Chapter 10

LISA: A proposed joint ESA–NASA gravitational-wave mission

Peter L. Bender, on behalf of the LISA Study and Mission Definition Teams
JILA, National Institute of Standards and Technology and
University of Colorado, Boulder, Colorado 80309-0440, USA

10.1 Description of the LISA mission

10.1.1 Introduction

The evidence for supermassive black holes at the centres of quasars and active galactic nuclei (AGNs) has been strong but not conclusive for several decades. Recently, the evidence for massive black holes in one particular AGN, in our own Galaxy, and in the Local Group Galaxy M32 has become extremely convincing. Thus, questions concerning gravitational waves generated by the interaction of massive black holes with smaller compact objects and with each other have become of strong interest. Signals from such sources also are likely to provide the strongest possible tests of general relativity.

Plans are being developed in both Europe and the USA for flying a dedicated gravitational wave mission called the Laser Interferometer Space Antenna (LISA). The antenna will measure gravitational waves in the frequency range from roughly 1 \mu Hz to 1 Hz, and thus will strongly complement the results expected from ground-based detectors. The primary objectives are to obtain unique new information about massive black holes throughout the universe, and to map the metric around massive black holes (MBHs) with much higher accuracy than otherwise would be possible. Other important objectives include studies of resolved signals from thousands of compact binary star systems in our Galaxy, and looking for a possible gravitational-wave cosmic background at millihertz frequencies.
The possibility of making sensitive gravitational-wave measurements at low frequencies by laser interferometry between freely floating test masses in widely separated spacecraft appears to have been first suggested in print about 1972 [1, 2]. More extensive discussions started in 1974. An initial proposal similar to that for the LISA mission was presented at the Second International Conference on Precision Measurement and Fundamental Constants in 1981 [3], and at the ESA Colloquium on Kilometric Optical Arrays in Space in 1984 [4]. Work on the concept was supported initially by the National Bureau of Standards, and later in the USA by NASA. However, the concept became much better defined and widely known in the 1993–1994 period, when more extensive studies were carried out in Europe under ESA support. Since then, further studies have been carried out by both ESA and NASA of a proposed joint ESA–NASA mission, that could fly as early as 2010 if all goes well.

The currently proposed mission is described in section 1 of this chapter. This includes the overall antenna and spacecraft design, the optics and interferometry system, the free mass sensors, the required micronewton thrusters for the spacecraft, and the mission scenario. The emphasis will be on aspects of the antenna that are quite different from those for ground-based detectors. Section 2 describes the main scientific results that seem likely to be obtained by LISA. This includes unique new information on three major astrophysical questions concerning MBHs and a nearly ideal test of general relativity, as well as the detection of thousands of compact binaries in our Galaxy. In addition, some speculations will be given on possible future prospects for gravitational-wave observations in space after the LISA mission.

Much more information on most of the above subjects can be found in 'Laser Interferometer Space Antenna', the Proceedings of the Second International LISA Symposium [5], in the LISA Pre-Phase A Study [6], and in a special issue of Classical and Quantum Gravity [7], which is the Proceedings of the First International LISA Symposium. More recent information from the 1999–2000 ESA Industrial Study of the LISA mission is being provided in the report of that study and in several papers in the Proceedings of the Third International LISA Symposium (Albert Einstein Institute, Potsdam, 11–14 July 2000).

10.1.2 Overall antenna and spacecraft design

The basic geometry is shown schematically in figure 10.1. Three spacecraft form an equilateral triangle 5000 000 km on a side, and laser beams are sent both ways along each side of the triangle. A Y-shaped thermal shield inside each spacecraft contains the sensitive parts of the scientific payload, consisting of two separate optical assemblies mounted in the two top arms of the Y, so that they are aimed along the two adjacent sides of the triangle. The spacecraft instrumentation and a cover over the sunward side of the spacecraft are not shown.

Each spacecraft is in a one year period solar orbit with an eccentricity $e$ of about 0.01 and an orbit inclination to the ecliptic of $3^0.5 \times e$ [8]. By choosing the phasing and the orientations of the orbits properly, the three spacecraft will form the desired nearly equilateral triangle, which is tipped at 60° to the ecliptic (figure 10.2). The plane of the triangle will precess around the pole of the ecliptic once per year, and the triangle will rotate in that plane at the same rate. The centre of the triangle is chosen to be about 50 000 000 km (20°) behind the Earth. The arm lengths for the triangle will stay constant to about 1% for a number of years. Locating the triangle 60° from the Earth would keep the arm lengths more constant, but at the expense of more propulsion required to reach the desired orbits and more telemetry capability to send the data back.

The layout of one of the two optical assemblies in each spacecraft is shown in figure 10.3 (see [9]). Each optical assembly contains an optical bench, a transmit/receive telescope, and a low-power electronics package. Each of these three subassemblies is mounted by low-thermal-conductivity struts from a stiffened support cylinder that forms the outside of the optical assembly. Because the distances between the spacecraft will change by up to 1% during the year, the angles between the sides will also change by roughly one degree. Thus, it is necessary to provide some adjustment for the angle between the axes of the two
Figure 10.2. Location of the LISA antenna, 50 000 000 km behind the Earth in orbit around the Sun. The plane of the antenna is tipped at 60° to the ecliptic.

Figure 10.3. Layout of one of the two optical assemblies in each spacecraft. The main components are the optical bench at the centre and the transmit/receive telescope at one end.

Figure 10.4. Mounting of optical assemblies inside main thermal shield. Each optical assembly must rotate slowly and smoothly during the year over about a one degree range.

Figure 10.5. Optical bench for one optical assembly. Light from the laser is mainly sent out to the telescope, with a little going to the photodiode. Received light from the other end of the arm passes through the beamsplitter to the test mass, and then goes to the photodiode. Changes in the phase of the beat signal determine the changes in the arm length.

optical assemblies in each spacecraft. This is accomplished by supporting the lower part of each optical assembly by a flexure from a common block near the base of the Y-shaped shield, and the upper part by two adjustable displacement mechanisms from an arm of the Y, as shown in figure 10.4 (see [10]). The mechanisms provide a very smooth one degree change in the angle between the arms with mainly a one year period.

At the centre of each optical bench, shown in figure 10.5 (see [9]), is a freely floating test mass that is protected as much as possible from sources of spurious accelerations. The basic measurements made by the LISA antenna
are the changes with time in the distances between test masses in the different spacecraft. Capacitive sensors on the inside of a housing surrounding each test mass sense the position of the spacecraft with respect to the test mass. This position information is fed via a proportional control servosystem to micronewton thrusters that keep the spacecraft essentially fixed with respect to the test mass. Each test mass plus the housing around it and the associated electronics will be referred to as a ‘free mass sensor’. The names ‘inertial sensor’, ‘drag-free sensor’, and ‘disturbance reduction sensor’ also have been used frequently, but each choice has some drawbacks. ‘Inertial sensor’ causes confusion with the gyroscopic sensors used in inertial navigation systems. ‘Drag-free sensor’ or ‘disturbance reduction sensor’ emphasizes the role of the sensor in permitting external forces on the spacecraft to be accurately compensated for, but do not indicate the importance of keeping the spurious accelerations of the test mass far below the residual accelerations of the spacecraft.

At frequencies below roughly 3 mHz, the threshold sensitivity of the LISA antenna will be determined mainly by the spurious accelerations of the test masses, despite the care that is taken to minimize such disturbances. However, at higher frequencies the main limitation will come from how well the relative changes in the 500000 km distances between the test masses can be measured. The optical interferometry system for measuring these distances will be described in the next section. The free mass sensors and the way they are used will be discussed in detail in a later section.

