may be greater than $(2\pi\tau_p)^{-1}$. However, no reasonably efficient system would have the pulse spectral width greater than about $1/2\pi\tau_{\mathrm{corr}}$, where τ_{corr} is the correlation time defined by the spread of velocities in the medium, $\tau_{\mathrm{corr}} = \lambda/2\sigma_{\mathrm{med}}$; thus, one would choose pulse durations $\tau_p > \tau_{\mathrm{corr}}$. When $\tau_p = \tau_{\mathrm{corr}}$ and $\delta\nu_{\mathrm{bw}} = (2\pi\tau_p)^{-1}$, the velocity-measurement error due to the pulse spectral width alone, $\sigma_{v,\mathrm{bw}}$, is equal to the spread σ_{med} in target velocities. If the pulse spectral widths were allowed to be as great as $(2\pi\tau_{\mathrm{corr}})^{-1}$, they would be 800 kHz and 80 kHz for the 1-μm and 10-μm systems, respectively; note that in Table 5.4.2, a pulse duration $\tau_p = 3.2$ μs and a pulse spectral width of approximately 50 kHz were assumed for both systems. Spectral widths of 50 kHz require stabilization of the frequencies of the reference lasers being used to generate the pulses to a part in 10^{10} and 10^9, respectively, for the 1-μm and the 10-μm system. In contrast, spectral widths of 800 kHz and 80 kHz for the 1-μm and 10-μm systems, respectively, would require reference-laser frequency stability of only 3 parts in 10^9 for each of the systems. With these larger pulse spectral widths, the velocity-measurement precision

for the two systems would become comparable and on the order of 0.45 m/s, still dominated by the A term in eqs. (4.2.17). Note that the velocity-measurement precision σ_v in Table 5.4.2 is the best possible precision for the given pulse width $\tau_p = 3.2$ μs (set by the desired range resolution) so long as the number of detected photons per integration, N_d, remains of order unity or higher. Improving the reference-laser frequency stability to better than 50 kHz will not improve velocity-measurement precision unless it occurs along with an increase in pulse duration (and corresponding narrowing of the pulse spectral width), which may be unacceptable in view of the desired range resolution.

Thus, in general, the frequency stability of the laser being used to generate pulses is a key ingredient to improving velocity-measurement precision. Looking back at the forms of the three terms in the expressions (4.2.17) and (4.2.18) for the measurement precision, this is almost obvious. The B and C terms depend on the signal strength N_d, but they grow rather slowly as the signal strength drops. By comparison, the A and C terms grow much more quickly with increasing $\sigma_{\rm bw}$, or pulse spectral width. Except in situations where the backscatter cross section is extremely small, the primary contributions to velocity-measurement error come therefore from the pulse spectral width and the spread of velocities of the scattering particles. However, it also has been pointed out that velocity-measurement precision cannot be improved indefinitely because of the upper limit set on pulse durations by a desired range resolution. Since the pulse spectral width $\delta\nu_{\rm bw} \gtrsim (2\pi\tau_p)^{-1}$ and the pulse width τ_p is constrained by the desired range resolution σ_L to satisfy $\tau_p \lesssim 2\sigma_L/(Mc)$ [$M \geq 1$, eq. (4.2.14)], the fractional frequency stability of the reference laser need not be better than

$$\frac{\delta\nu_{\rm bw}}{\nu} \sim \frac{M\lambda}{4\pi\sigma_L} \simeq 8 \times 10^{-11} M \left(\frac{\lambda}{1\ \mu\text{m}}\ \frac{1\ \text{km}}{\sigma_L}\right). \tag{5.4.1}$$

The requirements outlined at the start of this chapter for a global wind-sensing lidar system appear to be well within reach of systems operating at either 1 μm or 10 μm, given fairly modest requirements on frequency stability. Even with modest pulse energies (say 40 mJ for an eye-safe 1-μm system), the primary contribution to measurement error will come not from the received

signal strength, but from the spectral width of the pulses and the spread in velocities of the aerosol particles. Measurement error can be reduced only by narrowing the spectral width of the pulses. This requires improved frequency stability of the reference laser, and possibly also pulses of longer duration.

5.5 Light-scattering experiments with planetary spacecraft

Much of our knowledge about the dynamical and internal properties of planets and their satellites, as well as their atmospheres, ionospheres, and rings, has been inferred from radio tracking and communications sytems data obtained with planetary spacecraft, especially orbiters and landers. In particular, these data have provided information on gravity fields; rotational properties (spin-axis orientation and motion, spin rate); shapes; surface dielectric constants; pressure, temperature, and density profiles for atmospheres and ionospheres; altitude profiles for electron-number density; and refined ephemerides. Some of these measurements and their techniques are described qualitatively in chapter 6 below, with suggestions for analogous experiments that could be performed with optical (laser) tracking and communications data. This section looks at three kinds of experiments involving measurement of propagation effects on highly coherent laser light transmitted from a spacecraft: (1) atmosphere-occultation experiments, (2) ring-occultation experiments, and (3) surface-scattering experiments. Where possible, quantitative descriptions of the expected performance of laser systems are indicated and compared with information obtained from radio-frequency experiments performed with Voyager at Saturn.

A related category of experiments *not* included here is the study of planetary ionospheres and planetary magnetic fields through occultation experiments. Such experiments attempt to determine electron-density profiles and characterize turbulence in ionospheres, and measure magnetic fields. Magnetic-field strengths are inferred directly through Faraday rotation of the polarization. The degree of rotation is proportional to the magnetic-field strength, the interaction length, and a

wavelength-dependent factor proportional to λ^{-2}. Magnetic-field orientation can be inferred indirectly by studying plasma irregularities, since the latter are expected to be aligned with the field; scintillation data are used to deduce the orientation and anisotropy of the component of the plasma irregularities perpendicular to the direction of propagation (Tyler 1987).

Studies of plasma with occultation experiments are done more easily at radio frequencies than at optical frequencies, because the effects of propagation through regions containing charged particles — phase scintillation caused by decrease of the group velocity and increase of the phase velocity — are more pronounced (see section 3.4.4). Typically, dual-frequency coherent radio links are used to study these wavelength-dependent effects. For example, the Voyager spacecraft transmitted both S- and X-band signals, wavelengths approximately 13 and 3.6 cm, respectively. Addition of a coherent laser link, on which these propagation effects would be negligible, could aid the calibration of the radio-frequency data.

5.5.1 Atmosphere-occultation experiments

Traditional radio-frequency occultation experiments measure the intensity, frequency, polarization, and group delay of radio signals transmitted from a spacecraft to Earth as the spacecraft is occulted by a planet or satellite. Given adequately powerful sources and sufficient data-gathering capability on board the spacecraft, these experiments could be performed in reverse, *i.e.*, in an uplink mode. Some of the advantages and problems of performing the experiments in an uplink mode—Earth transmitting a signal to a spacecraft as the spacecraft passes behind a planet or satellite—are addressed in subsection 5.5.2d. Typical objectives of atmosphere-occultation experiments are to determine temperature and pressure as functions of altitude in the stratosphere and troposphere, determine composition (*e.g.*, methane and helium abundances), and investigate turbulence and other irregularities. For bodies with thin atmospheres (such as Mercury, Mars, or Io), occultation data can be used to improve estimates of their radii (*e.g.*, Howard, *et al.* 1974).

Measurements of the Doppler shift of the transmitted signal due to passage through an atmosphere give direct information about the angle of refraction in the atmosphere. These Doppler shifts are calculated from the total measured Doppler shifts by removing effects due to the orbital motion of the spacecraft, the Earth's rotation, and any other predictable sources. Using the notation defined in Figure 20a, the angle θ between the transmitter (spacecraft) velocity $\vec{v}_t$ and the wave vector along the ray path from the transmitter to receiver (Earth) is related to the Doppler shift ν_D and the wavelength λ by

$$v_t \cos\theta \;=\; \lambda\, \nu_D \;. \tag{5.5.1}$$

The refractive bending angle $\psi(a)$ is obtained as a function of ray-asymptote distance a by combining the measured Doppler shifts with knowledge of the spacecraft velocity and the Earth-planet and planet-spacecraft orientations and separations. Measurements of the refraction angle $\psi(a)$ then are used to estimate the refractive index $n(h)$, or refractivity $\mu \equiv n - 1$, as a function of height h above some reference value. Unfortunately, it is not possible in general to infer a unique refractivity profile $\mu(h)$ from measurements of the refraction angle; also, the inversion can be very complicated (*e.g.*, Fjeldbo and Eshleman 1965). A unique inversion does exist if the atmosphere is spherically symmetric. However, in general, the atmosphere is modeled as successive, thin concentric shells of constant refractivity in which conditions of hydrostatic equilibrium are assumed, and the refractivity profile calculated by numerical iteration.

If independent information exists on the composition or mean molecular mass $\bar{m}$, the refractivity profile can be used to determine the gas number density $n_g(h)$ (Tyler 1987). Then, assuming conditions of hydrostatic equilibrium, the density can be integrated down from the top of the neutral atmosphere to yield a pressure profile $p(h)$:

$$p(h) \;=\; \bar{m} \int_h^\infty g(h')n_g(h')dh' \;; \tag{5.5.2a}$$

here $g(h')$ is the acceleration due to gravity. If an initial temperature is assumed at the top of

the atmosphere, the temperature profile $T(h)$ can be inferred from the pressure profile through the perfect gas law:

$$T(h) = \frac{p(h)}{k_B \, n_g(h)} \, . \tag{5.5.2b}$$

Alternatively, in the absence of information on composition, one can begin with an independent measurement of temperature at a known altitude, *e.g.*, through infrared sounding measurements. The combination of temperature and the measured refractivity profile yields a value for $\bar{m}$, thus constraining the possible composition. An important potential source of uncertainty is the unknown departure of atmospheric gases from ideal behavior, which could be significant in the outer solar system where temperatures are low. Occultations as close as possible to the middle of the planet's disk (in the plane of the sky) and at small spacecraft-planet separation (D in Figure 20a) are desired in order to probe as deeply as possible into the atmosphere. Where multiple occultations occur at different parts of the planetary disk, the collection of thermal-structure profiles provides information on the general circulation and dynamics in the atmosphere. Obviously, it is of considerable value to compare data obtained in this way to *in situ* measurements of pressure and temperature. Occultation techniques are especially useful for regions of the atmosphere that are too high for direct probing, but too low for infrared sounding (Kliore and Patel 1980).

The most important error sources in estimations of the refractivity profiles from Doppler shifts are inadequate knowledge of the local vertical direction and uncertainty or instabilities in the frequency of the spacecraft oscillator. Knowledge of the local vertical must come from measurements of the gravity field and rotation rate, as well as measurements of local circulation velocities. Frequency stability of the onboard oscillator is particularly important for data obtained from deep within the atmosphere during the exit phases of the occultation, when an uplink frequency reference is not available. Errors in the local vertical and fractional frequency of the onboard oscillator both contribute linearly to errors in a derived temperature T. To see this, start from the premise (not derived here; see, *e.g.*, Eshleman, *et al.* 1977 and references therein) that the inferred temperature

T depends on the component of the projection of the spacecraft velocity on the plane of the sky v_s that is along the local vertical direction (defined at the distant planet) — call it $v_\perp \equiv v_s \cos\phi$ — in the following way:

$$\ln T \;=\; \frac{1}{\eta}\,\ln {v_\perp}^2 \;+\; \text{constant}\;. \tag{5.5.3a}$$

Here the subunity factor η is the signal-intensity attenuation due to extinction and refraction effects; typically it ranges from 10^{-4} to 10^{-2}. Hence the fractional error in determination of the temperature T is related to the error in inference of $v_\perp$ by

$$\frac{\delta T}{T} \;=\; \frac{2}{\eta}\,\frac{\delta v_\perp}{v_\perp}\;. \tag{5.5.3b}$$

Uncertainty in $v_\perp$ comes both from uncertainties $\delta\theta_{\text{vert}}$ in the direction of the local vertical and from uncertainties in measurement of the spacecraft velocity made from Doppler measurements. The next few paragraphs address each of these sources of error.

Since $v_\perp \equiv v_s \cos\phi$, where ϕ is the angle between the local vertical and the spacecraft velocity in the plane of the sky, the fractional error in $v_\perp$ is related to the error $\delta\theta_{\text{vert}} = -\delta\phi$ in the local vertical direction by

$$\frac{\delta v_\perp}{v_\perp} \;=\; \tan\phi\;\delta\theta_{\text{vert}}\;. \tag{5.5.4a}$$

Hence the fractional error in determination of temperature T is

$$\frac{\delta T}{T} \;\simeq\; \frac{2}{\eta}\,\tan\phi\;\delta\theta_{\text{vert}}\;. \tag{5.5.4b}$$

This expression for sensitivity to errors in the local vertical is appropriate at the maximum penetration level, when refractive bending is greatest (Eshleman 1975); at smaller penetration levels, the temperature estimate is somewhat less sensitive to uncertainty in the local vertical. The factor of $\tan\phi$ will be assumed to be approximately unity here ($\phi \simeq 45°$) in order to give worst-case numerical

estimates for the errors. Thus, to achieve a 1% error in temperature in the presence of a signal-intensity attenuation $\eta \simeq 10^{-3}$, the local vertical direction must be known to about 5 microradian (μrad). This requirement scales with different values of η and fractional temperature error as

$$\delta\theta_{\text{vert}} \simeq 5\ \mu\text{rad} \left(\frac{\eta}{10^{-3}}\right) \left(\frac{\delta T}{T} \cdot 100\right) . \tag{5.5.4c}$$

Uncertainty in the local vertical direction tends to be greater for the large outer planets than for Venus and Mars. For example, differential rotation between the zones and belts on Jupiter can introduce deviations in the local vertical of up to 5 milliradian, about three orders of magnitude higher than the allowable uncertainty for a 1% error in temperature estimation. If not correctable, such an error would invalidate profile data in the lower atmosphere. Note that if independent measurements (*e.g.*, by infrared sounding at various altitudes) can provide a more accurate determination of the temperature, they could be used to infer information about variations in the local vertical and thus information about atmospheric currents or variations in the gravity field.

Fractional errors in temperature associated with uncertainties or fluctuations in the frequency of the spacecraft oscillator arise primarily from uncertainty in the frequency drift rate. The measured Doppler shifts exhibit a linear drift with time due to motion of the spacecraft behind the planet, which correspond to an apparent drift $\dot{\nu}$ in the spacecraft oscillator's frequency ν of order

$$\dot{\nu} \simeq \nu \frac{{v_\perp}^2}{cD} . \tag{5.5.5a}$$

Here D is the distance of the spacecraft from the center of the planet at occultation (see Figure 20a), and c is the speed of light. The distance D typically is 3 to 4 planet radii for the terrestrial planets, and around 10 planet radii for the giant planets. The projected velocity $v_\perp$ is of order 5 km/s. Hence the factor ${v_\perp}^2/cD$ can be on the order of 3×10^{-9} s^{-1} for Venus, and 10 to 30 times smaller for the outer planets. Uncertainties in the determination of temperature arise from uncertainties in $v_\perp$, and hence from uncertainties $\delta\dot{\nu}$ in the inherent drift rate of the oscillator frequency. Over

measurement times t_m, typically on the order of 5 to 10 minutes, the Doppler shift will change, and the oscillator frequency appear to drift by an amount $\Delta\nu \simeq t_m\dot{\nu}$. This change can be measured no more accurately than the oscillator frequency ν is known over the measurement time; *i.e.*, the error associated with measurement of $\Delta\nu$ is on the order of $\delta\nu$ or larger. Thus, the fractional error in temperature caused by uncertainties $\delta\dot{\nu}$ in the modelled drift rate, given a signal-intensity attenuation η, is limited by uncertainties in the frequency of the spacecraft oscillator in the following way:

$$\frac{\delta T}{T} \simeq \frac{1}{\eta}\frac{\delta\dot{\nu}}{\dot{\nu}} \simeq \frac{1}{\eta}\left(\frac{cD}{t_m {v_\perp}^2}\right)\frac{\delta\nu}{\nu} . \qquad (5.5.5b)$$

Given that values for the measurement time t_m and the projected velocity $v_\perp$ for occultations of the inner planets do not change appreciably for occultations of the outer planets, better frequency stability will be required for the outer (giant) planets, where the distance D is larger. The factor $cD/t_m {v_\perp}^2$ is on the order of 10^6 for Venus occultations, and can be 10 to 30 times larger for the outer planets. Thus, a 1% error in temperature determination typically would require a fractional frequency stability of 10^{-11}, and likely better than that for the outer planets. Note that these stabilities refer to long-term (300 to 600 seconds) components of the unmodelled drift; however, more rapid fluctuations in frequency of comparable magnitudes would affect the shape of temperature profiles, and also are undesirable.

Fundamental to reconstruction of the detailed structure of atmospheres from occultation experiments is knowledge of the index of refraction, obtained by measuring the angle ψ through which rays are bent as they travel from the spacecraft to an Earth-vicinity receiver (or *vice-versa*, if the experiments are conducted in an uplink mode). This angle of refraction (ψ) is inferred from measurements of the Doppler residual of the transmitted signal, defined as the Doppler shift after all known contributions such as spacecraft-receiver relative velocity, oscillator-frequency drift rates, and relativistic effects have been subtracted. The error $\delta\psi$ associated with its measurement therefore

depends critically on knowledge (or stability) of the spacecraft oscillator's frequency ν:

$$\delta\psi \simeq \frac{c}{v_s}\left(\frac{\delta\nu}{\nu}\right) \simeq 30 \text{ nrad} \left(\frac{10 \text{ km/s}}{v_s}\right)\left(\frac{\delta\nu}{\nu}\, 10^{12}\right). \tag{5.5.6}$$

Here v_s is the projection of the spacecraft velocity onto the plane of the sky (perpendicular to the receiver-spacecraft line-of-sight). The "Ultra Stable Oscillator" used on Voyager 1 exhibited a fractional frequency stability of 1 to 4 $\times 10^{-12}$ over short times (1 to 600 sec) and $\simeq 5 \times 10^{-11}$ over periods of about a day, thus enabling measurements of the bending angle to an accuracy of about 10 nanoradians (Tyler 1987). Since typical values for v_s will be on the order of tens of kilometers per second, it can be stated generally that calibration or control of fluctuations in the spacecraft oscillator's frequency to one part in 10^N enables detection of a change in angle of refraction on the order of 10^{4-N} radians. Note that this relation is independent of the frequency of the spacecraft oscillator – *i.e.*, it holds whether the spacecraft oscillator is a laser or a microwave transmitter.

Doppler data offer the best potential accuracy for measurements of the refractive bending angle. However, simpler intensity measurements can provide a useful independent (though generally less precise) check, and they also can provide additional information, such as the intensity attenuation factor η required for temperature determination. Signal intensity is reduced primarily by two effects related to propagation through an atmosphere: extinction due to scattering and absorption, and differential refractive defocusing. The latter arises because refractive bending is greater for rays that penetrate deeper into the atmosphere. Absorption and scattering effects can be described by an opacity or optical depth τ; they cause a fractional loss $e^{-\tau}$ in signal. They can be distinguished from refractive losses to some extent by their spectral characteristics, but the distinction is more straightforward if Doppler measurements are available to determine the refractivity profile. Extinction profiles at two frequencies can provide information on the location and density of the clouds or other absorbing material. The fractional intensity loss due to refractive effects depends on the depth of the occultation $D\psi$ relative to the scale height H of the atmosphere and the planet radius R (see

Figure 20a, and assume $\psi << 1$; R is taken to be that radius where the atmospheric pressure is on the order of 0.1–1 bar). For a spherically symmetric, tenuous exponential atmosphere, the ratio of the intensities of the occulted and free-space waves, or the attenuation factor η, is

$$\eta \simeq e^{-\tau} \left(\frac{1}{1 + D\psi/H} \right) \left(\frac{1}{1 - D\psi/R} \right) \tag{5.5.7a}$$

(Tyler 1987). Usually differential refractive defocusing dominates other refractive losses, *i.e.*, $D\psi/H >> 1 >> D\psi/R$ in this approximate expression. Hence the signal-intensity reduction can be approximated by

$$\eta \simeq e^{-\tau} \left(\frac{H}{D\psi} \right) . \tag{5.5.7b}$$

The use of intensity measurements (as opposed to Doppler measurements) to measure the refractive bending angle ψ is most reliable for an approximately lossless atmosphere ($\tau = 0$), in which case the angle ψ or the scale height H can be inferred directly from measurements of the signal attenuation [eq. (5.5.7b)]. If both intensity and Doppler data are used, the optical depth τ can be inferred. Then measurements of the variation of τ with height and wavelength can be used to put constraints on atmospheric composition.

Accurate estimation of the refraction angle from intensity measurements may be hindered severely in practice by antenna-pointing errors. For such measurements, it is critical that the spacecraft antenna follow the refracted direction of Earth and keep the virtual Earth at the peak of the antenna beam, where signal intensity is maximum and sensitivity to small pointing fluctuations is minimal. For planets such as Venus, where the maximum bending angle can be as large as 15° to 20°, the necessary pointing control can be difficult to achieve. For Mars, where the maximum bending angle is only about 0.05° ($\simeq$ 1 mrad), the requirement is satisfied easily at radio wavelengths (beamwidths $\lambda/D \simeq 10$ mrad), but still might pose a challenge at optical wavelengths (beamwidths $\lambda/D \simeq 1$ μrad). For the outer planets, where maximum bending angles are on the order of 0.5°, these pointing errors currently limit intensity measurements made at radio frequencies (S- and X-bands)

to accuracies of 1% or worse, inferior to measurements based on Doppler shifts (Tyler 1987).

Table 5.5.1 summarizes the results described on the preceding pages, and gives numerical values of the relevant parameters for several solar-system bodies. The last two columns show the required knowledge of the local vertical and of the spacecraft oscillator's frequency in order to achieve a 1% error (roughly) in determination of temperature. For the outer planets, actual variations in the local vertical direction can be large enough to make temperature determination to this accuracy (by these methods) difficult or impossible in practice. The assumed value of 0.1 radian ($\approx 5°$) for the maximum refraction angle of the outer planets is an order of magnitude larger than Pioneers 10 and 11 were able to measure before signal attenuation became too great (Anderson, *et al.* 1974a,b; Null, *et al.* 1975) and 2 to 3 times larger than that measured for Jupiter and Saturn by the Voyager spacecraft (Tyler 1987). Since measurements corresponding to deeper penetration may be possible at shorter wavelengths or with higher transmitted powers, the larger maximum refraction angle is used here to suggest a desirable fractional frequency stability. For lasers operating at 1 μm, the most stringent requirement on fractional frequency stability is about 10^{-13} for an occultation by Jupiter's atmosphere, which implies control (or calibration) of the laser frequency to within 30 Hz over time scales of several minutes. As discussed in chapter 2, space-qualifiable solid-state lasers exhibiting this degree of frequency stability with average output powers of 1 watt or more are fast approaching reality (Day, *et al.* 1990; Byer 1988).

Table 5.5.1 Requirements on spacecraft laser fractional frequency stability (uncalibrated fractional drift $\delta\nu/\nu$ over a 5-minute integration time) and knowledge of local vertical direction ($\delta\theta_{\text{vert}}$) to achieve 1% error in temperature determination from atmospheric occultation experiments [see eqs. (5.5.3–5)]. Signal-intensity attenuation is assumed to arise primarily from differential refractive defocusing [$\eta \simeq H/(D\psi)$; see eqs. (5.5.7)], and spherical symmetry is assumed. The occultation distance D (separation between spacecraft and center of planet) is taken to be 3 to 4 planet radii for Venus, Mars, and Titan, and 10 planet radii for the outer planets. Planetary radii are taken from Allen (1973) and Titan's radius is from Tyler (1987). Scale heights H for Jupiter, Saturn, Uranus, and Neptune are based on assumed blackbody temperatures of 200, 170, 150, and 130 K, respectively. The maximum refractive bending angles ψ were supplied by Anderson (private communication 1989). For the calculation of required oscillator frequency stability, a projected spacecraft velocity $v_\perp$ of 5 km/s was assumed; thus $\delta\nu/\nu \simeq 0.01\eta/40D$.

Body	R (10^3 km)	D (10^3 km)	H (km)	ψ (rad)	η (10^{-3})	$\delta\theta_{\text{vert}}$ (μrad)	$\delta\nu/\nu$ (10^{-11})
Venus	6.1	25	10	0.3	1	5	1
Mars	3.4	10	10	0.001	10^3	5×10^3	2500
Jupiter	71.3	700	20	0.1	0.3	1.5	0.01
Saturn	60.1	600	40	0.1	0.7	3.5	0.03
Titan	2.6	10	25	0.05	50	250	125
Uranus	24.5	240	45	0.1	1.9	9.5	0.2
Neptune	25.1	250	25	0.1	1	5	0.1

Several other types of information can be obtained from atmospheric occultation data. Pressure profiles $p(h)$ over a range of altitudes h are needed to calculate chord lengths in the plane of the sky between specified pressure levels and infer the planet's shape at different pressure levels. The shape reveals information about atmospheric circulation and about the planet's interior structure. Occultation data from Voyager at Saturn were used in this way to conclude that Saturn winds persist to atmospheric depths of at least 10 scale heights (Tyler 1987).

Turbulence and other irregularities in planetary atmospheres can be studied by monitoring rapid fluctuations in signal amplitude (scintillation). This requires care in the removal of background effects, such as the average background atmospheric refractivity. These background effects cause variations in diffraction scale size with occultation depth, which in turn cause variations in signal amplitude. Internal gravity waves and other small-scale atmospheric structure can be de-

tected through comparison of observed intensity variations with predictions based on the observed or inferred vertical atmospheric structure.

5.5.2 Ring-occultation experiments

Ring-occultation experiments attempt to determine the following information: radial structure of the ring system; vertical structure of the rings (*e.g.*, whether they are cloudlike or single-layered); total amount of material in the rings; sizes, size distribution, and number density of the ring particles; and constraints on particle shapes, composition, and orientation. This information is inferred from measurements of the complex extinction (change in amplitude and phase of a signal propagating through the rings) and the forward-scattering cross section as functions of wavelength, polarization, and radial distance from the planet (or other occulting body). As in atmosphere-occultation experiments, at radio frequencies this is accomplished traditionally by transmitting two wavelengths coherently from the spacecraft to the ground, and recording the phase, intensity, and polarization of each. (Discussion of performing such an experiment in an "uplink" mode instead appears in subsection 5.5.2d below.) The received signals at each wavelength are of two types: (1) a coherent component consisting of light that has come directly from the spacecraft and which is attenuated in intensity and possibly phase-shifted because of differential refraction, but which retains its character as a quasimonochromatic sinusoidal signal; and (2) a Doppler-broadened incoherent component consisting of light that has been scattered into the line of sight by ring particles larger than the wavelength, the strength and spectral shape of which depend on the near-forward-scattering cross section as well as the occultation geometry.

This subsection describes the simplest models and some of the basic techniques that have been used in conjunction with frequency-stable radio-frequency transmitters on planetary spacecraft to obtain information about rings. These descriptions are applicable to the use of laser transmitters as well, and they will be so applied throughout this section. The techniques involve measurements of the

phase and intensity of the direct (coherent) signal, as well as spectral characteristics of the scattered (incoherent) signal. A description of the occultation geometry is given first, together with a simple model for the ring particles. Next, the observables associated with the coherent and incoherent signals are defined. Limitations to the amount and accuracy of information are discussed, arising from oscillator-frequency fluctuations, pointing errors, and photon statistics (or receiver thermal noise at radio frequencies).

The typical geometry for ring-occultation experiments is shown in Figure 20b. (See figure caption for elaboration.) In traditional experiments, the transmitter is on the spacecraft, near a ringed planet, and the receiver is on Earth. Throughout this discussion it is assumed that the ring particles move in circular orbits about the planet with mean velocities $v(r_p) = (GM/r_p)^{1/2}$, where M is the planet mass, r_p is the distance of the ring particles from the planet center of mass, and G is the universal gravitational constant. For Saturn ring particles at 4 Saturn radii, or about 240,000 km, this velocity is about 12 km/s. Since Saturn ring characteristics are known best from Voyager observations, they will be used frequently for illustration.

For simplicity, the ring particles can be modeled as dielectric spheres whose radii a are distributed over a range of values. Physically, the maximum particle size is determined by collisional fragmentation; sizes as large as several tens of kilometers are possible at Saturn (Marouf, *et al.* 1982), although observations indicate a predominance of particles in the centimeter-to-meter size range. The minimum particle size is determined primarily by radiation-pressure forces, which sweep the ring clear of objects smaller than a critical size. The distribution of particle sizes depends on the physical processes that produced the ring particles, as well as their evolutionary history. One of the simpler distributions used to model particle sizes is a power-law distribution $p(a) \propto a^{-q}$ for a in some range $(a_{\min}, a_{\max})$, with $p(a) \equiv 0$ outside this range. Power-law indices between 3 and 4 are consistent with a model in which rings are formed from the fragmentation of larger bodies. Smaller power-law indices (0 to 2) would be consistent with formation by condensation from smaller parti-

cles (Alfvén and Arrhenius 1976). In practice, more complicated distributions often are used (*e.g.*, Cuzzi and Pollack 1978). It is assumed here that the particles scatter independently, *i.e.*, that they are widely separated from one another. Equivalently, the fractional volume of the rings occupied by particles is much less than unity: $n_p 4\pi\langle a^3\rangle/3 << 1$, where n_p is the particle number density, a the average effective radius of the particles, and angle brackets denote a statistical average. This assumption of a small "volume fraction" is supported by observations; for example, volume fractions of order 1% or smaller are inferred from observations of Saturn's rings.

A radially-varying effective refractive index $n(r)$ or refractivity $\mu(r) = n(r) - 1$ is defined below [eqs. (5.5.9) and (5.5.10)]. It permits a useful analogy between ring-occultation experiments and atmosphere-occultation experiments, such as a description of the phase shift exhibited by the coherent component of the received signal in terms of differential refraction. This quantity will be used here to motivate certain results whose derivations are beyond the scope of this report.

5.5.2a Coherent (direct) signal

The direct, or coherent, component of the received light gives a detailed map of the distribution of particle sizes for particles smaller than the wavelength of the transmitted light, as well as information about the total amount of material traversed by the light. As the transmitted light passes through the ring material, it is attenuated because of scattering and absorption and bent by differential refraction effects associated with a nonuniform radial density of ring particles. The latter effect produces a phase shift with respect to the incident light (see Figure 21). The complex amplitudes E_c and E_0 of the received coherent signal and the transmitted signal, respectively, are related by

$$E_c = E_0\, e^{i\phi_0}\, \xi^{1/2}\, e^{-\tau_c/2}\, e^{i\phi_c}\ . \tag{5.5.8}$$

The factor ξ accounts for a loss (or gain) in signal power due to defocusing (or focusing) of the emerging wave, a consequence of the bending due to differential refraction. The "oblique" opacity

or optical depth, τ_c, represents the loss of energy from a light beam along its path from transmitter to receiver due to scattering and absorption by ring particles. The phase shift ϕ_0 arises from predictable Doppler frequency shifts due to effects such as the relative velocities of the transmitter, ring particles, and receiver, known oscillator frequency drift rates, and special and general relativistic effects. The phase shift ϕ_c is an additional phase shift experienced by the direct ray that would *not* arise if the ray were traversing free space (defined below). The ordinary opacity $\tau \equiv \tau_c \sin\theta_0$ and phase shift $\phi \equiv \phi_c \sin\theta_0$ describe the complex extinction that would be measured if the track of the transmitted ray were perpendicular to the ring plane, *i.e.*, if the ring plane were coincident with the plane of the sky as the spacecraft passed behind the rings. The angle θ_0 is the angle between the ring plane and the incident probe beam, commonly referred to as the "ring opening" (see Figure 20b). Ideally, one would like to arrange for the angle θ_0 to be 90°. Here convention will be followed by referring to the opacity τ and phase shift ϕ, but the reader is reminded that the quantities measured directly are τ_c and ϕ_c.

As noted above, the overall effect on the amplitude and phase of the direct ray due to propagation through the ring plane can be described in terms of an effective refractivity μ or refractive index $n \equiv \mu + 1$, by analogy with the analysis of atmosphere-occultation experiments. The imaginary part of the effective refractivity, μ_I, is inferred from measurements of the opacity τ through the relation

$$\tau = \frac{4\pi}{\lambda} \int_{-\infty}^{\infty} \mu_I(z)\, dz \,, \tag{5.5.9a}$$

where z denotes the direction perpendicular to the plane of the rings. If the refractivity were uniform across the rings so that no bending of the direct ray occurred, then the real part of the effective refractivity, μ_R, could be inferred in a similar way from measurements of the phase shift ϕ. Typically, the effective refractive index varies along the track followed by the transmitted rays (spacecraft) behind the ring plane, causing the phases of adjacent rays to be shifted relative to one another and the wave front of the transmitted ray to arrive slightly bent. The bending angle ψ is

related to the radial derivative of the real part of the refractivity by

$$\psi \simeq \int_{-\infty}^{\infty} \partial_r \mu_{\rm R}(\vec{r}_0, z) \frac{dz}{\sin\theta_0} . \qquad (5.5.9b)$$

The bending produces an additional change in the phase shift ϕ, and it also may contribute to a loss or gain in received signal power due to defocusing or focusing of the transmitted signal [the factor ξ in eq. (5.5.8)].

The phase shift ϕ is inferred from integration of the measured Doppler residuals ν_D, typically obtained with a closed-loop receiver that tracks the frequency of the coherent signal. Refractive bending through an angle ψ produces a contribution $\nu_D(\psi)$ to the Doppler residual:

$$\nu_D(\psi) \simeq v_s \psi/\lambda , \qquad (5.5.9c)$$

where v_s is the component of the spacecraft velocity in the plane of the sky [*cf.* eq. (5.5.6)]. The remainder of the Doppler residual can be described as arising from the real part of the effective refractivity, $\mu_{\rm R}$. The phase shift ϕ is thus

$$\phi \equiv \sin\theta_0 \int 2\pi\nu_D \, dt \simeq \frac{\pi D\psi^2}{\lambda} \sin\theta_0 + \frac{2\pi}{\lambda} \int_{-\infty}^{\infty} \mu_{\rm R}(z) \, dz , \qquad (5.5.9d)$$

where D is the distance between the transmitter and the point of intersection with the ring plane (Figure 20a; van de Hulst 1957; Marouf, *et al.* 1982). Thus, the inference of an effective refractivity profile $\mu(r)$ for the rings is analogous to that for atmospheres or ionospheres: measurements of Doppler residuals are used to calculate the bending angles $\psi(r)$, the residuals are averaged over time to calculate the phase shifts ϕ, and then this information is combined with measurements of signal-intensity extinction (opacity τ).

The direct ray is coherent with the transmitted signal because individual scattered waves very close to the forward direction interfere with each other coherently. The mathematical description for

this forward-scattered light is identical to that for a light beam scattered from a dielectric medium of refractive index n occupying the ring volume (Rayleigh 1899). The equivalent refractivity μ is related to the wavelength of the transmitted light λ, the average number density of particles n_p, and the average (dimensionless) forward-scattering amplitude A_0 by

$$\mu \simeq \frac{\lambda^3}{(2\pi)^2} n_p A_0 . \tag{5.5.10a}$$

(The dependences of μ, n_p, and A_0 on position in the ring plane have been suppressed.) In the absence of ray bending and for a 90° ring opening, the complex extinction is related in a simple way to the column density $n_c \equiv n_p d$ (where d is the thickness of the ring plane), the rms particle radius a, the wavelength λ, and the forward-scattering amplitude A_0:

$$\phi + i\frac{\tau}{2} \equiv \frac{2\pi}{\lambda} \int \mu(z)\, dz = \frac{\lambda^2}{2\pi} n_c A_0 . \tag{5.5.10b}$$

In the geometric-optics limit, where the average particle size a is much *greater* than the reduced wavelength $\lambda/2\pi$ of the illuminating beam, the forward-scattering amplitude $A_0 \simeq i\, 2\pi^2 a^2/\lambda^2$ (van de Hulst 1957), so there is negligible phase shift ($\phi \approx 0$), and the opacity τ is proportional to the average effective geometrical cross section of the particles, $\tau \approx 2\pi a^2 n_c$. The geometric-optics limits would be appropriate in most situations for laser transmitters (since wavelengths λ are on the order of 1 μm), so the optical depth τ would have this simple mathematical form and physical interpretation. Since the majority of particles in the rings of solar-system objects are not larger than several centimeters, the "radio" depth does not have this simple relation to the average geometrical cross section of the particles, and the opposite limit — the Rayleigh scattering limit — often applies. In that limit, where the average particle size a is much *smaller* than the reduced wavelength, if the particles are lossless (nonabsorbing), the opacity scales as a^6 and the phase shift scales as a^3. The opacity and phase shift are given more precisely by the Rayleigh formulas (Rayleigh 1899), which

include an equivalent refractive index n:

$$\tau = \frac{8\pi}{3}\left(\frac{n^2-1}{n^2+2}\right)^2 n_c\, a^6 \left(\frac{2\pi}{\lambda}\right)^4 \simeq 2.3\times 10^3\, n_c\, \frac{a^6}{\lambda^4}\,, \tag{5.5.11a}$$

$$\phi = \frac{4\pi}{3}\left(\frac{n^2-1}{n^2+2}\right) n_c\, a^3 \left(\frac{2\pi}{\lambda}\right) \simeq 11\, n_c\, \frac{a^3}{\lambda}\,. \tag{5.5.11b}$$

The last expressions apply to water ice with $n \simeq 1.78+i0.0$. Physically, the coherent-signal amplitude and phase constitute two independent pieces of information, whose combination allows inference of both the column density n_c and the average particle size a. If the measurements can be made at two different wavelengths (*e.g.*, S and X bands as done with the Voyager spacecraft), additional information can be inferred about the particle absorptivity, and hence composition.

The Rayleigh-limit behavior of lossy particles such as silicates with $n \simeq 2.32 + i0.03$ is quite different from that of lossless particles. Both τ and ϕ are proportional to a^3/λ (Campbell and Ulrichs 1969). In particular, when written in terms of the real and imaginary parts of the refractive index, $n \equiv n_{\rm R} + in_{\rm I}$, they take the form

$$\tau = \frac{8\pi}{3}\left[\frac{9n_{\rm I}n_{\rm R}}{|n^2+2|^2}\right] n_c\, a^3 \left(\frac{2\pi}{\lambda}\right) \simeq 0.6 n_c\, \frac{a^3}{\lambda}\,, \tag{5.5.12a}$$

$$\phi = \frac{4\pi}{3}(n_{\rm R}-1)\left[\frac{3}{2}\left|\frac{n+1}{n^2+2}\right|\right] n_c\, a^3 \left(\frac{2\pi}{\lambda}\right) \simeq 23.4\, n_c \frac{a^3}{\lambda} \tag{5.5.12b}$$

(Marouf, *et al.* 1982). In this situation, τ and ϕ are not independent quantities and cannot be combined to give complete information about the column density and average particle size.

Figure 22 shows curves relating τ and ϕ to the ratio a/λ. (See figure caption for further description.) These curves show quantitatively what was stated qualitatively above: for scattering from particles smaller than the transmitted wavelength, the optical depth and phase of the coherent wave depend strongly on particle size. In this regime, measurements of intensity and phase of the coherently transmitted ray provide detailed information on particle size and distribution. For

scattering from particles larger than the transmitted wavelength (the geometric-optics limit), the phase shift is zero for all particle sizes, and the opacity is proportional to the column density and geometric cross section of the particles. Inferences about particle sizes are discussed below, after discussion of the incoherent signal.

The direct ray is associated with a physical ring area approximately equal in area to the ellipse of the first Fresnel zone (Marouf, *et al.* 1982), with semimajor and semiminor axes of dimensions $(\lambda D)^{1/2}/\sin\theta_0$ and $(\lambda D)^{1/2}$, respectively (see Figure 23). Recall that θ_0 is the ring opening and D is the distance of the spacecraft behind the planet during occultation. For best resolution in mapping the radial structure of the particles in the rings, this ellipse would be lined up with the rings as shown in Figure 23, so that the semimajor axis of the ellipse falls along the projection on the ring plane of the spacecraft-receiver line of sight, and the spacecraft motion and semiminor axis are along the radial direction from the center of the planet. At Saturn ($D \sim 240{,}000$ km $\simeq 4$ Saturn radii), X-band measurements of the direct ray (τ and ϕ) gave ring widths to a resolution of ~ 6 kilometers for this optimum geometry. The analogous resolution obtainable with a laser ($\lambda = 1$ μm) ring-occultation experiment at Saturn could be ~ 30 meters. For ring widths $W >> (\lambda D)^{1/2}$, the ringlets behave locally as a plane slab, and ray-bending effects are negligible. For ring widths comparable to or smaller than the first Fresnel zone [$W \lesssim (\lambda D)^{1/2}$], ray-bending (focusing or defocusing of the beam) can become significant for scattering from particles much smaller than the wavelength. This ray-bending can be seen both in amplitude data by virtue of the gain or loss associated with focusing or defocusing effects, and in frequency data, where it will cause a shift in measured frequency (hence phase shift ϕ) for the direct ray. The absence of such frequency shifts in the data would indicate negligible ray bending and thus a predominance of particles larger than the wavelength. The Fresnel-zone width $(\lambda D)^{1/2}$ is the resolution with which contributions of an individual ringlet to the received signal can be isolated and is therefore also the resolution with which the radial profile of the optical depth and the scattering cross section of individual resolved ringlets can be reconstructed.