In describing the operation of the gravitational-wave antenna, it is useful to distinguish between the ‘main spacecraft’ and the ‘instrument package’. Here the instrument package will mean the Y-shaped thermal shield and everything inside it. In addition, it will include the lasers and a thermal radiation plate they are located on that is attached to the bottom of the Y. With this arrangement, the heat generated by the lasers can be radiated by the radiation plate out through a large hole in the bottom of the spacecraft to space. Everything else, including the spacecraft structure and all the other equipment mounted on it, will be referred to as the main spacecraft.

Probably the most important factor in the design of the LISA mission, after minimizing spurious forces on the test masses, is the need to keep the temperature distribution throughout each main spacecraft and instrument package as constant as possible. This is essential for two main reasons. One is that mass displacements due to local temperature changes within the desired measurement band, roughly $10^{-6}$ Hz to 1 Hz, will change the gravitational force on the test mass and look like a real signal [10,11]. The other is that a variation in the temperature difference between the two sides of the housing in a free mass sensor will cause a difference in the thermal radiation pressure force on the two sides of the test mass, and thus give a spurious acceleration.

The thermal design of the LISA spacecraft can be understood on the basis of the following rough model. The main spacecraft will have some thermal time constant for responding to changes in the incident solar flux or in the electrical power dissipated in its equipment. The Y-shaped thermal shield is mounted from the main spacecraft by low-thermal-conductivity flexural strips or stressed fibreglass bands, and forms a second passive thermal isolation stage. The third stage is the support cylinder around each optical assembly and the fourth stage is the optical bench that contains the free mass sensor. Since low-thermal-conductivity supports are used between each successive stage, and with low infrared emissivity coatings on as much of the surfaces as possible, quite long thermal time constants can be achieved for each stage.

With the above type of multistage passive thermal isolation approach, quite low levels of temperature fluctuations can be achieved at frequencies of 0.1 mHz and higher. But there are practical limits to how long the various thermal time constants can be made. To achieve high thermal stability at lower frequencies, some active thermal control will be needed. Even a fairly crude servosystem for the temperature of the Y-shaped thermal shield would help considerably, and some active control of the temperature of the main spacecraft probably will be desirable as well. Such active temperature control systems may be added to the baseline mission design later to meet a goal of improved antenna sensitivity at frequencies below 0.1 mHz, but are not included in the present mission requirements.

A fundamental feature of the mission design is the high stability of the solar energy dissipation in the spacecraft. In an ecliptic-based reference frame that rotates once per year about the pole of the ecliptic, the three spacecraft form a nearly equilateral triangle in a plane that is tipped at 60° to the ecliptic, as discussed earlier. Thus, the direction to the Sun makes a nearly constant 30° angle with the normal to the sunward side of each spacecraft, where the solar cells providing power for the spacecraft are located. Because of the stable solar illumination geometry, only fluctuations in the intrinsic solar intensity and in the electrical power dissipation in the various parts of the spacecraft produce significant temperature variations. Measures such as keeping the microwave transmitters operating at the same power level whether data is being transmitted or not are taken to reduce thermal changes. Also, the power needed in the optical assemblies and particularly on the optical benches is kept as low as possible, and is highly regulated.

### 10.1.3 Optics and interferometry system

The present parameters assumed for the LISA optical system are the use of 1 W of output power from cw Nd:YAG lasers operating at 1.064 μm wavelength, and of 30 cm diameter telescopes to transmit and receive the laser beams between different spacecraft. Perhaps the most impressive thing about the mission is that these fairly modest parameter values lead to roughly $2 \times 10^8$ photons per s being detected at the far end of a 500000 km arm, which would correspond to $2 \times 10^{-21}$ precision in measuring changes in the arm length in one second if there were no other sources of error. Since almost all expected types of LISA gravitational-wave sources can be observed for a year or more, the potential
measurement accuracy clearly is extremely high.

Each optical assembly has two lasers for redundancy, and a switching mechanism for switching between them. Each laser is capable of 2 W output over a mean lifetime of several years, but is operated at 1 W to extend the lifetime by a substantial factor. A schematic drawing of the optical bench for one optical assembly was shown earlier in figure 10.5. Light from the active laser is taken by a single mode fibre onto the optical bench, and then formed into a beam of roughly 5 mm diameter. A few per cent of the beam is split off for other purposes, and then a beamsplitter sends almost all of the power out to the telescope, where it is transmitted to the matching optical assembly on the other end of the arm. The return beam from the laser in the distant optical assembly is sent by the telescope to the test mass, where it is reflected from the face of the test mass, and then beat against a small amount of power from the local laser.

The beat signals detected in the six optical assemblies are the main data used to make the gravitational-wave measurements. In the present baseline scheme, one laser in spacecraft SC-1 is considered the master laser and locked to a stable reference cavity on its optical bench. The other laser in SC-1 is phase locked to the first with a small offset frequency of perhaps a couple of kilohertz. The two laser beams are transmitted along the adjacent arms of the triangle to SC-2 and SC-3, where they are reflected from the test masses in the receiving optical assemblies. The local lasers are then phase locked to the received beams, also with small frequency offsets, and their beams are sent back to SC-1. Finally, the second lasers in SC-2 and SC-3 are offset locked to the first lasers, their beams are sent over the third arm of the triangle, and the beats between them are recorded on both ends of the side, after reflection from the test masses.

The frequencies of the beat signals can range from well under 1 MHz to roughly 15 MHz, depending on the exact orbits of the spacecraft, or rather the orbits of the freely floating test masses in them. In one possible measurement scheme, the signals are beat down to a convenient frequency range for precise phase measurements, such as perhaps 10 kHz, using outputs from frequency synthesizers. The drive frequencies for the synthesizers come from ultra-stable crystal oscillators (USOs) on each spacecraft. In a second measurement scheme, the signals are first sampled at a high rate such as 40 MHz, and the results are analysed in the software. In this case, the sampling times are controlled by the USOs.

If only the signals from the two arms adjacent to SC-1 are considered, the antenna is like a single ground-based detector, with changes in the arm length difference being the quantity of interest. However, it does not appear practical to keep the difference in arm length nearly constant, since this would require applying quite large forces to the test masses to overcome the changing gravitational forces on them due to the Sun and to the planets. If such large forces were applied, it would be very difficult to avoid noise in those forces within the frequency band of interest from giving undesired changes that could not be distinguished from real signals. To avoid this problem, the test masses are left

nearly free, except for small differential applied forces on the two test masses in a given spacecraft that are necessary to compensate for the difference in spurious forces on them at dc and at frequencies below the useful measurement band.

Because of the up to 1% difference in arm lengths expected, the phase noise in the master laser has to be corrected for too high accuracy. The original suggestion that this was possible was made by Faller in the abstract for a 1981 conference [3]. The basic idea is to use the apparent changes in the sum of the lengths of the two arms to estimate the laser phase noise, and then to apply the corresponding correction to the measured difference in the arm lengths. Approximate algorithms for doing this were published by Giampieri et al in 1996 [12]. More recently, Armstrong, Tinto and Estabrook [13, 14] have given what are believed to be rigorously correct algorithms, and it is planned to use these during the LISA mission.

So far, only the use of the signals from the two detectors on SC-1 has been discussed. In addition, the signal from one of the detectors on each of SC-2 and SC-3 is used to phase lock a laser to the received laser beam from SC-1. So the signals measured on SC-2 and SC-3 from the beams sent between them can be combined to give the changes in length of the third arm. As discussed by Cutler [15], the length of the third arm minus the average of the lengths of the other two is an observable that gives the other polarization from the one determined by the difference in lengths of the first two. Thus, having the measurements over the third arm gives a valuable addition to the scientific information that can be obtained from a mission like LISA.

In the recent papers of Armstrong, Tinto and Estabrook [13, 14], the analysis is done in a way that assumes all the lasers are running independently, presumably stabilized by locking to their own stable reference cavities, but not locked to each other. This appears to put some additional requirements on the mission measurement system, but they point out that there is additional information in the resulting signals at the shorter gravitational wavelengths that is not available with the current baseline measurement system. It appears that this may be equivalent to the information that would be obtained from running LISA as a Sagnac interferometer, but this possibility has not yet been considered in detail.

In addition to providing the second polarization if all is going well, having the capability of making measurements over the third arm has another very important benefit for the LISA mission. With two fully functional optical assemblies on each of the three spacecraft, if something in one of the six were not operating properly, the antenna would still provide two-arm gravitational-wave data with the full planned sensitivity. Even if a second optical assembly were out of operation, provided the two were not on the same spacecraft, the two-arm data could still be obtained. Thus, the third arm capability provides an essential level of redundancy for the mission.

In addition to the shot noise in the photocurrents from the received laser beams, other noise sources in measuring changes in the arm length differences also have to be considered. One such source is fluctuations in the attitude of
each spacecraft. If the transmitted beam gave a perfectly spherical wavefront in the far field, attitude changes for the transmitting spacecraft would not give phase changes in the received signal at the distant spacecraft. However, the combination of diffraction due to the finite size transmitting aperture plus imperfections in the optical system cause the wavefronts at the distant spacecraft to be somewhat distorted. Thus, the attitude of the spacecraft and any other sources of jitter in the pointing of the transmitted beam have to be controlled quite closely.

There are two methods under consideration for measuring jitter in the spacecraft attitude. In one, perhaps 10% of the received light from the distant spacecraft is picked off and focused on a CCD array or quadrant diode via a fairly long effective focal length optical system. Changes in attitude then give changes in the position of the focal spot on the detector, which are used in a servo system to control the spacecraft attitude, along with the similar signals from the second optical assembly on the spacecraft. In the second approach, all of the light goes to the main detector for measuring changes in the arm length, but the detector is replaced by a quadrant device. If the spacecraft tips slightly with respect to the received wavefronts, the differences in phase of the four detected signals will change, and these changes are used as the inputs to the attitude control system.

Even if the attitude jitter is made small, there still is a need to set the mean beam pointing direction carefully. Diffraction plus a defocus of the transmitted beam would result in the phase of the received wavefront varying only quadratically with the angular offset from the optical axis of the transmitting system. However, non-axisymmetric defects in the transmitted wavefronts can make the change in the received phase vary linearly with the attitude change, even on what otherwise would have been the optical axis. To avoid this increased sensitivity to beam pointing, the pointing along each axis of each optical assembly is modulated at a known frequency, and the resulting apparent changes in the arm length differences at the modulation frequencies and their second harmonics are detected. This information permits the outputs at the modulation frequencies to be minimized by small offsets in the dc pointing directions, which corresponds to having just the quadratic variations in the received phase due to attitude jitter.

As a measure of the remaining requirement on pointing jitter, a convenient test case is to assume only astigmatic error in the transmitted wavefronts with an error of a tenth of a wavelength rms. For this example, if the distance error due to pointing is allowed to be equal to that from shot noise, the errors in the dc pointing offset and in the pointing jitter can be as large as 10 milliradians and 4 milliradians/rtHz, respectively. (Here and later, rtHz stands for 'per square root Hz', and the given error or noise is the spectral amplitude of the error, which is the square root of the power spectral density.) These error allocations are each about three times larger than those given in section 3.1.8 of [6], since the error allocation in that case was assumed to be only 10% of the error due to shot noise.

Further information on the optical path error allocation budget for LISA is given in sections 4.2.1 and 4.2.2 of [6]. The total error allocation for the measurement of the difference in the round-trip path lengths for two arms of the antenna is 40 nm/rtHz. With only shot noise considered, the corresponding value would be 22 nm/rtHz, and with beam pointing jitter included also, the value becomes 30 nm/rtHz. Other error sources considered explicitly are as follows: residual laser phase noise after the correction procedure discussed earlier is applied; noise in the ultra-stable oscillators after a similar correction procedure; noise in the laser phase measurements and phase locks; and scattered light effects. The 40 nm/rtHz total error allocation for the difference in round-trip paths for two arms corresponds to 20 nm/rtHz for determining the difference in lengths for the two arms.

### 10.1.4 Free mass sensors

Historically, the first free mass sensor was developed jointly by the Johns Hopkins University Applied Physics Laboratory and by Stanford University, and was flown on the TRIAD satellite in 1972 [16]. A spherical test mass was contained in a spherical cavity with three opposing pairs of capacitive electrodes on the inside of the housing for sensing the relative position of the test mass. Whenever the atmospheric drag caused the housing to move a few millimetres with respect to the test mass, pulsed thrusters on the satellite fired to keep it centred on average on the test mass. Care was taken to keep forces on the test mass other than those due to external gravitational fields as small as possible. Thus, the orbit of the satellite was nearly drag-free, and was much more predictable than normal. The theory of such drag-free systems had been given earlier in 1964 by Lange [17].

There is a close connection between free mass sensors and the high-performance force-rebalance accelerometers developed over the last three decades by the Office National d’Etudes et de Recherches Aérospatiales (ONERA) in Paris for flight on various missions. The basic approach in the accelerometers is to measure the position of a free or nearly free test mass inside a housing by means of capacitive electrodes on the inside of the housing. Forces are then applied to the test mass by means of voltages on the electrodes to keep the test mass centred in the housing. The required voltages are measures of the accelerations along the three perpendicular axes.

The first such accelerometer designed and built by ONERA was flown from 1975 to 1979 on the French CASTOR-D5B satellite. It had a spherical test mass. However, later ONERA-designed accelerometers have used test masses in the form of rectangular parallelepipeds. The first such design, called the GRADIO accelerometer, was for possible use in a proposed gravity gradiometer mission (ARISTOTLES) to map the Earth’s gravity field, and has been the basis for a number of later designs.

The test mass for the GRADIO accelerometer is 4 cm by 4 cm by 1 cm in dimensions. The material used for the test mass is a Pt-Rh alloy with a density of about 20 g cm⁻³. The housing is made of ultra-low expansion glass (ULE), with a gold coating on the inside that is carefully patterned to form the capacitive plate electrodes. The gaps between the plates are recessed and kept very small.
to reduce the charge buildup on the underlying material. The position sensing is done with 100 kHz transformer bridge circuits. Extra pairs of plates are included to permit relative rotation of the test mass with respect to the housing to be sensed and controlled.

The gaps between the capacitive plates and the test mass is 30 μm for the large square faces and 300 μm for the smaller rectangular ones. With the test mass horizontal, the factor four smaller vertical dimension and the 30 μm gap permit the test mass to be electrostatically suspended for laboratory tests. The entire accelerometer is contained in a vacuum enclosure. Extensive tests have been carried out at ONERA with two GRADIO accelerometers mounted on a single pendular support platform in a special sub-basement room to reduce the effects of tilts in the floor. The results for the differential horizontal accelerations have been very encouraging, but are still several orders of magnitude worse than the theoretical sensitivity of roughly $10^{-12}$ m s$^{-2}$/μHz from a few millihertz to 1 Hz in a quiet and constant temperature zero-g environment.