Contributions from individual ringlets also can be identified in contour maps of constant Doppler frequency shift obtained from spectral measurement of the incoherent signal, although typically with poorer resolution (see Figure 24 and discussion below).

5.5.2b. Incoherent (scattered) signal

Next, consider the scattered or incoherent signal. The scattered ray arrives with a range of frequencies relative to the frequency of the direct ray, depending on the velocities $\vec{v}_p$ of the scattering particles (see Figure 21). In the notation of Figure 20b, the classical Doppler shift ν_D between the direct ray arriving from a particle at point O on the ring plane and a ray that has been scattered into the receiver's line of sight from a particle at point P with velocity $\vec{v}_p$ is

$$\nu_{\rm D} \equiv \frac{1}{\lambda}\,(\vec{v}_t - \vec{v}_p)\,\cdot\,(\hat{r}_{or} - \hat{r}_{pr}) \;=\; \frac{4\sin^2\beta}{\lambda}\,(\vec{v}_t - \vec{v}_p)\,\cdot\,\hat{r}_{op}\;, \tag{5.5.13}$$

where $\hat{r}_{or}$ and $\hat{r}_{pr}$ are unit vectors pointing to the receiver from points O and P, respectively; $\hat{r}_{op}$ is a unit vector pointing from O to P; and $\beta << 1$ is the angle at the receiver between the direct and scattered rays.

The measured power spectral density $S(\nu)$ of the incoherent signal can be used to infer the near-forward scattering cross section $\sigma(\beta, \vec{r}_p)$, a function of the small angle β and the scattering-particle position in the ring plane. Azimuthal symmetry will be assumed, so $\vec{r}_p$ will be replaced by $r_p \equiv r$. In general the inference is not straightforward, because the incoherent signal is a weighted average over an extended ring area defined by the width of the illuminating beam. The power spectral density of the incoherent signal can be shown to have the form of a "Fredholm equation of the first kind" (Marouf, *et al.* 1982):

$$S(\nu) \;=\; \int dr\;\frac{\sigma(\beta;r)\;G^2(\beta)}{2\;{r_{pr}}^2}\;\delta_{\rm D}[\nu - \nu_{\rm D}(r)]\;; \tag{5.5.14a}$$

Here r_{pr} is the distance between the scattering particle and the receiver, $\nu_{\rm D}$ is the deterministic Doppler shift given by eq. (5.5.13), $\delta_{\rm D}[\ldots]$ denotes the Dirac delta function, and $G^2(\beta)$ is the

antenna power gain in the scattered direction $\hat{r}_{pr}$, *i.e.*, at a small angle β relative to the direct ray (see Fig. 20b). The cross section σ used here is dimensionless and is defined as the effective scattering cross section per unit physical area of the rings. Note that the antenna power gain determines the ring volume that contributes to $S(\nu)$. In general, this spectral density must be convolved with a probability-density function for the random component of the total Doppler shift caused by a random component to the particle velocities over and above the Kepler velocity; however, in practice, this random component of the particle velocity typically is 10^{-6} to 10^{-7} of the Keplerian component (Tyler 1987), so it can be neglected. It is useful to rewrite expression (5.5.14a) for the spectral density of the received incoherent signal as a contour integral:

$$S(\nu) \;=\; \oint_C dl\, \frac{\sigma(\beta;r)\; G^2(\beta)}{2{r_{pr}}^2}\, \frac{1}{|\vec{\nabla}\nu_{\rm D}(r)|}\;, \tag{5.5.14b}$$

where $\vec{\nabla}\nu_{\rm D}(r)$ is the two-dimensional gradient (with respect to scattering particle position) of the deterministic Doppler shift defined above, C is the contour of constant Doppler shift on which $\nu = \nu_{\rm D}$, and dl is the element of arc length along C. Thus, the area bounded by any two contours ν and $\nu + d\nu$ of a Doppler map (defined as a collection of contours of constant Doppler frequency) identifies the physical portion of the ring(s) that contributes a power $S(\nu)d\nu$ to the total received incoherent power. Ideally, one would like the Doppler frequency to be constant for a given radial distance r over the illuminated portion of the ring. Then the contours of constant Doppler frequency would be aligned with those of constant cross section σ, and expressions (5.5.14) could be used to recover the cross section σ from the measured Doppler spectra (see Figure 24). For a given ring opening, this is accomplished by optimizing the spacecraft trajectory. Of course, for a given spacecraft trajectory, ideal alignment cannot occur globally; it will occur only for a narrow stretch of illuminated ring area along the projected line of sight, the area over which the kernel $G^2/2{r_{pr}}^2$ is maximum.

As noted earlier, the Doppler maps also enable isolation of the contribution from individual ringlets, though typically with poorer resolution than achievable with the coherent signal. If the

forward-scattering cross section σ is uniform over the illuminated area of the rings, and the Doppler contours are aligned with contours of constant radial distance (as indicated in Figure 24), each Doppler frequency would correspond uniquely to a particular radial distance from the planet center of mass. Any deviation from ideal alignment causes the contribution to the spectral power of the incoherent signal of any isolated ringlet to fall below what it would be in the case of ideal alignment. The narrower the ringlet, the greater is this decrease in spectral power, and the narrower is the frequency range over which the ringlet contributes to the received power spectrum $S(\nu)$. The spectral power for a ringlet of a given width (narrower than the illuminated beam width) must be computed using expression (5.5.14b), with some assumption for the radial profile of the ringlets (*e.g.*, Gaussian). Voyager 1 ring-occultation experiments at Saturn were able to achieve a resolution of a few hundred kilometers in this way, with a 400-km-wide ringlet showing about a 1-dB drop in its contribution to the power spectrum relative to what would have been expected from a uniform scattering cross section over the illuminated portion of the rings (Marouf, *et al.* 1982; Eshleman, *et al.* 1977).

The width and shape of the spectra [graphs of $S(\nu)$ *vs.* ν] for a given illuminated ringlet are determined by the narrower of the antenna beam width at the rings [equivalently, the gain G^2 in eqs. (5.5.14)] or the scattering cross section σ. For particle sizes much smaller than the spacecraft antenna (diameter D_t), the forward-scattering cross section σ is dominated by diffraction effects and has a characteristic angular width $\sim \lambda/2a$. In this case the scattering is nearly isotropic over the narrower angular width $\sim \lambda/D_t$ of the transmitted beam, so the shape of the spectral envelope is that of the transmitted beam. In this regime, little information is obtained about particle size distributions. For particles larger than the transmitter ($2a > D_t$), the spectral envelope is determined by the forward-scattering cross section of the particles. For particles whose sizes are comparable to or larger than the transmitter antenna, there is a one-to-one relation between particle size and observed envelope width. Under these circumstances it can be shown (Tyler 1987) that the largest

particles have the greatest influence on the shape of the spectral envelope. The observed envelope reveals more than just an average equivalent size for the scattering particles; information also can be gained about the first few moments of an unknown size distribution or about the parameters of an assumed distribution (Marouf, *et al.* 1982). Inference of this information and limitations on obtainable information are discussed further below.

5.5.2c Measurement errors

Consider now the limits to information due to measurement errors. The primary error sources are frequency fluctuations of the transmitting oscillator, antenna-pointing errors, and photon statistics (or receiver thermal noise at radio frequencies). Oscillator frequency fluctuations limit the accuracy with which the phase shift ϕ of the coherent component of the transmitted beam can be measured. Antenna-pointing errors make it difficult to distinguish scattering-induced changes in received intensity from changes induced by pointing fluctuations, and therefore lead to a required minimum opacity τ. Photon statistics, on the other hand, set a maximum opacity τ beyond which the coherent signal will be attenuated so much by ring scattering that it will not be detectable. These constraints translate into bounds on the measurable scattering cross section σ and thence on the measurable ranges of particle column density n_c and size a.

The precision with which the phase shift ϕ of the coherent component of the transmitted light can be measured is limited by the spacecraft oscillator frequency stability $\delta\nu$ over a measurement time t_m as follows: $\delta\phi \geq 2\pi\ t_m\ \delta\nu$. The Ultra Stable Oscillator on Voyager 1 at Saturn exhibited a short-term ($\lesssim$ 1000 sec) frequency stability of 1 to 4 $\times 10^{-12}$, or $\delta\nu \sim 10$ mHz at X band (Eshleman, *et al.* 1977). This would cause a phase-measurement error of at least one cycle for measurement integration times longer than about 100 seconds. Comparable fractional frequency stability for a laser oscillator would require a linewidth $\delta\nu \lesssim 300$ Hz, resulting in phase-measurement ambiguities of at least one cycle over measurement times longer than a few milliseconds. Accurate measurement

and monitoring of the phase shift ϕ over integration times of seconds or more would require extremely stable laser transmitters, with $\delta\nu/\nu \lesssim 10^{-15}$.

Fluctuations $\delta\phi$ in the measured phase shift result from fluctuations δn_c in the column density of a slab of ring particles:

$$\delta\phi \;=\; \frac{\lambda^2}{2\pi}\,\mathrm{Re}(A_0)\;\delta n_c \tag{5.5.15}$$

[eqs. (5.5.10)]. A measurement integration time t_m corresponds to a sampled distance $\Delta r = v_r t_m$, where v_r is the component of the spacecraft velocity parallel to the ring plane, here assumed perpendicular to the Earth-spacecraft line of sight. Hence the smallest detectable gradient $\delta n_s(r)/\Delta r$ in column density (assuming Rayleigh scattering by particles of size $a << \lambda/2\pi$) is related to the spacecraft oscillator frequency stability $\delta\nu/\nu$ by

$$\frac{\delta n_s(r)}{\Delta r} \;\geq\; \frac{4\pi^2 c}{v_r}\left(\frac{\delta\nu}{\nu}\right)\frac{1}{\lambda^3\,\mathrm{Re}(A_0)} \;=\; \frac{c}{v_r}\,\frac{3}{4\pi a^3}\left(\frac{n^2+2}{n^2-1}\right)\left(\frac{\delta\nu}{\nu}\right) \;\simeq\; 0.6\,\frac{c}{v_r}\,\frac{1}{a^3}\left(\frac{\delta\nu}{\nu}\right), \tag{5.5.16}$$

where the last relation holds for lossless particles of water ice with $n \simeq 1.78$ [eqs. (5.5.10) and (5.5.11)]. If the spacecraft velocity $v_r \simeq 10$ km/s and the average particle size $a \approx 1$ cm, this gradient would be on the order of 2.5×10^{-3} cm^{-2} km^{-1} for a fractional frequency stability $\delta\nu/\nu \simeq 10^{-11}$. As noted above, the extinction caused by ring scattering must be great enough that the effects of ring scattering can be distinguished from intensity fluctuations caused by unstable pointing of the transmitting antenna toward the receiver. If the minimum detectable change in received power due to pointing errors is 1%, the corresponding lower bound on the oblique optical depth τ_c (opacity integrated along the light-travel path) is about 0.01, since $e^{0.01} \simeq 1.01$. Ring material of greater thickness will attenuate the coherent signal power more and more, until it is no longer detectable. The upper bound on measurable opacity is determined by how weak a received signal can be detected by the receiver. This is determined by photon statistics at optical frequencies and by receiver thermal noise at radio frequencies. Consider again the Voyager spacecraft at Saturn. The receiving antenna (on Earth) was 64 meters in diameter with a system noise temperature of 25 K, so that

$k_B T \simeq 3.5 \times 10^{-22}$ J. A transmitted power of 20 W at X band from a 3.7-m spacecraft antenna at Saturn (10 AU or 1.5×10^9 km), assuming an overall efficiency factor of 10%, would result in a received power of 3.8×10^{-17} W [$P_r = P_t \eta (D_t D_r / \lambda r)^2$], about 50 dB above the receiver noise level for a 1-s integration time. A detectability threshold of 6 dB then sets the upper bound on the measurable oblique opacity at about 10 ($e^{-\tau_c} \geq 4 k_B T / P_r$). This upper bound scales with transmitter power P_t, overall efficiency η, transmitter and receiver antenna diameters D_t and D_r, distance r between transmitter and receiver, receiver temperature T, and wavelength λ as follows:

$$\tau_c \lesssim 10 + \ln\left[\frac{P_t}{20\ \mathrm{W}} \frac{25^\circ K}{T} \frac{\eta}{0.1} \left(\frac{D_t}{3.7\ \mathrm{m}} \frac{D_r}{64\ \mathrm{m}} \frac{10\ \mathrm{AU}}{r} \frac{3.6\ \mathrm{cm}}{\lambda}\right)^2\right] . \tag{5.5.17a}$$

For Voyager 1 at Saturn, the lower and upper bounds on the ordinary opacity $\tau \equiv \tau_c \sin\theta_0$ were on the order of 0.001 and 1, respectively, the latter consistent with the above parameter values and a ring opening θ_0 of approximately 6°. For Voyager 2 at Uranus, the rings were almost fully open with $\theta_0 \simeq 82^\circ$ (Tyler, *et al.* 1986; Tyler 1987). The corresponding upper limit on measurable ordinary or oblique opacity for the Uranus ring-occultation experiments ($r \sim 20$ AU) was about 9.

For comparison, consider a 10-m optical receiving telescope limited primarily by photon statistics. A power of 1 W at 1-μm wavelength transmitted through a 1-m telescope on a spacecraft at Saturn, assuming an overall efficiency factor of 0.1, would result in a received power of about 4.4×10^{-12} W. With $h\nu \simeq 2 \times 10^{-19}$ J and assuming Poisson photocount statistics, this corresponds to a signal-to-noise ratio of about 37 dB in 1 second. Assuming a detectability threshold of 6 dB, this sets an upper bound on the oblique opacity of about 7, which scales with parameters as

$$\tau_c \lesssim 7 + \ln\left[\frac{D_t}{1\ \mathrm{m}} \frac{D_r}{10\ \mathrm{m}} \frac{10\ \mathrm{AU}}{r} \left(\frac{P_t}{1\ \mathrm{W}} \frac{\eta}{0.1} \frac{\lambda}{1\ \mu\mathrm{m}}\right)^{1/2}\right] . \tag{5.5.17b}$$

With more modest parameters such as a laser transmitter power of 100 milliwatts, a 1-m receiver aperture and 30-cm transmitter aperture, and a 2% overall efficiency, the upper bound on measurable opacity would be about 1.5.

To translate these constraints on the complex extinction (τ and ϕ) into bounds on the particle column density n_s and effective particle radius a, return to expressions (5.5.11) for the complex extinction. Assume here for illustration that the particles all are roughly the same size, and that their composition is lossless water ice with a refractivity $\mu \approx 0.78$. Figure 25 is a plot of the column density against the ratio of particle radius to wavelength for two different wavelengths: 3.6 cm (X band) and 1 μm (optical). Four curves are drawn on each plot: one is the curve for $\phi = 2\pi$, which exists only for $a/\lambda << 1/2\pi$ (recall that as the geometric-optics limit $a \gtrsim \lambda/2\pi$ is approached, $\phi \rightarrow 0$). The second and third bound a region defined by $0.001 \leq \tau \leq 1$. A fourth curve delineates a region corresponding to detection of the incoherent signal derived from specific constraints on SNR for given integration time, and sets a lower bound on the forward-scattering cross section (see figure caption for more detail). The changing slopes for the two opacity curves reflect the different dependences on particle size in the Rayleigh and the geometric-optics limits, as described above. If the lower and upper bounds on opacity for detection of a coherent signal are 0.001 and 1, as was the case for Voyager 1 at Saturn, then one can identify the different regions on these plots as follows: Above the $\tau = 1$ curve, the coherent signal has been scattered so strongly that it is too weak to be detected by the receiver. Below the $\tau = 0.001$ curve, the coherent signal differs so little from its "free- space" value (the signal reaching the receiver that did not pass through the ring material) that it cannot be distinguished from intensity changes due to fluctuations in the pointing of the spacecraft transmitter toward the receiver. Between these two curves, in the lined region, the coherent signal can be detected, and the measured amplitude information gives information about particle size, particularly for particles much smaller than the reduced wavelength of the light being transmitted. For these small particles, the phase shift of the coherent wave due to differential refraction effects can be measured provided it is on the order of one cycle or larger, which corresponds to the region above the $\phi = 2\pi$ curve. Both phase changes and amplitude changes are detectable in the intersecting region marked by cross-hatching. If the particles are lossy (*e.g.*, silicates with refractive index $n = 2.32 + i0.03$), the Rayleigh-limit behavior differs from that shown in Figure 25

in that the opacity curves rise less steeply with decreasing particle size [see eqs. (5.5.12)].

For scattering by particles whose sizes are comparable to the reduced wavelength, neither the Rayleigh-scattering nor the geometric-optics limit applies. Information about particle sizes is obtained best from differencing measurements made with light transmitted at different wavelengths, since in this regime the ratio of opacities at different wavelengths is maximally sensitive to particle size (Eshleman 1973; Eshleman, *et al.* 1977).

How is information inferred from the incoherent signal? If the opacity is not too large ($\tau_c \lesssim 1$ or $\tau \lesssim \sin\theta_0$, the strength and spectrum of the incoherent signal can be modeled fairly accurately by assuming that the transmitted light is scattered only once before traveling on to the receiver. In this case, the incoherent received power exhibits strong dependence on particle size only for sizes much larger than the reduced wavelength of the transmitted light, and the strength and spectral properties are determined by diffraction effects, which produce a forward-scattering lobe of angular width $\sim \lambda/a$. In this regime, the forward-scattering cross section $\sigma(\beta = 0) \equiv \sigma(0)$ scales with the fourth power of particle size a (see, *e.g.*, van de Hulst 1980). Physically, this says that a ring particle larger than the wavelength intercepts from the incident wave a total power proportional to its geometric cross-sectional area, but rescatters that power in the near-forward direction in proportion to the *square* of its cross-sectional area. Away from the forward diffraction lobe, the scattering cross section does not have as simple an analytical or physical interpretation. The forward-scattering cross section reaches a maximum when the opacity τ is on the order of $\sin\theta_0$ (van de Hulst 1980), but for larger τ, multiple scattering processes begin to dominate, and calculation of the scattered signal strength and spectrum becomes considerably more complicated. For increasing opacity, the cross section σ becomes less sensitive to the angle β between the spacecraft-receiver and the particle-receiver lines of sight. For opacities larger than about $10\sin\theta_0$, the scattering has spread out significantly from the forward direction, and σ decreases rapidly toward its isotropic limit $\sigma_{\rm iso} \simeq \sin\theta_0$. Tractable approximate solutions exist for the near-forward scattering cross section $\sigma(0)$ as a function of opacity

when multiple scattering dominates (van de Hulst 1980), provided the opacities are small enough that $\sigma(0) >> |\sin\theta_0|$; this inequality typically holds for the following ranges of particle sizes a, forward scattering angle β, and opacity τ: $a \gtrsim 3\lambda$; $\beta \lesssim 0.6\lambda/a \lesssim 12.7$ arcseconds ($\simeq 3.5$ millidegrees); $\tau \lesssim 10 \sin\theta_0$ (Marouf, *et al.* 1982). Given bounds on opacity of $0.001 \leq \tau \leq 1$ for signal detectability (as applied, for example, to the Voyager 1 ring-occultation experiments at Saturn), Figure 25 shows that the forward-scattering cross section would be large enough for the incoherent signal to be detectable only for $a \gtrsim 10\lambda$.

Thus, in summary, detailed information about particles whose sizes are much smaller than the wavelength of the transmitted light can be obtained from measurements of both the amplitude and the phase shift of the coherent component of the transmitted light. Information about particles whose sizes are on the order of the wavelength is more elusive; the best approach uses differential measurements of the phase of the coherent light at different frequencies, since in this particle-size regime the opacity is strongly dependent on wavelength. Absence of strong differential effects would indicate a predominance of particles much larger than the wavelength. Finally, information about particles much larger than the wavelength of the transmitted light can be obtained from the strongly forward-scattered, Doppler-broadened, incoherent component of the transmitted light. Two pieces of information generally are available in this regime of large particle sizes: (1) the opacity inferred from extinction of the coherent signal and (2) the forward-scattering cross section estimated from spectral measurements of the incoherent signal. Together, these data determine a unique particle size. Even if extinction of the coherent signal is not detectable clearly, it may be possible to obtain size information from the incoherent signal. In this large-particle regime (the geometric-optics limit), the scattering becomes so tightly directed in the forward direction that enough power may be collected by the receiver to detect the otherwise relatively weak incoherent signal. If the particle sizes are much larger than the transmitter antenna, so that the forward lobe is entirely contained within the antenna beam, size information can be extracted directly from the (spectral) width of

the spectrum. For the Voyager spacecraft transmitting at 13 cm and 3.6 cm through a $\sim$ 3.7-m antenna, it was possible to obtain information about ring particle sizes ranging from $\sim$ 1 mm to $\sim$ 20 m. Addition of a laser transmitter on board such a spacecraft could provide information about particle sizes in the submicron regime from measurements of the amplitude and phase of the coherent signal (extinction and differential refraction) and about particles in the meter-or-larger regime from analysis of the spectral characteristics of the incoherent signal as well as the extinction of the coherent signal. Use of two or more laser wavelengths could provide information about particle sizes between these extremes. In view of the apparent prevalence of centimeter-to-meter size particles in planetary rings, there is no strong overall argument for the use of laser transmitters *instead* of radio-frequency transmitters for obtaining information about particle sizes and size distribution; however, their use clearly could complement and enhance radio measurements.

5.5.2d Occultation experiments in an uplink mode

As promised at the start of section 5.5.1, brief consideration is given now to the advantages and possibilities of performing atmosphere- and ring-occultation experiments in an *uplink* mode, with the occulted spacecraft functioning primarily in a receiving and data-processing capacity. Operation in a downlink mode has offered the advantage that the data handling could be done on the ground. The major disadvantage to operating in a downlink mode is the limited transmitter power available from a spacecraft — 1 to 20 W of S- or X-band power, perhaps 1 to 5 W of laser power. (Note that relative output powers of radio and optical transmitters do not indicate relative performance of microwave and laser systems for experiments such as those described in this report!) In contrast, ground-based radio-frequency transmitters routinely radiate 20 to 100 kW, and ground-based solid-state lasers pumped by diode arrays are projected to be capable of frequency-stabilized operation with average (as opposed to peak) output powers in the vicinity of 1 kW by the mid-1990s (Byer

1988). Tyler (1987) has pointed out that operation in an uplink mode, even if it provided only a modest improvement (a factor of 100, say) in sensitivity, would enable much finer study of Titan's surface and the structure of Saturn's "Ring B," as well as studies of other solar-system objects. Generally speaking, use of an Earth-vicinity transmitter together with a spacecraft receiver orbiting a planet or satellite could enhance measurements of atmospheres, rings, and surface properties significantly, because of the greater flexibility for studying temporal and spatial variations, the greater sensitivity, and the improved oscillator stability. The primary obstacle to operation in an uplink mode is the large volume of data that must be processed on the spacecraft ($\sim 10^{11}$ to 10^{12} bits for the Voyager 1 experiments at Saturn, according to Tyler 1987). The handling of this volume on the spacecraft is problematic because of limitations on power consumption, size and mass of hardware, and data storage capacity. Recent advances in integrated circuit technology and digital signal processing techniques may make uplink experiments at radio or optical frequencies more practical in the future. As progress in high-power frequency-stabilized lasers continues, uplink experiments at optical frequencies may emerge quickly as a powerful complement to traditional downlink radio-frequency occultation experiments.

5.5.3 Surface-scattering experiments

Scattering of electromagnetic waves off the surface of a planet or satellite permits determination of the surface dielectric constant and surface roughness. (Here "surface" refers to the upper layers that contain the material responsible for reflection, typically on the order of a few centimeters in depth for radio-wave reflection.) Historically, such experiments have been performed with the transmitter on the spacecraft and the receiver on Earth. The dielectric constant is inferred from the surface reflectivity, which in turn is inferred from the total power in the "echo," the spectrally broadened signal obtained from specular reflection at a point on the surface. The surface roughness

is indicated by the spectral width of the echo; roughly speaking, an increase in surface roughness increases the spectral width of the echo. More specifically, for a homogeneous and isotropic surface described by Gaussian statistics, the spectral width of the echo is proportional to the rms slope, or tilt with respect to the mean surface (Simpson, *et al.* 1979). Homogeneity of the surface means that the width of the echo spectrum is independent of reflectivity; and, in general, the reflectivity is independent of roughness to second order. Thus, the roughness and dielectric constant can be derived independently.

Measurement of the dielectric constant gives information on the density and composition of the surface material [Tang, *et al.* (1977) discuss this for Mars]. If the surface is modeled as a packed powder, the Rayleigh mixing formula (Gold, *et al.* 1970) can be used to relate the dielectric constant to the bulk density of the reflecting material. The dielectric constant ϵ can be determined directly from measurement of the Brewster angle θ_B by the relation $\epsilon = \tan^2\theta_B$, where the Brewster angle is that angle of incidence for which the reflected wave is linearly polarized in the plane perpendicular to the plane of incidence, thus displaying equal power in oppositely rotating circular polarizations. Composition, density, and roughness are characteristics relevant both to geologic models of surface structure and to practical aspects of placing landers and operating rovers.

Typically in radio surface-scattering experiments, the beam from the spacecraft is aimed toward the point on the planet or satellite surface that results in a mirror-like (specular) reflection toward Earth. The roughness of the surface causes multiple rays to be reflected from a small region about the mean specular point. Spectral broadening of the received signal (echo) arises because of the small differences in the rates of change of the individual path lengths. Underestimation of the surface roughness can occur either because of noise in the data or because of uneven illumination of the scattering area. Underillumination can result at radio frequencies because of the tapered pattern of high-gain radio-frequency antennas. Often the problem of underillumination can be minimized by proper choice of geometry. Overestimation of the surface roughness will result from any artificial

broadening of the echo. This might occur, for example, because of uncorrected frequency drift in the transmitted signal over the averaging interval, which typically is on the order of 10 seconds.

As an alternative to tracking the region of specular reflection, scattering off a particular point on the surface can be studied. Here, underillumination is not a problem; in fact, a narrow beam width is an advantage in that it permits suppression of the echo from neighboring regions. However, more accurate pointing is required. At radio frequencies, typical beam widths used for this purpose are 20° with pointing tolerances of 10% (Anderson, private communication 1989). (The diffraction-limited beam width for a 10-cm aperture at X band is 20° or 0.36 rad.)

The size of the surface structure sensed is determined by the wavelength used; shorter wavelengths respond to smaller-scale roughness. At X band, most of the scattering results from structure between 3 cm and about 30 meters in size, 5 meters being an average effective scale size (Tyler, *et al.* 1971). At optical frequencies ($\lambda \simeq 1$ μm), most of the scattering could be expected to result from structure between 1 μm and $\sim$ 1 mm in size. For engineering studies of the surface for lander/rover application, laser scattering would likely have a finer resolution than required, and X-band or K_a-band (32-GHz) scattering studies might be more appropriate. Studies performed at multiple wavelengths simultaneously (to ensure that reflection occurs from the same physical surface area) could provide useful information on the size and density distributions of scattering particles.

Requirements on oscillator frequency stability are not severe for surface-scattering experiments. It is only necessary that the oscillator not drift appreciably relative to the broadening inherent in the specular reflection. For a transmitter of diameter D_t located a distance D from the surface of a planet which is at a distance $r >> D$ from the receiver, the frequency spread $\delta\nu$ due to the spread in path lengths is at least as large as $2L\theta\nu/r$, where θ is the beam width of the transmitter. For a diffraction-limited beam, $\theta \simeq \lambda/D_t$, hence the requirement on laser fractional frequency stability is

$$\frac{\delta\nu}{\nu} \lesssim \frac{2\,D\,\lambda}{rD_t} \,. \tag{5.5.18}$$

Hence oscillator fractional frequency fluctuations $\delta\nu/\nu$ must be known or controlled to less than $2D\theta/r$. Using values for D from Table 5.5.1, one finds that for surface-scattering experiments at Venus, assuming a closest-approach distance of 0.3 AU, a frequency stability $\delta\nu/\nu \lesssim 10^{-5}$ would be required at X band ($D_t = 3.7$ m) and $\delta\nu/\nu \lesssim 2 \times 10^{-9}$ at optical wavelengths ($\lambda = 1$ μm, $D_t = 0.5$ m). At Jupiter, the corresponding stability requirements would be about a factor of 5 more stringent — $\delta\nu/\nu \simeq 2 \times 10^{-6}$ at X band and 4×10^{-10} at 1 μm. If the optical beam width is widened, a broader range of Doppler shifts will be produced, and the requirement on laser frequency stability can be eased. However, a narrower beam width ($\theta_{\rm opt} \sim \lambda/D_t \sim 2 \times 10^{-6}$, *vs.* $\theta_{\rm X} \sim 0.01$) enables finer probing of the surface. Measurements using synthetic aperture techniques would put more stringent requirements on frequency stability of both transmitter and receiver; this author has not examined this possibility nor found any such experiments proposed.

For surface-scattering experiments, intensity fluctuations of the oscillator must be kept reasonably small, since estimates of dielectric constant and inferences about density from surface reflectivity depend sensitively on the accuracy of intensity measurements made on the reflected signal. For radio-frequency surface-scattering experiments on Mars, for example, errors of 1 dB in total received power result in errors in estimates of the dielectric constant of about 14%, and these propagate to errors on the order of 17% in density (Anderson, private communication 1989; Gold, *et al.* 1970).

C-3

6. ADDITIONAL SCIENCE APPLICATIONS AND BENEFITS

This report has described the general kinds of measurements enabled by frequency-stabilized lasers and related technology (*e.g.*, phase-matching laser transponders), and analyzed several prominent scientific applications that require or would benefit from the measurement precision achievable with stable, optical-frequency oscillators. With the groundwork laid in chapters 3 and 4 and the detailed examples given in chapter 5, it is possible to imagine and evaluate numerous additional scientific applications involving frequency-stabilized lasers. Most of these will be left to the reader. However, a few obvious ones will be pointed out in this chapter. Each of the ideas suggested below could be analyzed and evaluated more carefully, as was done for the applications discussed in chapter 5.

6.1 Solar-system mission enhancements

Much of the science to be gained from using frequency-stabilized laser technology in space will come from missions throughout the solar system. Before discussing additional specific science applications or experiments, consider some of the general ways in which frequency-stabilized laser technology can be used to enhance such missions. These fall into three areas: telemetry (communications), navigation and tracking, and instrumentation. They are discussed in turn below.

6.1.1 Coherent optical communications

The potential benefits of optical over microwave systems for deep-space communications have been discussed in many other places (*e.g.*, Sokolowski and Lesh 1987). Chief among the telemetry advantages are higher data rates and wider communication bandwidths. For example, with modest coding techniques, a 1-W laser firing through a 30-cm telescope from Saturn could provide data rates of 200 kilobits per second or more with bit-error rates of 10^{-6}. Compelling additional benefits include reduced size and mass for the spacecraft transmitter and receiver hardware, and reduced

power requirements on the spacecraft transmitter, which can translate to reduced requirements on total spacecraft power.

These advantages apply to both incoherent and coherent communication systems. Incoherent communication systems use "direct" photodetection of the incoming laser signal, while coherent communication systems combine the incoming laser signal with a strong local-oscillator laser before photodetection. The performance of incoherent systems depends primarily on the number of signal photons detected; it is insensitive to the phase of the incoming light and only coarsely sensitive to the frequency. Typical filter bandwidths for deep-space optical direct-detection systems would be 5 to 10 Angstroms, or 150 to 300 GHz at 1-μm wavelength. These provide a limited capability for rejection of background light. In contrast, the process of combining the incoming signal with a frequency-stable local-oscillator signal in coherent, or heterodyne, detection enables a high degree of spatial and spectral discimination. Coherent detection can perform well under high-background conditions, where signal-to-noise ratios would be unsuitably low for direct detection, both because of this high degree of background rejection and because of the signal amplification that results from combination with the strong local-oscillator signal. For planetary missions, coherent detection could enable optical communications with a spacecraft to perform well even when the spacecraft is within a few degrees of the Sun or a bright planet.

Key ingredients to full exploitation of the possibilities of coherent optical communications are a frequency-stabilized local oscillator and companion frequency-stabilized transmitting laser, at both the spacecraft and the receiving station. The degree to which the frequencies can be controlled has important impact on the kind of information-coding schemes that can be used. For example, phase- and frequency-shift keying can provide much better communication performance than the on-off keying or pulse-position modulation schemes commonly used for direct detection; but they require stability and control of the phase and frequency of the lasers (Yamamoto 1980; Okoshi 1986).

6.1.2 Remote optical navigation and tracking

The attraction of improved telemetry performance with the use of laser communication links to spacecraft, as well as reduced size, mass, and power requirements over traditional radio-frequency systems, provide strong impetus to consider missions involving only laser communication and navigation systems or hybrid systems using both laser and radio-frequency technology. Hence studies and development of remote optical tracking techniques are under way (see, *e.g.*, Schumaker 1988, Folkner and Finger 1990) that would provide angular as well as range and Doppler information, with performance comparable to that achieved currently by NASA's Deep Space Network for spacecraft carrying S- and X-band transmitters. Optical tracking and navigation presents both advantages and disadvantages compared to radio-frequency operation. One benefit is the potential for making real-time angular measurements from a single station. Another is the possibility of direct tracking of spacecraft relative to their targets when the targets are sunlit solar-system bodies. An obvious disadvantage, at least for ground-based operation, is atmospheric degradation and limited performance during daylight hours.

Most of the optical tracking techniques currently being studied for use with planetary missions would be appropriate for use with incoherent communication links. One of the strongest candidates is a "filled-aperture" astrometric technique (as opposed to dilute-aperture, or interferometric, techniques) that uses a telescope with a finely ruled grating at its focal plane. Such instruments could be used to measure the angular separations between a laser-carrying spacecraft (or orbiter, lander, or rover) and nearby reference stars or other laser-carrying spacecraft. They show promise of angular measurement accuracies of 10 to 20 nanoradians (a few milliarcseconds) from the ground and subnanoradian accuracies from space, being limited primarily by atmospheric effects on the ground and laser faintness from space. These ground-based accuracies have been demonstrated for stars of apparent visual magnitude $m_v = 11$ or brighter (the brightness of the Sun if it were 150 parsecs = 3×10^7 AU away from Earth) (Monet and Dahn 1983; Gatewood, *et al.* 1980; Gatewood

1987; Buffington and Geller 1990). A 1-W, 0.5-μm wavelength laser firing through a 30-cm telescope at the distance of Saturn would also have an apparent visual magnitude of 11; that is, viewed through a detector whose response function is similar to that of the human eye, the laser, despite its extreme monochromaticity compared to a star, would produce approximately the same detected photon count as a star of apparent magnitude 11 (Schumaker 1988). These instruments typically have fields of view no wider than about 1°. Since there are on the average only about 1.3 stars as bright as $m_{|rmv} = 11$ per square degree in the sky (Allen 1973), highly accurate (better than 10 to 20 nanoradians) remote optical navigation may require instruments with wider fields of view or better sensitivities. However, for many solar-system missions, remote navigation accuracies of several tens of nanoradians is adequate when augmented by onboard optical techniques.

As was pointed out in section 5.1, optical interferometers being developed for precision astrometry and astronomical imaging may provide accuracies of several tens of picoradians for space-based angular measurements among optical sources. Interferometric techniques can measure angular separations over wide fields (30 to 90°), which is difficult or impossible with filled-aperture techniques. Unlike telescopes, they are free from the requirement that a reference star be in the field of view simultaneously with the target, which might be a laser-carrying spacecraft. Sophisticated laser metrology systems are required to achieve the high levels of accuracy possible with such instruments.

While navigation to a planet or satellite may not require nanoradian or better accuracies, the ability to track spacecraft this accurately relative to stars or solar-system objects, or landers and rovers relative to orbiters, would enable a wealth of possible science experiments. Some of these are outlined in section 6.2 below. Frequency-stabilized lasers would play an important part in enabling precise tracking of laser-carrying spacecraft in at least two different ways: First, they are required for coherent range- and range-rate (Doppler) tracking (see section 4.2.1); second, they are a key element for instrument metrology in optical interferometers that would be used for high-accuracy angular tracking (see section 5.1). The former application also would benefit from the development

of coherent laser transponders, which would receive, amplify, and retransmit laser signals, thus reducing signal losses due to propagation (*i.e.*, received power would fall as $1/r^2$ for transponders instead of $1/r^4$ as is the case for reflectors; see subsection 4.1.4a).

6.1.3 Improved instrumentation

Certainly one of the most important benefits associated with the use of lasers and laser communication links with solar-system missions is the capacity for higher data rates. For example, the high-resolution imaging spectrometer (HIRIS) planned for the Earth Observing System will be unable to operate at full capacity because, while its internal data rate is on the order of 500 Mbits/s, its maximum external data rate is only 300 Mbits/s (Butler, *et al.* 1987b). The higher data rates afforded by laser telemetry links would also permit simultaneous operation of several instruments.

Several kinds of onboard instruments related to such tasks as navigation, pointing, and maneuvering would be enhanced by coherent (*i.e.*, frequency-stabilized) laser-based systems. For example, coherent lidar could be used both to aid space-based operations and general vehicle maneuvering and to perform science experiments in planetary atmospheres and rings. Extremely precise optical rotation sensors (*e.g.*, laser gyros) could be used for a variety of engineering and scientific purposes.

6.2 Solar-system science, astronomy, and astrophysics

This broad category encompasses most of the applications discussed in this report. An enormous number and variety of important scientific measurements become possible with the existence of coherent, quasi-monochromatic laser links between spacecraft at planetary distances, or even just in Earth orbit. This section reviews briefly some additional applications in this category and some extensions of those already discussed. The grouping of these applications is somewhat arbitrary, since there can be significant overlap among them. One potentially rich field of scientific measurements

that is *not* discussed here consists of experiments involving the simultaneous use of laser and radio links. In such experiments, what would constitute "noise" with the radio or laser link alone, can become "signal" when the possibility of using both links simultaneously exists. For example, the immunity of laser links to plasma-induced phase scintillation would enable them, when used in conjunction with radio links, to calibrate these and other charged-particle effects, and thus provide information about electron densities in planetary ionospheres and the solar corona.

6.2.1 Improved ephemerides

Since this chapter began by considering mission enhancements, it seems appropriate to begin this discussion by noting the important benefit of improved ephemerides of solar-system bodies. This improvement could arise both from improvements in optical galactic and extragalactic reference frames achieved through high-accuracy optical astrometry and from precise optical tracking of laser-carrying spacecraft near solar-system objects. Frequency-stabilized laser technology would contribute to both by enabling accurate, coherent Doppler tracking of spacecraft throughout the solar system or by enabling high-precision optical astrometric interferometers, which could be used both to track the angular motion of the spacecraft and to improve the reference-frame catalog through systematic astrometric measurements.

6.2.2 Planet and satellite masses, gravity fields

Radio-tracking observations have long been used to gain information about the gravitational fields of planets and satellites [see, *e.g.*, Anderson, *et al.* 1974b (Jupiter); Sjogren, *et al.* 1975 (Mars); Sjogren, *et al.* 1980 (Venus); Tyler, *et al.* 1982 (Saturn); Tyler, *et al.* 1986 (Uranus)]. Total masses and higher mass moments are inferred from fly-by trajectories, and more detailed information is gained with orbiters. The description given in section 5.3 of the use of a laser interferometer

among orbiting spacecraft to map a planet's gravity field is appropriate here. In general, given comparable tracking accuracies (both angular and line of sight), optical-tracking observations and radio-tracking observations would provide similar information. Given comparable accuracies, the advantages of making these observations optically would stem primarily from the reduced size, mass, and power requirements, and reduced atmospheric drag effects (see, *e.g.*, section 5.3 and section 6.1). Prior to recent technology advances leading to high-power, frequency-stabilized lasers suitable for autonomous operation in space, radio-frequency measurements have, in general, offered greater potential measurement accuracies. But, as the planet gravity-field-mapping example in section 5.3 showed, given laser power and frequency stability enabled by current technology, laser measurements could provide better resolution and sensitivity in characterizing nonuniformities in a planet's gravity field. Of course, there also would be significant scientific advantage to making measurements simultaneously at *both* radio and optical frequencies in order to calibrate the errors for the two systems. For example, plasma-induced phase noise may dominate the radio measurement error, whereas photon statistics or laser frequency stability could dominate the optical measurement error.

For terrestrial planets, both gravitational and topographic data are used to determine the degree of isostatic compensation of the surface topography (Sjogren, *et al.* 1975, 1980). In isostatic equilibrium, which holds approximately for Earth, gravity "roots" would exist under mountains, and "antiroots" would exist under the basins. Topographic data are used to subtract the estimated "external" contributions to the gravity field, and gravity data are then used to determine the degree of isostasy. Or, if isostasy is assumed, then the presence of gravity anomalies could indicate convection currents or "primordial" density variations deep within the planet. If gravity data indicate that the topography is not compensated, inferences can be made about the elastic strength of the interior and thence about constraints on the radial temperature profile. For example, if the inferred temperature profile is unreasonable (*e.g.*, temperatures are too low), this would indicate that viscous stresses associated with convection play a part in supporting the topography; the accuracy with which their

magnitudes can be estimated depends directly on the accuracy of the gravity data.

Topographic data typically are obtained with Earth-based photography and radar, and stellar occultations. Laser (lidar) altimeters used on orbiters or during fly-by missions could improve on all of these methods. And, if it were used together with radar, information could be obtained about the surface composition by comparing both the light-travel time (depth of penetration) and the backscatter cross sections.

For the giant planets, gravity data can be used in two ways: to deduce the density distribution and to put constraints on composition. In these larger planets, deviations from a perfectly spherically symmetric gravity field arise primarily from rotation. The deviation from sphericity is greatest if the planet is homogeneous (constant density). Variations in rotation rate with latitude and depth into the atmosphere reflect nonuniformities in the density distribution within the planet. When combined with independent information on rotation rate, gravity data can be used to make inferences about the density distribution. Precise tracking of spacecraft during close fly-bys of the giant planets can give information about several higher moments of the gravity field, in addition to the total mass; these, together with observed radii and inferred density distributions, put constraints on the composition. For example, it was determined in this way that Jupiter is not of solar composition, since its large mass moments require larger fractions of elements heavier than helium than are consistent with solar abundances (Eshleman, *et al.* 1977).

Close fly-bys of satellites, especially the more massive satellites (*e.g.*, the Galilean satellites and Titan), could permit determination of the second-degree harmonics in their gravity fields and hence determination of the differences among their principal moments of inertia. By comparing the deformations due to rotation and to tidal forces, one can estimate the degree of central condensation and determine the degree to which the satellite is in hydrostatic equilibrium. This, in turn, puts constraints on possible interior models (Hubbard and Anderson 1978).

Finally, cometary masses could be determined during rendezvous with a spacecraft by combining (1) Earth-derived astrometric information on the relative positions of the comet, spacecraft and Sun (hence the relative solar tidal pull between spacecraft and comet); (2) measurements of the spacecraft-comet range obtained with lidar or radar on board the spacecraft; and (3) measurements of the comet-induced deceleration (and any other nongravitational accelerations) on the spacecraft obtained with onboard accelerometers. The comet-spacecraft acceleration is determined by subtracting the inferred solar tidal acceleration and nongravitational accelerations; this, combined with knowledge of the spacecraft-comet range and spacecraft mass, determines the comet mass (Anderson, private communication 1989).

6.2.3 Atmospheric studies

Elsewhere in this report, techniques involving frequency-stabilized lasers have been discussed that are applicable to studies of Earth's atmosphere. These include coherent Doppler lidar measurements of velocity-vector components of particulates (sections 4.2.2 and 5.4); coherent forward-scattering measurements of particle sizes and size distributions, number densities, and shapes of particulates (section 5.5.2); and coherent forward-scattering measurements to obtain atmospheric refractivity profiles through measurements of the complex extinction (sections 5.5.1 and 5.5.2). The last technique is of use in monitoring variability in vertical temperature and pressure profiles for weather forecasting, as well as for studying varying chemical content (*e.g.*, ozone levels) in the atmosphere.

Spectroscopic measurements benefit in an obvious way from improved frequency stability, particularly absorption measurements. It is equally important for such measurements that the laser sources be tunable, so that they can be tuned to resonance with particular molecular transitions and then tuned away from resonance to establish zero-concentration baselines. The narrower the line width of the laser, the more sensitive is the information obtainable about absorption or emis-

sion line shapes, which in turn can provide more detailed information on atmospheric pressure and temperature profiles. Many of these types of measurements are being, or have been, performed in Earth's atmosphere, while others will be enabled by the development of space-qualifiable, compact, frequency-stabilized lasers. For example, the use of stable lasers on geostationary or low-Earth-orbit platforms would enable measurements of long-path transmission through the atmosphere, including various experiments involving atmospheric occultation of the Sun to aid in monitoring the chemical composition of the atmosphere. For detailed descriptions of many proposed Earth-related experiments and a rich source of ideas for further uses of frequency-stabilized laser technology, the reader is referred to the series of study reports compiled by the Science Steering Committee for the Earth Observing System (Butler, *et al.* 1987a–f).

Detailed analyses of various experimental approaches for studying Earth's (or other planetary) atmosphere, some of which use lasers and others that use radar but could benefit from extension to coherent laser techniques, can be found in the following references, in addition to those already cited throughout this report: Churnside and Clifford 1988 (refractive turbulence profiling); Hansen 1980 (laser scattering off atmospheric particles); Ismail and Browell 1989 (water-vapor measurements using lidar); Kavaya and Menzies 1985 (lidar aerosol backscatter measurements); Kazovsky 1984a and 1984b (laser forward scattering off particles); Kazovsky and Kopeika 1983 (coherent laser techniques through turbid media such as rain and snow); Kyle 1989 (high-resolution laser imaging); Lindberg and Gillespie 1977 (atmospheric dust refractivity measurements); Lindberg, *et al.* 1984 (lidar determination of extinction through clouds); and Megie and Menzies 1980 [DIfferential Absorption Lidar (DIAL) for measurements of atmosphere composition]. A layman's overview of some of the kinds of measurements of particular relevance to Earth studies also can be found in Covault (1989).

6.2.4 Cloud studies

Techniques for probing cloud microstructure include laser scattering measurements to gain information on the sizes of cloud droplets, optical imaging to gain size and structural information about precipitation elements, polarization measurements to differentiate liquid from ice in clouds, and measurements of liquid water content. The most revealing information about the microstructure of clouds is knowledge of the sizes, size distribution, number density, and shapes of liquid or solid cloud and precipitation particles ("hydrometeors"). Cloud particles vary widely in size (0.1 μm to 10 mm) and number density (10^{-4} to $10^3/cm^3$), so a single instrument or measuring technique typically is insufficient. Observations from aircraft or Earth orbit are required for clouds well above ground level, because of the large size of cloud forms and the fast motion of the particles. However, measurements from aircraft are difficult because of the electrically and acoustically noisy environment.

Optical measuring techniques offer a favorable alternative to conventional *in situ* "physical sampling" measurements in clouds, in which particles are captured on substrates or instruments are inserted into the cloud and particles aspirated. This kind of sampling disturbs the airstream, it can provide data only through a few sections of a cloud, and its analysis is slow and tedious. In contrast, light-scattering measurements are efficient, relatively nonintrusive ways to count and size cloud particles in the 1- to 50-μm range; larger particles can be studied using *in situ* imaging techniques such as photography and holgraphy. The only important characteristic not readily accessible with optical techniques is liquid-water content, although this can be inferred indirectly through integrations of measured size distributions. Fog measurements require maximum sensitivity at sizes near 1 μm, while cloud-droplet measurements require maximum sensitivity around 10 to 20 μm. Thus, the instruments for these two purposes generally do not overlap. Meteorological emphasis historically has been placed on the analysis of cloud droplets, although laser-scattering methods have been proposed that would be useful for both regimes (Silverman, *et al.* 1964; Knollenberg 1981).

The principles behind laser scattering measurements on cloud or fog droplets are similar to those

described in section 5.5 for atmosphere- and ring-occultation experiments. For the droplet sizes typical of clouds and fogs, forward scattering predominates (Knollenberg 1981). Since the refractive index for clouds is fairly well known, a great deal of information can be obtained about sizes in the micron range with forward-scattering laser measurements. Typical laser scattering instruments used in cloud-physics measurements are described in Knollenberg (1976). These instruments are designed to measure velocities between 10 and 125 m/s and measure particle sizes in the range 2 to 30 μm. To date, almost all these instruments have used gas lasers, particularly helium-neon (He/Ne) devices (Knollenberg and Luehr 1975). Gas lasers require high voltages and have very limited lifetimes, and reliability has been a problem. For example, the flight of the ATMOS Fourier Transform Spectrometer on SPACELAB 3 in 1985 was terminated by a failure in the He/Ne laser power supply after barely three days in orbit. Instruments proposed for Earth-orbiting platforms, such as the Earth Observing System, must be able to operate continuously for several years. It is likely that diode-pumped solid-state lasers will replace these gas lasers because they require lower voltage, are smaller in size, more reliable, longer-lived, and have the potential for greater output power and better frequency stability. Improved frequency stability translates to improved spectral resolution for spectrometer measurements. Shorter wavelengths (*e.g.*, 0.5 μm instead of 1 μm) are preferred because sampling intervals are defined as equal intervals of wavelength; hence, better control can be obtained with equivalent frequency stability at shorter wavelengths.

6.2.5 Cosmic and interplanetary dust

Interplanetary dust particles (IDP) have several possible origins (*e.g.*, Carey and Walker 1986a). Cosmic sources include comets and asteroids as well as interstellar grains. In the last twenty years, man-made debris in Earth orbit has become a major contributor to the local flux of dust particles, creating both a hazard for spacecraft and an error signal in studies of the properties of the "natural" dust. This subsection summarizes the properties of the dust from different origins, including the

man-made particles. Trajectory determination and possible differentiation among the origins is then described briefly. Finally, estimates are made of the properties of a Doppler lidar system needed to measure the particle trajectories and the laser frequency stability needed to make the measurements.

6.2.5a Particle origins and properties

Since the orbital parameters of several major meteor streams match very closely those of observed comets (Kresàk 1980), there are clear parent-daughter orbital relationships between some meteoroids and the comets they derive from. It is commonly held that comets consist in part of interstellar material that originally existed in the solar nebula. Arguments based on mass balance (Whipple 1978) and analysis of the orbital parameters obtained from photographic data on meteors (Jacchia and Whipple 1961) support the view that much of the interplanetary dust comes from comets. Laboratory measurement of IDP collected from Earth's stratosphere shows substantial enrichment of deuterium and hydrogen relative to other solar-system materials in about 60% of the collected particles, thus supporting the view of their earlier origin; and midinfrared absorption spectra on some IDP show strong similarities to infrared sources thought to be protostars (McKeegan, *et al.* 1985).

Radar studies (Sekanina 1973) have also indicated asteroids as the parent bodies for some meteoroids. Dust bands observed with the Infrared Astronomical Satellite suggest that asteroids are sources of a significant fraction of the IDP. Some fraction of the IDP likely is interstellar dust that has penetrated the inner solar system. Calculations indicate that such particles would be on the order of 10 μm or larger in size (Morfill and Grün 1979). However, no measurements of the orbital parameters of particles in the 10-μm size range have ever been made. (Photographic and radar measurements are limited to much larger particles.) Until measurements are made of the orbital parameters of particles in the 10-μm size range, it is not possible to rule out the possibility that a significant fraction of the IDP is interstellar material. Nevertheless, recent calculations based on

estimates of the gas density and gas-to-dust ratio of the local interstellar medium suggest that at most 1% of the dust flux in the 1 to 100 μm size range is interstellar (Sandford 1985).

A capture apparatus in Earth orbit, such as the Cosmic Dust Collection Facility proposed recently for the Space Station Freedom (see *e.g.*, Horz 1986), will encounter these particles as well as man-made debris (paint flecks, for example). Figure 26 is a graph (due to Kessler 1985) of the observed flux of natural objects (meteoroids) and man-made orbital debris as a function of particle size. At any one time, there is an estimated 200 kg worth of meteoroid material moving through Earth's atmosphere at altitudes below 2000 km, at an average speed of about 20 km/s. Most of the mass is found in particles of about 100 μm diameter (Zook, *et al.* 1970). Man-made particulates, such as rocket-exhaust products and collisional or explosive fragments of spacecraft, form an increasing fraction of the "dust" particles encountered in Earth orbit and are a growing concern for manned and unmanned orbiting structures. Particles with diameters between 100 μm and about 10 mm make up the primary collisional hazard, since smaller particles can be shielded against and larger particles are infrequent. It is estimated that the total mass of debris in orbit is now approximately 2×10^6 kg at altitudes below 2000 km (Kessler 1986). The average relative velocity of the debris particles is only 10 km/s, about half that of the meteoroids; and most of the debris mass is found in objects several meters in diameter, as opposed to 100 μm for meteoroids. Due to explosions, collisions, deterioration, and rocket-motor firings, it is expected that a large amount of material in the 100-μm to 10-mm range has been generated from this orbital debris. Discrimination between "natural" and man-made particles is achieved primarily through collection and subsequent compositional study. Ongoing detailed examination and analysis of pieces of the Solar-Max satellite suggest that there may be twice as many orbital debris impacts as meteoroid impacts, and that there likely are billions of 100-μm debris particles in Earth orbit.

6.2.5b Trajectory information

Information from chemical, isotopic, and other analyses alone is inadequate to provide insight into the composition and evolution of the parent bodies and particularly to distinguish between comets and asteroids as parent bodies. Even particles known to come from comets (because of their association with particular meteor showers) differ substantially in character from comet to comet. Measurement of the orbital parameters of specific dust particles is essential to correlate physical characteristics with specific sources, and to answer questions about the diversity of particles of cometary origin that could not be answered, for example, by sample-return missions to single comets.

Unfortunately, dust grains seldom travel in orbits identical to their parent bodies. Grains emitted from a comet are emitted in a variety of directions, due to the outflowing gas. Grains emitted from asteroids as a result of meteoroid impact also are ejected with a variety of velocities. Solar radiation pressure decreases the effective gravitational pull felt by the small grains and leads to increased orbital periods and semimajor axes relative to those of the larger grains; Poynting-Robertson drag gradually shrinks these orbits again (Zook 1986). Different gravitational perturbations, added to the Poynting-Robertson drag, cause increasing divergence among orbital parameters of the parent body, the larger grains, and the smaller grains. The change in orbital parameters depends strongly on particle size and on the time since separation from the parent body. These orbital evolution processes are topics of detailed theoretical investigations, for their understanding is essential to identification of collected meteoroids with specific source bodies or types thereof. Because it is not known even what percentage of the smaller dust particles are associated uniquely with a specific object or even a general class of objects, it is considered important that *in situ* determinations of orbital elements be made to the highest precision possible.

A system to measure the desired trajectory information must be capable of measuring relative velocities in the range of 10 to 100 km/s for a wide range of particle sizes, from about 10 microns to $\sim$1 meter. Studies indicate the velocity components of impacting meteoroids must be measured to

a precision better than 1% in order to enable separation of objects from asteroids and from comets. A similar level of measurement precision is necessary to distinguish different populations of Earth-orbiting spacecraft debris from each other and to determine their sources, *e.g.*, whether they are lunar ejecta or man-made Earth-orbiting debris generated in geosynchronous transfer orbits (Kessler 1985, 1986; Alexander and Corbin 1980). Photographic meteor trajectories can be obtained with accuracies of 0.1% to 0.4% (Jacchia and Whipple 1961; Knollenberg 1981). The situation for smaller meteoroids, at present, is not nearly as good.

Any system to measure trajectories of dust particles must be capable of reliable, autonomous operation for periods of a year or longer in order to obtain statistically significant results. Further, because of the small flux of particles, collecting areas must be at least as large as 1 m^2 (Carey and Walker 1986a,b). Several different techniques have been considered seriously for making velocity measurements on dust grains. One uses a grid of wires in which an electrical signal is induced when a charged particle passes; several grids of wires, spaced about 10 cm from each other, are used to determine the position coordinates of a particle at two different crossing planes separated by $\sim$ 10 cm (Auer 1975). Another approach uses a thin metal foil at the top of a capture cell. Here a particle's arrival causes a plasma pulse on a system of grids just below the top foil; time of flight is recorded as the particle is detected again at a second foil followed by a second system of grids (Carey and Walker 1986b). Acoustic sensors are a third possibility; use of several sensors and measurement of arrival times would enable trajectory determination. A fourth possibility, examined below, is the use of coherent Doppler lidar.

6.2.5c Doppler lidar system requirements

Using the analysis presented in section 4.2.2 of this report, one can assess the utility of coherent Doppler lidar for obtaining trajectory information on dust particles in Earth's atmosphere (and, by extension, in other environments). Assume an average velocity of 10 to 20 km/s for the particles (typically 10 km/s for orbital debris, 20 km/s for natural objects), with a spread $\sigma_{\text{med}} \simeq 3$ km/s. With the low fluxes indicated in Figure 26, the scattering cross sections, either forward or backward, are likely to be extremely small; therefore the measurement accuracy most likely would be limited by photon statistics. Assume, therefore, that the beam width, detector area, pulse duration, and integration time have values such that at least one scattered photon is detected in a single measurement integration time. At the impact rates of 1000/m^2-yr for 1 to 10 μm-diameter particles indicated in Figure 26, detector collecting areas should be on the order of 100 m^2 if integration times are on the order of 5 to 10 minutes. One can then calculate the relative magnitudes of the terms in eqs. (4.2.17) for the velocity measurement precision and get a rough idea of how well this technique might work.

Assume, for illustration, laser pulses of duration 1 μs and spectral width $\delta\nu_{\text{bw}} \simeq 3$ MHz, at an operating wavelength of about 1 μm. Note that these pulse spectral widths imply a modest fractional frequency stability (of order 10^{-8}) for the reference laser that generates the pulses and serves as a local oscillator for heterodyne detection of the returning pulses. The range resolution $\sigma_L \equiv c\tau_p/2 \simeq 150$ m. Ignoring the error due to photon statistics for the moment, consider the contribution to velocity-measurement error from the quantity A defined in eqs. (4.2.17). With the numbers assumed above, this contribution is on the order of 63 m/s, dominated by the effect of the velocity spread of the particles. (In the language of section 4.2.1, $\sigma_{v,\text{bw}} \simeq 1.5$ m/s and $\sigma_{v,\text{med}} \simeq 63$ m/s.) If no bandwidth-reduction techniques were used, the sampling frequency required in order to be sensitive to maximum velocities of about 25 km/s would be $f \simeq 50$ GHz [eq. (4.2.15)]. Assume here that the detection process uses a bank of approximately 1000 spectral filters, so that

the effective sampling frequency associated with each filter is only 50 MHz. Then the quantity ξ that contributes to measurement error through the term B in eqs. (4.2.17) is of order 100, and the contribution to measurement error from term B (assuming a single detected photon per integration time) is about 190 m/s. The contribution from term C is only about 0.75 m/s. The total velocity-measurement error is therefore approximately 200 m/s, or 1 to 2% of the average particle velocity for meteoroids and orbital debris, respectively. If the reference-laser fractional frequency stability were improved to 10^{-10}, so that the pulse spectral widths were only about 30 kHz, and pulse durations of about 10 μs were used, the final velocity-measurement precision would be approximately 35 m/s, or 0.2 to 0.4% of the average particle velocity, and the range resolution would be on the order of 1.5 km. Thus, provided adequate collector areas and detector sensitivity are employed to ensure a minimum of one detected scattered photon per integration time, coherent pulsed Doppler lidar techniques might be useful in obtaining precise trajectory information.

6.2.6 Solar studies

In the early 1980s, serious study was given to a proposed mission called "Starprobe," which would send a spacecraft on a trajectory over the pole of the Sun to a perihelion distance of approximately four solar radii (Underwood and Randolph 1982). With both S- and X-band capabilities on such a spacecraft, numerous scientific experiments could be carried out, including studies of the Sun (*e.g.*, its mass moments, density distribution, shape and dimensions, surface composition, and luminosity), studies of energetic particle and plasma phenomena, and relativity experiments and other tests of gravitation theories. Many solar studies seek to characterize the solar wind and solar corona, as well as interplanetary turbulence (*e.g.*, Woo 1978; Volland, *et al.* 1977; Tyler, *et al.* 1977; Winn, *et al.* 1977). This information is obtained through measurement of the differential effect of the solar plasma on the group delay of radio waves at two different wavelengths. As noted earlier, this kind of information clearly cannot be obtained with laser beams alone, since optical-frequency radiation is relatively unaffected by passage through plasma. However, simultaneous transmission of light waves

with dual-frequency radio waves could improve the calibration of plasma effects, and thus enhance the information gained from radio experiments.

A variety of measurements related to the mass, shape, and gravity field of the Sun could be aided significantly with the addition of laser links to solar spacecraft. This is because virtually all of these measurements rely on precise tracking of a spacecraft as it nears the Sun, ideally as near as four solar radii ($\simeq 1^\circ$ as viewed from Earth vicinity). Dual-frequency S- and X-band links likely would not be able to provide the desired measurement accuracy. Even if plasma effects were small and the transponder turnaround ratios identical for the two bands, only first-order effects (plasma effects that scale with the square of the frequency, but not those that scale with the fourth power) would be cancelled. In practice, even first-order cancellation is imperfect. Coronal scintillation causes the ground-receiver's phase-locked loops to lose lock intermittently, and dissimilar transponder turnaround ratios result in imperfect cancellation of the plasma-induced phase delay. Ideally, of course, the turnaround ratios would be made identical. In practice, this is not always possible; at the time Starprobe was being considered, expected transponder turnaround ratios were $240/221 = 1.085$ for S band and $880/749 = 1.175$ for X band (Underwood and Randolph 1982). With improved calibration of the plasma effects through use of a laser link, high accuracy could be maintained despite cycle slips at S band for a spacecraft in close proximity to the Sun.

The remainder of this section concentrates on measurements of harmonics of the solar gravity field. Other measurements that are of primary interest for testing general relativity and other theories of gravitation are discussed in section 6.2.6 below. A very important scientific objective of a solar mission is to determine the quadrupole moment (J_2) of the solar gravity field. This would provide information about the state of rotation of the solar interior, particularly the core, which in turn could provide insight into star formation by indicating how angular momentum was distributed among the Sun and the planets, as well as improve estimates of the solar mass distribution and the Sun's total angular momentum J. When combined with measurements of surface rotation, which

could be considerably slower than the core's rotation rate, knowledge of J_2 would shed light on possible mechanisms for slowing of the rotation rate (such as magnetic braking due to the solar wind) and their possible connection to the surface activity seen in middle-aged stars such as the Sun. If the value of J_2 is found to be large enough that strong differential rotation is likely, a variety of mixing mechanisms become possible, all of which lead to predictions for such observables as neutrino flux that differ substantially from the "canonical" theories of star formation that assume no mixing of envelope material into the core. Measurement of the J_2-axis orientation also would be of significant value for understanding the solar magnetic field. If the Sun has a large, centered, dipole magnetic field whose axis is aligned with the solar rotation axis, then distortion of the gravitational potential induced by the magnetic field would be indistinguishable from that caused by rotation. However, if the J_2 axis is displaced from the rotation axis, as it is expected to be, a magnetic dipole field then could be inferred (Underwood and Randolph 1982). Finally, accurate measurement of J_2 would remove the solar oblateness as a source of error in solar-system tests of gravitation theories.

Measurement of J_2 would be accomplished through the analysis of remote spacecraft tracking data. Frequency-stabilized laser technology would be required because coherent (heterodyne-detection) links, as opposed to incoherent or direct-detection links, are necessary (1) to provide adequate filtering of the sunlight when the spacecraft is within a few degrees of the Sun, and (2) to provide highly accurate range and velocity information. The angular motion of the spacecraft could be tracked both at radio frequencies, *e.g.*, with very-long-baseline interferometry or differenced-range techniques, and at optical frequencies with filled-aperture or interferometric optical astrometric techniques.

Analyses suggest that a Doppler velocity-measurement accuracy of 0.1 mm/s would be required to infer J_2 to an accuracy of 2×10^{-8}, or approximately 10% of its estimated value [see Figure 27 and Anderson 1988, and Mease, *et al.* 1984]. For measurement integration times of about 60 seconds, this would require oscillator fractional frequency stabilities of about 3×10^{-14} over periods at least

as long as several hours, corresponding to stable line widths of about 10 Hz at laser wavelengths ($\lambda \simeq 1$ μm) or 300 μHz at X-band [eq. (4.2.9c)]. A drag-free system accurate to $10^{-10}g$ also would be required to reach this accuracy for J_2, although compensation to only $10^{-9}g$ still would give an accuracy of approximately 3×10^{-8} in J_2. Compensation to $10^{-10}g$ corresponds to a reduction by a factor of about 10^5 in the nongravitational accelerations experienced by a spacecraft in the proximity of the Sun. These numbers have assumed that the error in the drag-free system comes primarily from the zero-frequency (DC) component; if instead the drag-free system has significant noise components in the 10^{-4} Hz range, the accuracy would be worse than indicated here (Anderson 1988). Implications of this accuracy for tests of general relativity are noted in section 6.2.7.

With the kind of precise tracking capability enabled by a combination of dual-frequency microwave links and a coherent optical link to a solar spacecraft, other measurements somewhat more demanding than the measurement of J_2 could be considered. For example, the next rotational moment, J_4, might be determined. For a uniformly rotating Sun $J_3 \simeq 0$, and $J_4 \simeq {J_2}^2$ (dimensionless units). Studies indicate that for a differentially rotating Sun, J_4 could be considerably larger than ${J_2}^2$, perhaps as large as $0.1J_2$ (Ulrich and Hawkins 1981). Other even zonal harmonics, especially J_6, also may be large for a differentially rotating Sun. However, even if J_6 were as large as $0.01J_2$ ($\simeq 2 \times 10^{-9}$), it probably would require finer measurement accuracy than indicated in Figure 27 for definitive measurements to be made. The time variation of J_2 also may be accessible with tracking accuracies of 0.1 mm/s or better. It is estimated that J_2 has a sinusoidally varying term with amplitude of order 7×10^{-8} (Christensen-Dalsgaard and Gough 1980), which arises from observed 160-minute oscillations of the Sun (presumably of acoustic and gravity-wave origin). Comparison of the amplitudes of optical and gravitational oscillations, the latter inferred from J_2 measurements, would give valuable information about the Sun's internal structure. Finally, the Sun's total angular momentum may be inferred from measurement (by means of precise tracking) of a general relativistic effect known as "inertial-frame dragging," described in section 6.2.7.

6.2.7 Tests of general relativity and gravitation theories

Gravitation is the weakest of the four fundamental interactions in nature, and the most difficult to study experimentally. But because it acts over very long ranges and couples to inertial mass, it dominates the large-scale structure of the universe. Measurements of relativistic gravitational effects are of profound importance for furthering an understanding of this most basic interaction, and for understanding physical properties of the strong gravitational fields associated with astrophysical sources and their implications for cosmology. Several relativistic theories of gravitation exist, in addition to general relativity, that are consistent with the limited experimental results currently available. Discrimination among the theories is difficult because of the relative weakness of the different predicted effects on observable scales. While it is possible that the failure of one or more of these theories may be apparent only on cosmological or subatomic scales, at present the most effective experiments that can be performed are on solar-system scales.

To date, the correctness of general relativity and other gravitation theories over the Newtonian theory of gravity has been tested only in the weak-field limit, and to at best about 0.2% accuracy. The advent of ultrastable oscillators, microwave and optical, enables more accurate measurements than previously possible for many classic tests of relativistic gravitation theories, some of which are described below. In order to compare the various theories with experiments made on a solar-system scale and with Newtonian predictions, a parametrized formalism has been developed, called the Parameterized Post-Newtonian, or PPN, formalism (see Misner, *et al.* 1973). This formalism is valid only in the weak-field limit, where the dimensionless gravitational potential $\phi \equiv GM/rc^2$ is smaller than the solar poential $GM_\odot/R_\odot c^2 \simeq 2 \times 10^{-6}$, and velocities are nonrelativistic, $v \lesssim c\sqrt{\phi}$.* In this formalism, terms are grouped according to combined powers of ϕ, or M, and v^2. First-order terms give the predictions of Newtonian theory, second-order terms are called post-Newtonian,

* The symbol $\odot$ refers to quantities for the Sun; M denotes mass and R the radius of an equivalent uniform, spherical body, and G is the universal gravitational constant $\simeq 6.7 \times 10^{-11}$ m^3/kg-s^2.

and higher-order terms are neglected. (Post-post-Newtonian formalisms also exist, which take into account higher-order terms.) The PPN formalism has 10 parameters in all; this discussion will be limited to six. Their physical significance for a particular metric theory of gravitation lies in the roles they play in the part of the metric that describes the deviations of the local space-time curvature from Euclidean flat space.

To understand the roles played by the other PPN parameters, it is helpful to see where they appear in the metric. The flat-space four-dimensional metric $\eta_{\mu\nu}$ (μ, ν take on values 0, 1, 2, or 3, corresponding to the time and three spatial components, respectively) is diagonal, with components $\eta_{00} \equiv -1$ and $\eta_{ij} \equiv \delta_{ij}$, where i, j = 1, 2, or 3 δ_{ij} is the Kronecker delta ($\delta_{ij} = 1$ for i=j, $\delta_{ij} = 0$ for i$\neq$j). The space-time metric $g_{\mu\nu}$ at a distance r and angular position θ from a spherical object of mass M, physical radius R, angular momentum $\vec{J}$ (whose direction defines $\theta = 0$), and quadrupole moment J_2, is given by $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, where $h_{\mu\nu}$ has the following nonzero components, up to second order in $\hat{M} \equiv GM/c^2$:

$$
\begin{aligned}
h_{00} &= \frac{2\alpha\hat{M}}{r}\left[1 - J_2\left(\frac{R}{r}\right)^2 P_2(\cos\theta)\right] - 2\beta\left(\frac{\hat{M}}{r}\right)^2 ; \\
h_{0j} &= \frac{-(\eta\Delta_1 + \Delta_2)}{8}\,\frac{\vec{J}\times\vec{r}}{r^3} ; \\
h_{ij} &= \delta_{ij}\left\{\gamma\frac{2\hat{M}}{r}\left[1 - J_2\left(\frac{R}{r}\right)^2 P_2(\cos\theta)\right] + \frac{3}{2}\Lambda\left(\frac{\hat{M}}{r}\right)^2 + X\right\} .
\end{aligned}
\tag{6.2.1}
$$

Here the normalized mass $\hat{M}$ has units of length (*e.g.*, for the Sun, $GM_\odot/c^2 \equiv M_\odot \simeq 1.48$ km); $P_2(x) \equiv (3x^2 - 1)/2$ is the second-order Legendre polynomial; and X represents terms that vanish in the theory of general relativity but in other theories depend on the object's baryon density, pressure, and specific internal energy density (Misner, *et al.* 1973). The parameter α is equal to unity for all metric theories of gravitation, by virtue of the empirical definition of the unit of mass and the gravitational constant; in practice, α is sometimes retained as a parameter that describes the gravitational redshift. The parameter γ, which appears in the diagonal spatial components of the metric multiplied by the gravitational potential $\hat{M}/r$, is the first-order non-Euclidean contribution

to spatial curvature. The parameters β and Λ describe second-order contributions to the time and spatial components of the metric, respectively; they appear in the purely temporal and purely spatial components of the metric multiplied by $(\hat{M}/r)^2$. The parameters Δ_1 and Δ_2 are responsible for an effect known as the "dragging of inertial frames;" they appear in the off-diagonal time-spatial components of the metric multiplied by the angular momentum $\vec{J}$ of the effective mass producing the local gravitational field. In the metric theory of general relativity, all of these parameters take on the value unity.

Before discussing particular experiments and their significance, a few general comments are in order about technology requirements for solar-system tests of relativistic gravitation theories. Precision tracking of solar-system spacecraft is essential. Desired capabilities include ranging accuracy of a few centimeters or better, range-rate accuracy of 0.1 mm/s or better, and angular accuracy to a milliarcsecond or better. These levels of accuracy require that the spacecraft (or test masses within the spacecraft) be kept relatively drag-free by use of active disturbance compensation systems (see discussions in sections 3.4.5, 5.2, and 5.3). Alternatively, planetary landers should be used. The buffeting of solar-orbiting spacecraft by nongravitational forces (solar-wind intensity fluctuations, for example), if uncompensated, would degrade the accuracies by up to four-to-five orders of magnitude. For example, a spacecraft in the vicinity of the Sun (such as the Starprobe mission mentioned in section 6.2.5) would experience nongravitational accelerations on the order of $10^{-5}g$, or about 10^{-4} m/s^2. The fluctuations associated with these nongravitational accelerations must be reduced to an effective level of $10^{-10}g$ or smaller over time intervals on the order of a minute in order to enable meaningful tests of gravitational theories. Planetary orbiters are better, in that their nongravitational accelerations are small, and their orbits can be calibrated frequently. Landers offer the best stability, particularly if they are used for measurements spanning several years. Several significant tests of post-Newtonian gravitation theories could be performed with precise ranging to orbiters or landers on Mercury or Mars, for example, including tests of the principle of equivalence, deflection of light by the Sun's gravitational field, post-Newtonian effects on the orbital motions of

planets, a possible time variation of the gravitational constant G, and others, some of which are described below.

Brief descriptions follow of several proposed tests of gravitation theories that would benefit from the use of coherent laser links and frequency-stabilized laser technology with planetary spacecraft.

6.2.7a Advance of the perihelion of Mercury

This classic test of general relativity, originally proposed by Einstein, measures the secular advance of the perihelion of Mercury's orbit (approximately 43 seconds of arc per century). The predicted advance $\delta\phi_M$ (in radians per revolution) depends on the PPN parameters β and γ, and the solar quadrupole moment J_2, as follows:

$$\delta\phi_M = \frac{6\pi\hat{M}_\odot}{p}\left[\frac{2(1+\gamma)-\beta}{3} + J_2\frac{{R_\odot}^2}{2p\hat{M}_\odot}\right], \tag{6.2.2}$$

where $p \equiv a(1-e^2)$ is the semilatus rectum of Mercury's orbit, the parameters β and γ are defined above, $R_\odot \simeq 7 \times 10^5$ km, and $\hat{M}_\odot \equiv GM_\odot/c^2 \simeq 1.48$ km. The term dependent on the Sun's quadrupole moment J_2 is not a relativistic gravitational effect; it is a Newtonian effect on Mercury's orbit arising from the Sun's asphericity. If the contribution of J_2 is negligible, measurements of the perihelion advance indicate that the term $(2+2\gamma-\beta)/3 = 1 \pm 0.005$, in agreement with general relativity to 0.5 (Shapiro, *et al.* 1976). However, uncertainty in J_2 leaves open the possibility of disagreement with general relativity. Accurate measurement of the Sun's J_2 and of its total angular momentum J also have great importance to solar and stellar physics, as indicated in section 6.2.5.

Placement of an orbiting satellite around Mercury has been proposed (to provide a means of testing several aspects of general relativity and other gravitation theories (Bender 1988). These include (1) measurements of the perihelion precession and of the predicted relativistic time-delay (see below) that would be two orders of magnitude more accurate than currently known; (2) measurement

of the solar quadrupole moment J_2 to an accuracy of 10^{-9}; and (3) determination of the rate of change of the quantity $GM_\odot$ to an accuracy of 10^{-14}/yr [see Bender, *et al.* (1989), Vincent and Bender (1990), and Ashby, *et al.* (1991) for more complete enumerations and descriptions]. These results require ranging accuracies of approximately 3 cm (Bender 1988, Vincent and Bender 1989) and Doppler measurement accuracies for 10-minute observation times of about 10^{-14}. This proposed mission would use dual-frequency (X- and K-band) two-way microwave links. Addition of a coherent laser link could be an important aid in calibrating effects of Earth's ionosphere and atmosphere. And, if a stable laser transmitter and transponder of comparable fractional frequency stability (10^{-14}) could be included on the orbiter, optical tracking measurements could constitute significant tests of their own, in addition to verifying measurements made with the radio links.

6.2.7b Relativistic time-delay and light-deflection experiments

General relativity and other theories of gravitation predict that electromagnetic signals passing near a massive object will be pulled slightly from their trajectory, resulting in an observed time delay and ray deflection analogous to what would be produced by passage through a refractive "atmosphere" surrounding the massive object, with refractivity $\mu_g \equiv 2M/r$ at distance r from the massive object. These effects can be measured best with spacecraft at conjunction with the Sun. The relativistic contribution to the time delay Δt_d measured at Earth depends on the PPN parameter γ as follows (Misner, *et al.* 1973):

$$\Delta t_d \;=\; 2GM_\odot/c^3\,(1+\gamma)\,\ln\!\left(\frac{(1+\cos\theta_e)\,(1+\cos\theta_s)}{\sin\theta_e\;\;sin\theta_s}\right), \tag{6.2.3}$$

where θ_e is the angle subtended at Earth between the Sun and the spacecraft, θ_s is the angle subtended at the spacecraft between the Sun and Earth, and "ln" denotes the natural logarithm. This time delay can be as large as 250 μs for signals that just graze the solar limb ($\sin\theta_e \simeq \sin\theta_s \simeq$

$R_\odot/1$ AU ≈ 0.0047). Measurements made by radio-tracking of the Mariner spacecraft gave results consistent with general relativity to within the estimated errors, about 2% (Anderson, *et al.* 1975), and more recent measurements with the Viking landers on Mars indicate agreement with general relativity to an estimated uncertainty of 0.2% (Shapiro, *et al.* 1977, Reasenberg, *et al.* 1979). This small uncertainty was enabled by a ranging accuracy of a few meters, made possible by combining range measurements to the lander with range measurements to one of the Viking orbiters, which had X-band as well as S-band downlink capability.

The deflection of light due to passage near the Sun has been measured with radio-frequency Very Long Baseline Interferometry (VLBI), in which the relative angular position of two quasars was monitored as one of them passed behind the Sun. These measurements have confirmed that the PPN parameter γ is equal to 1 to within 2% (Fomalont and Sramek 1976, 1977). Accuracy was limited by plasma effects in the solar corona and Earth ionosphere, as well as atmospheric effects. The deflection due to first-order effects (the parameter γ) is about 1.75 arcseconds for light just grazing the Sun (see, for example, Misner, *et al.* 1973). As shown in Figure 28, a microarcsecond astrometric capability opens up the possibility of measuring second-order effects, as well as the predicted "dragging of inertial frames" due to the Sun's total angular momentum $\vec{J}$. Space-based optical interferometers (see discussion in section 5.1), relying on sophisticated laser metrology systems and stable materials and mechanical design, are expected to be capable of such measurements. In the future, light-deflection experiments performed with coherent links to spacecraft carrying frequency-stabilized lasers might provide even more sensitive tests of these subtle effects. As noted several times already, the immunity of the laser links to plasma effects offers a significant advantage, which for this kind of experiment could make up for difficulties associated with optical tracking close to the Sun.

6.2.7c Tests of the principle of equivalence

A third test of general relativity originally proposed by Einstein (after the precession of Mercury's orbit and light deflection by the Sun) is the gravitational redshift of light, which actually is a test of the principle of equivalence and as such applies to all relativistic theories of gravitation. According to this prediction, a clock at a higher gravitational potential (*e.g.*, nearer a massive body) should run more slowly than one at a lower gravitational potential. Thus, signals transmitted from a stable oscillator with frequency ν that are received in a region of lower gravitational potential will appear Doppler-shifted by an amount $\delta\nu/\nu = \alpha\Delta\phi/c^2$, where $\Delta\phi$ is the difference in gravitational potential, and the PPN parameter $\alpha = 1$ for all metric theories. This is a necessary consequence of the equivalence principle, which asserts that all forms of mass and energy are equivalent in producing a gravitational acceleration, and that this gravitational acceleration is indistinguishable from a mechanical acceleration (equivalence of gravitational and inertial mass). The first reliable measurement of this effect was by Pound and Snider (1965) using the Mossbauer effect; they reported a 1% uncertainty in their measurement of the redshift of photons falling 20 meters in the Earth's gravity field. A measurement of the combined effects of the relativistic second-order Doppler shift (time dilation, predicted by special relativity) and the gravitational redshift was accomplished by Vessot, *et al.* (1980), by comparing a hydrogen maser on board a spacecraft in suborbital trajectory at 10,000-km altitude ("Gravity Probe A") with a counterpart on the ground, using multilink radio-frequency Doppler tracking and communications. A separate two-way coherent radio link was used to calibrate the Doppler shift and atmospheric and ionospheric effects, which were then subtracted from measurements of the received signal frequency. With this cancellation scheme, accuracy in comparison of the frequencies was shown to be on the order of 10^{-14} for 100-second integration times. Agreement with the predictions of general relativity was obtained with an uncertainty of 7×10^{-5}. It has been estimated (Vessot 1984) that a direct measurement of the gravitational redshift to second order in the solar mass *i.e.*, to 10^{-9} (nearly five orders of magnitude more sensitive

than the current best measurements), could be made with a hydrogen maser exhibiting a fractional frequency stability on the order of 10^{-14} or better over several hours. This maser would be placed on a spacecraft passing near the Sun. Use of a coherent laser link could aid in calibrating the plasma effects sufficiently to reach the desired measurement accuracy; and, if the onboard laser had a fractional frequency stability comparable to that of the hydrogen maser, the test could be made both more definitive, and, perhaps, more precise.