A later but lower sensitivity version of the GRADIO accelerometer has been flown a few times on the space shuttle, and other versions will fly soon on other missions. These include the German–French CHAMP mission scheduled for launch in 2000, a USA–German mission called GRACE scheduled for a 2001 launch, and a later ESA mission called GOCE. GRACE is a satellite-to-satellite tracking mission that will map time variations in the Earth’s gravity field over a five-year period with high accuracy, and an accelerometer on each satellite will monitor drag and other non-gravitational forces on the satellites. GOCE is a gravity gradiometer mission to map very short wavelength spatial structure in the gravity field, using the differential accelerations between pairs of accelerometers in the same spacecraft.

For the free mass sensors for the LISA mission, the requirements are quite different from those for accelerometers. The purpose of the accelerometers is to measure the non-gravitational accelerations of the spacecraft to which they are attached, and thus the readout sensitivity has to be high. For the GRADIO accelerometer, with 300 μm gaps along the two most sensitive axes, the position measurement sensitivity is $6 \times 10^{-12}$ m/μHz at frequencies above 5 mHz. However, the small gaps lead to increases in some of the time-varying spurious accelerations of the test mass, as discussed later. For LISA, it turns out that the necessary sensitivity of the capacitive position measurements is only about $10^{-9}$ m/μHz, so gaps of 2 mm or larger and different values of some of the other measurement parameters can be used to reduce the spurious accelerations.

An initial ONERA study of a proposed free mass sensor for LISA, called CAESAR, was published by Touboul, Rodrigues and Le Clerc in 1996 [18]. The test mass is a cube 4 cm on a side, made of a 10% Pt–90% Au alloy to achieve low magnetic susceptibility and a density of 20 g cm$^{-3}$ (see figures 10.6 and 10.7). The gaps assumed were 4 mm along the most sensitive axis and 1 mm for the other two axes. An error budget for the sensor was developed, based on the high thermal stability expected for the optical bench during the LISA mission.

![Figure 10.6. Early ONERA design of LISA free mass sensor called CAESAR. Up to four capacitor plates on each face of the housing around the test mass allow the relative position and angular orientation to be measured accurately. Electrical forces also can be applied, if desired.](image)

After completion of the initial CAESAR design, additional studies of free mass sensor designs and of the LISA requirements have been carried out, influenced strongly by the CAESAR design. A number of the factors influencing the design were discussed by Speake [19] and by Vitale and Speake [20]. A design with a considerably different geometry for the capacitor plates was investigated by Josselin, Rodrigues and Touboul [21], and has since been constructed and tested in the laboratory and the LISA free mass sensor requirements were simplified and somewhat relaxed [6].

One current version of the error allocation budget for the free mass sensor is described in section 4.2.3 of [6]. The total acceleration error allowed for one sensor is $3 \times 10^{-15}$ m s$^{-2}$/μHz over the frequency range 0.1–30 mHz. The six largest allocations to individual error sources are as follows: thermal distortion of the main spacecraft; temperature difference variations across the
test mass housing; electrical force on the charge on the test mass; Lorentz force on the charged test mass from the fluctuating interplanetary magnetic field; residual gas impacts on the test mass; and fluctuating forces due to electrical field dissipation in the test mass housing. These sources are each allocated an error of $1 \times 10^{-15} \text{ m s}^{-2}\text{/rtHz}$.

In view of the performance level desired for the LISA free mass sensors, it is planned to have a technology demonstration flight for them at the earliest possible date. Efforts to arrange for such a flight currently are under way both in Europe and the USA. Two sensors would be flown on a thermally isolated optical bench on a small spacecraft, with an interferometer to measure changes in the separation of the two test masses. To keep the rate of charging of the test masses comparable with what it would be well away from the Earth, and to reduce the effect of the Earth’s gravity gradient, a perigee altitude of 10000 km or higher is desired. To reduce the cost, a tentative goal of demonstrating only $3 \times 10^{-14} \text{ m s}^{-2}\text{/rtHz}$ performance from 1–10 mHz currently is assumed. It is believed that cautious engineering extrapolation from tests at this level plus some in-flight tests with intentionally increased disturbances will provide a sound basis for proceeding with the LISA mission.

10.1.5 Micronewton thrusters

The main non-gravitational force on a LISA spacecraft is expected to be from the solar radiation pressure and to have a magnitude of about 20 $\mu$N. If not compensated for, it would cause an acceleration of the spacecraft of roughly $10^{-7} \text{ m s}^{-2}$. Since the test masses are shielded from this source of acceleration, it is necessary to apply force to the spacecraft to keep it from moving with respect to them. The spectral amplitude of the fractional fluctuations in the solar pressure force [22] over the 0.1–10 mHz frequency range is approximately

$$1.3 \times 10^{-3} (f/1 \text{ mHz})^{-1/3}.$$  

Comparable fluctuations in force but with a more reddened spectrum and a much lower dc level will be present from the solar wind.

The type of thrusters that are planned for use on the LISA mission are field emission electric propulsion (FEEP) thrusters. They operate by accelerating ions through a potential drop of 5–10 kV and ejecting them to provide thrust. The ejection velocity is roughly 60–100 km s$^{-1}$, corresponding to a specific impulse of 6000 to 10000 s. In view of this high specific impulse and the low thrust level needed, the fuel required per thruster for a ten year extended mission lifetime is only a few grams.

Historically, most of the development of FEEP thrusters has been based on the use of Cs ions. This work has been done mainly at the European Space Research and Technology Centre in Noordwijk, The Netherlands and at Centrospazio in Pisa, Italy (see section 7.3 of [6]).

A schematic drawing of a thruster is shown in figure 10.8. In each thruster, liquid Cs metal at a temperature somewhat above the melting point of 29 °C is contained in a small reservoir. It is drawn by capillary forces through a narrow channel between two polished metal plates spaced 1 or 2 $\mu$m apart. The accelerating voltage is applied by a plate with a slot in it, located on the outer edge of the channel.

The high field at the surface of the Cs metal causes an instability, and Taylor cones roughly a micrometre in diameter and a few micrometres apart form on the surface. The tips of these cones are very sharp, and the high field around the tips causes Cs ions to be drawn out by field emission and accelerated away in a beam perhaps 30° wide. The one substantial drawback to the use of Cs ions is that any water vapour that is present will react with the Cs to form CsOH. Thus, the thrusters are kept in vacuum containers roughly 5 cm in dimensions until in space and ready for use. A spring-loaded cover with an O-ring seal is then released.

An alternative to Cs is the use of In. Field emission of In ions for the neutralization of positive charge on spacecraft has been pioneered by the Austrian...
10.1.6 Mission scenario

A single launch with the Delta II launch vehicle is assumed in current studies of the LISA mission [7]. Each of the three spacecraft has a thin cylindrical propulsion module attached to it. The three composite vehicles are stacked on top of each other for launch, and then separate from the launch vehicle and from each other a short time later. Since the spacecraft will be separated by 5000 000 km and have different orbit planes in their final configuration, they travel separately and each has its own onboard guidance system.

It is currently planned to use solar electric propulsion for the propulsion modules. The top area of the main spacecraft is covered with solar cells, and enough power is generated to provide 15 or 20 mN of thrust with either Hall effect or another type of Xe ion thruster. The time necessary for the three spacecraft to reach their proper locations and velocities about 20° behind the Earth in orbit about the Sun is 14 or 15 months.

After the nominal orbits have been achieved, the spacecraft will be tracked for a week or two with NASA’s Deep Space Tracking Network. Then, any desired minor corrections to the orbits can be made. The next step is separation of the propulsion modules from the final spacecraft. After that, the orbits will be almost completely gravitational, with each spacecraft servocontrolled to follow either the average of the two test masses inside it or just one of the two. The difference of the spurious accelerations of the two test masses will be roughly $10^{-10}$ m s$^{-2}$ or less, and can be corrected for by applying weak and stable electrical forces to the test masses at frequencies below the measurement range via the capacitor plates.