Violations of the equivalence principle should show up as measurable anomalies in the orbital motions of planets and satellites. Lunar laser ranging has been used to test this, by looking for changes in the Earth-moon separation that would result if the mass equivalent of the gravitational binding energy of the Earth-moon system in the Sun's gravitational field does not behave like ordinary mass (Williams, *et al.* 1976; Shapiro, *et al.* 1976). Specifically, the laser-ranging experiments can be shown to constitute a measurement of the quantity $\eta \equiv 4\beta - \gamma - 3$ (Vessot 1984). A detailed description of these measurements and the variety of relativistic effects that can be tested can be found in Alley (1983). The discrepancy between the prediction of general relativity and the measured value was found to be zero to within the measurement uncertainty of 2%. Recent calculations (Nordtvedt 1988) indicate that laser-ranging experiments to the moon and to the Lageos satellite have in fact provided some verification (by virtue of an absence of predicted anomalous motion) of "gravitomagnetic," or frame-dragging, effects predicted by general relativity. The lunar laser-ranging measurements use fast electronic circuits capable of subnanosecond timing, and are accurate to about 30 cm. Their accuracy has resulted to a great extent from the short, high-energy pulses available from lasers and the capability for sensitive single-photon detection. Use of laser transponders on the lunar surface in place of the retroreflectors now used, which would improve the signal-to-noise ratio ($1/r^2$ fall-off instead of $1/r^4$ fall-off), could enhance the accuracy and broaden considerably the implications for tests of gravitation theories. If the current visions of a lunar outpost become a reality, one could imagine similar experiments performed with frequency-stabilized lasers and coherent laser transponders. The implications for various tests of relativity and gravitation

theories could be significant.

This section has barely touched on the ways in which frequency-stabilized lasers and coherent laser links to or among planetary spacecraft could be used to test general relativity and other fundamental theories. The subject is rich and ripe to be explored. Perhaps this report will help that exploration come about.

7. SUMMARY

This report has investigated aspects of frequency-stabilized lasers that are important for a variety of scientific applications in space. Many error sources and frequency-stabilization issues have been considered in chapters 2, 3, and 4. Chapter 2 examined the characterization of frequency stability and some techniques for enhancing the stability. The impact of frequency stability on measurement precision was investigated in chapter 3 and compared to other factors that limit measurement accuracy in a variety of circumstances. Techniques for making measurements with lasers were examined in chapter 4.

Throughout chapters 2 through 4, heuristic derivations were used to motivate the dependence of measurement precision on a variety of quantities. Rigorous derivations in all the aspects considered were beyond the scope of this report; however, every attempt was made to make the simple, heuristic arguments sound and complete. Many of the derivations were worked out specifically for this report and do not appear in the literature, at least not in easily accessible form. When appropriate, errors or omissions in the available publications were pointed out. The potential scientific applications have been kept in mind throughout and have been cited where necessary to show the need for specific levels of frequency stability.

Chapter 5 gave more detailed descriptions of several important applications of frequency-stabilized lasers in space. Those applications include metrology for optical astrometric and imaging interferometers, gravitational-wave detection, gravity-field mapping, atmospheric wind-sensing using coherent pulsed lidar, and a variety of experiments involving the propagation of laser light through planetary atmospheres and rings and scattering off planetary surfaces. These subsections drew greatly on the general analyses done in the earlier chapters.

Chapter 6 outlined qualitatively some additional scientific applications of frequency-stabilized lasers in space. Most of these applications could be investigated further by using the formalism

developed in chapters 3 and 4 and/or by extending in a straightforward way the analyses of specific applications in chapter 5.

Clearly there is a wealth of scientific applications for space-qualified frequency-stabilized lasers. Detection of low-frequency gravitational waves and high-resolution mapping of Earth's gravity field are two examples of experiments that demand a high level of control or calibration of laser frequency fluctuations. Other applications such as atmospheric wind-sensing and laser metrology systems for microarcsecond astrometry and imaging with optical interferometers put more modest demands on laser frequency stability. Table 7-1 summarizes the requirements on knowledge (control or, where possible, calibration) of laser frequency fluctuations required to enable the scientific applications discussed in this report. Further details and the same table are provided in the Executive Summary.

This report has provided some general background and a convenient framework for analyses of the potential performance and science return of a variety of science experiments in space that use stable oscillators and lasers in particular. The author hopes that it will help or encourage others to expand and improve these analyses and perform new ones, in order to guide technology development and the planning for future missions which someday may bring these and other science applications to fruition.

Table 7-1. Requirements on spacecraft laser frequency stability for a variety of scientific applications. Conservative estimates are given. For astrometry, gravitational-wave detection, and gravity-field mapping, it is assumed that dual-arm interferometers or their equivalent are used to provide cancellation of errors due to laser frequency fluctuations to 0.1%. L denotes interferometer arm length.

Scientific application	Measurement technique	Observable	Information sought	$\delta\nu/\nu$	Reference
Microarcsecond astrometry	interferometry (space) $L \sim 2$ to 20 m	optical-path difference	star angular position and distance	5×10^{-9}	eq. (5.1.3)
Gravitational waves, low frequency	interferometry (space – solar orbit) $L \sim 10^7$ km	optical-path change	strain $\gtrsim 10^{-21}$	10^{-13*}	eq. (3.1.3)
Gravitational waves, pulses	interferometry (ground) $L \sim 5$ km	optical-path change	"	10^{-17}	eq. (3.4.6d)
High-resolution gravity-field mapping (1 mgal, resolution 40 to 50 km, 10^6 1-s measurements)	interferometry (in orbit)	change in range rate	gravity anomalies, geoid undulation	10^{-10} to 10^{-9}	eq. (5.3.10)
Atmospheric wind-sensing	pulsed Doppler lidar	velocity, range	wind speeds	10^{-10}	eq. (5.4.1)
Atmospheric occultation, Venus	one-way Doppler	Doppler shift, intensity	temperature, pressure profiles	10^{-11}	Table 5.5.1
Atmospheric occultation, Jupiter	one-way Doppler	"	"	10^{-13}	"
Ring occultation	one-way Doppler	"	phase shift complex extinction	10^{-15}	§5.5.2c
Ring occultation	one-way Doppler	"	forward-scattering cross section, column-density gradient	10^{-11}	eq. (5.5.16)

*With active laser phase calibration and appropriate tracking filters, this may be relaxed to 10^{-13} and perhaps made even less stringent.

FIGURES

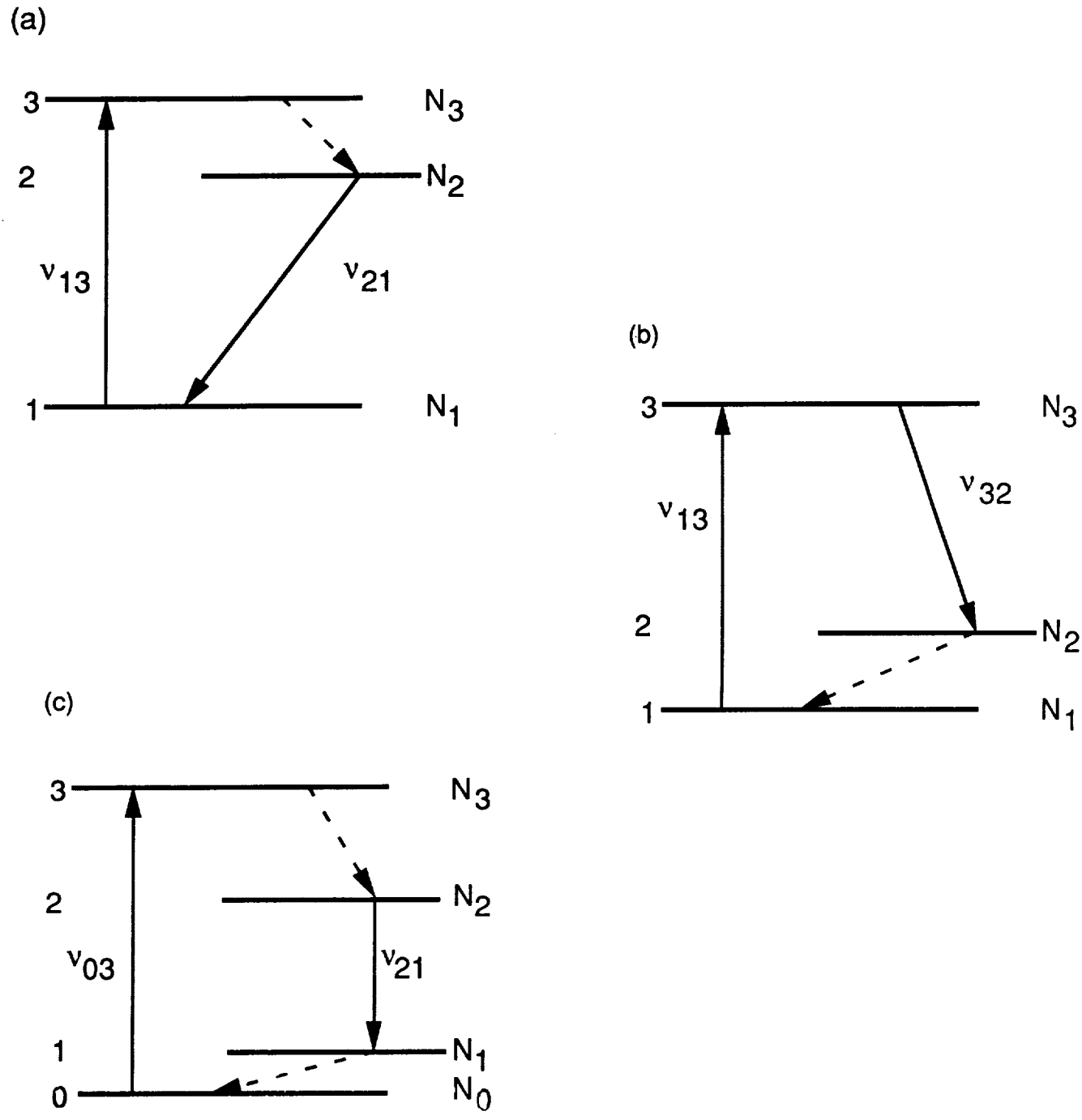


Figure 1. Three- and four-level systems of quantum states for a laser. N_0, N_1, N_2, and N_3 represent the relative populations of the levels. In (1a), the 3–2 transition is rapid (highly probable) and nonradiative, and an inversion is established between 2 and 1 ($N_3 < N_1 < N_2$). In (1b), the 2–1 transition is rapid and nonradiative, leaving an inversion between 3 and 2 ($N_2 < N_3 < N_1$). In (1c), the 3–2 transition is rapid and nonradiative, but a direct transition 2–0 is forbidden. Instead, the metastable level 2 decays to 1, from which it returns rapidly and nonradiatively to 0. The population N_1 is negligible; hence the inversion between 2 and 1 can be reached more easily than in a three-level system ($N_1 < N_3 < N_2 < N_0$). Levels and 2 and 1 must both be relatively narrow.

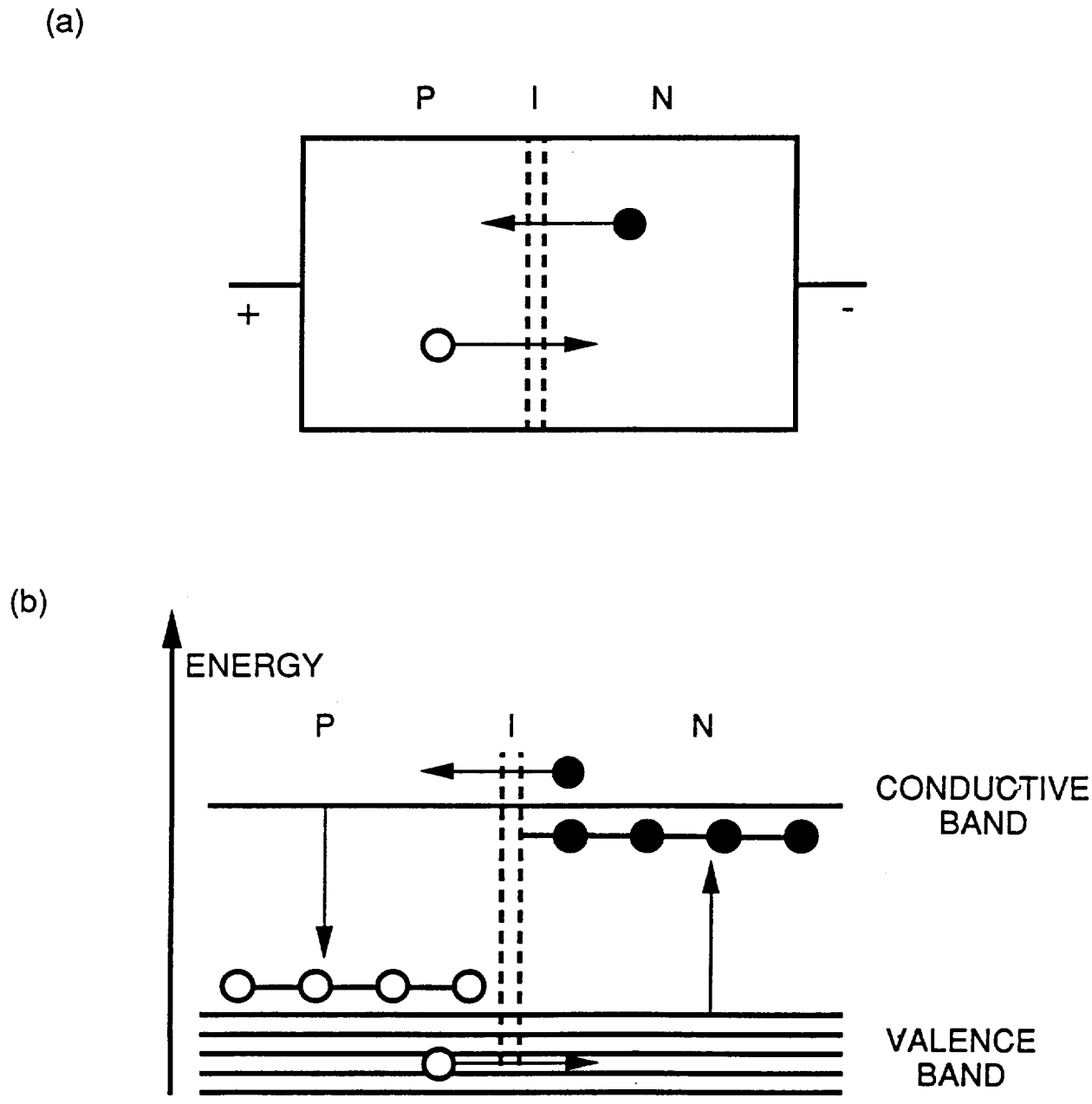


Figure 2. Two-level system for a semiconductor laser. Electrical pumping produces an inversion at the P-N junction, where electrons in the conduction band and holes in the valence band come in contact. Maintenance of the necessary electrical bias ensures a constant flow of electrons and holes toward the junction.

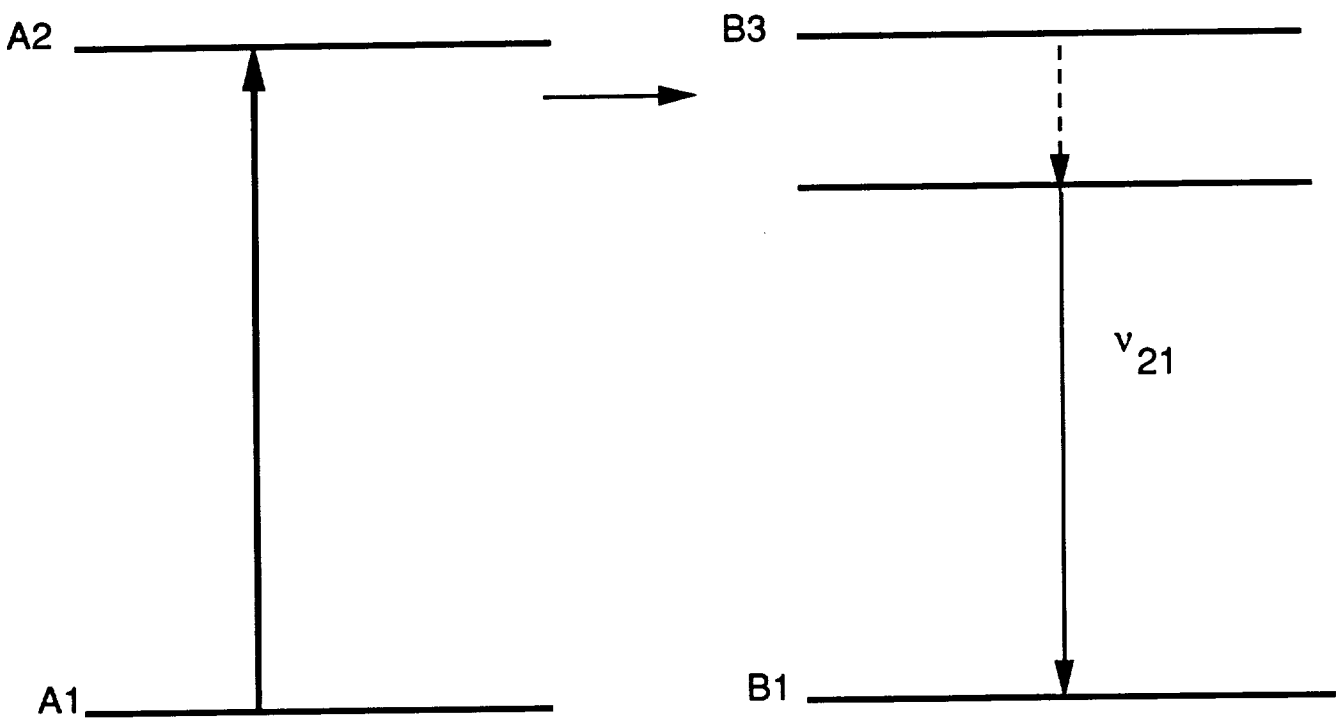


Figure 3. Energy levels for a gas laser with a mixture of two gases. Gas A is pumped to populate level A2. Energy is transferred to gas B by inelastic collisions, bringing gas B to level B3. The B3–B2 transition is rapid and nonradiative, producing an inversion between B2 and B1.

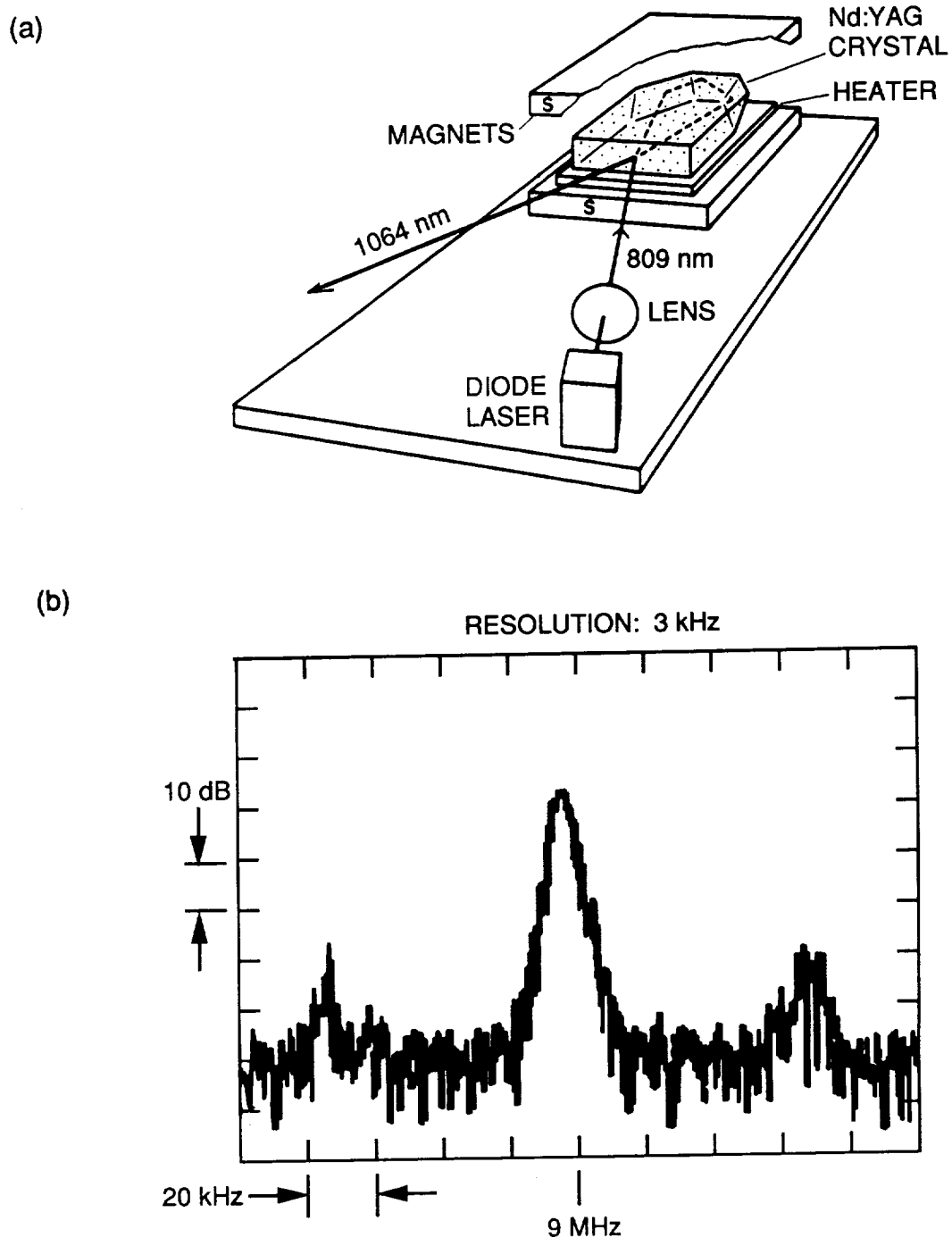
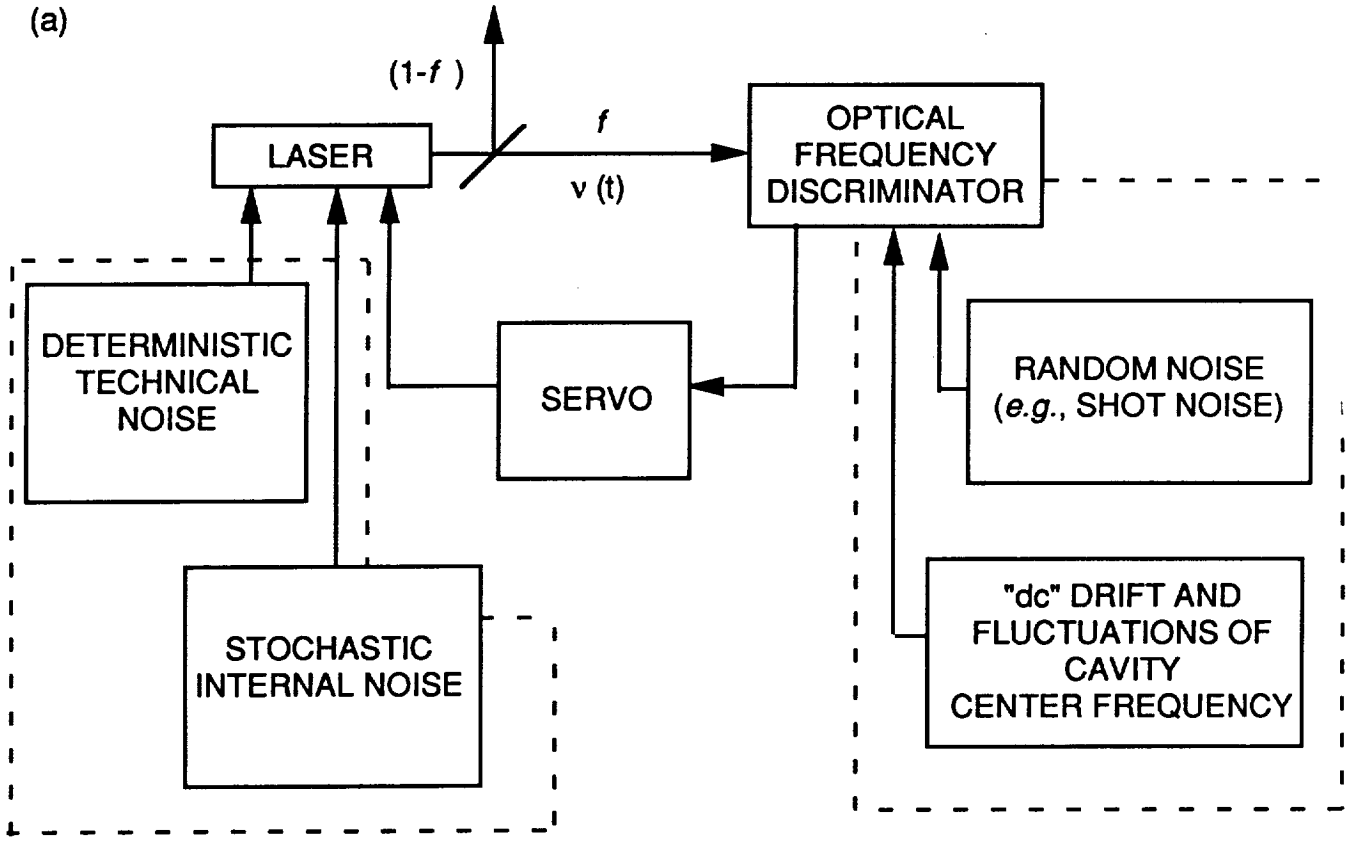


Figure 4. Laser-diode-pumped Nd:YAG nonplanar ring oscillator (NPRO). Figure 4a is a schematic. The output wavelength is 1.064 μm. Figure 4b is a spectrum-analyzer trace of the heterodyne beat signal between two free-running diode-pumped Nd:YAG NPROs, showing a central peak at the 9-MHz offset frequency and relaxation-oscillation sidebands. FWHM of the central peak is 3 kHz, limited by the resolution bandwidth of the spectrum analyzer (from Kane, *et al.* 1987).



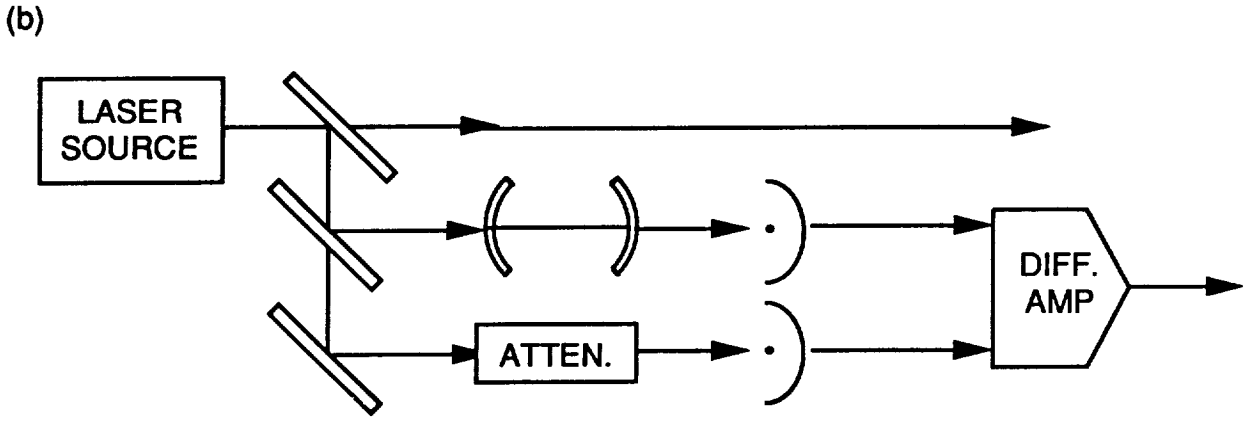
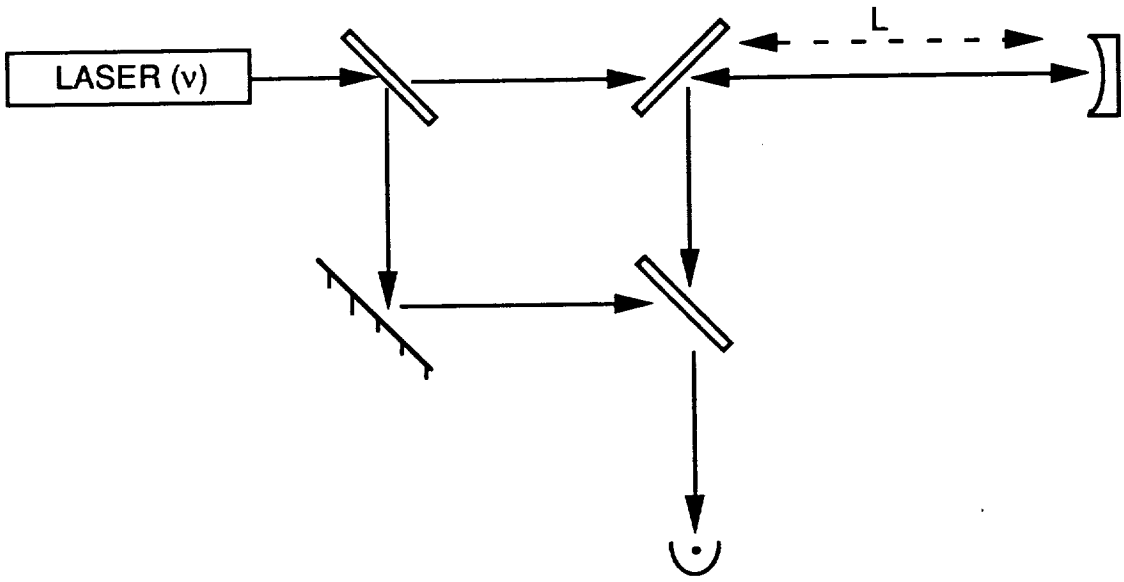


Figure 5. Two common methods of active-feedback laser frequency stabilization. In (5a), a fraction f of the laser light is directed to an optical frequency discriminator, whose error signal is amplified and fed back to the laser. The laser's free-running frequency stability is affected by deterministic technical noise (such as thermal or acoustic noise from the environment) and by internal, stochastic noise. Performance of the optical frequency discriminator is affected by drift and fluctuations of the center frequency of the cavity and by measurement noise within the feedback loop, such as the random "shot" noise associated with photoelectron counting statistics. Figure 5b depicts a simple "fringe-side" optical frequency discriminator (from Hall 1986). See section 2.2.3 of text.



$$\delta L = (\lambda/2\pi)\,\delta\phi :$$

$$\delta L_\nu = L\,\frac{\delta\nu}{\nu} \qquad \delta L_{ph} = \left(\frac{\lambda}{2\pi}\,\frac{h\nu}{P_0\tau}\right)^{1/2}$$

$$\delta L_n = L\,\delta n \qquad \delta L_{sc} = \varepsilon_{sc}\,\Delta L_{sc}\,\frac{\delta\nu}{\nu}$$

Figure 6. Simplistic drawing of a single-arm laser interferometer, in which the relative phase is monitored between light emitted from a reference laser at time t and light emitted from the laser a time $\Delta t = 2L/c$ earlier, the latter having travelled an additional distance $2L$. Sources of error in measurement of the optical path L (or $2L$, as shown here) include fluctuations in the laser frequency (δL_ν), photon statistics (δL_{ph}), scattered light (δL_{sc}), and fluctuations in the refractive index of the optical medium (δL_n). Errors that grow linearly with arm length L can be suppressed in a dual-arm interferometer in which the arm lengths L_1 and L_2 are held nearly equal and the difference in arm lengths, $\ell \equiv L_1 - L_2$, is the quantity measured. The fractional error caused by those error sources is suppressed by the factor ℓ/L, which may be of order 10^{-3} or smaller in practice.

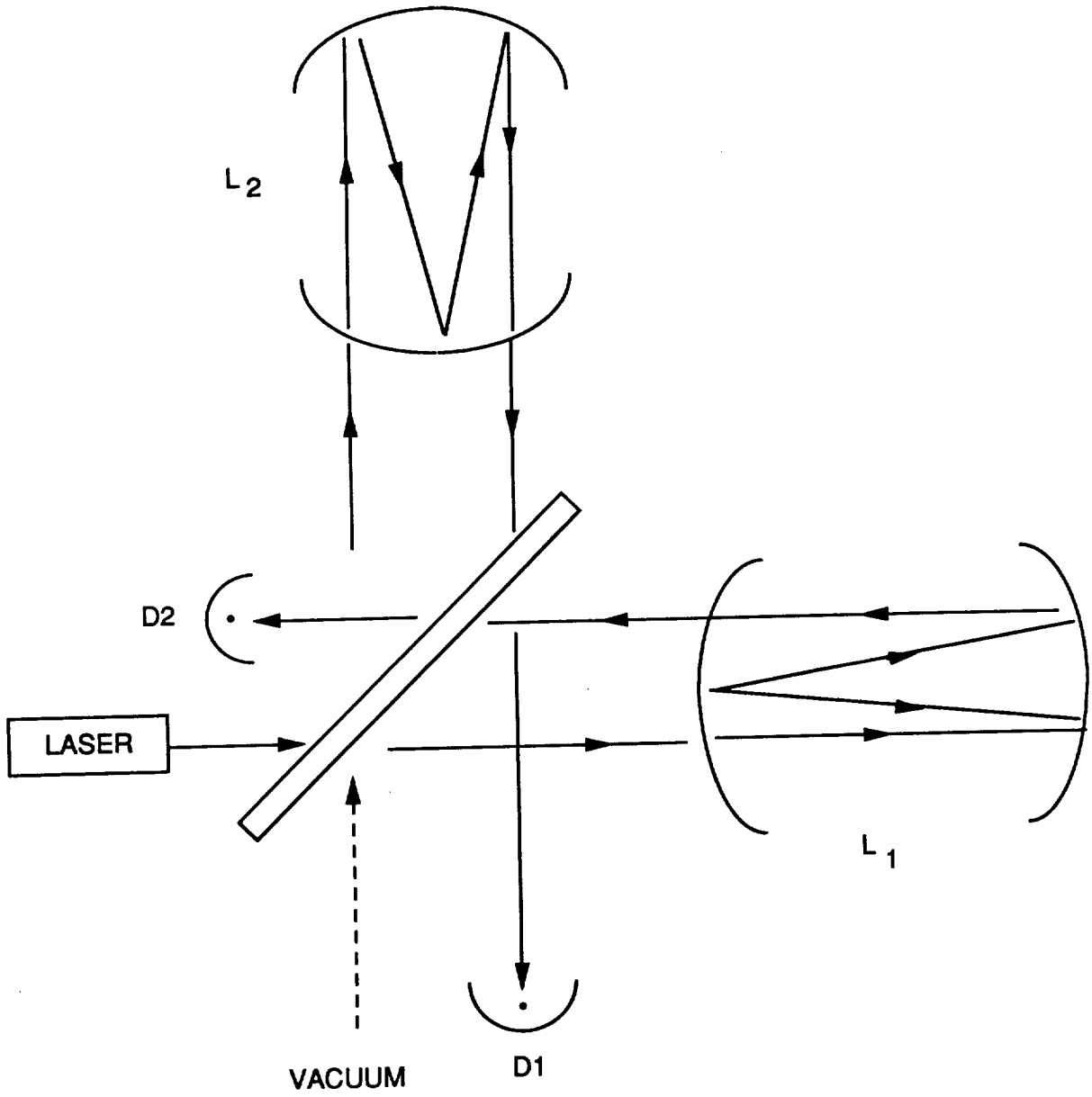


Figure 7. Schematic of a dual-arm multireflection Michelson interferometer. (Only two round trips are shown in the delay lines depicted here.) The reference laser beam is split in two at a 50-50 beam splitter; the second input port of the beam splitter typically is unused, *i.e.*, the field entering that port is the vacuum field. The arms can also be operated as optical (Fabry-Perot) cavities, instead of delay lines.

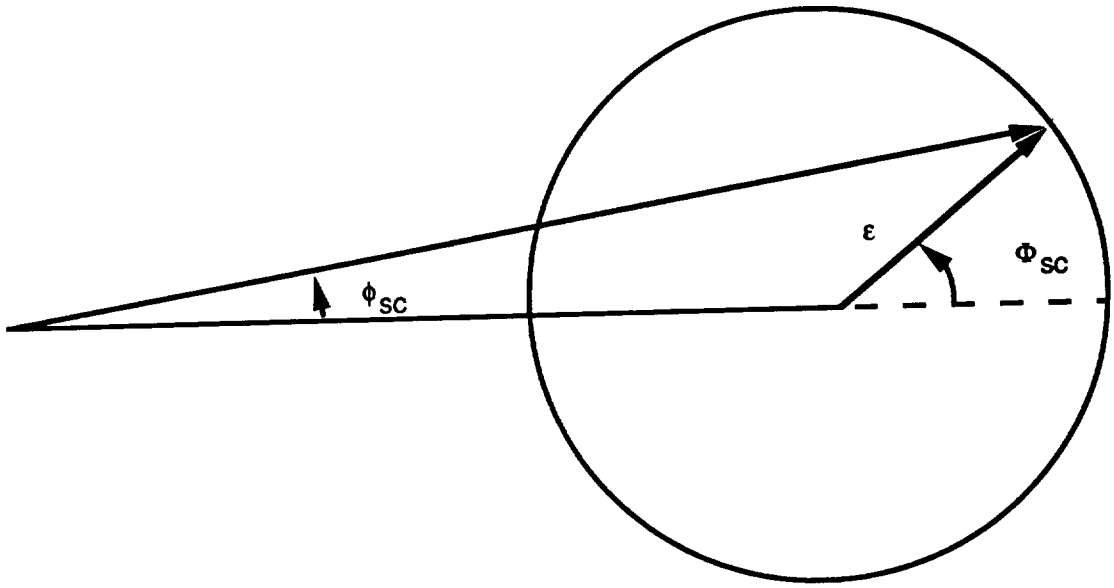


Figure 8. Phase shift ϕ_{sc} of main beam due to interference with scattered light of relative amplitude ϵ; see discussion in §3.4.2 (after Schnupp, *et al.* 1985).

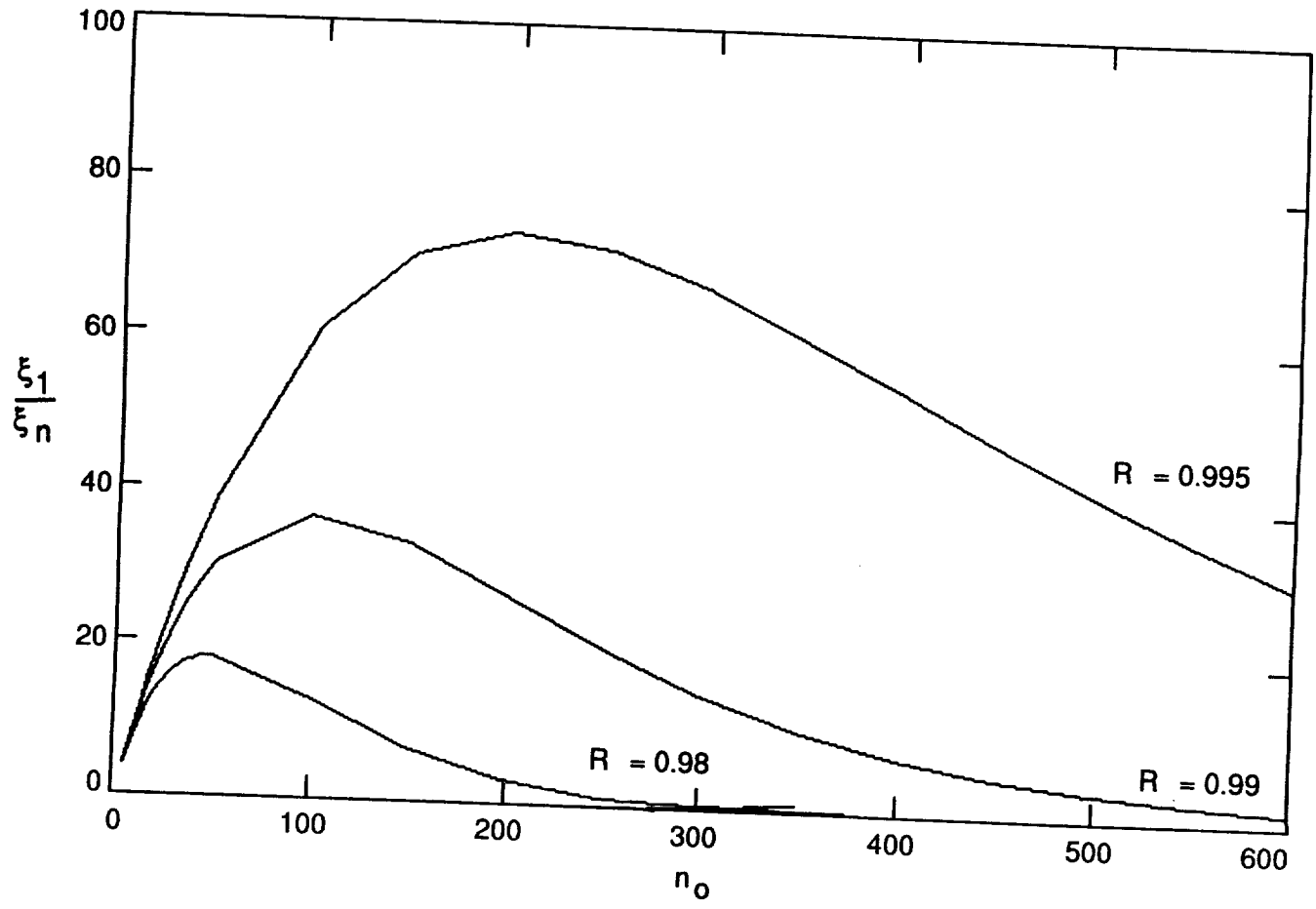


Figure 9. Graph showing the improvement in sensitivity with multiple reflections over a simple, single-reflection Michelson. The graph plots the ratio of fractional measurement error $\xi \equiv \delta L/L$ for one reflection to that for n reflections off the end mirrors (n round trips), for various mirror reflectivities R. See discussion in §4.1.2a [eqs. (4.1.1) and (4.1.2)].