The next step is for each spacecraft to acquire the other two optically. Star trackers aboard each spacecraft give the attitude to a few arcseconds, and the transmitted laser beams are defocused by roughly a factor ten so that they can be detected on the other end of an arm even with relatively poor beam pointing. The received laser beam normally is bright enough to look like a magnitude-3 star, so there will be a substantial brightness level even with some defocusing. If necessary, an angular scan pattern on the transmitted beam could be used to find the beam, since the stability of the star tracker will be considerably better than its absolute accuracy.

As soon as a receiving spacecraft finds the beam, the output from the CCD array or quadrant diode in its angle tracker can be used to lock the attitude to the direction of the distant spacecraft, and the defocus in the transmitted beam can be removed to give a high S/N ratio. The local laser then is turned on, and its frequency adjusted to give a fairly low beat frequency with the received beam. The resulting heterodyne signal back at the original transmitting end provides a higher S/N ratio, and makes detecting the defocused return beam easier. After the beam pointing signals are locked in and the focus is corrected on both ends of each arm, the system is ready for operation.

Even with careful phase locking of the master laser to a stable cavity, or possibly the locking of more lasers, the phase noise will still be significant up to...
frequencies of a few hundred hertz to 1 kHz. Also, the amplitude of the phase noise at frequencies near 1 Hz and below will be large. To avoid aliasing the higher frequency noise into the band of interest, below perhaps 3 Hz, it may be desirable to make the phase measurements at a rate approaching a kilohertz, and then filter the data in software. For low-frequency noise, it is planned to combine all the signals at a single spacecraft and perform the algorithm for correcting for laser phase noise there. The resulting signals then can be compressed before being sent to the ground.

The number and amplitude of mechanical motions aboard the spacecraft are kept as small as possible to reduce gravitational attraction changes and excitation of vibrations. Roughly 30 cm diameter X-band antennae probably will be used to send data down to 34 m Deep Space Network antennae on the ground, and will have to be repositioned about once a week. Step changes in the angle between the axes of the two optical assemblies probably will be made at the same time with coarse adjustment mechanisms, but smooth adjustments over roughly a range of 5 arcmin are required between the steps. It is expected that offsets in the position, velocity, and dc acceleration of the test masses will have to be solved for at the times of the steps.

The nominal mission lifetime after the antenna geometry is established probably will be three years. However, there are no expendables required except for the In or Cs fuel for the micronewton thrusters, and that will be made adequate for a ten year or longer lifetime. Thus, a rolling three year approval period hopefully can be established, provided that the antenna continues to operate properly.

10.2 Expected gravitational-wave results from LISA

10.2.1 LISA sensitivity and galactic sources

The two types of noise sources for the LISA antenna have been discussed earlier. The one that dominates at frequencies below about 3 mHz is spurious accelerations of the test masses, and the level of error allocated for frequencies $f$ between 0.1 and 30 mHz is $3 \times 10^{-15} \text{ m s}^{-2} \text{rtHz}$ for each free mass sensor. Considering only two arms of the antenna for simplicity, and adding the errors from the four sensors quadratically, the resulting noise level for the difference in (one-way) length of the two arms is

$$\Delta(L_2 - L_1) = 1.520 \times 10^{-16} \left(\text{1 Hz}/f\right)^2 \text{ m/rtHz}.$$ 

This should be combined quadratically with the error allocation of $2 \times 10^{-11} \text{ m/rtHz}$ for measuring the difference in the distances between the test masses along the two arms.

For low frequencies, where the wavelength is long compared with the arm length, the response of the antenna to a gravitational wave [24] can be given simply for the optimum direction of propagation and polarization of the wave. The change in $L_2 - L_1$ is just the amplitude $h$ of the wave times the cosine of the angle between the two arms. Averaging over the direction of propagation and the polarization gives a factor $50.5$ lower signal amplitude. However, at higher frequencies, the signal is reduced by a factor that depends in a complicated way on the direction of propagation and the polarization. The transfer function giving the rms value of this reduction factor, and including the cosine of the angle between the two arms, has been given by Schilling [25]. It was used in calculating the LISA threshold sensitivity, as shown in figure 10.9 and discussed below. Essentially the same curve has been calculated independently by Armstrong, Tinto and Estabrook [13], but for a slightly different choice of arm length.

For almost all of the gravitational-wave signals LISA will see, the frequency, amplitude and polarization of the wave will be very stable, and will not change substantially over a few years of observation. Thus, it is desirable to plot the LISA threshold sensitivity as a curve such that there is a good chance of seeing a source in a reasonable observing time if its rms amplitude lies above the curve. To accomplish this, the threshold sensitivity curves in figures 10.9 and 10.11 are plotted as the rms amplitude of a stable signal needed in order to have a S/N ratio of five for one year of observation, when averaged over source direction and polarization. The S/N ratio of five is chosen because almost all the sources will be unknown from optical observations, and roughly this S/N ratio is needed to determine the reality of sources with unknown frequencies and with only
statistical information on their directions in the sky. When averaged over a year, the LISA antenna response is remarkably independent of the source direction in the sky [26], so no allowance for the ecliptic latitude or longitude of the source is needed in using the LISA threshold sensitivity curve.

The dominant type of signal LISA will see is gravitational waves from binary stars systems in our galaxy. Mironowskii [27] pointed out in 1966 that there would be roughly $10^8$ W UMa type binaries in the galaxy giving substantial signals. These binaries consist of two main sequence stars so close together that their Roche lobes are in contact, and material can flow between them. Their frequencies lie in a fairly narrow band near 0.1 mHz. With so many sources at low frequencies, it is clear that even many years of observations would not permit individual signals to be resolved, except for a few that happen to be close to the Sun and thus have high signal strengths. There are even larger numbers of non-contact binaries composed of normal stars at lower frequencies.

In 1984 Iben [28] pointed out the expected existence of a very large number of close white dwarf binaries (CWDBs) in our galaxy, and the strength of signals expected from them. Because of their radii being only about 10,000 km, they can give gravitational-wave signals at frequencies up to roughly 30 mHz. However, they had not been observed directly until the last few years, and their space density in the galaxy is only poorly known. On the other hand, since the orbits of the interesting CWDBs evolve mainly by gravitational radiation, the number per hertz will decrease as the $1/3$ power of the frequency, and thus the frequency above which most of them can be resolved is only fairly weakly dependent on the space density.

Recent observations by Marsh and collaborators [29–32] have found a number of examples of CWDBs, so their abundance must be moderately close to the astrophysical estimates. However, the criteria used in selecting their targets of observation make it difficult to derive a space density. Based on the best astrophysical studies available at present [33,34], it is expected that most CWDBs will be resolved at frequencies above roughly 3 mHz, even for those near the galactic center. The signals are strong enough so that the frequencies of and directions to a few thousand such sources will be determined by LISA.

The mean strength of signals from CWDBs at the galactic center as a function of frequency is shown in figure 10.9 as a solid curve. It hoks sharply upward at about 15 mHz because the majority of the CWDBs below that frequency contain either one or two He white dwarfs, and coalesce just above 15 mHz. This leaves only binaries containing mainly white dwarfs composed almost completely of carbon and/oxygen, which are called CO white dwarfs. The CO white dwarfs are roughly a factor two more massive and are smaller in radius, so they can reach somewhat higher frequencies before coalescing.