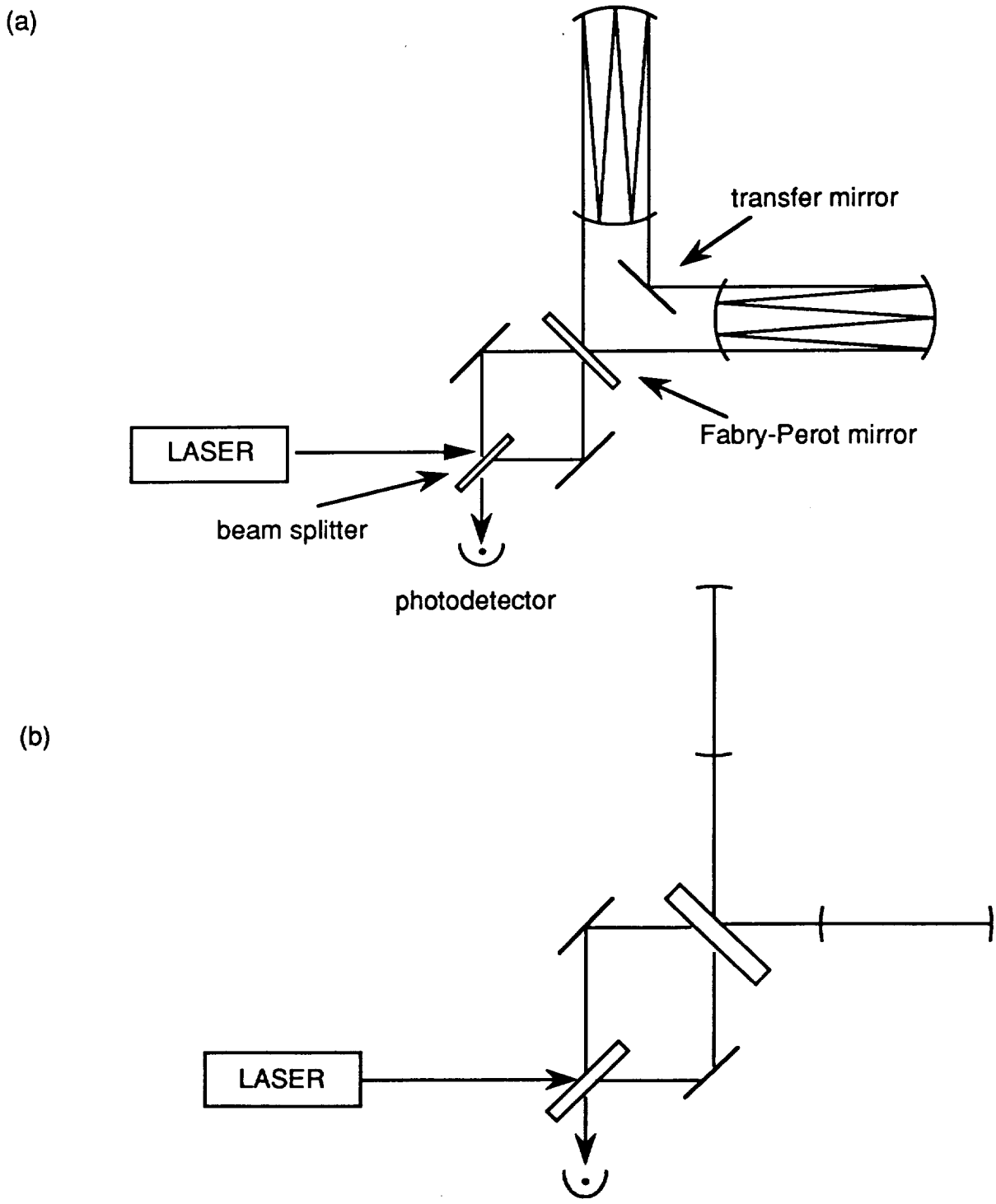


Figure 10. Schematics of optical configurations that could be used for resonant recycling of laser light in (a) multireflection Michelson interferometers and (b) optical-cavity interferometers. See discussion in subsection 4.1.2c.

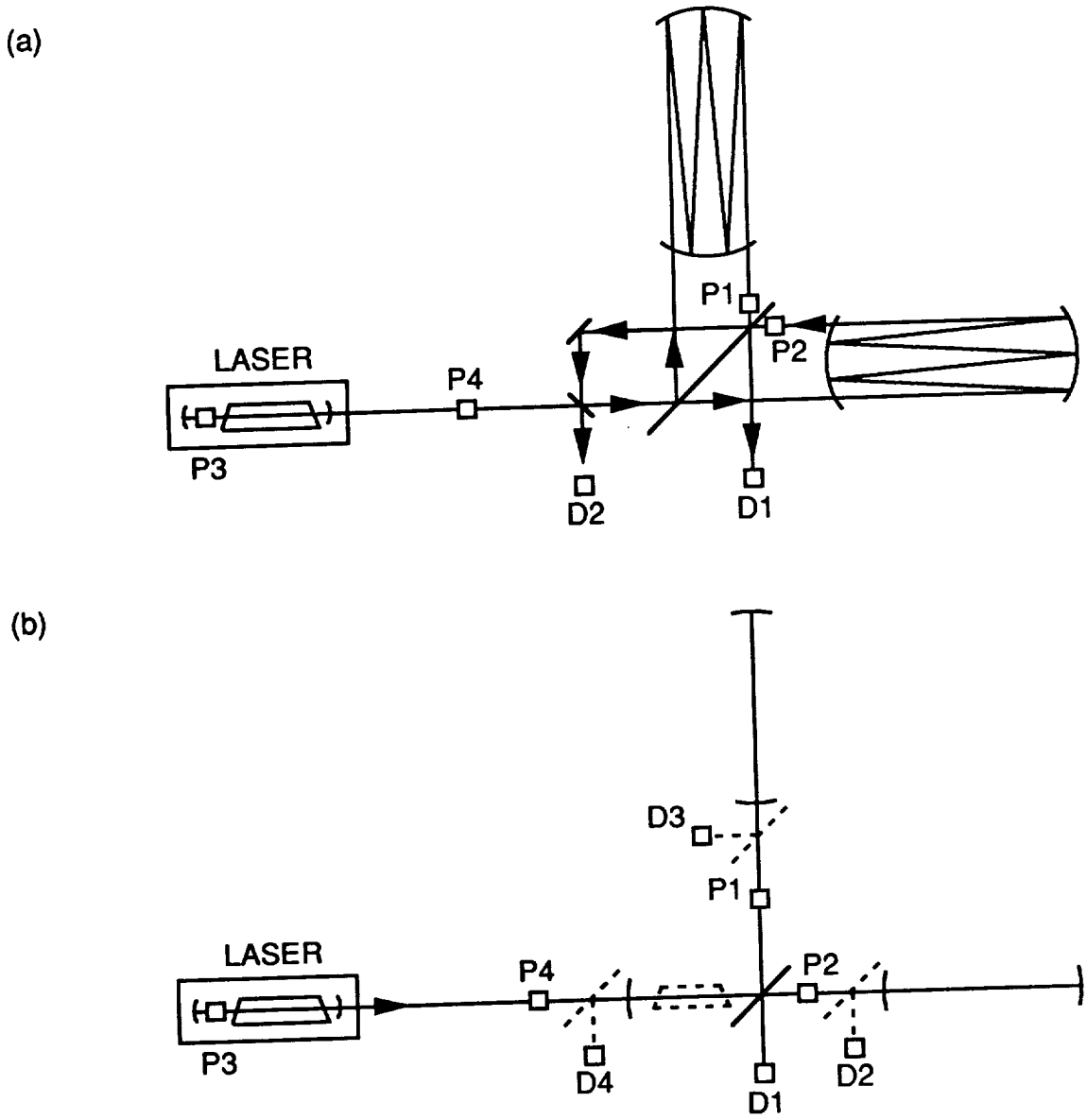


Figure 11. Power recycling schemes for laser-interferometer gravitational-wave detectors, used to increase the effective circulating laser power and thus drive down the photon-statistics limit to measurement precision. See discussion in subsection 4.1.4b. Figure 11a shows the technique with a multireflection Michelson interferometer; Figure 11b shows it with an optical-cavity interferometer. Laser light bounces back and forth in the two arms and recombines at the beam splitter, after passing through optional Pockels cells P1 and P2. (A gravitational-wave detector might drive P1 and P2 in antiphase at a modulation frequency high compared to the measurement frequency, for example.) The optical path difference between the arms is controlled with feedback to maintain an intensity minimum at photodetector D1. Output laser light from the other port of the beam splitter (which exhibits an intensity maximum) is fed back coherently to reinforce the input laser light. The laser frequency is adjusted for resonance by Pockels cell P3, which is controlled by feedback from D2 (11a) or D2, D3, and D4 (11b). P4 is used to phase modulate the laser light, and the error signal used to adjust the laser frequency is obtained by demodulating the output from D2. Near resonance, considerable buildup of light flux can occur (from Drever, *et al.* 1983a).

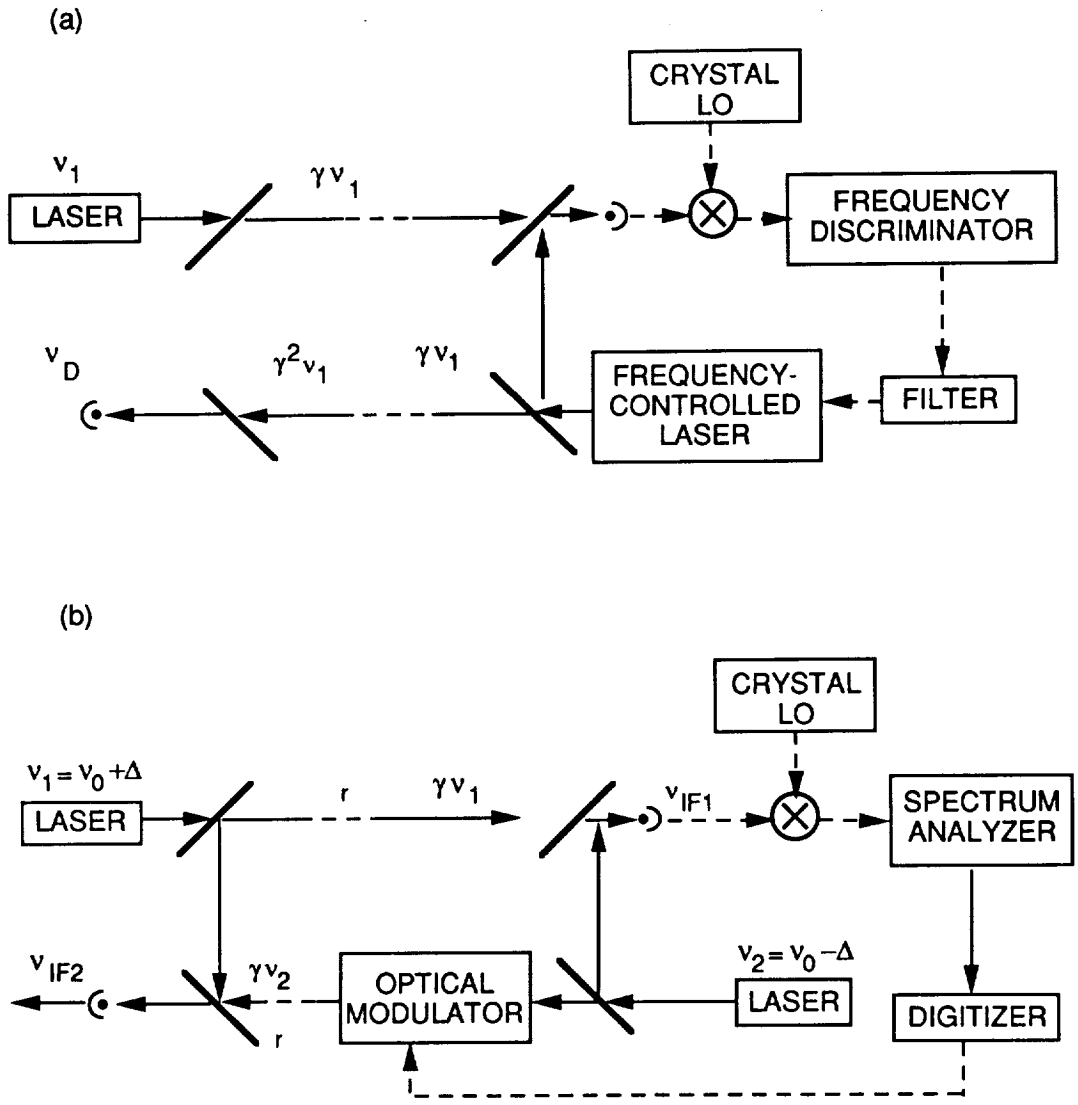


Figure 12. Two-way coherent range-rate (Doppler) measurement systems (see discussion in section 4.2.1). In figure 12a, there is a stabilized laser transmitter and local-oscillator (LO) at spacecraft 1 and a phase- or frequency-matching laser transponder at spacecraft 2. In figure 12b, there are stabilized laser transmitters and LOs at each spacecraft. In each case, the frequencies of the lasers at the two spacecraft are offset from one another by at least the expected one-way Doppler shift. In case (b), processing of the heterodyne beat-frequency measurements made on both spacecraft (ν_{if1} and ν_{if2}) provides the functional equivalent of the two-way operation in (a). Both schemes benefit from cancellation of low-frequency laser phase fluctuations, provided $r/c < \tau_{coh}$.

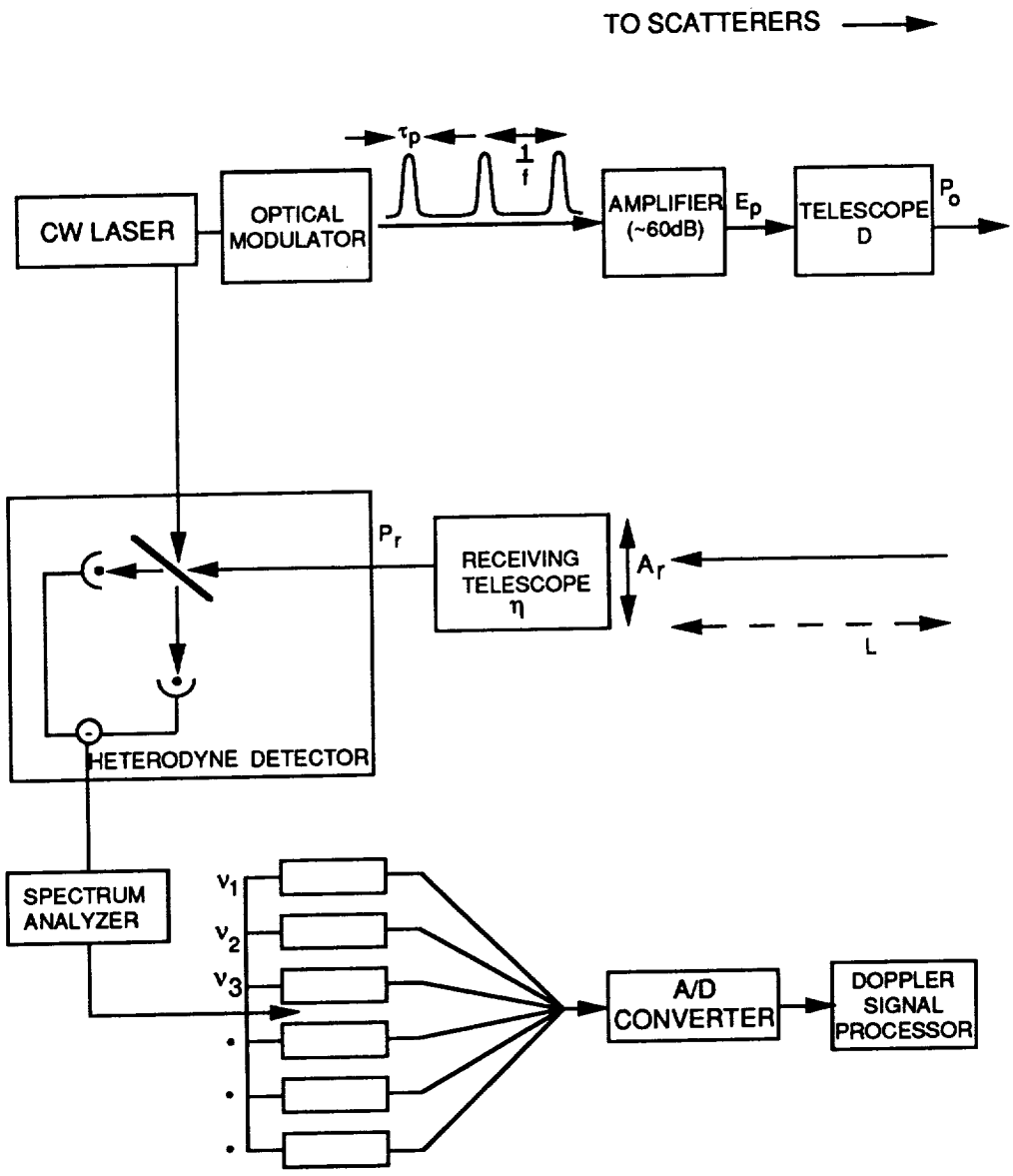


Figure 13. Schematic of coherent Doppler lidar system for real-time velocity and range measurements on stochastic targets. See discussion in section 4.2.2.

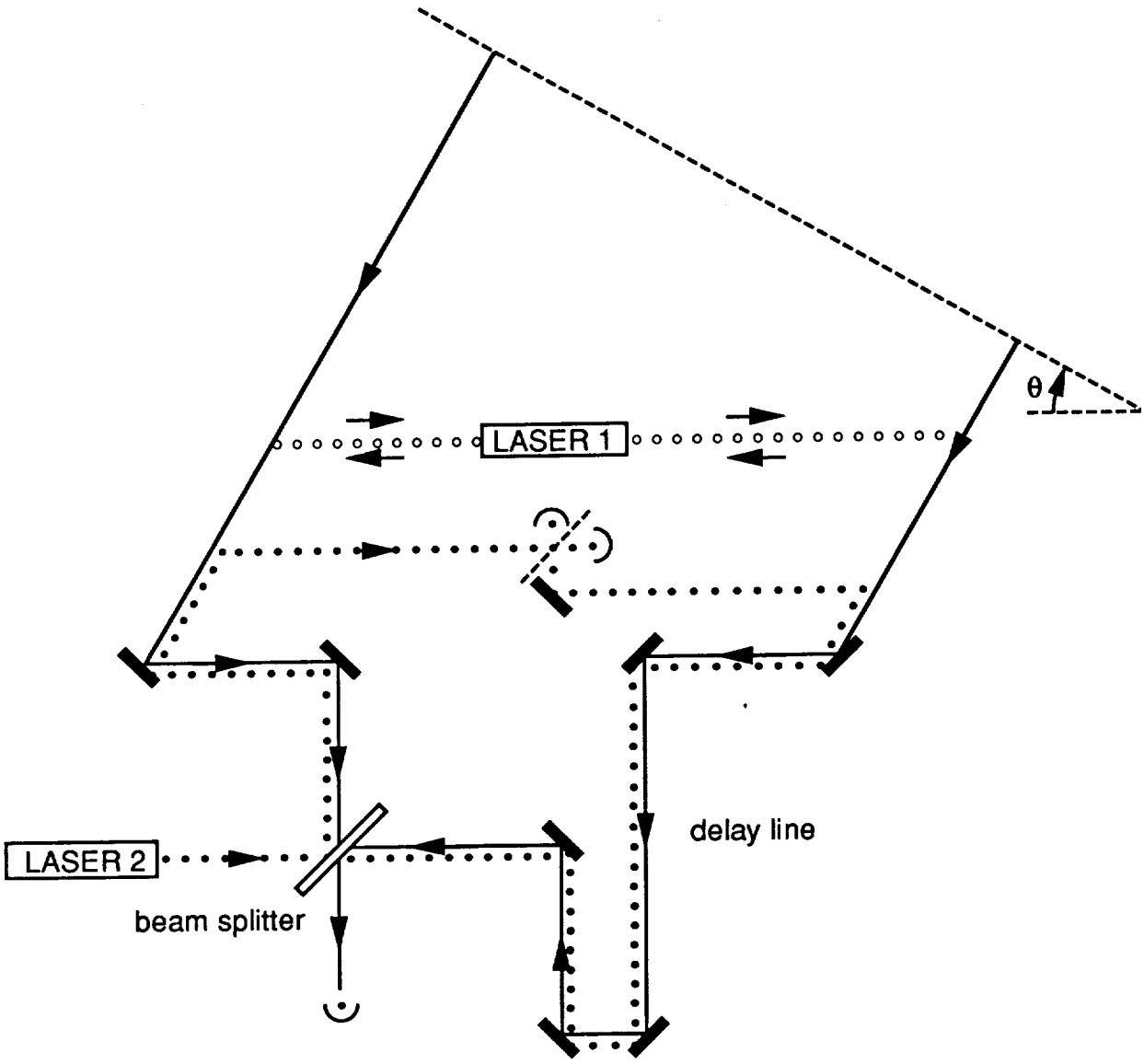


Figure 14. Generic, simplified laser metrology systems for astrometric and astronomical imaging interferometers. (See discussion in section 5.1.) The difference in path lengths travelled by the signal (starlight, denoted by the solid, arrowed lines) to detectors 1 and 2 is determined by precise measurement of interference fringes between the detector outputs. The angular spacing between fringes is $\Phi = \lambda/L \cos\theta$, where L is the baseline separation between the detectors, λ is the wavelength of the starlight, and θ is the angle between the interferometer baseline and the incoming wavefront. In this figure, two laser metrology interferometers are used to measure and control the relative path lengths. Light from laser 1, whose path is indicated with open circles, is used to measure the baseline between the two telescopes. Light from laser 2 is used to monitor the average positions of all optical elements encountered by the starlight (a simplistic rendition of what is known as "full-aperture metrology"). This light, whose path is indicated by filled circles, is injected through the primary beam splitter, traces the path of the input signal in the opposite direction (including transversal of the delay line), and is recombined with itself at the front end of the system. Details of the reflections in front of the primary mirrors are not shown. A sophisticated version of this metrology system has been proposed for an orbiting astrometric optical interferometer ("POINTS"); see Reasenberg, *et al.* (1988) for discussion and a more detailed configuration.

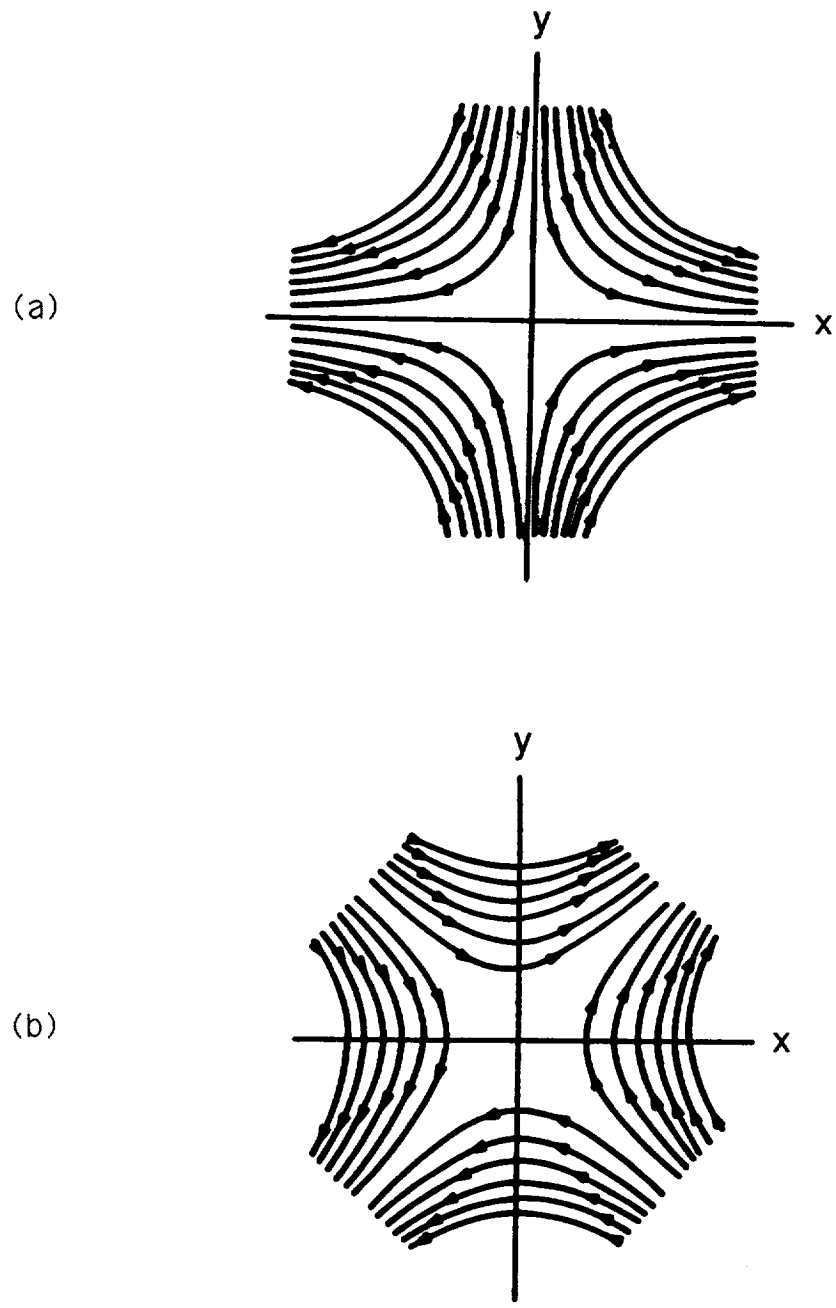


Figure 15. Lines of force associated with the two possible polarizations for gravitational waves predicted by general relativity.

Figure 16. Spectral densities of expected strain amplitudes for various astrophysical sources of gravitational waves and sensitivities of different types of detectors. Figures 16a and 16b show expected strains *vs.* frequency for periodic sources of gravitational waves; Figure 16a considers all kinds of detectors, while Figure 16b considers only space-based long-baseline laser interferometers. Figures 16c and 16d do the same, but for burst sources of gravitational waves. Figure 16e shows expected strains *vs.* frequency for a stochastic background. Figures 16a, 16c, and 16e are adapted from Thorne 1987: "beams" refer to laser-interferometric detectors, "LIGO" refers to ground-based laser-interferometric detectors of the sort currently being developed by Caltech, MIT, and the Max-Planck Institut, and "bars" refer to resonant-bar detectors such as those used by groups at Maryland, Stanford, and Tokyo. Figures 16b and 16d are from Bender (private communication). The symbols SN, BH, WD, and NS stand for supernova, black hole, white dwarf, and neutron star, respectively. The symbol E_{gw} is the energy radiated in the form of gravitational waves. In these figures M refers to the mass of a black hole, and $M_\odot$ is the solar mass. In Figure 16e, the symbol Ω_{gw} represents the ratio of the gravitational-wave energy (in the bandwidth of interest) to the energy required to close the universe, assuming a Hubble constant $H_0 = 100$ km s^{-1} Mpc^{-1}. See section 5.2 and references therein for further discussion.

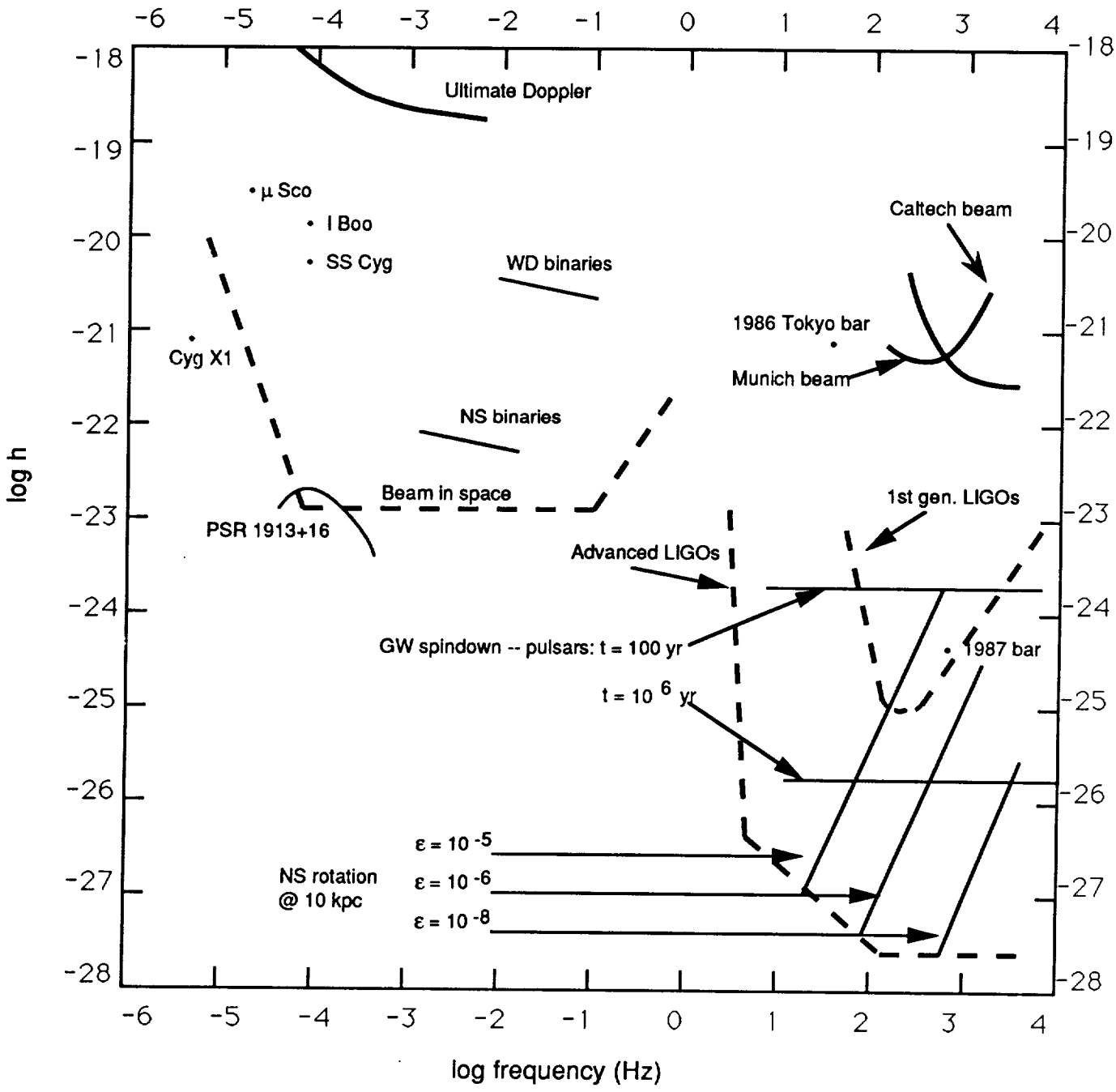

Ultimate Doppler
• μ Sco
• I Boo
• SS Cyg
WD binaries
Caltech beam
1986 Tokyo bar
Munich beam
Cyg X1
NS binaries
Beam in space
PSR 1913+16
1st gen. LIGOs
Advanced LIGOs
GW spindown -- pulsars: t = 100 yr
1987 bar
t = 10 6 yr
ε = 10 -5
ε = 10 -6
ε = 10 -8
NS rotation
@ 10 kpc
log h
log frequency (Hz)

Figure 16a.

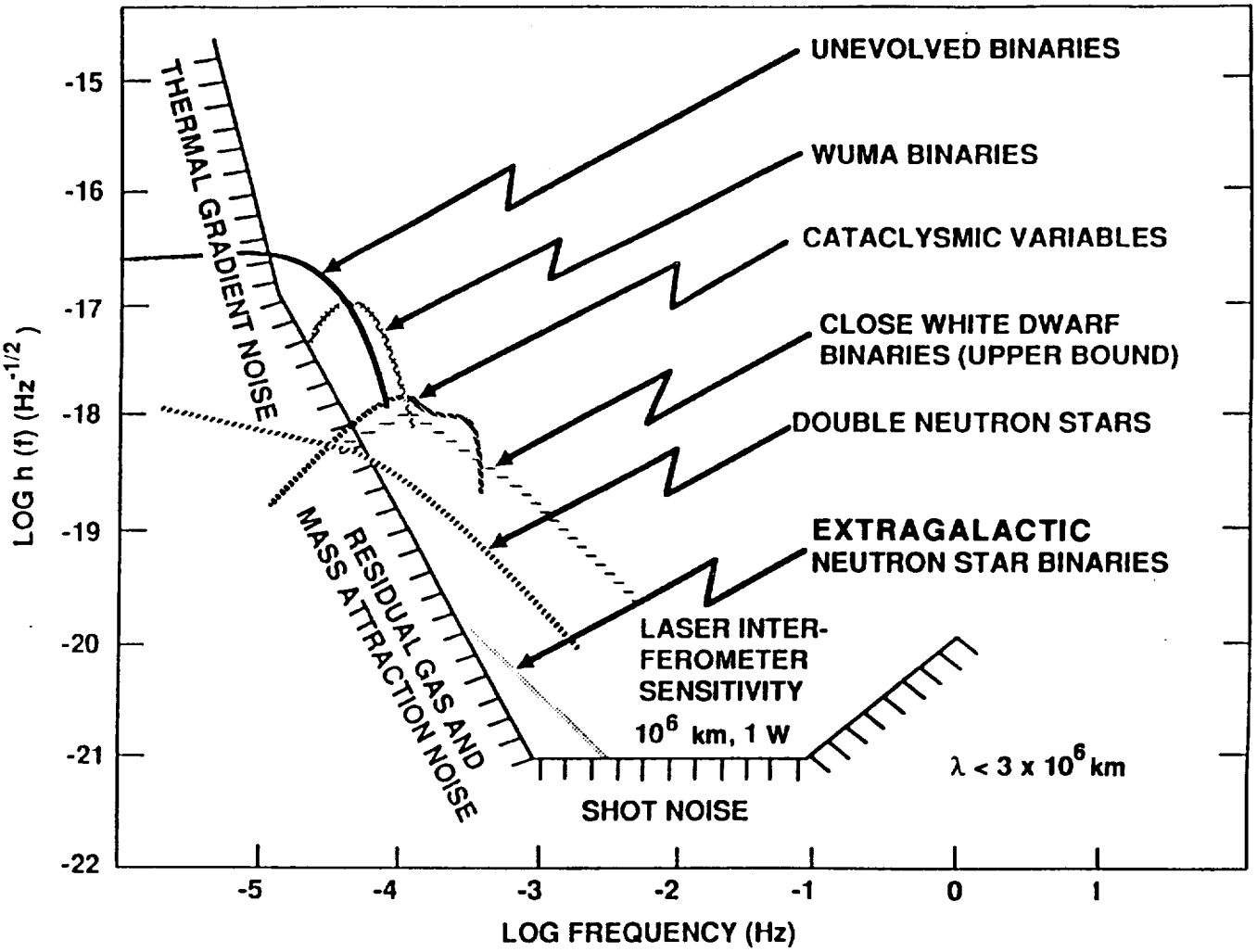

UNEVOLVED BINARIES
WUMA BINARIES
CATACLYSMIC VARIABLES
CLOSE WHITE DWARF BINARIES (UPPER BOUND)
DOUBLE NEUTRON STARS
EXTRAGALACTIC NEUTRON STAR BINARIES
LASER INTER-FEROMETER SENSITIVITY
10^6 km, 1 W
λ < 3 x 10^6 km
THERMAL GRADIENT NOISE
RESIDUAL GAS AND MASS ATTRACTION NOISE
SHOT NOISE
LOG h (f) (Hz^-1/2)
LOG FREQUENCY (Hz)
-15
-16
-17
-18
-19
-20
-21
-22
-5
-4
-3
-2
-1
0
1

Figure 16b.

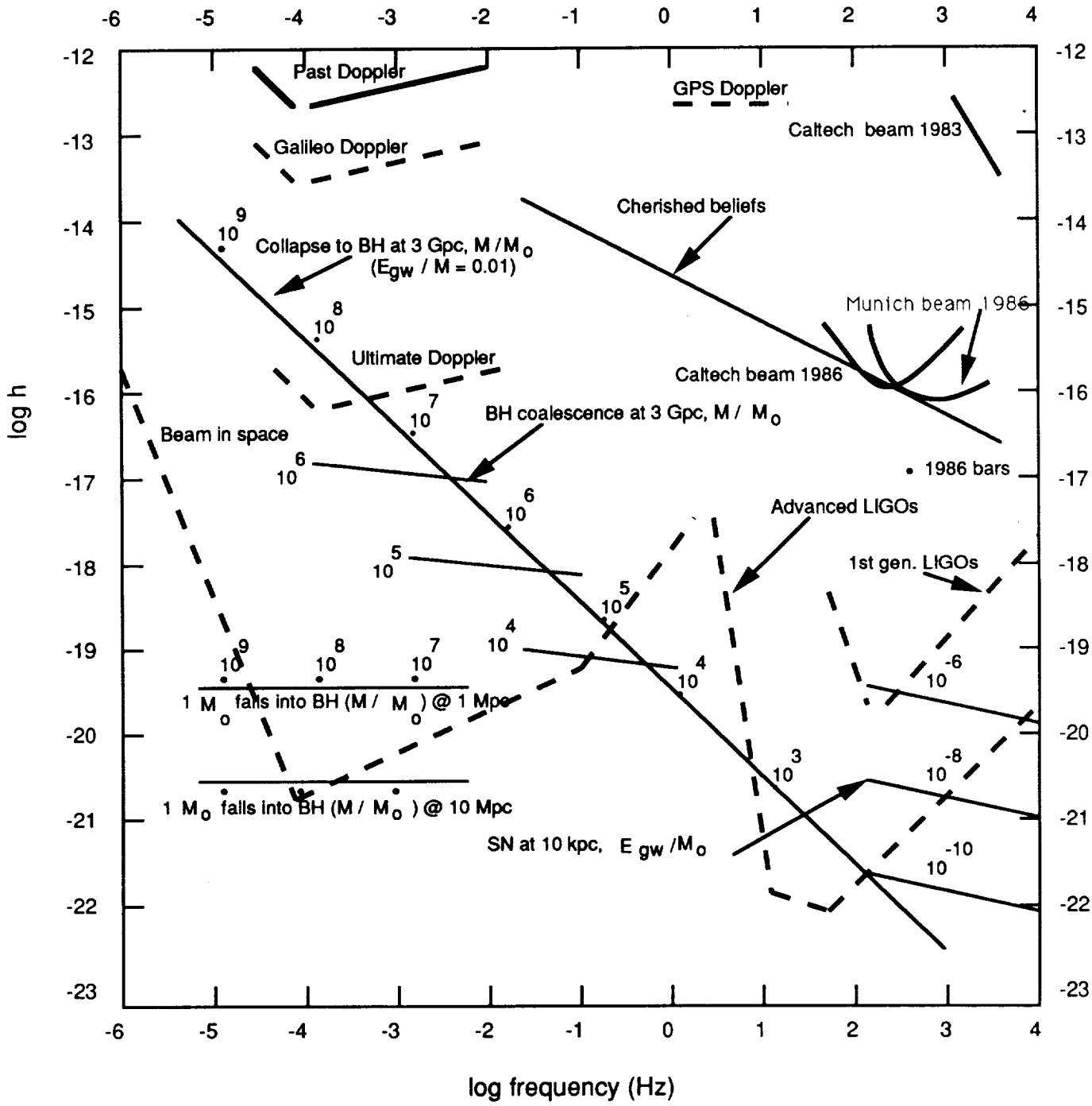

Past Doppler
GPS Doppler
Galileo Doppler
Caltech beam 1983
Cherished beliefs
Collapse to BH at 3 Gpc, M / M o
(Egw / M = 0.01)
Munich beam 1986
Ultimate Doppler
Caltech beam 1986
BH coalescence at 3 Gpc, M / M o
Beam in space
• 1986 bars
Advanced LIGOs
1st gen. LIGOs
1 M o falls into BH (M / M o) @ 1 Mpc
1 M o falls into BH (M / M o) @ 10 Mpc
SN at 10 kpc, E gw /M o
log h
log frequency (Hz)

Figure 16c.

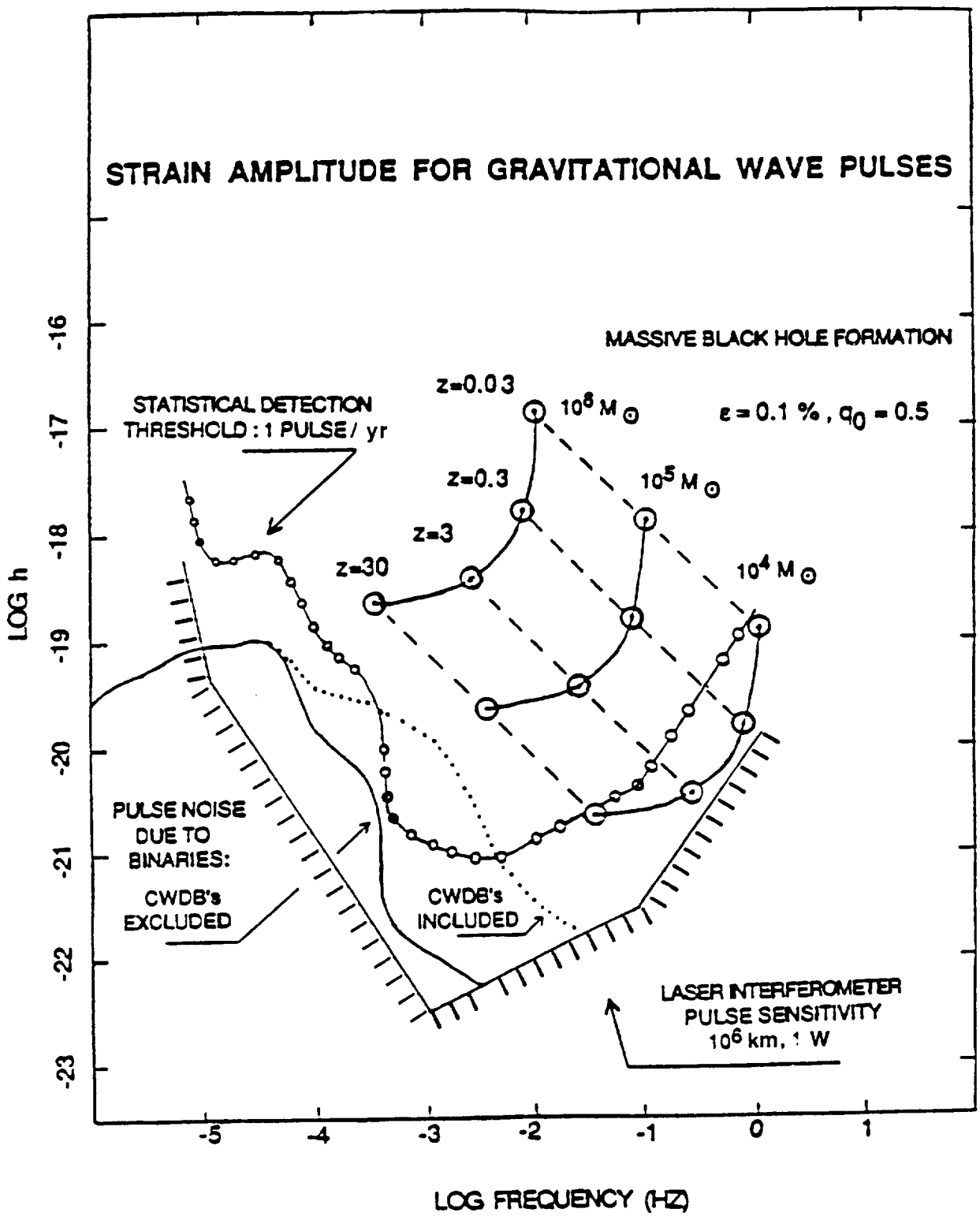

STRAIN AMPLITUDE FOR GRAVITATIONAL WAVE PULSES
MASSIVE BLACK HOLE FORMATION
STATISTICAL DETECTION
THRESHOLD : 1 PULSE / yr
z=0.03
$10^6 M_\odot$
$\varepsilon = 0.1\ \%$, $q_0 = 0.5$
z=0.3
$10^5 M_\odot$
z=3
z=30
$10^4 M_\odot$
PULSE NOISE
DUE TO
BINARIES:
CWDB's
EXCLUDED
CWDB's
INCLUDED
LASER INTERFEROMETER
PULSE SENSITIVITY
10^6 km, 1 W
LOG h
-16
-17
-18
-19
-20
-21
-22
-23
-5
-4
-3
-2
-1
0
1
LOG FREQUENCY (HZ)

Figure 16d.

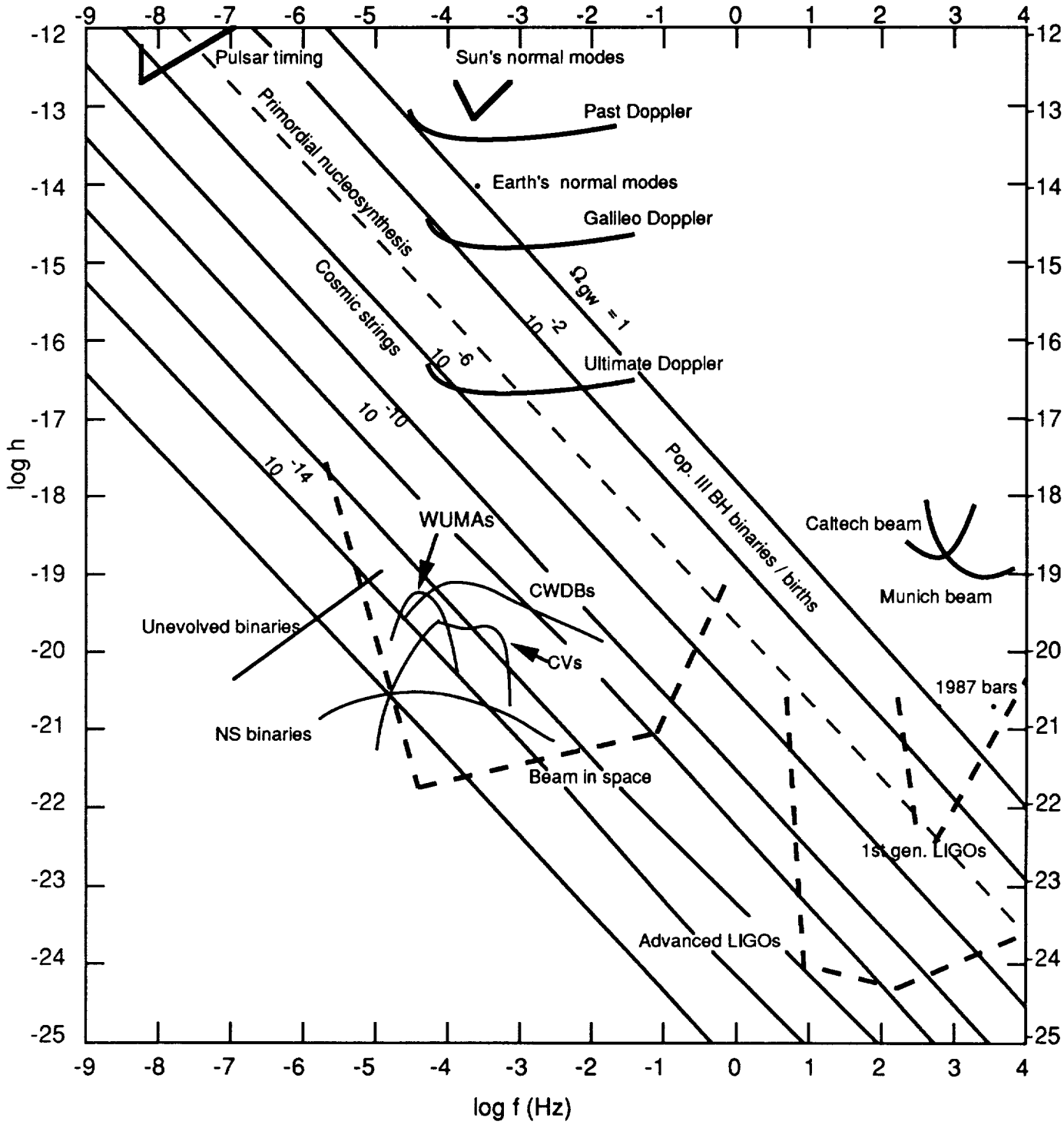

Pulsar timing
Sun's normal modes
Past Doppler
Primordial nucleosynthesis
Earth's normal modes
Galileo Doppler
Cosmic strings
Ωgw =1
10 -2
10 -6
Ultimate Doppler
10 -10
10 -14
Pop. III BH binaries / births
WUMAs
Caltech beam
Munich beam
CWDBs
Unevolved binaries
CVs
1987 bars
NS binaries
Beam in space
1st gen. LIGOs
Advanced LIGOs
log h
log f (Hz)

Figure 16e.

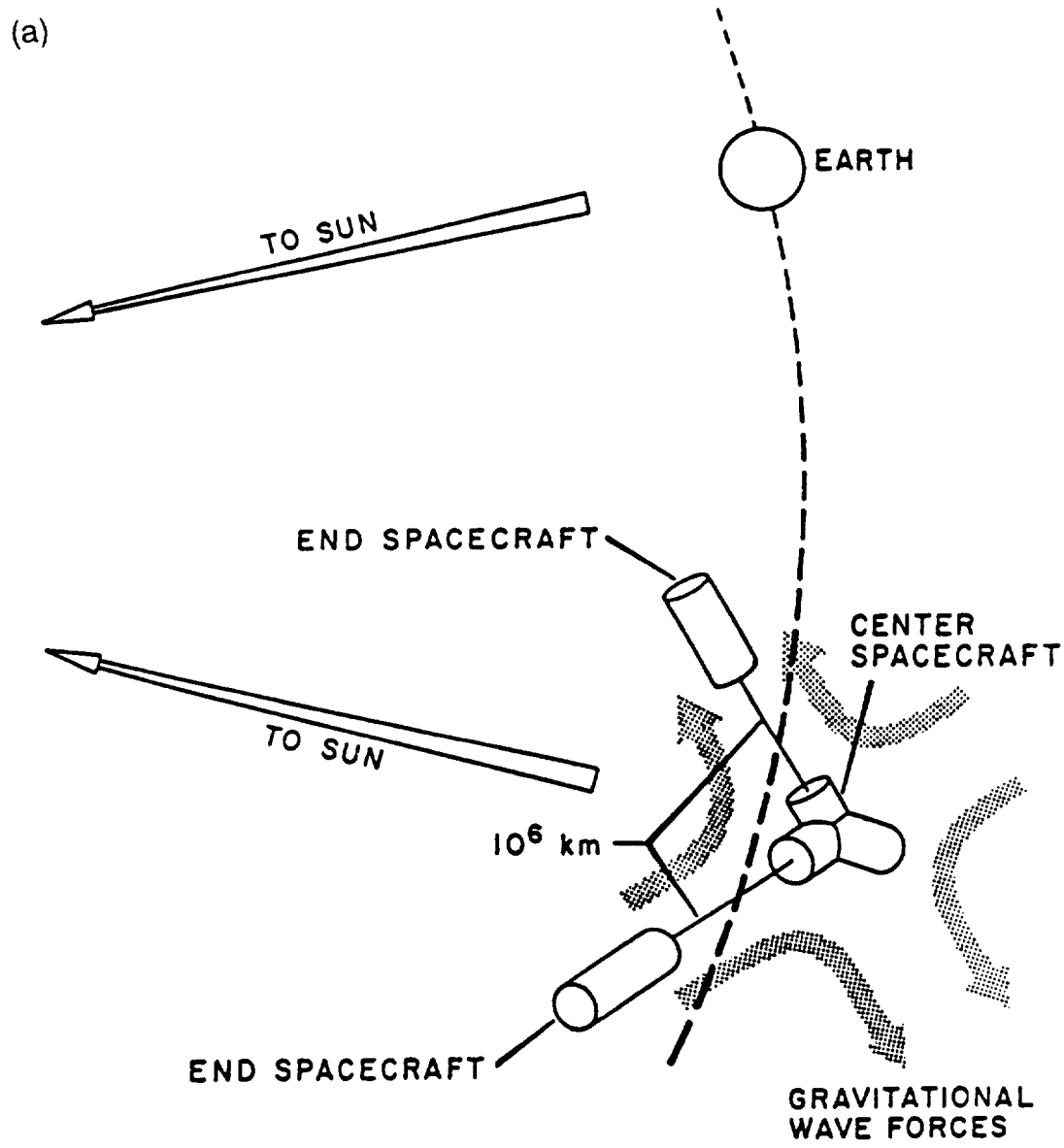


Figure 17. Space-based low-frequency laser-interferometer gravitational-wave detector. Figure 17a depicts a possible laser-interferometer gravitational-wave detector between spacecraft in solar orbit separated from one another by $L \sim 10^7$ km. Figure 17b shows the noise spectral density, in terms of strain sensitivity $h \sim \delta\ell/L$, for such an interferometer. (This instrument sensitivity also appears superposed on Figures 16b and 16d.)

(b)

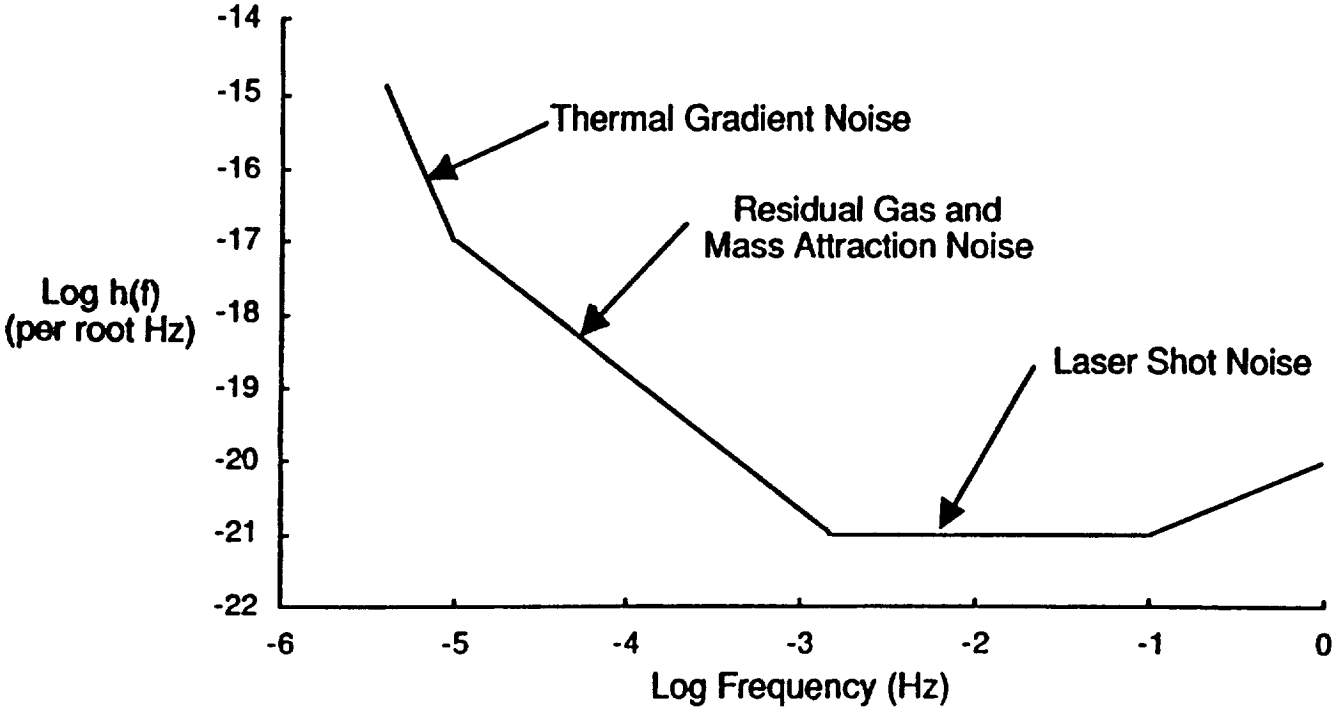

-14
-15
-16
-17
-18
-19
-20
-21
-22
Log h(f)
(per root Hz)
Thermal Gradient Noise
Residual Gas and
Mass Attraction Noise
Laser Shot Noise
-6
-5
-4
-3
-2
-1
0
Log Frequency (Hz)

Figure 17b.

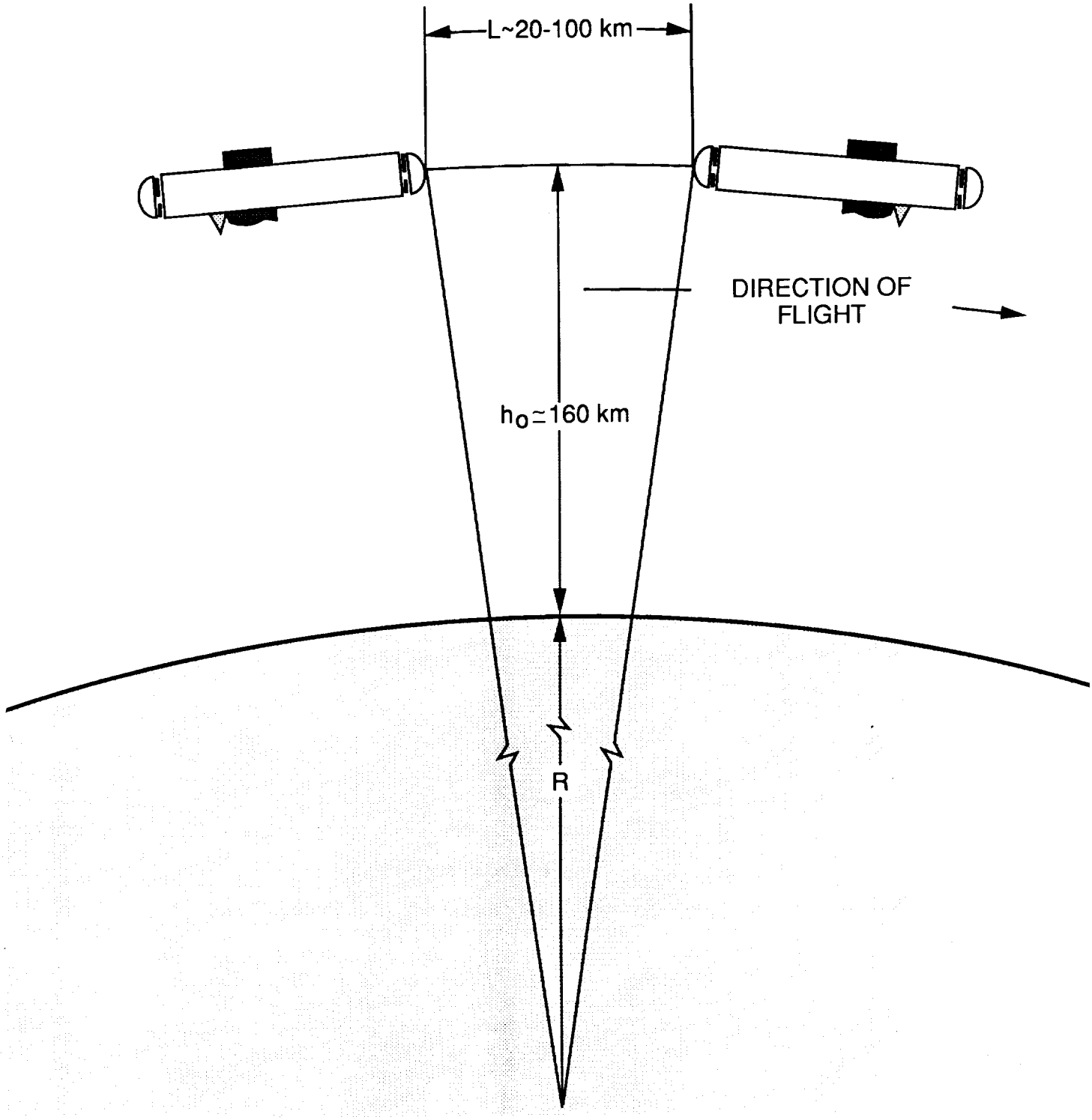


Figure 18. Possible single-arm laser interferometer between coorbiting spacecraft for gravity-field mapping (after Bender 1980 and Keating, *et al.* 1986). As discussed in text (see section 5.3), performance could be improved considerably by adding a third spacecraft in the same orbit to form a collinear dual-arm interferometer.

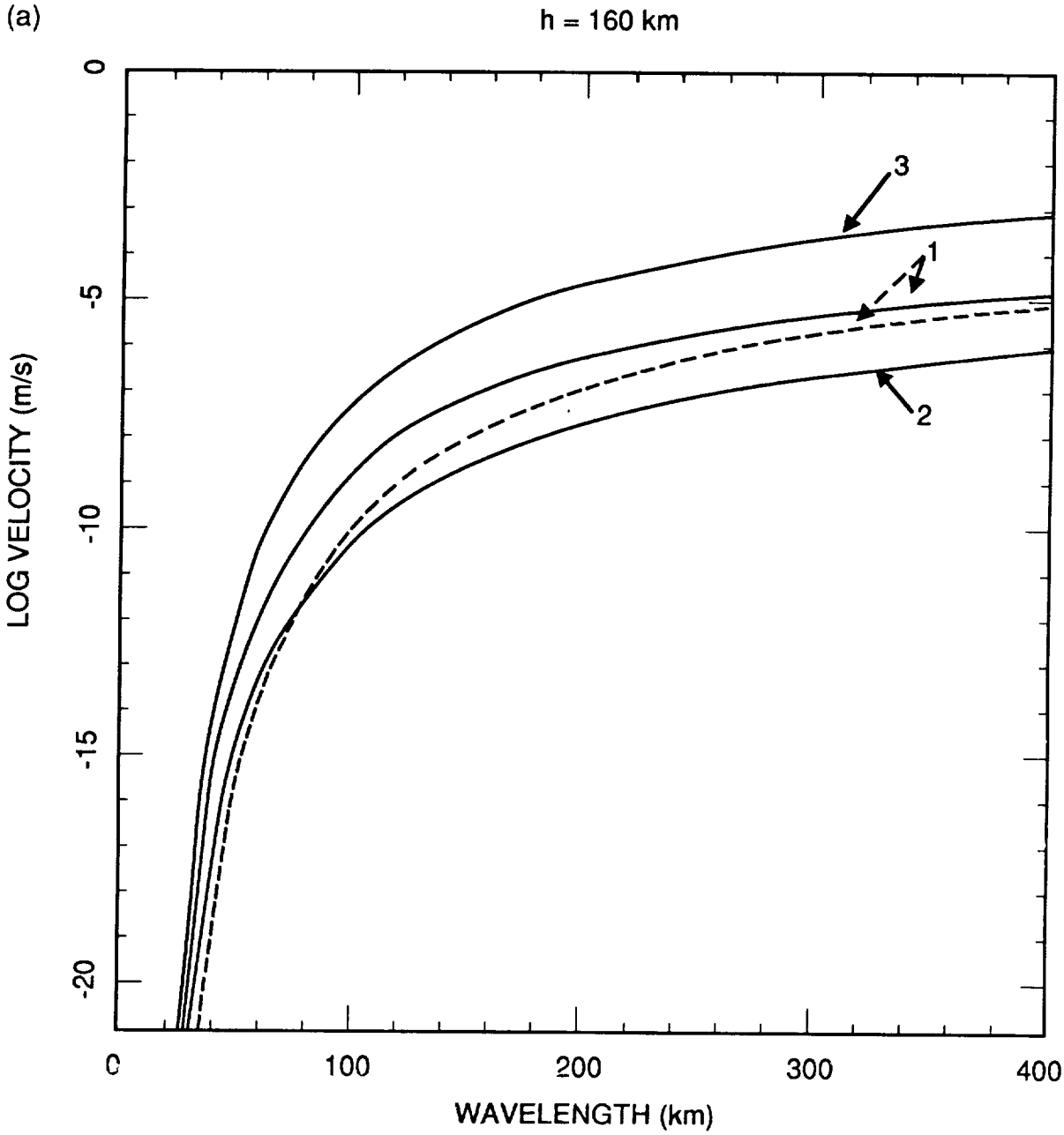


Figure 19. Planet gravity-field mapping with Doppler measurements between co-orbiting spacecraft. Figure 19a is for a 160-km orbit altitude, and Figure 19b is for a 200-km altitude. For comparison, the "1" curve for 200-km orbit altitude is also shown in Figure 19a as a dashed curve. The curves labeled "1" show the maximum change in range rate between co-orbiting spacecraft induced by passage over a 1-mgal anomaly in a planet's gravity field, as a function of spatial wavelength $\lambda_l \equiv 2\pi R/l$. The curves labeled "2" and "3" show the required velocity-measurement accuracy (or maximum tolerable measurement error) for detection of a 1-mgal anomaly with a single measurement or with a combination of 10^6 measurements, respectively. Mapping resolution is defined as $\lambda_l/2$. The integer l is the highest order of the terms in a spherical-harmonic expansion of the gravity field; knowledge of the field to order l corresponds to knowledge of the (spatial) Fourier spectrum down to wavelengths λ_l. See discussion in section 5.3. For wavelengths shorter than about 100 km, or resolutions smaller than about 50 km, the maximum change in range rate and maximum tolerable velocity-measurement error decrease sharply, but their magnitudes are quite sensitive to orbit altitude. For example, for 40-km resolution, the required accuracy with 10^6 1-second measurements (curves "3") is approximately 3 nm/s for 160-km orbit altitude, a factor of $e^{\pi} \simeq 25$ less demanding than the requirement at 200-km orbit altitude. For 20-km resolution, the required accuracy is about 4 fm/s at 160-km orbit altitude, which is a factor of $e^{2\pi} \simeq 535$ less demanding than the requirement at 200-km orbit altitude. Other numerical values are given in Table 5.3.2.

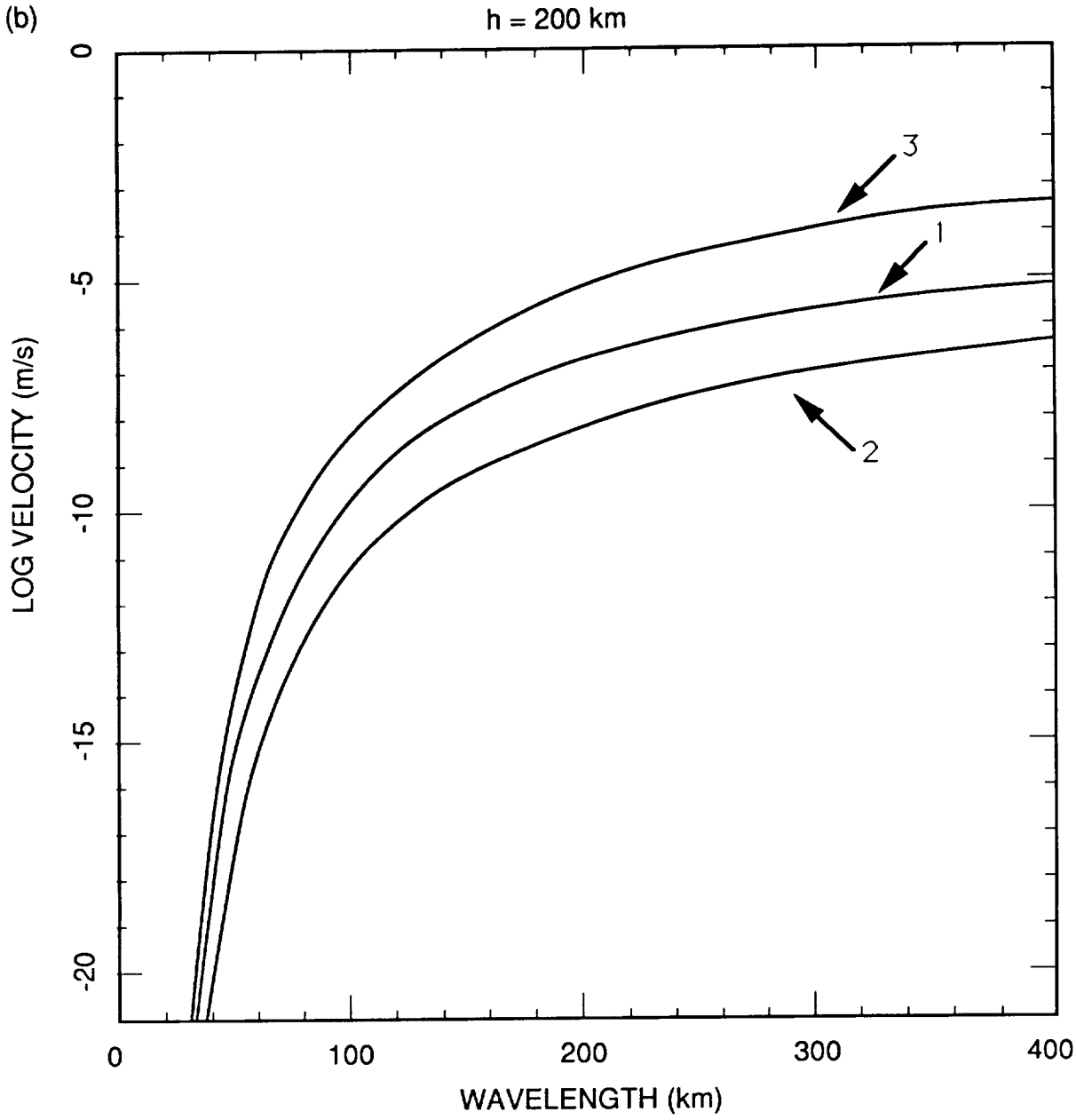

(b)
h = 200 km
LOG VELOCITY (m/s)
0
-5
-10
-15
-20
0
100
200
300
400
WAVELENGTH (km)
1
2
3

Figure 19b.

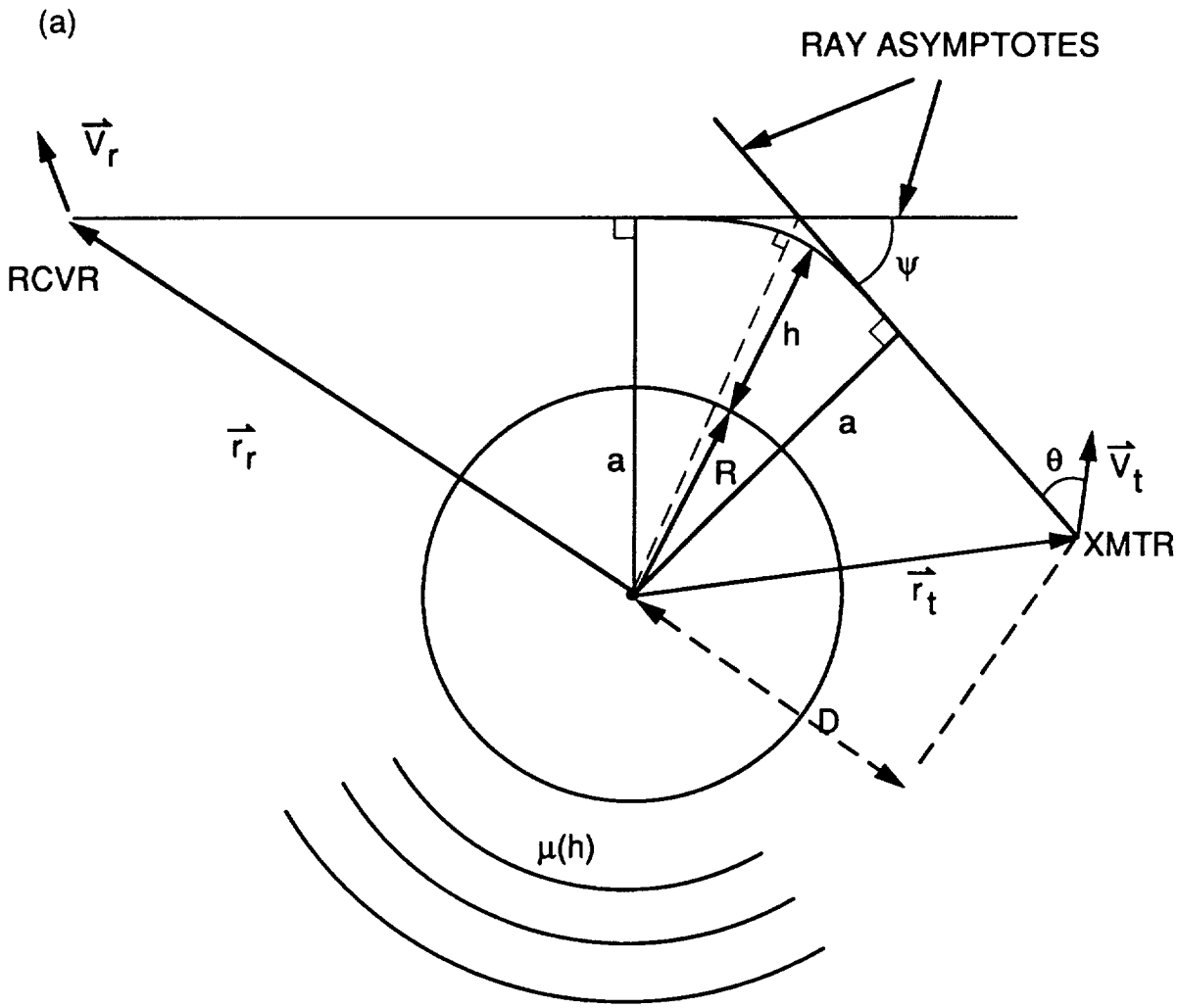


Figure 20. Geometry of (a) atmosphere- and (b) ring-occultation measurements. (a) The light from transmitter (spacecraft) to receiver (Earth) is refracted by an angle ψ. Measurements of the bending angle ψ as a function of ray asymptote "miss distance" a are used to deduce a refractivity profile $\mu(h)$. If spherical symmetry exists, the refractivity $\mu(h)$ is related to a and $R+h$, the closest distance of the refracted ray to the planet center, by Bouger's rule: $\mu = a/(R+h)$. See discussion in section 5.5.1. (b) Signals from a transmitter below the ring plane (x, y-plane) to a receiver above the plane intercept the ring plane at points along the y-axis; for occultation near the equator, the y-axis is a radial line from the planet center. The received signal consists of a direct ray, which has passed through point O (corresponding to the central point of the antenna beamwidth) and which is coherently phase-shifted with respect to the transmitted signal, and a scattered ray, which comes from particles at various points P on the ring plane and which is incoherently (randomly) phase-shifted with respect to the transmitted signal. The receiver is assumed far enough away from the ring system that the rays $\vec{r}_{or}$ and $\vec{r}_{pr}$ from points O and P are essentially parallel. The distance between the transmitter and the ring plane is $D \equiv |\vec{r}_t - \vec{r}_o| \equiv |\vec{r}_{to}|$; the ring opening θ_0 is the angle between the ring plane and the vector $\vec{r}_{to}$. The direct ray is bent by an angle ψ due to differential refraction. Transmitter and particle velocities are $\vec{v}_t$ and $\vec{v}_o \simeq \vec{v}_p$. See discussion in section 5.5.2.

(b)

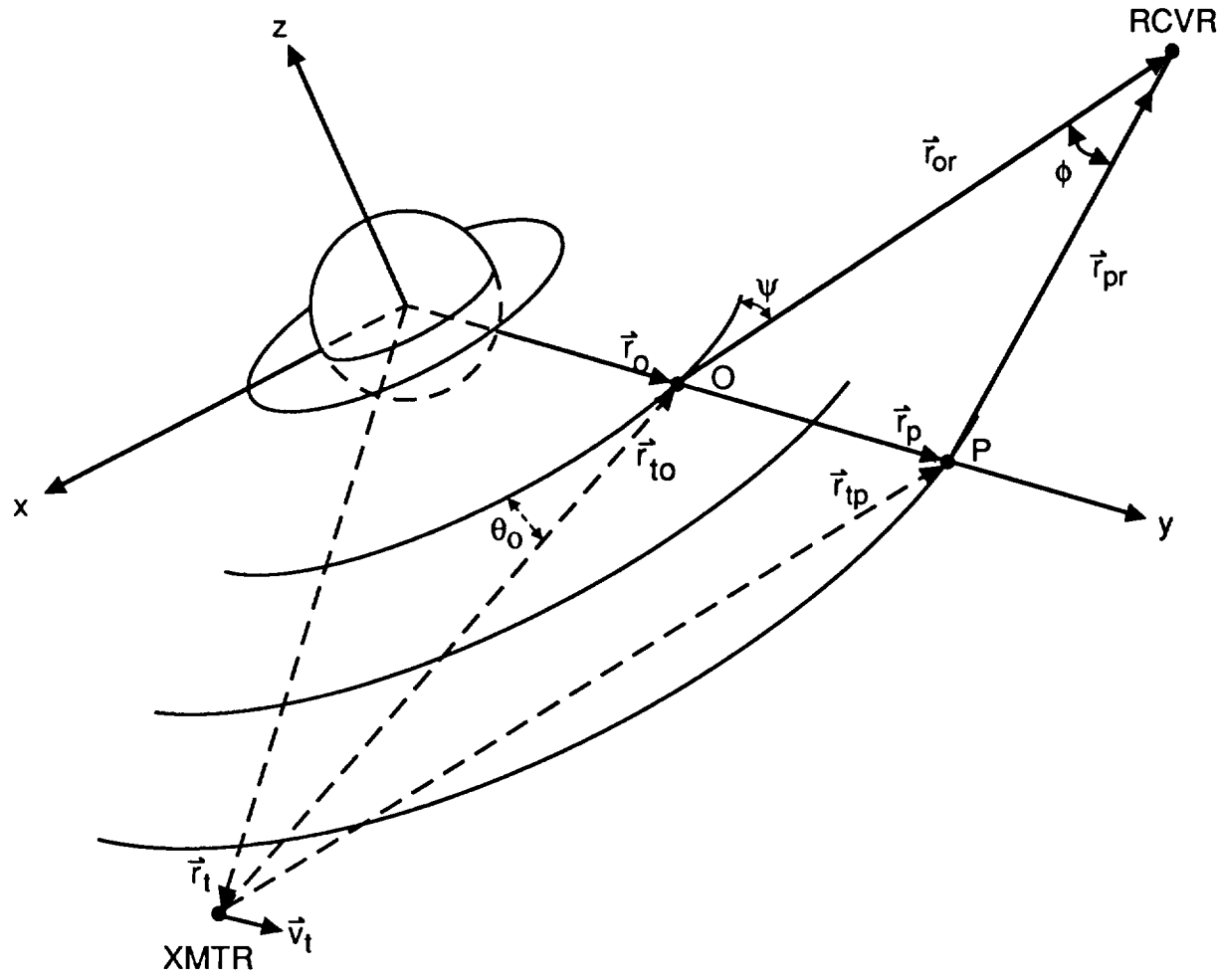

z
RCVR
$\vec{r}_{or}$
ϕ
$\vec{r}_{pr}$
ψ
$\vec{r}_o$
O
$\vec{r}_p$
P
$\vec{r}_{to}$
$\vec{r}_{tp}$
x
y
θ_o
$\vec{r}_t$
$\vec{v}_t$
XMTR

Figure 20b.

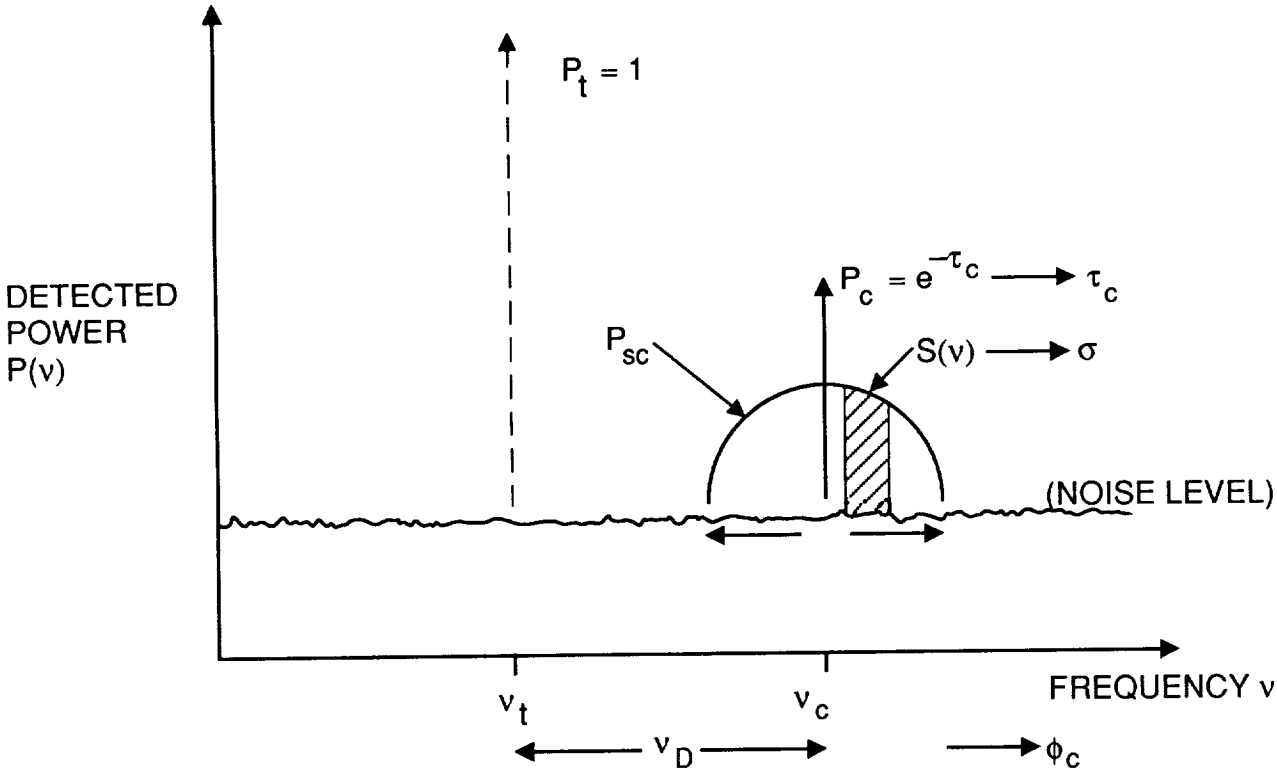


Figure 21. Spectral characteristics of the received signal during an occultation of a spacecraft by planetary rings. See discussion in subsection 5.5.2a. The received signal consists of two components. The first is a coherent, directly transmitted part (power $P_{\rm coh}$, frequency ν_c), attenuated in power by a factor $e^{-\tau_c}$ and possibly slightly shifted in frequency (Doppler shift ν_D) and hence phase (ϕ_c) caused by differential refraction through the rings. The frequency shift is exaggerated in this figure. The second component is a Doppler-broadened incoherent part [power $P_{\rm sc}$, noise spectral density $S(\nu)$] due to transmitted light that has been scattered into the receiver's line of sight by particles moving with various velocities relative to the transmitting spacecraft and the receiver. The near-forward scattering cross section σ is inferred from measurements of the spectral power in the incoherent signal, while the complex extinction $\phi_c + i\tau_c/2$ is inferred from measurements of the coherent signal.