Above and below the curve for the galactic center are dot-dashed curves labelled 5% and 95%. The 5% curve shows the mean strength of the signals from CWDBs at a distance from the Earth such that 5% of all the ones in the galaxy are closer, and the definition for the 95% curve is similar. Thus roughly 90% of the CWDBs in the galaxy will give signal strengths in the region between the two curves. Since this region is substantially above the LISA threshold sensitivity curve for S/N = 5, most of these sources will have S/N ratios of 10–50.

Several other types of binaries also can contribute signals above the instrumental sensitivity level for LISA [35–43]. When all of these types are considered together, there is some frequency at which the expected number of galactic sources per frequency bin for one year of observations drops below a value of very roughly $1/3$. Below that frequency, very little information can be obtained about the scientifically very important extragalactic sources, unless they are stronger than the sum of signals from the galactic binaries. Above that frequency, some information is lost because of having to fit out the galactic signals, but some survives [44–46]. The effective binary noise level then drops by a factor of about 3–10, at which point it is coming from all the binaries in other galaxies throughout most of the universe.

A simple calculation shows that the extragalactic contribution from each shell around our galaxy of a given thickness will give the same average contribution out until redshifts comparable with one, provided the effects of galaxy evolution are not too large [47]. A recent estimate of the effective noise level introduced by all of the galactic and extragalactic binary star systems, but not allowing for evolution [46], is shown in figure 10.9. It has been adjusted to correspond to S/N = 5. This confusion noise curve lies above the instrumental sensitivity curve for frequencies from about 0.1–3 mHz, and has a major effect on determining whether LISA can reliably detect particular sources from one year of observations. This affects some galactic sources, as well as extragalactic ones.

A few individual known sources also are shown on figure 10.9. Since the frequencies, as well as directions in the sky, are known for most of these sources, the S/N ratio needed to detect them is only about two. Thus, the ones somewhat below the threshold sensitivity or confusion noise curves probably are detectable from one to two years of observations. i Boo is a non-contact binary composed of normal stars, and WZ Sge is a very short period dwarf nova binary. The four Am CVn binaries shown are each believed to consist of an extremely low mass helium dwarf losing mass to a CO dwarf primary [39]. They also are called helium cataclysmics, and effects due to such binaries and their possible progenitors have been included in the recent confusion noise estimates [46]. The final known binary shown is 4U 1820-30, an x-ray binary believed to contain a neutron star or stellar mass black hole.

Before discussing the important questions about massive black holes and about relativity that LISA hopefully will contribute unique new information on, it is useful to say something about how the data will be analysed [48]. First, any large and easily recognizable signals will be fitted and removed from the data. Next, a convenient data set such as a one year data record will be analysed carefully to determine as many as possible of the galactic binary signals. This search will be easy for frequencies above perhaps 10 mHz, where the signals are quite well separated. However, the sidebands on the signals due to rotation
and motion of the antenna will overlap strongly as the frequency decreases to near 3 mHz, and a simultaneous fit of all the signals in each roughly 100 or 200 cycle/year band is likely to be necessary in order to fit the signals well. At still lower frequencies, only the strongest sources whose signals rise substantially above the large number from the galactic center will be detectable. All of the identified sources will then be removed from the data record, before looking for the swept frequency signals expected from extragalactic sources involving massive black holes.

10.2.2 Origin of massive black holes

An important astrophysical question is how the seed black holes that later grew to be the massive and supermassive black holes observed today were formed. To aid in identifying different mass ranges for black holes, we refer to those from roughly 1.5 to 30 solar mass (M⊙) that are thought to be capable of being formed by the evolution of very massive stars as stellar mass black holes, larger ones up to about 3 x 10^7 M⊙ as massive black holes (MBHs), and still larger ones as supermassive black holes.

As is well known, supermassive black holes were invoked first to provide the energy source necessary to explain quasars. However, more recent optical and other observations have provided strong evidence for the existence of MBHs in AGNs much closer to the Earth [49]. In one case, the evidence is from observations of OH masers in Keplerian orbits around the MBH, and is essentially conclusive [50,51].

For some normal galaxies, the evidence also is very strong. One case comes from extremely high-resolution infrared observations of stars moving around a 2.6 ± 0.2 x 10^5 M⊙ object at the centre of our galaxy, that really can only be a MBH [52,53]. Two other cases come from optical observations of the motions of stars or gas around the centres of the M31 (Andromeda) and M32 galaxies in the Local Group at distances of about 0.6 Mparsec. The evidence for stellar mass black holes comes from observations of x-ray binaries.

For MBH masses of 10^6 M⊙ or less, the only fairly convincing observations so far are two OH maser observations indicating about 10^6 M⊙ central objects. Except for a few more observations of this kind, it is not apparent whether electromagnetic observations are likely to tell us much about the fraction of normal galaxies containing MBHs that are this low in mass. Because of this, the information potentially available from gravitational wave observations with LISA concerning the existence and masses of MBHs in other normal galaxies and how they formed is likely to be quite valuable.

There are two main types of theories concerning how seed black holes were formed. In one, collisions of stellar mass objects in dense galactic nuclei led to the formation of higher mass objects, and these sank down toward the centre (mass segregation), where their collision rates were enhanced. If the mass got to be a few hundred times the solar mass and an object was not already a black hole, it would evolve quickly to form one. When the largest objects got to be roughly a thousand solar mass in size, they would then be able to continue growing fairly rapidly by absorption of gas in the galactic nucleus and by tidal disruption of stars. At some point, the largest black hole would grow enough faster than the others that it would swallow up the ones of comparable size and become the seed for growth of a perhaps 10^5 M⊙ or larger MBH.

The alternate type of theory involves the evolution of a dense cloud of gas and dust to the point where it becomes optically thick, and radiation pressure plus magnetic fields can prevent further fragmentation of the cloud to form stars. At that point, if energy and angular momentum can be dissipated fairly rapidly, there are two options. In one, a supermassive star possibly 10^7–10^8 M⊙ in size is formed, and quickly evolves to the point of relativistic collapse and forms a MBH. In the other option, the cloud can become dense enough to reach the point of relativistic instability and collapse directly to a MBH without going through the supermassive star stage. There also is the possibility of a relativistic star cluster becoming unstable and collapsing to a MBH.

For the collisional growth scenario, quite detailed calculations starting from 1M⊙ stars were carried out by Quinlan and Shapiro [54] (see this 1990 paper for earlier references). They found that roughly 100 M⊙ objects could form in a few times 10^8 years, starting from plausible conditions in a dense galactic nucleus, and including the effects of mass segregation. However, it was not possible at that time to follow the process further.

In an alternate approach, Lee [55] started from assuming that 1% of the mass in a dense galactic nucleus was in the form of 7 M⊙ black holes that resulted from evolution of stars at the high end of the initial mass function (i.e. initial mass distribution). The rest of the material was in the form of 0.7 M⊙ normal stars. Dynamical friction led to segregation of the black holes to the core, and core collapse among the black holes occurred on a time scale much shorter than for a single component cluster. For rms stellar velocities above 100 km s⁻¹ and for plausible densities in the nucleus, it was shown that many black hole binaries formed and merged to produce 14 M⊙ black holes within about two billion years. The process was not followed further, but it seems likely that most of the black holes would have merged rapidly to form a substantial sized seed MBH.

The main objection raised to the collisional growth scenario is that it seems difficult to produce the seed black holes and have them grow much further to the supermassive black hole size before the appearance of quasars as early as a redshift of four [56–58]. Instead, it was suggested that the inefficiency of star formation would leave most of the material in a dense cloud in the form of gas and dust. The cloud would cool and condense toward the centre until angular momentum support became important. Gravitational instabilities and other effects would help to remove energy and angular momentum and permit the density to become high enough for collapse to a supermassive star or directly to a MBH perhaps 10^5 M⊙ or larger in size. A related argument made is that a self-gravitating gaseous object of more than 10^8 M⊙ does not appear to have any
stationary non-relativistic equilibrium state that can be supported for very long.

Under the cloud collapse scenario, an important question is whether signals are likely to be produced that LISA could see. For example, if a supermassive star forms and then evolves to the relativistic instability, the final collapse to a MBH could be slow enough that most of the gravitational-wave radiation would be at such low frequencies that LISA would have poor sensitivity. Also, if the collapse were nearly spherically symmetric, the radiative efficiency would be poor. However, recent fully relativistic calculations by Baumgarte and Shapiro [59] of the evolution of a rotating supermassive star up to the onset of collapse provide some basis for a more optimistic view. The later evolution can be determined reliably only by a numerical, three-dimensional hydrodynamics simulation in general relativity. However, estimates of what will happen indicate that most of the mass will go into a MBH, and that a bar instability which radiates efficiently at frequencies observable by LISA may form.

Even if the supermassive black holes in quasars at high redshifts are formed initially by cloud collapse, it still seems quite possible that the collisonal growth scenario may contribute substantially to the formation of seed black holes for more modest sized MBHs, like the one in our galaxy. Under the collisional growth scenario, if a number of 500$\: M_\odot$ seed black holes are formed before runaway growth occurs and the largest has already swallowed the others, then the coalescence of two of these seeds could be seen by LISA even at a substantial redshift. For such a coalescence at $z = 5$, it can be shown that the signal strength as a function of frequency during the last year before coalescence, for a circular orbit, would just about at the LISA threshold sensitivity curve level, so the event would be detectable with $S/N = 5$.

A remaining question, even if many intermediate sized MBHs are produced by the collisional growth of seeds, is whether the time of runaway growth of the largest seed black hole would be delayed to high enough mass for LISA to observe the coalescences. As pointed out by Lee [60], the calculations of Quinlan and Shapiro are based on the Fokker-Planck approach, and that approach does not allow for the proper statistical treatment of a runaway instability. This is true for the calculations of Lee [55] also. In considering this issue, it should be noted that only the chirp mass for the binary is important for determining the signal strength and frequency as a function of time. Thus the coalescence of a 100$M_\odot$ black hole with a 4000$M_\odot$ one would be as observable as for two 500$M_\odot$ black holes. Further work on the runaway growth question, as well as on the overall collisional growth and cloud collapse scenarios, certainly would be valuable.

### 10.2.3 Massive black holes in normal galaxies

Another important astrophysical question concerns the abundance of intermediate size MBHs of roughly $10^5$–$10^6M_\odot$. From observations based almost entirely on galaxies containing larger MBHs and SMBHs, various authors have estimated that the mass of the central object is about 500 times smaller than the mass of the spheroid (central bulge) of the galaxy. An even tighter relationship to the velocity dispersion in the spheroid has been reported recently [61]. However, the reasons for these relationships are not yet known [56, 62–64]. It appears likely that LISA data can address whether a relation something like this extends to galaxies with smaller spheroids, which constitute the majority of all galaxies.

MBHs in galactic centres are expected to usually have an increased density of stars around them in the region where the potential of the MBH dominates that of the galaxy. This density cusp is usually taken to be a power law cusp, with a $-7/4$ power dependence on radius if the distribution of stellar motions is relaxed and a $-3/2$ power for some unrelaxed cusps. It is generally expected that there will be large numbers of compact stars, i.e. white dwarfs and neutron stars, in the cusp. Occasionally, one of them that is on a nearly radial orbit and passes close to the MBH may be deflected enough by the other stars so that it comes within five or so gravitational radii of the MBH and loses significant amounts of energy and angular momentum by gravitational radiation. If so, and if further deflections by the other stars are not important, the compact star orbit will continue to shrink gradually until coalescence with the MBH occurs.

Unfortunately, in almost all cases for white dwarfs and neutron stars, the above gradual approach scenario is interfered with by interactions with other stars. Hils and Bender [65] have simulated what happens for a particular model which assumes $1M_\odot$ for both the compact stars and the normal stars in the cusp. After the first pass near the MBH, the orbit of the compact star is modified by interactions with the other stars more rapidly than by the gravitational radiation, unless the compact star is bound very tightly to the MBH initially. Thus the compact star usually will plunge rapidly into the MBH, or its point of closest approach will wander far enough away to not give appreciable interaction. A rapid plunge will not provide enough integration time for detection. In the remaining favourable cases, the signal typically will be observable by LISA from a one year data record starting up to roughly 100 years before coalescence.

Despite a loss of several orders of magnitude in the event rate due to plunging, the study by Hils and Bender [65] gave some hope of LISA seeing such signals. However, studies by Sigurdsson and Rees [66] and by Sigurdsson [67], plus an unpublished extension of the above study by Hils and Bender, indicated that the prospects were considerably better for observing gradual approaches to coalescence with galactic centre MBHs for roughly 5 or 10$M_\odot$ black holes. The effect of mass segregation was included. Such events would be detectable at a redshift of $z = 1$ during the last year before coalescence for MBH masses from $5 \times 10^4$ to $2 \times 10^6M_\odot$.

Estimates of the event rate are certainly model dependent, but still offer encouragement that multiple signals of this kind will be observed. As an example, results obtained by Hils and Bender for one particular model are shown in figure 10.10. About 1% of the mass in the cusp was assumed to be in 7$M_\odot$ black holes, and mass segregation was included. For each factor two range in the mass of the central MBH about a nominal value $M$, and for each factor two
The third of the important astrophysical questions that LISA has a good chance of giving new information on concerns the development of structure in the universe. It is currently believed that objects considerably smaller than galaxies formed first, and that continuing sequences of interactions and mergers led to the present distribution of galaxy types and sizes, and to galaxy clusters and superclusters. If MBHs were already present in pre-galactic structures that merged, this could lead to the MBHs sinking down to the centre of the new structure by dynamical friction and getting close enough together to coalesce due to gravitational radiation before the next merger [58,69].

The above scenario is quite attractive, and could give a high event rate for LISA. If so, valuable new constraints on the merger process and on the conditions after mergers would be obtained. However, there are a number of issues that affect the event rate that have to be considered.

One question is how effective dynamical friction will be in bringing MBHs of a given size to the centre of the new structure in less than the time between mergers. For galaxies like ours, the time required is less than the Hubble time for MBHs of roughly $3 \times 10^6 M_\odot$ or larger (see, e.g., Zhu and Ostriker [70]). However, the stars in the cusp around a MBH before merger will stay with it for some time, and will affect the dynamics. Thus, whether the mergers of pre-galactic structures in galaxies like ours with fairly modest spheroid and MBH masses are likely to have produced coalescences of perhaps $10^5$ or $10^6 M_\odot$ MBHs appears to be an open one.

There also is a question about whether the MBHs will modify the star distribution near the centre of the new structure enough so that the dynamical friction will be decreased considerably. Calculations indicate that this may occur when the MBHs are fairly close together, but not yet close enough to coalesce by gravitational radiation in less than a Hubble time (see, e.g., Makino [71]). On the other hand, conditions such as Brownian motion of the MBH and tri-axiality in the structure may affect the results, and there also may be complex enough motions going on soon after a merger to change the conclusions. Thus the effectiveness of this “hang-up” in the coalescence process is not known.

Another question arises if another merger occurs before the MBHs from an earlier merger have coalesced. In principle, one, two or all three of the MBHs could be thrown out of the new structure by the slingshot effect. However, the new merger could even have a beneficial effect, if the resulting disturbances in the central star density overcame the possible hang-up for the earlier two MBHs and allowed them to coalesce.