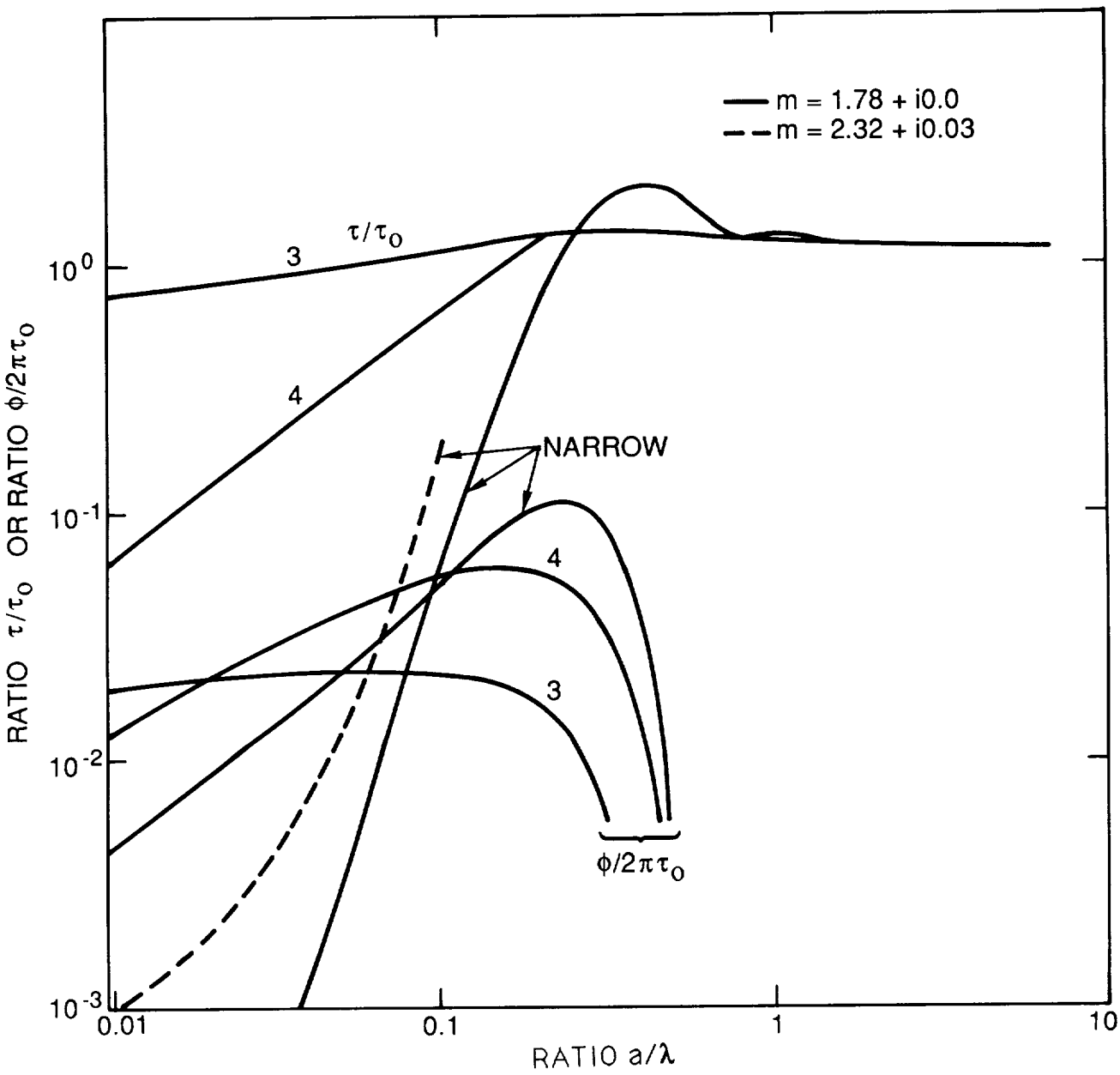


Figure 22. Complex extinction (opacity $\tau \equiv \tau_c \sin\theta_0$ and phase shift $\phi \equiv \phi_c \sin\theta_0$, where θ_o is the ring opening) for the coherent component of the received signal as a function of the ratio of particle size to wavelength, a/λ. See discussion in subsection 5.5.2a, especially eqs. (5.5.10–12). The dimensionless quantity $\tau_0 \equiv 2\pi a^2 n_c$, where n_c is the particle column density. The curves marked "narrow" describe the situation when all particles are approximately the same size. (A spread in sizes of about 5% was used in this figure, adapted from an analysis for S- and X- band wavelengths by Marouf, *et al.* 1982.) The curves marked "3" and "4" describe situations where the distribution of sizes obeys a power law $p(a) \propto a^{-q}$, with $q = 3$ and $q = 4$, respectively for a in some range $(a_{\min}, a_{\max})$ with $a_{\max} \simeq 30\lambda$. For these curves, the abscissa represents the ratio $a_{\min}/\lambda$. The flattening of the opacity curves for broader size distributions (smaller q) indicates a dominant effect from particles larger than the wavelength. The solid curves assume lossless water-ice particles, while the dashed curve is for (lossy) silicates. The difference in behavior between lossy particles and lossless water-ice is most pronounced for particles much smaller than the wavelength (deep into the Rayleigh scattering limit). In the geometric-optics limit, where $a >> \lambda/2\pi$, the phase shift vanishes, and the opacity $\tau \simeq \tau_0$.

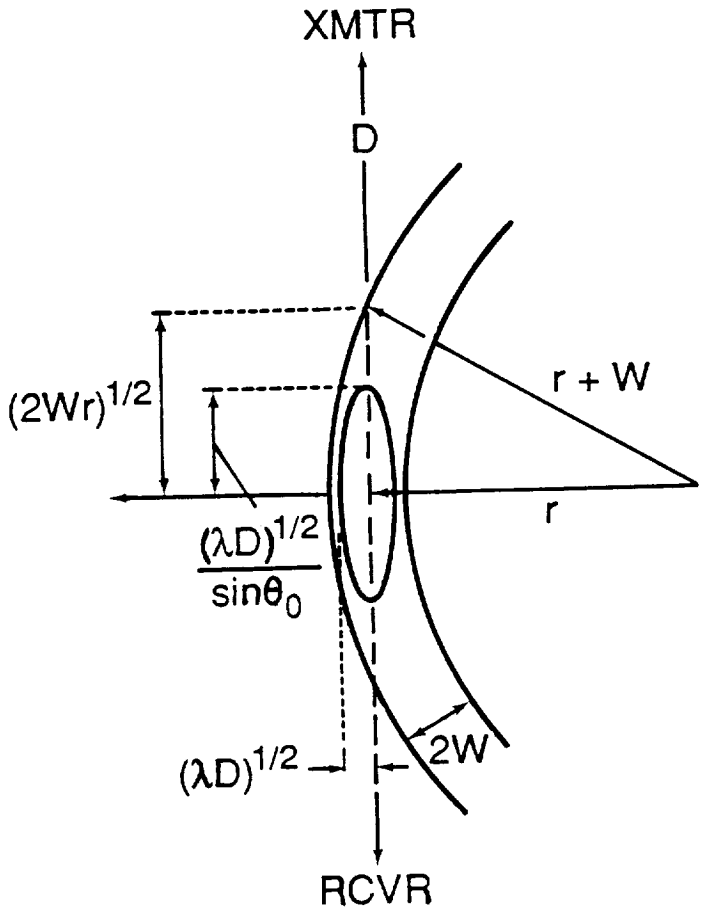


Figure 23. Geometry of the first Fresnel-zone ellipse relative to a ringlet of width $2W$, not to scale (adapted from Marouf, *et al.* 1982). See discussion in subsection 5.5.2a. The transmitting spacecraft is located out of the page, and the receiver is directly below the spacecraft (into the page). This is the optimum geometry for achieving high resolution, with the semimajor axis of the ellipse aligned with the rings, and the spacecraft motion and semiminor axis along the radial direction. For ring widths W greater than the width $(\lambda D)^{1/2}$ of the ellipse (D is the distance of the spacecraft behind the planet center of mass) ray-bending effects are neglible, *i.e.*, there is little or no focusing or defocusing of the ray. At Saturn, with $D \simeq 4$ Saturn radii $\simeq 240{,}000$ km, the resolution at X band (3.6 cm) is about 6 km, and at optical wavelengths (1 μm) it could be on the order of 30 m. Individual contributions from isolated ringlets can also be isolated in the Doppler data from the incoherent signal, although typically with poorer resolution (see Figure 24 and discussion in subsection 5.5.2a).

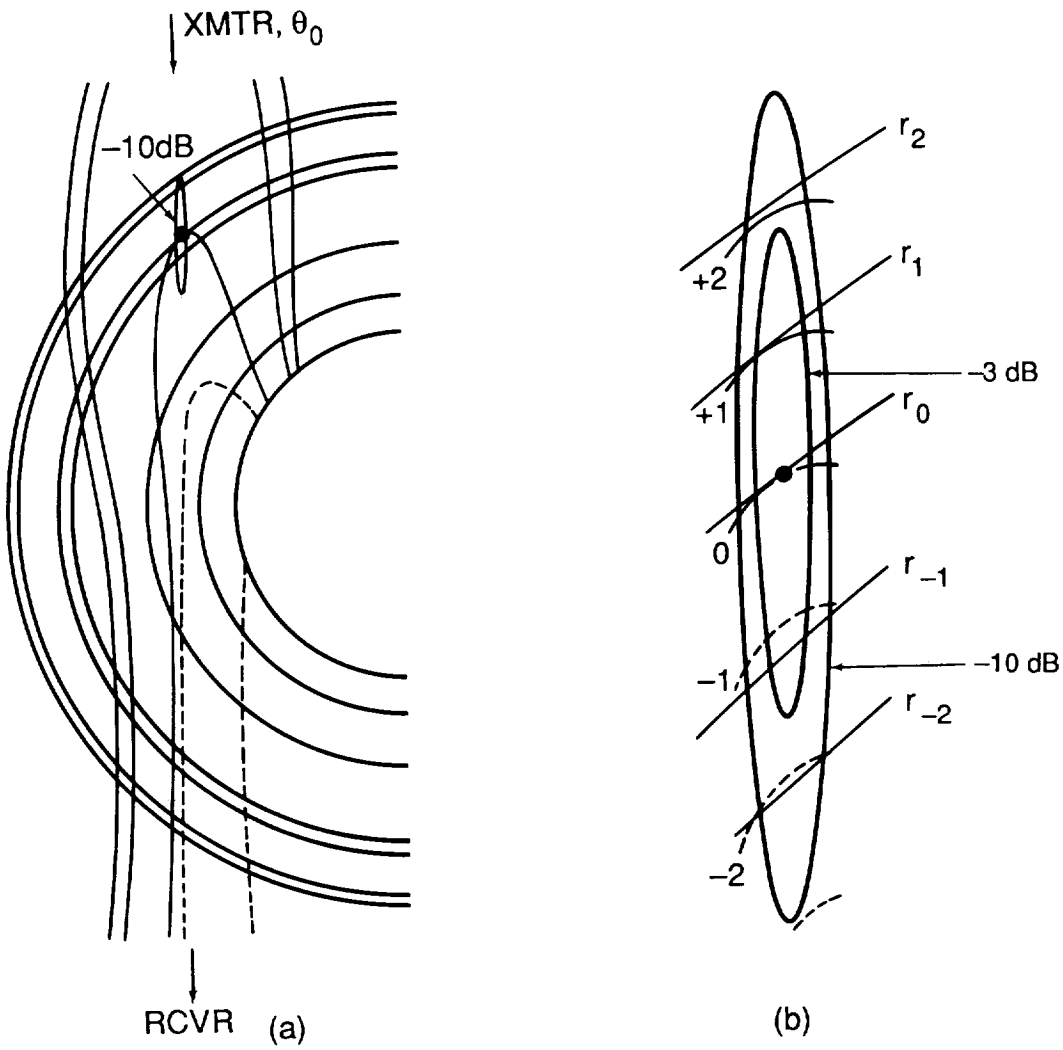


Figure 24. Doppler map showing contours of constant Doppler shift for a ring-occultation experiment (adapted from figures in Marouf, *et al.* 1982 and Tyler 1987 describing Voyager 1 at Saturn). (a) One of a series of Doppler contour maps created as a spacecraft moves radially behind the rings of a planet, transmitting to a receiver in front of the rings. (b) Close-up of the illuminated portion, showing optimum alignment of Doppler contours with arcs of constant radial distance. This alignment permits straightforward reconstruction of the scattering cross section from the spectra of the received signal, as well as identification of contributions from individual ringlets. The ellipses correspond to the −3 and −10 dB antenna-gain contours.

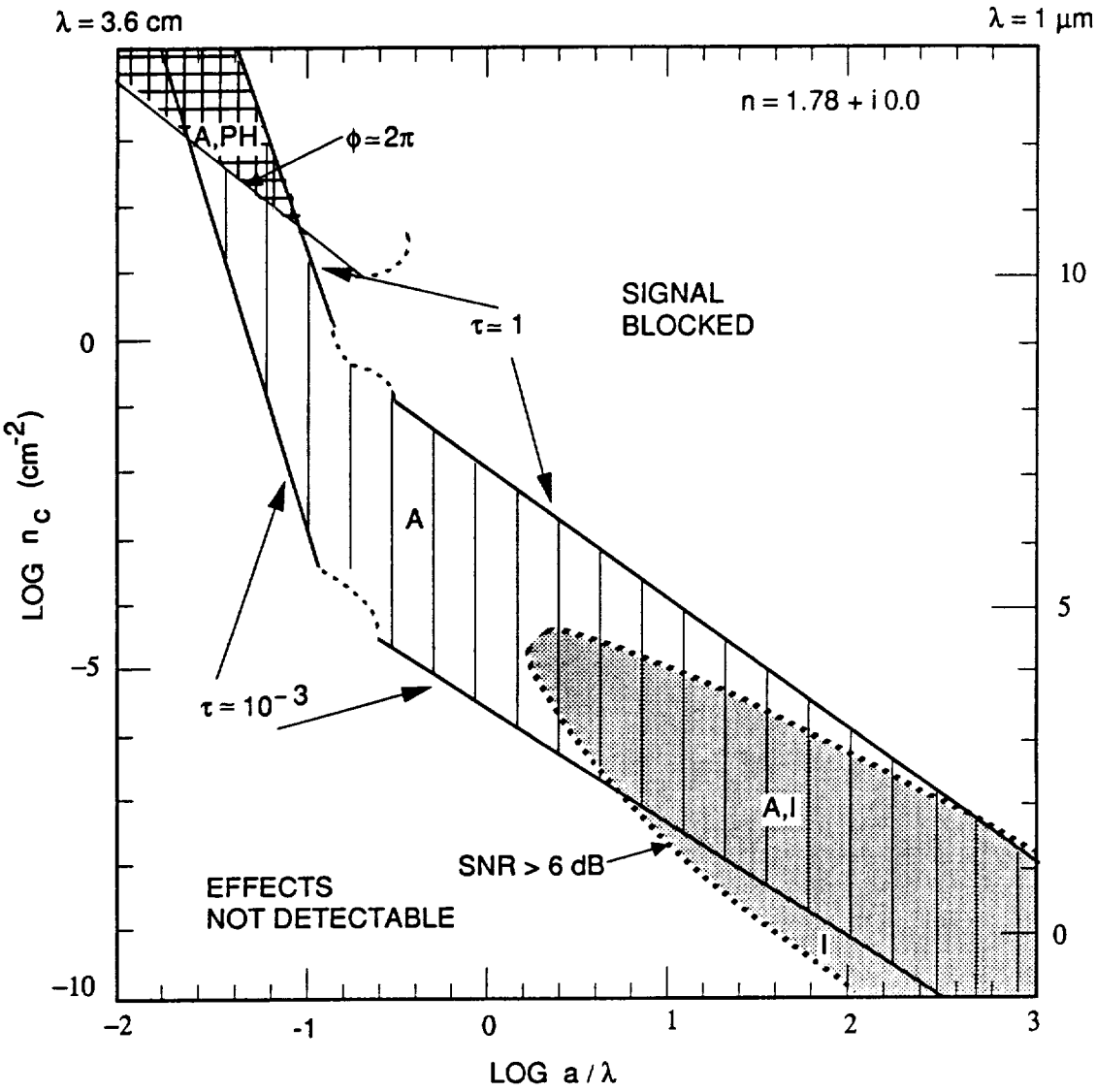


Figure 25. Bounds on detectable particle column densities and particle sizes as a function of wavelength, for lossless (water-ice) particles with a very narrow distribution in size. See discussion following eqs. (5.5.14–16) in subsection5.5.2a. Each curve represents a threshold for measurement of the phase shift ϕ or opacity τ from the coherent signal or of the forward-scattering cross section from the incoherent signal. In the region above the curve labelled $\phi = 2\pi$, the phase change is at least one cycle; this curve does not extend beyond $a/\lambda \sim 0.1$ because there is no phase shift in the geometric-optics limit. The other solid-line curves bound the region in which the opacity is small enough to ensure adequate received power but large enough to ensure that the extinction caused by scattering through the rings can be distinguished from intensity fluctuations caused by pointing jitter; here these bounds are taken to be $0.001 < \tau < 1$. The opacity curves have different slopes in the Rayleigh-scattering and the geometric-optics regimes [eqs. (5.5.10 and (5.5.11)]. The dashed portions of each of these curves are approximations to the true curves for this intermediate regime between Rayleigh scattering and geometric optics, which have been calculated for X-band wavelengths by Marouf, *et al.* (1982). Numerical values for the column density corresponding to X-band wavelengths (3.6 cm) and optical wavelengths (1 μm) are on the left and right sides of the graph, respectively. The lined region between the two opacity curves, labeled "A," is the region in which scattering-induced amplitude changes in the coherent signal are detectable. In the double-hatched region labeled "A, PH," both amplitude and phase information are accessible from the coherent signal. The shaded region labeled "A, I" is the regime in which the incoherent signal is detectable, based on the requirement that the SNR in a 10-kHz bandwidth exceed 6 dB for integration times of 100 sec; this requirement puts a lower bound on the measurable forward-scattering cross section $\sigma(0)$, hence lower and upper bounds on measurable opacity as a function of particle size. In the region labeled "I," the opacity is so small that only the incoherent signal is useful.

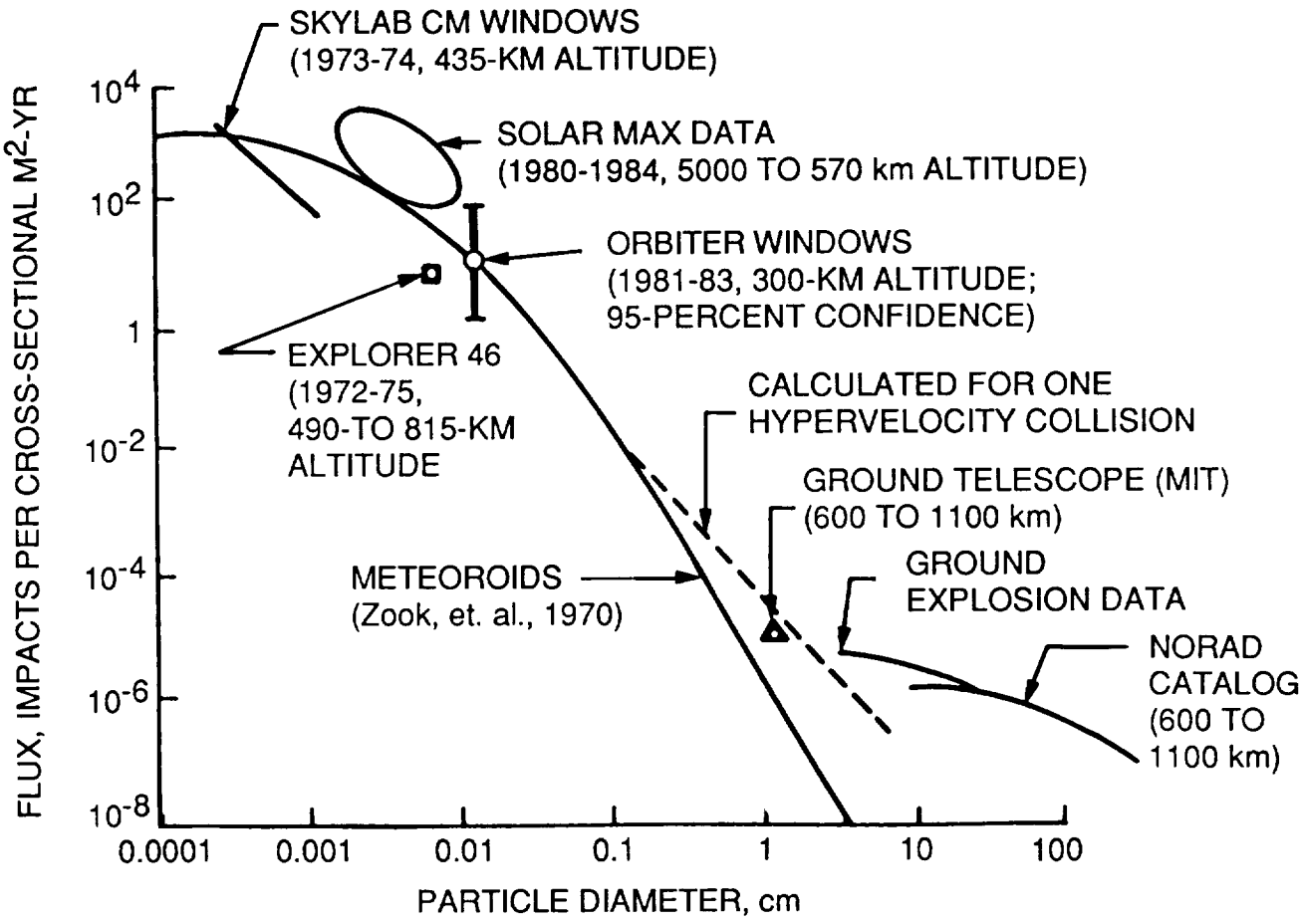


Figure 26. The observed flux in Earth orbit of natural objects, primarily meteoroids (solid curve), compared with that of man-made debris, as measured by several different techniques. This graph is taken from an article by Kessler, in Horz (1985). The flux of man-made debris is observed to exceed that of meteoroids at both large and small masses (*e.g.*, the SOLAR MAX data). (See discussion in section 6.2.5.) Few observations exist for the millimeter-size range; however, extrapolations from collisional and explosive fragmentation events indicate an abundance of these smaller particles.

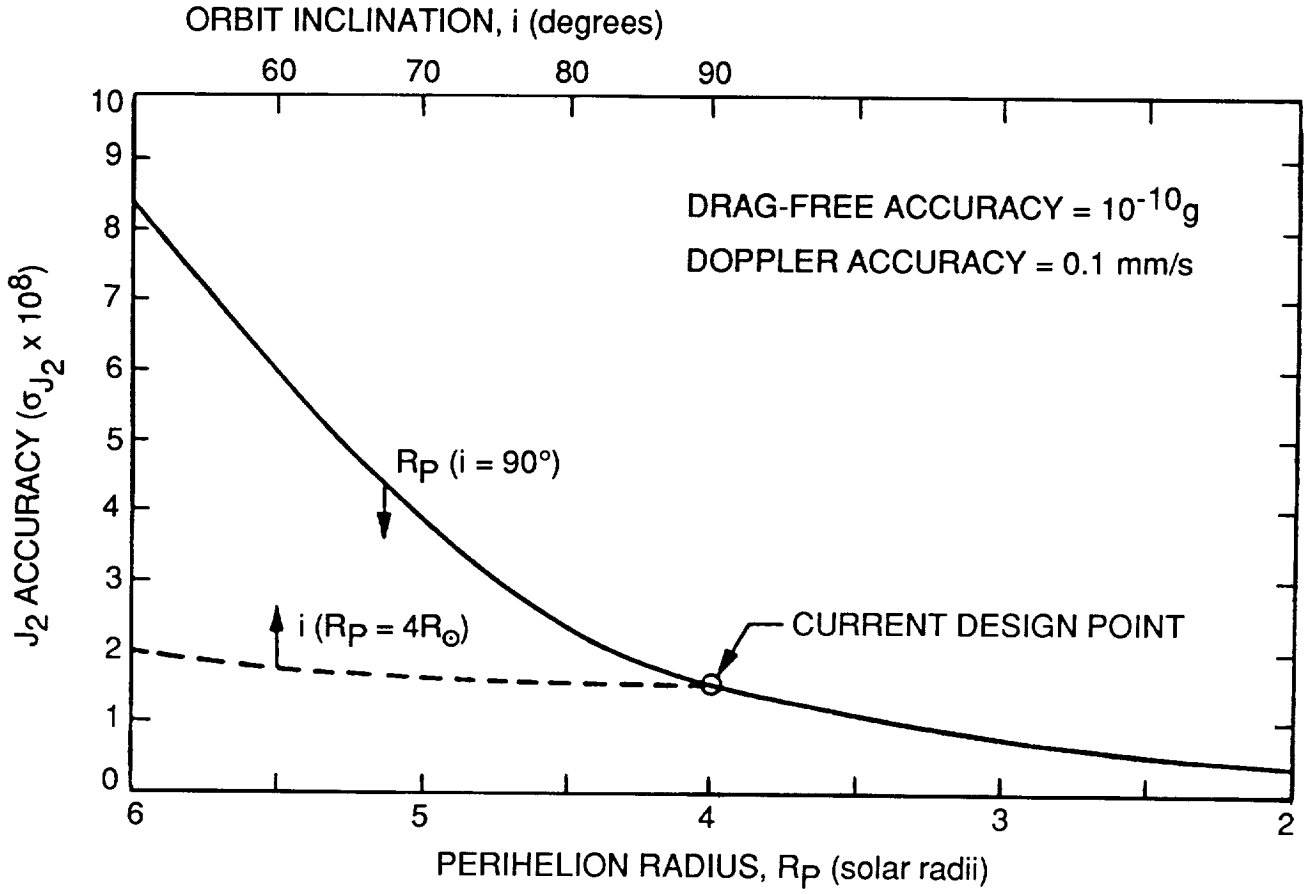


Figure 27. Estimated accuracy of a determination of the solar quadrupole moment J_2 from Doppler tracking of a spacecraft located approximately four solar radii ($R_\odot$), or 1° (viewed from Earth), from the Sun (from Anderson 1988). See discussion in §6.2.6. The solid curve shows the loss of accuracy for a perihelion distance R_p larger than about 4 $R_\odot$, for an orbit that is perpendicular to the ecliptic. The dashed curve indicates little sensitivity to orbit inclination for a given perihelion distance ($4R_\odot$, in this case), which is consistent with the assumption that all multipole moments of the Sun's gravity field except J_2 are very small. The assumed Doppler accuracy of 0.1 mm/s refers to the one-sigma accuracy of range-rate measurements taken at 60-sec intervals, assuming white noise during the solar encounter. The goal of 1–2 $\times 10^{-8}$ accuracy in J_2 can be met only for a perihelion radius of order $4R_\odot$ or smaller and an inclination greater than about 70°.

Light-deflection tests of general relativity (GR)

Max. contribution to θ	Sun	Jupiter
$4\frac{M}{R}\left(\frac{1+\gamma}{2}\right)$	$1.75''$	20 mas
$\pm\frac{4J}{R^2}\left(\frac{7\Delta_1+\Delta_2}{8}\right)$	± 0.6 μas	± 0.2 μas
$4J_2\frac{M}{R}\left(\frac{1+\gamma}{2}\right)$	$\sim$1–10 μas	300 μas
$4\pi\left(\frac{M}{R}\right)^2\left[\frac{1+\gamma}{2}+\frac{1}{4}\left(\frac{3\Lambda}{4}-\beta\right)\right]$	10 μas	10^{-3} μas

- γ: Linear contribution to space curvature
- Λ, β: 2nd-order contributions to space and time curvatures
- Δ_1, Δ_2: "Frame-dragging" effect due to angular momentum J
- GR: $\gamma = \Delta_1 = \Delta_2 = \Lambda = \beta = 1$

Figure 28. Angular measurement accuracy required for light-deflection tests of general relativity, using the Sun or Jupiter. Microarcsecond-accuracy astrometric measurements are required to measure effects that are of second order in the solar potential M/R (dimensionless, *i.e.*, defined as GM/c^2R, with $G \equiv 1$ here). Note that $M/R \simeq 2 \times 10^{-6}$ for the Sun, roughly 100 times larger than for Jupiter.

8. APPENDIX. ESTIMATION THEORY FOR GRAVITY-FIELD MAPPING

Here, some simple aspects of estimation theory are described that were used to derive velocity errors given in section 5.3.1. The mean-square uncertainty σ_q in a quantity q is related to the noise spectral density S_q by

$$\sigma_q{}^2 = \int \frac{d^2k}{(2\pi)^2} S_q(\vec{k}) , \tag{8.1}$$

where $d^2k \equiv dk_x dk_y$ denotes the two-dimensional Fourier surface element. In this discussion, q will represent perturbations in the local average gravitational acceleration, δg, or in the local geoid height, δn_g. The Fourier components of q are linearly related to the Fourier components of the perturbations in the potential (dependences on k are suppressed):

$$\tilde{q} \equiv \tilde{f}_q \delta\tilde{U}_0 ; \qquad \tilde{f}_g = k, \quad \tilde{f}_{n_g} = \frac{1}{g_0} . \tag{8.2a}$$

Measured changes in the relative velocity v arise both from perturbations in the potential, which have the signal of interest, $\delta\tilde{v}_s$, and from measurement error, $\delta\tilde{v}_m$ (for simplicity here, all other error sources are either ignored or included in $\delta\tilde{v}_m$ as white noise, following Breakwell 1979) :

$$\delta\tilde{v} = \delta\tilde{v}_s + \delta\tilde{v}_m \equiv \tilde{H}\delta\tilde{U}_0 + \delta\tilde{v}_m . \tag{8.2b}$$

The transfer function $\tilde{H}(k)$, given by eq. (8.2b), relates signal-induced velocity perturbations δv_s to perturbations $\delta\tilde{U}_0$ in the surface potential.

Denote an estimate of q by $\hat{q}$, assumed to be linearly related to changes in the relative velocity of the spacecraft by some measurement transfer function $\tilde{\psi}(k)$:

$$\hat{q} \equiv \tilde{\psi}\, \delta\tilde{v} . \tag{8.3a}$$

The estimation error is

$$\Delta\hat{q} \equiv \hat{q} - \tilde{q} = (\tilde{\psi}_q \tilde{H} - \tilde{f}_q)\delta\tilde{U}_0 + \tilde{\psi}_q \, \delta\tilde{v}_m \, . \tag{8.3b}$$

The error spectral density for q is related to the spectral density of the surface potential fluctuations, S_{U_0}, and the measurement error spectral density $S_{v_m} \equiv \langle|\delta\tilde{v}_m|^2\rangle$ by

$$S_q \equiv \langle|\Delta\hat{q}|^2\rangle = |\tilde{\psi}_q \tilde{H} - \tilde{f}_q|^2 \, S_{U_0} + |\tilde{\psi}_q|^2 \, S_{v_m} \tag{8.4}$$

(angle brackets denote an ensemble average). The function $\tilde{\psi}_q$ that minimizes the error spectral density S_q is the Fourier transform of the statistically best weighting of the measurements for estimating q. With that optimal measurement transfer function, the error spectral density S_q takes the form

$$S_q = \frac{S_{q_0}}{1 + S_{v_s}/S_{v_m}} \, . \tag{8.5a}$$

Here

$$S_{q_0} \equiv \langle|\delta\tilde{q}|^2\rangle = |\tilde{f}_q|^2 \, S_{U_0} \tag{8.5b}$$

is the spectral density corresponding to the *a priori* (without velocity measurements) knowledge of the actual fluctuations in q;

$$S_{v_s} \equiv \langle|\delta\tilde{v}_s|^2\rangle = |\tilde{H}|^2 \, S_{U_0} \tag{8.5c}$$

is the spectral density of signal fluctuations in v due to perturbations in the potential, and

$$S_{v_m} \equiv {\sigma_v}^2 \, \frac{2\pi^2 R^2}{N_m} \tag{8.5d}$$

is the measurement error spectral density. Here, the (white noise) measurement error spectral density is taken to be the mean-square uncertainty (variance) of a single measurement of velocity, ${\sigma_v}^2$, divided by the number of measurements per unit area, where N_m is the total number of measurements and $2\pi^2 R^2$ is the area of a Mercator enlargement of the Earth over which the measurements are

essentially independent of latitude. The error spectral density in an estimate of q, based on N_m measurements of the relative velocity between spacecraft, is therefore*

$$S_q = \frac{{\sigma_v}^2 \langle |\delta\tilde{q}|^2 \rangle}{{\sigma_v}^2 + N_m \, (2\pi^2 R^2)^{-1} \, \langle |\delta\tilde{v}_s|^2 \rangle} \, . \tag{8.6}$$

In the limit that the second term in the denominator of eq. (8.6) dominates the first, the error spectral density takes the form

$${\sigma_q}^2 \simeq \frac{{\sigma_v}^2}{N_m} \, 2\pi^2 R^2 \int \frac{d^2k}{(2\pi)^2} \, \frac{\langle |\delta\tilde{q}|^2 \rangle}{\langle |\delta\tilde{v}_s|^2 \rangle} \, . \tag{8.7a}$$

A minimum for this quantity can be obtained by taking the maximum value for the spectral density of the signal fluctuations $S_{v_s}(\max) \equiv \langle |\delta\tilde{v}_{s_{max}}| \rangle$ [eqs. (5.3.7)]. This approximation also removes the angular dependence of the integral in eq. (8.7a). This defines a quantity $\tilde{\sigma}_q^2$, equal to the Fourier component at harmonic order l of the mean-square uncertainty ${\sigma_q}^2$ in an estimate of q, based on measurements of v whose precision is σ_v:

$${\sigma_q}^2 \gtrsim \frac{{\sigma_v}^2}{N_m} \, \pi R^2 \int dk \; k \, \frac{\langle |\delta\tilde{q}|^2 \rangle}{\langle |\delta\tilde{v}_{s,\max}|^2 \rangle} \equiv R \int dk \; \tilde{\sigma}_q^2 \, ; \tag{8.7b}$$

$$\tilde{\sigma}_q^2 \equiv \frac{{\sigma_v}^2}{N_m} \, \pi \, l \, \frac{\langle |\delta\tilde{q}|^2 \rangle}{\langle |\delta\tilde{v}_{s,\max}|^2 \rangle} \, . \tag{8.7c}$$

This result is used to give the rms velocity errors for gravity anomalies and geoid undulations in section 5.3.3 [eqs. (5.3.8) and (5.3.9)].

* A note on dimensions: the velocity measurement precision σ_v has units of velocity, but the Fourier quantities are defined in units of wave number; hence the denominator of eq. (8.6) is dimensionally consistent.

9. REFERENCES

V. J. Abreu, "Wind measurements from an orbiting platform using a lidar system with incoherent detection: an analysis," *Applied Optics,* vol. 18, pp. 2992–2997, 1979.

W. M. Alexander and J. D. Corbin, "Interaction of lunar ejecta and the magnetosphere of the earth," in *Solid Particles in the Solar System,* eds. I. Halliday and B. A. McIntosh, Dordrecht: Reidel, pp. 425–428, 1980.

H. Alfvén and G. Arrhenius, *Evolution of the Solar System,* NASA SP-354, 1976.

D. W. Allan, "Statistics of atomic frequency standards," *Proceedings of the IEEE,* vol. 54, pp. 221–230, 1966.

C. W. Allen, *Astrophysical Quantities,* London: Athlone Press, 1973.

C. O. Alley, "Laser ranging," in *Quantum Optics, Experimental Gravitation, and Measurement Theory,* eds. P. Meystre and M. O. Scully, New York: Plenum Press, pp. 459–475, 1983

J. D. Anderson, private communication, 1989.

J. D. Anderson, "Gravitational Experiments on a Solar Probe mission: scientific objectives and technology considerations," in *Relativistic Gravitational Experiments in Space,* NASA Conference Publication 3046, pp. 148–154, NASA, Washington, D.C., 1988.

J. D. Anderson, P. B. Esposito, W. Martin, C. L. Thornton, and D. O. Muhleman, "Experimental test of general relativity using time-delay data from Mariner 6 and Mariner 7," *Astrophysical Journal,* vol. 200, pp. 221–233, 1975.

J. D. Anderson, W. B. Hubbard, and W. L. Slattery, "Structure of the Jovian envelope from Pioneer 10 gravity data," *Astrophysical Journal,* vol. 193, pp. 149–150, 1974a.

J. D. Anderson, G. W. Null, and S. K. Wong, "Gravity results from Pioneer 10 Doppler data," *Journal of Geophysical Research,* vol. 79, pp. 3661–3664, 1974b.

N. Ashby, P. L. Bender, and J. M. Wahr, "Accuracy of gravitational physics tests and of the solar quadrupole moment from ranging to a Mercury orbiter," *Astrophysical Journal,* to be submitted, 1991.

S. Auer, "Two high resolution velocity vector analyzers for cosmic dust particles," *Rev. Sci. Inst.,* vol. 46, pp. 127–135, 1975.

R. L. Barger, J. S. Sorem, and J. L. Hall, "Frequency stabilization of a CW dye laser," *Applied Phys. Lett.,* vol. 22, pp. 573–575, 1973.

J. A. Barnes, *et al.,* "Characterization of frequency stability," *IEEE Transactions on Instrumentation and Measurement,* vol. IM-20, pp. 105–120, 1971.

P. L. Bender, private communication, 1989.

P. L. Bender, "Distance measurements in space: Gravitational physics tests and a proposed laser gravitational wave antenna," in *Proc. Eleventh Int. Conf. on Atomic Physics,* eds. S. Haroche, J. C. Gay, and G. Grynberg, Singapore: World Scientific Pub. Co., 1988.

P. L. Bender, "Laser interferometer method for mapping ultra-short-wavelength structure in the Earth's gravitational field," proposal to NASA, 1980.

P. L. Bender, *et al.*, "Optical interferometer in space," in *Relativistic Gravitation Experiments in Space*, ed. R. W. Hellings, NASA CP-3046, pp. 80–88, 1988.

P.L. Bender, N. Ashby, M. A. Vincent, and J. M. Wahr, "Conceptual design for a Mercury relativity satellite," *Advances in space research*, vol. 9, pp. 113–116, 1989.

H. Billing, *et al.*, "The Munich gravitational wave detector using laser interferometry, Part II: scattered light," in *Quantum Optics, Experimental Gravity, and Measurement Theory*, eds. P. Meystre and M. O. Scully, New York: Plenum, pp. 533–543, 1983.

G. Björk and Y. Yamamoto, "Generation and amplification of number states by nondegenerate parametric oscillators with idler-measurement feedback," *Phys. Rev. A*, vol. 37, pp. 125–147, 1988.

R. S. Bondurant, "Reduction of radiation-pressure-induced fluctuations in interferometric gravity-wave detectors," *Phys. Rev. A*, vol. 34, pp. 3927–3931, 1986.

R. S. Bondurant and J. H. Shapiro, "Squeezed states in phase-sensing interferometers," *Phys. Rev. D*, vol. 30, pp. 2548–2556, 1984.

V. B. Braginsky, Y. I. Vorontsov, and K. S. Thorne, "Quantum nondemolition measurements, *Science*, vol. 209, pp. 547-557, 1980.

C. Brans and R. H. Dicke, "Mach's principle and a relativistic theory of gravitation," *Phys Rev.*, vol. 124, pp. 925–935, 1961.

J. V. Breakwell, "Satellite determination of short wavelength gravity variations," *Journal of the Astronautical Sciences*, vol. 27, pp. 329–344, 1979.

A. Buffington and M. R. Geller, "A photoelectric astrometric telescope using a Ronchi ruling," *Publications of the Astronomical Society of the Pacific*, vol. 102, pp. 200-211, 1990.

S. P. Bush, A. Gungor, and C. C. Davis, "Studies of the coherence properties of a diode-pumped Nd:YAG ring laser, *Appl. Phys. Lett.*, vol. 53, pp. 646–647, 1988.

D. M. Butler, *et al.*, *From pattern to process: the strategy of the Earth Observing System, Earth Observing System science steering committee report*, vol. II, National Aeronautics and Space Administration, 1987a.

D. M. Butler, *et al.*, *HIRIS, High-Resolution Imaging Spectrometer: Science opportunities for the 1990s, instrument panel report, Earth Observing System*, vol. IIc, National Aeronautics and Space Administration, 1987b.

D. M. Butler, *et al.*, *LASA, Lidar Atmospheric Sounder and Altimeter, instrument panel report, Earth Observing System*, vol. IId, National Aeronautics and Space Administration, 1987c.

D. M. Butler, *et al.*, *SAR, Synthetic Aperture Radar, instrument panel report, Earth Observing System*, vol. IIf, National Aeronautics and Space Administration, 1987d.

D. M. Butler, *et al.*, *LAWS, Laser Atmospheric Wind Sounder, instrument panel report, Earth Observing System*, vol. IIg, National Aeronautics and Space Administration, 1987e.

D. M. Butler, *et al., Altimetric system, Earth Observing System, instrument panel report,* vol. IIh, National Aeronautics and Space Administration, 1987f.

R. L. Byer, private communication, 1989.

R. L. Byer, "Diode laser-pumped solid-state lasers," *Science,* vol. 239, pp. 742–747, 1988.

M. J. Campbell and J. Ulrichs, "The electrical properties of rocks and their significance for lunar radio observations," *Journal of Geophysical Research,* vol. 74, pp. 5867–5881, 1969.

W. C. Carey and R. M. Walker, "Interplanetary dust: the interstellar connection," in *Trajectory Determinations and Collection of Micrometeoroids on the Space Station,* LPI Technical Report 86-05, ed. F. Horz, pp. 47–48, 1986a.

W. C. Carey and R. M. Walker, "Prospects for an orbital determination and capture cell experiment," in *Trajectory Determinations and Collection of Micrometeoroids on the Space Station,* LPI Technical Report 86-05, ed. F. Horz, pp. 49–51, 1986b.

B. J. Carr, "The ABC of Population III," in *Inner Space/Outer Space,* ed. E. W. Kolb, *et al.,* Chicago: University of Chicago Press, pp. 83–102, 1986.

P. Carruthers and M. M. Nieto, "Phase and angle variables in quantum mechanics," *Rev. Mod. Phys.,* vol. 40, pp. 411–440, 1968.

C. M. Caves, "Quantum-mechanical noise in an interferometer," *Phys. Rev. D.,* vol. 23, pp. 1693–1708, 1981.

C. M. Caves, "Defense of the standard quantum limit for free-mass position," *Phys. Rev. Lett.,* vol. 54, pp. 2465–2468, 1985.

C. M. Caves, "Laser stabilization using squeezed light," in *Squeezed and nonclassical light,* eds. P. Tombesi and E. R. Pike, New York: Plenum, pp. 29–38, 1989.

C. M. Caves, *et al.,* "On the measurement of a weak classical force coupled to a quantum-mechanical oscillator. I. Issues of principle, *Rev. Mod. Phys.,* vol. 52, pp. 341-392, 1980.

S. Chandrasekhar, "Solutions of two problems in the theory of gravitational radiation," *Phys. Rev. Lett.,* vol. 24, pp. 611–615, 1970.

J. Christensen-Dalsgaard and D. O. Gough, "Perturbatiosn in gravitational potential associated with solar oscillations," in *Nonradial and Nonlinear Stellar Pulsation,* eds. H. A. Hill and W. A. Dziembowski, New York: Springer-Verlag, pp. 369–380, 1980.

J. H. Churnside and S. F. Clifford, "Refractive turbulence profiling using stellar scintillation and radar wind profiles," *Applied Optics,* vol. 27, pp. 4884–4890, 1988.

S. F. Clifford and L. Lading, "Monostatic diffraction-limited lidars: the impact of optical refractive turbulence," *Applied Optics,* vol. 22, pp. 1696–1701, 1983.

S. F. Clifford and S. Wandzura, "Monostatic heterodyne lidar performance: the effect of the turbulent atmosphere," *Applied Optics,* vol. 20, pp. 514–516, 1981.

C. Covault, "Mission to planet Earth," *Aviation Week & Space Technology,* pp. 34–49, March 13, 1989.

P. C. Crane and P. J. Napier, "Sensitivity," in *Synthesis Imaging*, eds. R. A. Perley, F. R. Schwab, and A. R. Bridle, Green Bank, WV: National Radio Astronomy Observatory, pp. 87–108, 1986.

J. N. Cuzzi and J. B. Pollack, "Saturn's rings: particle composition and size distribution as constrained by microwave observations. I. Radar observations," *Icarus*, vol. 33, pp. 233–262, 1978.

T. Day, E. K. Gustafson, and R. L. Byer, "Active frequency stabilization of a 1.062-μm Nd:GGG diode-laser-pumped nonplanar ring oscillator to less than 3 Hz of relative linewidth," *Optics Letters*, vol. 15, pp. 221–223, 1990.

N. Deruelle and T. Piran, eds., *Gravitational Radiation*, Amsterdam: North Holland, 1983.

B. C. Douglas, C. C. Goad, and F. F. Morrison, "Determination of the geopotential from satellite-to-satellite tracking data," *Journal of Geophysical Research*, vol. 85, pp. 5471–5480, 1980.

R. J. Doviak and C. T. Jobson, "Dual Doppler radar observations of clear air wind perturbations in the planetary boundary layer," *Journal of Geophysical Research*, vol. 84, pp. 697–702, 1979.

R. W. P. Drever, "Interferometric detectors for gravitational radiation," in *Gravitational Radiation*, eds. N. Deruelle and T. Piran, Amsterdam: North Holland, pp. 321–338, 1983.

R. W. P. Drever, *et al.*, "Gravitational wave detectors using laser interferometers and optical cavities: ideas, principles and prospects," in *Quantum Optics, Experimental Gravitation, and Measurement Theory*, eds. P. Meystre and M. O. Scully, New York: Plenum Press, pp. 503–514, 1983a.

R. W. P. Drever, *et al.*, "Laser phase and frequency stabilization using an optical resonator," *Applied Physics B*, vol. 31, pp. 97–105, 1983b.

D. M. Eardley, "Theoretical models for sources of gravitational waves," in *Gravitational Radiation*, eds. N. Deruelle and T. Piran, Amsterdam: North Holland, pp. 257–296, 1983.

A. Einstein, "Zur Quantentheorie der Strahlung," *Phys. Z.*, vol. 18, p. 121, 1917; English translation, "On the quantum theory of radiation," in *Sources of Quantum Mechanics*, ed. B. L. Van der Waerden, North-Holland, Amsterdam, 1967, paperback reprint, Dover, New York, 1968.

D. S. Elliott, R. Roy, and S. J. Smith, "Extracavity laser band-shape and bandwidth modification," *Phys. Rev. A*, vol. 26, pp. 12–26, 1982.

V. R. Eshleman, "The radio occultation method for the study of planetary atmospheres," *Planetary and Space Science*, vol. 21, pp. 1521–1531, 1973.

V. R. Eshleman, "Jupiter's atmosphere: problems and potential of radio occultation," *Science*, vol. 189, pp. 876–878, 1975.

V. R. Eshleman, *et al.*, "Radio science investigations with Voyager," *Space Science Reviews*, vol. 21, pp. 207–232, 1977.

C. W. F. Everitt, B. W. Parkinson, and J. P. Turneaure, "The Gravity Probe B relativity gyroscope program," *Relativistic Gravitational Experiments in Space*, ed. R. W. Hellings, NASA CP-3046, pp. 118–125, 1988.

J. E. Faller and P. L. Bender, "A possible laser gravitational wave experiment in space," *Precision Measurement and Fundamental Constants II*, NBS Spec. Publ. 617, pp. 689–690, 1984.

J. E. Faller, P. L. Bender, J. L. Hall, D. Hils, and M. A. Vincent, "Space antenna for gravitational wave astronomy," in *Kilometric optical arrays in space,* Paris: European Space Agency, pp. 157–163, 1984.

T. Y. Fan and R. L. Byer, "Diode laser-pumped solid-state lasers," *IEEE Journal of Quantum Electronics,* vol. 24, pp. 895–911, 1988.

G. Fjeldbo and V. R. Eshleman, "The bistatic radar-occultation method for the study of planetary atmospheres," *Journal of Geophysical Research,* vol. 70, pp. 3217–3225, 1965.

W. M. Folkner and M. H. Finger, "Preliminary error budget for an optical ranging system: range, range rate, and differenced range observables," *TDA Progress Report* 42-101, vol. January–March 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 121-135, May 15, 1990.

E. B. Fomalont and R. A. Sramek, "Measurements of the solar gravitational deflection of radio waves in agreement with general relativity," *Phys. Rev. Lett.,* vol. 36, pp. 1475–1479, 1976.

E. B. Fomalont and R. A. Sramek, "The deflection of radio waves by the Sun," *Comments on Astrophysics,* vol. 7, pp. 19–33, 1977.

J. L. Friedman and B. F. Schutz, "Secular instability of rotating Newtonian stars," *Astrophysical Journal,,* vol. 222, pp. 281–296, 1978.

G. Gatewood, "The multichannel astrometric photometer and atmospheric limitations in the measurement of relative positions," *Astronomical Journal,* vol. 94, pp. 213–224, 1987.

G. Gatewood, *et al.,* "On the astrometric detection of neighboring planetary systems. II.," *Icarus,* vol. 41, pp. 205–231, 1980.

J. Gea-Banacloche, "Squeezing of spontaneous emission in a laser," *Phys. Rev. Lett.,* vol. 59, pp. 543–546, 1987.

T. Gold, M. J. Campbell, and B. T. O'Leary, "Optical and high-frequency electrical properties of the lunar sample," *Science,* vol. 167, pp. 707–709, 1970.

P. Grangier, R. E. Slusher, B. Yurke, and A. LaPorta, "Squeezed-light–enhanced polarization interferometer," *Phys. Rev. Lett.,* vol. 59, pp. 2153–2156, 1987.

J. L. Hall, "Stabilizing lasers, for applications in quantum optics," *International Conference on Quantum Optics,* vol. IV, eds. J. Harvey and D. Walls, pp. 273–284, 1986.

M. Z. Hansen, "Atmospheric particulate analysis using angular light scattering," *Applied Optics,* vol. 19, pp. 3441–3448, 1980.

H. F. Harmuth, *Sequency Theory, Foundations and Applications,* New York: Academic Press, 1977.

R. C. Harney, "Coherent infrared radar systems and applications II," *Proceedings of SPIE,* vol. 415, pp. 53–76, 1983.

P. B. Hays, V. J. Abreu, J. Sroga, and A. Rosenberg, "Analysis of a 0.5 micron spaceborne wind sensor," presented at the American Meteorological Society Conference on Remote Satellite Sensing, Clearwater, FL, 1984.

W. A. Heiskanen and H. Moritz, *Physical geodesy,* San Francisco: W. H. Freeman, 1967.

J. Helmcke, S. A. Lee, and J. L. Hall, "Dye laser spectrometer for ultrahigh spectral resolution: design and performance," *Applied Optics,* vol. 21, no. 9, pp. 1686–1694, 1982.

D. Hils, P. L. Bender, and R. F. Webbink, "Gravitational radiation from the galaxy," *Astrophysical Journal,* vol. 360, pp. 75–94, 1990.

E. D. Hinkley, ed. *Laser monitoring of the atmosphere,* Topics in Applied Physics, vol. 14, New York, Springer-Verlag, 1976.

C. J. Hogan, "Gravitational radiation from cosmological phase transitions," *Mon. Not. R. Astron. Soc.,* vol. 218, pp. 629–636, 1986.

F. Horz, ed., *Trajectory determinations and collection of micrometeoroids on the space station,* Houston: Lunar and Planetary Institute, No. 86-05, 1986.

J. Hough, *et al.,* "Gravitational wave detectors using laser interferometers and optical cavities: some practical aspects and results," in *Quantum Optics, Experimental Gravitation, and Measurement Theory,* eds. P. Meystre and M. O. Scully, New York: Plenum Press, pp. 515–524, 1983.

H. T. Howard, *et al.,* "Mercury: results on mass, radius, ionosphere, and atmosphere from Mariner 10 dual-frequency radio signals," *Science,* vol. 185, pp. 179–180, 1974.

D. A. Howe, *Frequency Domain Stability Measurements: A Tutorial Introduction,* Boulder: National Bureau of Standards, 1976.

W. B. Hubbard and J. D. Anderson, "Possible flyby measurements of Galilean satellite interior structure," *Icarus,* vol. 33, pp. 336–341, 1978.

R. M. Huffaker, ed., *Feasibility study of satellite-borne lidar global wind monitoring system,* Boulder: U.S. National Oceanic and Atmospheric Administration, No. PB-293-879, 1978.

J. R. Ipser and R. A. Managan, "On the emission of gravitational radiation from inhomogeneous Jacobi configurations," *Astrophysical Journal,* vol. 282, pp. 287–290, 1984.

S. Ismail and E. V. Browell, "Airborne and spaceborne lidar measurements of water vapor profiles: a sensitivity analysis," *Applied Optics,* vol. 28, pp. 3603–3651, 1989.

L. G. Jacchia and F. L. Whipple, "Precision orbits of 413 photographic meteors," in *Smithsonian Contributions to Astrophysics,,* vol. 5, pp. 97–129, 1961.

S. F. Jacobs, "How monochromatic is laser light?," *American Association of Physics Teachers,* vol. 47, pp. 597–601, 1979.

S. K. Jordan, "Statistical model for gravity, topography, and density contrasts in the Earth," *Journal of Geophysical Research,* vol. 83, pp. 1816–1824, 1978.

T. J. Kane, A. C. Nilsson, and R. L. Byer, "Frequency stability and offset locking of a laser-diode-pumped Nd:YAG monolithic nonplanar ring oscillator," *Optics Letters,* vol. 12, pp. 175–177, 1987.

T. J. Kane, B. Zhou, and R. L. Byer, "Potential for coherent Doppler wind velocity lidar using neodymium lasers," *Applied Optics,* vol. 23, pp. 2477–2481, 1984.

W. M. Kaula, *Theory of satellite geodesy,* Waltham, MA: Blaisdell Publishing Co., 1966.

W. M. Kaula, "Inference of variations in the gravity field from satellite-to-satellite range rate," *Journal of Geophysical Research,* vol. 88, pp. 8345–8349, 1983.

M. J. Kavaya and R. T. Menzies, "Lidar aerosol backscatter measurements: systematic, modeling, and calibration error considerations," *Applied Optics,* vol. 24, pp. 3444–3453, 1985.

L. G. Kazovsky, "Particle analysis using forward scattering data," *Applied Optics,* vol. 23, pp. 448–454, 1984a.

L. G. Kazovsky, "Estimation of particle size distributions from forward scattering data," *Applied Optics,* vol. 23, pp. 455–564, February 1984b.

L. G. Kazovsky and N. S. Kopeika, "Heterodyne detection through rain, snow, and turbid media: effective receiver size at optical through millimeter wavelengths," *Applied Optics,* vol. 22, pp. 706–710, 1983.

T. Keating, P. Taylor, W. Kahn, and F. Lerch, *Geopotential Research Mission, science, engineering and program summary,* NASA Technical Memorandum 86240, NASA Goddard Space Flight Center, 1986.

G. S. Kent, G. K. Yue, U. O. Farrukh, and A. Deepak, "Modeling atmospheric aerosol backscatter at CO_2 laser wavelengths. 1: aerosol properties, modeling techniques, and associated problems," *Applied Optics,* vol. 22, pp. 1655–1665, 1983.

D. J. Kessler, "Orbital debris issues," *Adv. Space Res.,* vol. 5, no. 2, pp. 3–10, 1985.

D. J. Kessler, "Orbital debris measurements," in *Trajectory Determinations and Collection of Micrometeoroids on the Space Station,* LPI Technical Report 86-05, ed. F. Horz, pp. 61–63, 1986.

H. J. Kimble and D. F. Walls, eds., "Squeezed states of the electromagnetic field," *Journal of the Optical Society of America B,* special issue, vol. 4, pp. 1450–1741, 1987.

A. J. Kliore and I. R. Patel, "Vertical structure of the atmosphere of Venus from Pioneer Venus Orbiter radio occultations," *Journal of Geophysical Research,* vol. 85, pp. 7957–7962, 1980.

R. G. Knollenberg, "Three new instruments for cloud physics measurements: The 2-D spectrometer, the forward scattering spectrometer probe, and the active scattering aerosol spectrometer," *Am. Meteorological Society,* International Conf. on Cloud Physics, pp. 554–561, 1976.

R. G. Knollenberg, "Techniques for probing cloud microstructure," in *Clouds, Their Formation, Optical Properties, and Effects,* New York: Academic Press, pp. 15–92, 1981.

R. G. Knollenberg and R. Luehr, "Open cavity laser 'active' scattering particle spectrometry from 0.05 to 5 microns," in *Fine Particles, Aerosol, Generation Measurement, Sampling and Analysis,* ed. B. Y. H. Liu, New York: Academic Press, pp. 669–696, 1975.

K. Koch and B. U. Witte, "Earth's gravity field represented by a simple-layer potential from Doppler tracking of satellites," *Journal of Geophysical Research,* vol. 76, pp. 8471–8478, 1975.

Y. Kojima and T. Nakamura, "Gravitational radiation from a particle with orbital angular momentum plunging into a Kerr black hole," *Progress of Theoretical Physics,* vol. 71, pp. 79–90, 1984.

L. Kresàk, "Sources of interplanetary dust," in *Solid Particles in the Solar System,* eds. I. Halliday and B. A. McIntosh, Boston: Reidel, pp. 211–222, 1980.

T. G. Kyle, "High resolution laser imaging system," *Applied Optics,* vol. 28, pp. 2651–2656, 1989.

J. D. Lindberg and J. B. Gillespie, "Relationship between particle size and imaginary refractive index in atmospheric dust," *Applied Optics,* vol. 16, pp. 2628–2630, 1977.

J. D. Lindberg, W. J. Lentz, E. M. Measure, and R. Rubio, "Lidar determinations of extinction in stratus clouds," *Applied Optics,* vol. 23, pp. 2172–2177, 1984.

R. Loudon, *The quantum theory of light,* New York: Oxford University Press, 1983.

W. H. Louisell, *Quantum statistical properties of radiation,* New York: Wiley, 1973.

J. L. MacArthur and A. S. Posner, "Satellite-to-satellite range-rate measurement," *IEEE Transactions on Geoscience and Remote Sensing,* vol. GE-23, pp. 517–523, 1985.

S. Machida and Y. Yamamoto, "Ultrabroadband amplitude squeezing in a semiconductor laser," *Phys. Rev. Lett.,* vol. 60, pp. 792–794, 1988.

I. S. McDermid, J. B. Laudenslager, and D. Rees, "UV-excimer laser based incoherent Doppler lidar system," in *Global Wind Measurements,* eds. W. E. Baker and R. J. Curran, Hampton, VA: A. Deepak, 1985.

K. D. McKeegan, R. M. Walker, and E. Zinner, "Ion microprobe isotopic measurements of individual interplanetary dust particles," *Geochim. Cosmochim. Acta,* vol. 40, pp. 1971–1987, 1985.

T. H. Maiman, "Stimulated optical radiation in ruby," *Nature,* vol. 187, pp. 493–494, 1960.

C. Man, A. Brillet, and P. Cerez, "Suppression of optical feedback effects on saturated absorption signals by phase modulation of the reflected light," *J. Phys. E: Sci. Instrum.*, vol. 11, pp. 19-21, 1978.

E. A. Marouf, G. L. Tyler, and V. R. Eshleman, "Theory of radio occultation by Saturn's rings," *Icarus,* vol. 49, pp. 161–193, 1982.

K. D. Mease, J. D. Anderson, L. J. Wood, and L. K. White, "Tests of general relativity using Starprobe radio metric tracking data," *Journal of Guidance, Control, and Dynamics,* vol. 7, pp. 36–44, 1984.

B. J. Meers, "Recycling in laser-interferometer gravitational-wave detectors," *Phys. Rev. D,* vol. 38, pp. 2317-2326, 1988.

G. Megie and R. T. Menzies, "Complementarity of UV and IR differential absorption lidar for global measurements of atmospheric species," *Applied Optics,* vol. 19, pp. 1173–1183, 1980.

R. Meneghini, J. Eckerman, and D. Atlas, "Determination of rain rate from a spaceborne radar using measurements of total attenuation," *IEEE Transactions on Geoscience and Remote Sensing,* vol. GE-21, no. 1, pp. 34–43, 1983.

R. T. Menzies, "A comparison of Doppler lidar wind sensors for Earth-orbit global measurement applications," in *Global Wind Measurements,* eds. W. E. Baker and R. J. Curran, Hampton, VA: A. Deepak, 1985.

R. T. Menzies, "Doppler lidar atmospheric wind sensors: a comparative performance evaluation for global measurement applications from Earth orbit," *Applied Optics,* vol. 25, pp. 2546–2552, 1986.

E. Merzbacher, *Quantum Mechanics*, New York: Wiley, 1970.

P. Meystre and M. O. Scully, eds., ***Quantum optics, experimental gravitation, and measurement theory***, New York: Plenum Press, 1983.

C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, San Francisco: W. H. Freeman and Company, 1973.

D. G. Monet and C. C. Dahn, "CCD astrometry I. preliminary results from the KPNO 4-m parallax program," *Astronomical Journal*, vol. 88, pp. 1489–1507, 1983.

G. E. Morfill and E. Grün, "The motion of charged dust particles in interplanetary space: Interstellar grains," *Planetary and Space Science*, vol. 27, pp. 1283–1292, 1979.

F. Morrison, "Algorithms for computing the geopotential using a simple density layer," *Journal of Geophysical Research*, vol. 81, pp. 4933–4936, 1976.

P. F. Moulton, "Tunable solid-state lasers targeted for a variety of applications," *Laser Focus/ Electro-optics*, pp. 56–69, 1987.

D. Mozurkewich, *et al.*, "Preliminary measurements of star positions with the Mark III stellar interferometer," *Astronomical Journal*, vol. 95, pp. 1269–1277, 1988.

E. Müller, "The collapse of rotating stellar cores: the amount of gravitational radiation predicted by various numerical models," in *Problems of Collapse and Numerical Relativity*, eds. D. Bancel and M. Signore, pp. 271–286, Reidel: Dordrecht, 1984.

C. D. Nabors, *et al.*, "Injection locking of a 13-watt cw Nd:YAG ring laser," *Optics Letters*, vol. 14, pp. 1189–1191, 1990.

A. C. Nilsson, E. K. Gustafson, and R. L. Byer, "Eigenpolarization theory of monolithic nonplanar ring oscillators," *IEEE Journal of Quantum Electronics*, vol. QE-25, pp. 767–790, 1989.

K. Nordtvedt, "Gravitomagnetic interaction and laser ranging to Earth satellites," *Phys. Rev. Lett.*, vol. 61, pp. 2647–2649, 1988.

G. W. Null, J. D. Anderson, and S. K. Wong, "Gravity field of Jupiter from Pioneer 11 tracking data," *Science*, vol. 188, pp. 476–477, 1975.

T. Okoshi, "Ultimate performance of heterodyne/coherent optical fiber communications," *Journal of Lightwave Technology*, vol. LT-4, pp. 1556–1562, 1986.

V. R. Pandharipande, D. Pines, and R. A. Smith, "Neutron star structure: theory, observation, and speculation," *Astrophysical Journal*, vol. 208, pp. 550–566, 1976.

E. M. Patterson, *et al.*, "Global measurements of aerosols in remote continental and marine regions: concentrations, size distributions, and optical properties," *Journal of Geophysical Research*, vol. 85, pp. 7361–7376, 1980.

R. V. Pound and J. L. Snider, "Effect of gravity on gamma radiation," *Phys. Rev.*, vol. 140, pp. B788–B803, 1965.

M. J. Post, "Aerosol backscattering profiles at CO_2 wavelengths: the NOAA data base," *Applied Optics*, vol. 23, pp. 2507–2509, 1984.

M. J. Post, R. A. Richter, R. M. Hardesty, T. R. Lawrence, and F. F. Hall, Jr., "National Oceanic and Atmospheric Administration's (NOAA) pulsed, coherent, infrared Doppler lidar – characteristics and data," *Proc. Soc. Photo-Opt. Instrum. Eng.*, vol. 300, pp. 60–65, 1981.

W. H. Press and K. S. Thorne, "Gravitational-wave astronomy," *Annual Review of Astronomy and Astrophysics*, vol. 10, pp. 335–374, 1972.

W. Priedhorsky, L. Stella, and N. E. White, *International Astronomical Union Circular* No. 4247, August 28, 1986.

R. H. Rapp, "Geos 3 data processing for the recovery of geoid undulations and gravity anomalies," *Journal of Geophysical Research*, vol. 84, pp. 3784–3792, 1979.

P. S. Ray, *et al.*, "Dual-Doppler observation of a tornadic storm," *Journal of Applied Meteorology*, vol. 14, pp. 1521–1530, 1975.

Lord Rayleigh, "On the transmission of light through an atmosphere containing small particles in suspension and on the origin of the blue of the sky," *Philosophical Mag.*, vol. 47, pp. 375–384, 1899.

R. D. Reasenberg, "Microarcsecond astrometric interferometry," in *Astrometric Techniques*, eds. H. K. Eichhorn and R. J. Leacock, Dordrecht: Reidel, pp. 321–330, 1986.

R. D. Reasenberg, *Optical interferometers for tests of relativistic gravity in space*, Annapolis: Smithsonian Astrophysical Observatory, 1988.

R. D. Reasenberg, *et al.*, "Viking relativity experiment: Verification of signal retardation by solar gravity," *Astrophysical Journal (Letters)*, vol. 234, pp. L219–L221, 1979.

R. D. Reasenberg, *et al.*, "Microarcsecond optical astrometry: an instrument and its astrophysical applications," *Astronomical Journal*, vol. 96, pp. 1731–1745, 1988.

M. K. Reed, W. J. Kozlovsky, R. L. Byer, G. L. Harnagel, and P. S. Cross, "Diode-laser-array-pumped neodymium slab oscillators," *Optics Letters*, vol. 13, pp. 204–206, 1988.

M. J. Rees, "Extragalactic sources of gravitational waves," in *Gravitational Radiation*, eds. N. Deruelle and T. Piran, Amsterdam: North Holland, pp. 297–320, 1983.

D. O. Richstone, "Evidence for massive black holes in galaxies," in *Supermassive Black Holes*, ed. M. Kafatos, Cambridge: Cambridge University press, pp. 87–95, 1988.

B. J. Rye, "Power radio estimation in incoherent backscatter lidar: direct detection with Gaussian noise," *Applied Optics*, vol. 28, pp. 3639–3646, 1989.

C. Salomon, D. Hils, and J. L. Hall, "Laser stabilization at the millihertz level," *J. Opt. Soc. Am. B.*, vol. 5, pp. 1576–1587, 1988.

S. Sandford, "Laboratory infrared transmission spectra from 2.5 to 25 microns of individual interplanetary dust particles," Ph.D. thesis, Washington University, 1985.

W. L. W. Sargent, *et al.*, "Dynamical evidence for a central mass concentration in the galaxy M87," *Astrophysical Journal*, vol. 221, pp. 731–744, 1978.

P. R. Saulson, "Terrestrial gravitational noise on a gravitational wave antenna," *Phys Rev. D*, vol. 30, pp. 732–736, 1984.

A. L. Schawlow and C. H. Townes, "Theory of infrared and optical masers," *Phys. Rev.*, vol. 112, pp. 1940–1949, 1958.

L. Schiff, *Quantum Mechanics,* New York, McGraw-Hill, 1968.

R. Schilling, *et al.*, "A method to blot out scattered light effects and its application to a gravitational-wave detector," *J. Phys. E: Sci. Instrum.*, vol. 14, pp. 65-70, 1981.

L. Schnupp, W. Winkler, K. Maischberger, A. Rüdiger, and R. Schilling, "Reduction of noise due to scattered light in gravitational wave antennae by modulating the phase of the laser light," *J. Phys. E: Sci. Instrum.*, vol. 18, pp. 482–485, 1985.

B. L. Schumaker, "Correction for laser phase fluctuations," contribution to NASA-sponsored workshop on a LAser Gravitational-wave Observatory in Space (LAGOS), Annapolis, MD, April 1990.

B. L. Schumaker, unpublished manuscript, 1989.

B. L. Schumaker, "Apparent brightness of stars and lasers," *TDA Progress Report* 42-93, vol. January-March 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 111-130, May 15, 1988.

B. L. Schumaker, S. H. Perlmutter, R. M. Shelby, and M. D. Levenson, "Four-mode squeezing," *Phys. Rev. Lett.*, vol. 58, pp. 357–360, 1987.

B. L. Schumaker, "Quantum mechanical pure states with Gaussian wave functions," *Physics Reports,* vol. 135, pp. 317–408, 1986.

B. L. Schumaker, "Validity of the standard quantum limit for monitoring free-mass position," in *Theoretical Investigations in Nonlinear Quantum Optics, Theory of Measurement, and Pulsations of General Relativistic Models of Neutron Stars,* Ph. D. Thesis, California Institute of Technology, 1985.

B. L. Schumaker, "Noise in homodyne detection," *Optics Letters,* vol 9, pp. 189–191, 1984.

B. L. Schumaker and K. Thorne, "Torsional oscillations of neutron stars," *Mon. Not. R. Astr. Soc.*, vol. 203, pp. 457–489, 1983.

Z. Sekanina, "Statistical model of meteor streams. III. Stream search among 19303 meteors," *Icarus,* vol. 62, pp. 253–284, 1973.

M. Shao, private communication, 1988.

M. Shao, *et al.*, "The Mark III stellar interferometer," *Astronomy and Astrophysics,* vol. 193, pp. 357–371, 1988.

M. Shao, *et al.*, "Application of interferometry to optical astrometry," *Astronomical Journal,* vol. 93, pp. 1280–1286, 1987.

J. H. Shapiro, *et al.*, "Theory of light detection in the presence of feedback," *J. Opt. Soc. Am.*, vol. B4, pp. 1604-1619, 1987.

I. I. Shapiro, *et al.*, "The Viking relativity experiment," *Journal of Geophysical Research,* vol. 82, pp. 4329–4334, 1977.

I. I. Shapiro, C. C. Counselman, III, and R. W. King, "Verification of the principle of equivalence for massive bodies," *Phys. Rev. Lett.*, vol. 36, pp. 555–558, 1976.

B. A. Silverman, B. J. Thompson, and J. H. Ward, "A laser fog disdrometer," *J. Appl. Met.*, vol. 3, pp. 790–801, 1964.

R. A. Simpson, G. L. Tyler, J. P. Brenkle, and M. Sue, "Viking bistatic radar observations of the Hellas basin on Mars: preliminary results," *Science*, vol. 203, pp. 45–56, 1979.

W. L. Sjogren, J. Lorell, L. Wong, and W. Downs, "Mars gravity field based on a short-arc technique," *Journal of Geophysical Research*, vol. 80, pp. 2899-2908, 1975.

W. L. Sjogren, R. J. Phillips, P. W. Birkeland, and R. N. Wimberly, "Gravity anomalies on Venus," *Journal of Geophysical Research*, vol. 85, pp. 8295–8302, 1980.

D. Sliney and M. Wolbarsht, *Safety with lasers and other optical sources*, New York: Plenum Press, 1982.

M. Sokolowski and J. Lesh, "Deep-space optical communications," *Proc. Soc. Photo-Opt. Inst. Eng.*, vol 810, pp. 172–177, 1987.

R. Spero, "The Caltech laser-interferometric gravitational wave detector," in *Proc. of the Fourth Marcel Grossman Meeting on Recent Developments of General Relativity*, ed. R. Ruffini, Amsterdam: North Holland, pp. 615–620, 1986.

R. Stachnik, ed., *Proceedings of the Workshop on Technologies for Space Optical Interferometry*, Lanham, MD, unpublished, 1989.

Staff of the Space Dept., Johns Hopkins Applied Physics Lab., and the Staff of the Guidance and Control Lab., Stanford Univ., "A satellite freed of all but gravitational forces: TRIAD I," *Journal of Spacecraft and Rockets*, vol. 11, pp. 637–644, 1974.

R. T. Stebbins, *et al.*, "A laser interferometer for gravitational-wave astronomy in space," in *Proc. 5th Marcel Grossman Conference*, eds. D. G. Blair and M. J. Buckingham, Cambridge: Cambridge University Press, pp. 1759–1767, 1989.

G. A. Tammann, "Supernova statistics and related problems," in *Supernovae: A Survey of Current Research*, eds. M. J. Rees and R. J. Stoneham, Dordrecht: Reidel, pp. 371–403, 1981.

C. H. Tang, T. I. S. Boak III, and M. D. Grossi, "Bistatic radar measurements of electrical properties of the Martian surface," *Journal of Geophysical Research*, vol. 82, pp. 4305–4315, 1977.

J. H. Taylor and J. M. Weisberg, "A new test of general relativity: gravitational radiation and the binary pulsar PSR 1913+16," *Astrophysical Journal*, vol. 253, pp. 908-920. 1982.

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

K. S. Thorne, "Gravitational radiation," in *300 Years of Gravitation*, eds. S. W. Hawking and W. Israel, Cambridge: Cambridge University Press, pp. 330–458, 1987.

K. S. Thorne, "How to test gravitation theories by means of gravitational-wave measurement," *Ondes et radiations gravitationnelles*,, ed. Y. Choquet-Bruhat, Paris: C.N.R.S, pp. 213–223, 1974.

G. L. Tyler, "Radio propagation experiments in the outer solar system with Voyager," *Proceedings of the IEEE,* vol. 75, pp. 1404–1431, 1987.

G. L. Tyler, *et al.,* "Voyager 2 radio science observations of the Uranian system: Atmosphere, rings, and satellites," *Science,* vol. 233, pp. 79–84, 1986.

G. L. Tyler, *et al.,* "Radio science with Voyager 2 at Saturn: Atmosphere and ionosphere and the masses of Mimas, Tethys, and Iapetus," *Science,* vol. 215, pp. 553–558, 1982.

G. L. Tyler, J. P. Brenkle, T. A. Komarek, and A. I. Zygielbaum, "The Viking solar corona experiment," *Journal of Geophysical Research,* vol. 82, pp. 4335–4340, 1977.

G. L. Tyler, R. A. Simpson, and H. J. Moore, "Lunar slope distributions: a comparison of bistatic-radar and photographic results," *Journal of Geophysical Research,* vol. 76, pp. 2790–2795, 1971.

R. K. Ulrich and G. W. Hawkins, "The solar gravitational figure–J_2 and J_4," *Astrophysical Journal,* vol. 246, pp. 985-988, 1981 (erratum in vol. 249, p. 831, 1981).

J. H. Underwood and J. E. Randolph, *Starprobe scientific rationale: a report of the ad hoc working groups,* Pasadena: Jet Propulsion Laboratory, no. 82-49, 1982.

H. C. van de Hulst, *Light Scattering by Small Particles,* New York: Wiley, 1957.

H. C. van de Hulst, *Multiple Light Scattering: Tables, Formulas and Applications,* New York: Academic Press, 1980.

R. F. C. Vessot, "Tests of gravitation and relativity," *Contemporary Physics,* vol. 25, pp. 355–380, 1984.

R. F. C. Vessot, *et al.,* "Test of relativistic gravitation with a space-borne hydrogen maser," *Phys. Rev. Lett.,* vol. 45, pp. 2081–2084, 1980.

M. A. Vincent and P. L. Bender, "Orbit determination and gravitational field accuracy for a Mercury transponder satellite," *Journal of Geophysical Research,* vol. 95, pp. 21357–21361, 1990.

J. -Y. Vinet, B. J. Meers, C. N. Man, and A. Brillet, "Optimization of long-baseline optical interferometers for gravitational-wave detection," *Phys. Rev. D,* vol. 38, pp. 433-447, 1988.

H. Volland, M. K. Bird, G. S. Levy, C. T. Stelzried, and B. L. Seidel, "Helios-1 Faraday rotation experiment: results and interpretations of the solar occultations in 1975," *Journal of Geophysics,* vol. 42, pp. 659–672, 1977.

J. A. Weinman, "Derivation of atmospheric extinction profiles and wind speed over the ocean from a satellite-borne lidar," *Applied Optics,* vol. 27, pp. 3994–4001, 1988.

R. Weiss, private communication to Schnupp, *et al.,* 1982.

F. L. Whipple, in *Cosmic Dust,* ed. J. A. M. McDonnell, New York: Wiley and Sons, pp. 1–70, 1978.

J. G. Williams, *et al.,* "New test of the equivalence principle from lunar laser ranging," *Phys. Rev. Lett.,* vol. 36, pp. 551–554, 1976.

F. B. Winn, *et al.*, "A solar plasma stream measured by DRVID and dual-frequency range and Doppler radio metric data," *Deep Space Network Progress Report,* Pasadena: Jet Propulsion Laboratory, vol. 42-37, pp. 43–54, 1977.

M. Wolff, "Direct measurements of the earth's gravitational potential using a satellite pair," *Journal of Geophysical Research,* vol. 74, pp. 5295–5300, 1969.

R. Woo, "Radial dependence of solar wind properties deduced from Helios 1/2 and Pioneer 10/11 radio scattering observations," *Astrophysical Journal,* vol. 219, pp. 727–739, 1978.

L. Wu, H. J. Kimble, J. L. Hall, and H. Wu, "Generation of squeezed states by parametric down conversion," *Phys Rev. Lett.*, vol. 57, pp. 2520–2523, 1986.

Y. Yamamoto, "Receiver performance evaluation of various digital optical modulation-demodulation systems in the 0.5–10 μm wavelength region," *IEEE Journal of Quantum Electronics,* vol. QE-16, pp. 1251–1259, 1980.

P. J. Young, J. A. Westphal, J. Kristian, C. P. Wilson, F. P. Landauer, "Evidence for a supermassive object in the nucleus of the galaxy M87 from SIT and CCD area photometry," *Astrophysical Journal,* vol. 221, pp. 721–730, 1978.

H. P. Yuen, "Contractive states and the standard quantum limit for monitoring free-mass positions," *Phys. Rev. Lett.*, vol. 51, pp. 719–722, 1983.

M. Zimmermann, "Revised estimate of gravitational radiation from Crab and Vela pulsars," *Nature,* vol. 271, pp. 524–525, 1978.

H. A. Zook, "Precision requirements on cosmic dust trajectory measurements," in *Trajectory Determinations and Collection of Micrometeoroids on the Space Station,* LPI Technical Report 86-05, ed. F. Horz, pp. 97–99, 1986.

H. A. Zook, R. E. Flaherty, and D. J. Kessler, "Meteoroid impacts in the Gemini windows," *Planetary and Space Sciences,* vol. 18, pp. 953-964, 1970.

D. S. Zrnic, "Estimation of spectral moments for weather echoes," *IEEE Transactions on Geoscience Electronics,* vol. GE-17, no. 4, pp. 113–128, 1979.

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No. JPL Publication 90-50	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Scientific Applications of Frequency-Stabilized Laser Technology in Space		5. Report Date December 1, 1990 6. Performing Organization Code
7. Author(s) Bonny L. Schumaker		8. Performing Organization Report No.
9. Performing Organization Name and Address JET PROPULSION LABORATORY California Institute of Technology 4800 Oak Grove Drive Pasadena, California 91109		10. Work Unit No. 11. Contract or Grant No. NAS7-918 13. Type of Report and Period Covered JPL Publication
12. Sponsoring Agency Name and Address NATIONAL AERONAUTICS AND SPACE ADMINISTRATION Washington, D.C. 20546		14. Sponsoring Agency Code RE 4 BP-646-00-00-02-00

15. Supplementary Notes

16. Abstract

This report is a synoptic investigation of the uses of frequency-stabilized lasers for scientific applications in space. It begins by summarizing properties of lasers, characterizing their frequency stability, and describing limitations and techniques to achieve certain levels of frequency stability. Limits to precision set by laser frequency stability for various kinds of measurements are investigated and compared with other sources of error. These other sources include photon-counting statistics, scattered laser light, fluctuations in laser power and intensity distribution across the beam, propagation effects, mechanical and thermal noise, and radiation pressure. Methods are explored to improve the sensitivity of laser-based interferometric and range-rate measurements. Several specific types of science experiments that rely on highly precise measurements made with lasers are analyzed, and anticipated errors and overall performance are discussed. Qualitative descriptions are given of a number of other possible science applications involving frequency-stabilized lasers and related laser technology in space, applications that will warrant more careful analyses as technology develops.

17. Key Words (Selected by Author(s))	18. Distribution Statement
Spacecraft Communications, Command, and Tracking Optics Space Sciences (General) Astronomy	Unclassified -- Unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	276	