The signals arising from MBH coalescences after mergers of galaxies or of pre-galactic structures would be very large. Figure 10.11 shows several cases
of the signal strength as a function of frequency during the last year before coalescence for events at a redshift of $z = 1$ and for circular orbits. The square symbols show the time at 0.2 year intervals, so the first is one year before, the second is 0.8 year before, etc. The last symbol is shifted slightly to be 0.5 week before coalescence, instead of at that time. The cases shown are for equal MBH masses, and only the spiral-in part of the event before the last stable orbit is reached is considered. However, it should be remembered that the detectability of the signal for the spiral-in phase depends mainly on the chirp mass, so the curves given can be used to estimate conditions under which unequal mass coalescences would be observable also. The case of possible coalescence of 500$M_{\odot}$ seed black holes during the early growth of MBHs that was discussed earlier is shown for $z = 1$ for comparison.

Since the LISA threshold sensitivity curve and the confusion noise estimate are defined on the assumption that the frequency and signal strength are fixed for the one year period of the observation, and the MBH–MBH coalescence signals change dramatically during the year, it is necessary to integrate the square of the amplitude $S/N$ ratio during the year and then take the square root to obtain the effective $S/N$ ratio. The way in which the effective $S/N$ ratio builds up during the year, with the LISA instrumental noise and the confusion noise combined quadratically, is shown in figure 10.12. What is shown for each of the cases from

figure 10.11 is the accumulated $S/N$ ratio after the first week, the second week, etc, up until the end of the year. The large jump in the last week or so, in all but the 500$M_{\odot}$ case, is due to the frequency of the signal having swept up to where the LISA sensitivity is much higher.

It is clear by extrapolating roughly from figures 10.11 and 10.12 that MBH–MBH coalescences resulting from structure mergers could be seen clearly even at redshifts of 10 for $10^7M_{\odot}$ MBHs and 20 or more for less massive ones. Thus, any such events at any plausible occurrence time will be observable. Particularly because of the large uncertainties concerning what happens for roughly $10^6M_{\odot}$ or smaller MBHs after mergers of smaller black holes, the actual rate for LISA apparently could range from less than one per decade to quite a large value.

### 10.2.5 Fundamental physics tests with LISA

Of similar importance to the astrophysical questions about MBHs discussed in the last three sections is the fundamental physics question of whether Einstein’s general relativity theory is correct. Very strong tests of the theory can be provided either by MBH–MBH coalescences with high $S/N$ ratio, if they are observed, or by highly unequal mass coalescences of 5 or 10$M_{\odot}$ black holes with MBHs in galactic centres. Only the latter case will be discussed here, since the much larger number of orbital periods observable and the substantial orbit eccentricity expected are likely to be more important for testing relativity than a higher $S/N$ ratio.
It has been emphasized by Cutler et al [72] that the strongest test of a theory is likely to come from the observed phase of the signal, rather than the amplitude. This is because the cross-correlation of a theoretical template with the observed signal will be reduced substantially if the two get out of phase by even half a cycle or less during the entire data record. For our case of a perhaps 10$^M$ black hole orbiting around a 10$^5$ or 10$^6$M$\odot$ MBH during the last year before coalescence, there will be roughly 10$^5$ cycles. Even a very small error in the metric will lead to a continuously increasing error in the orbit, and thus in the phase of the calculated signal. While some type of metric error could, in principle, affect the orbit in the same way as slightly different values of parameters in the problem, such as the spin and mass of the MBH, it seems unlikely that a conceptual breakdown in the theory would be nearly equivalent to just having different parameter values. Thus, correctly predicting the signal phase over 10$^5$ cycles would be an extremely strong test of the theory.

For this case, it is important that the orbit started out being nearly radial. The later evolution of the eccentricity from its initial value very close to unity can be calculated from the rates of loss of energy and angular momentum due to gravitational radiation. The results obtained by Tanaka et al for a Schwarzschild MBH [73] show for a typical case that the orbit never becomes circular, but instead remains substantially eccentric (0.5 ?) up until the final plunge begins. However, rigorous calculations will be required in order to fit the data well [74].

During the last year, the periastron distance changes only very slowly, but the apoapsis distance decreases substantially. At periastron, the speed is about half that of light, so the dynamics are highly non-Newtonian. In fact, the precession of periastron during one radial motion period can be about a whole cycle. With relativistic beaming of the gravitational radiation and variation of the strength of the radiation with time during a radial period, the amplitude observed in a given direction can vary in a quite complex way. For an orbit plane that is not perpendicular to the spin axis for a rapidly rotating MBH, rapid Lense-Thirring precession also will be present. In view of the complexity of the relativistic motion and the large number of cycles over which the phase of the signal can be followed, such a signal would give a nearly ideal test of the predictions of general relativity.

It is of course necessary to be able to detect the signal in order to test the theory. It was mentioned earlier that a S/N ratio of about ten would be needed in order to detect such a complex signal. But there also is the question of how difficult the search for such a signal would be. A very crude estimate of the number of templates needed for a brute force search with one year of data gives, within a couple of orders of magnitude, something like 10$^{18}$. Even with rapid advances in computing power, such a search probably would not be possible. Some improvement could be made by a hierarchical search strategy, but it is not clear that this approach would be sufficient. On the other hand, more powerful search algorithms such as 'genetic algorithms' and 'stimulated annealing' have been developed and demonstrated for substantially more challenging search problems. Consideration of such algorithms for use with LISA data is just starting, but it currently appears unlikely that the difficulty of the search problem will be a real limitation in the use of the LISA data, even for the case of highly unequal mass binaries. For all other expected types of LISA sources, the search problem is much easier than for ground-based detectors because of the roughly 10$^4$ times lower number of data points that have to be handled for a year of observations.

The other important question is whether the ability to calculate theoretical templates will have improved enough by the time LISA data is available to carry out a thorough test of general relativity. For comparable mass MBHs, there is a very long way to go to accomplish this. However, for the highly unequal mass case, the chances for fairly rapid progress seem much better. It is hoped that numerical methods can be developed that start from the test mass approximation, and converge moderately well. Still, since small changes in the initial conditions can lead to very large changes in the motion a year later, it is important that this problem as well as the search problem for the highly unequal mass case be pursued vigorously in the next few years.

For either detailed tests of general relativity with LISA or for studies of astrophysical questions concerning MBHs, the first requirement is that some signals involving MBHs be seen. While it seems likely that several of the types of MBH sources discussed earlier will be observed, this certainly is not guaranteed. Thus, the situation is somewhat like that for ground-based observations, where the detection of signals within the next decade seems quite likely, but is not certain. Still, a reasonable prediction is that both ground-based detectors in the next decade and LISA not too much later will detect the desired types of signals, and extremely strong tests of general relativity will be among the most important scientific results.

Another fundamental physics test that LISA may contribute to concerns the possible existence of a primordial gravitational-wave background [75, 76]. Standard inflation theory predicts a nearly scale-invariant spectrum, in which the spectral amplitude for the gravitational waves falls off at about the $-1.5$ power of the frequency. However, the observed COBE microwave background spectrum is believed to be determined mainly by the large-scale density fluctuations at the time of decoupling, and with a scale invariant spectrum would predict a very low amplitude in the frequency range of LISA and of ground-based detectors.

There are theories that could give a non-scale-invariant spectrum and detectable amplitudes for a cosmic background at LISA or ground-based frequencies. LISA could detect an isotropic cosmic background near 10 mHz if its energy density were roughly $10^{-11}$ of the closure density. A candidate for giving a detectable background is a phase transition at about the electro-weak energy scale, but it would have to be strongly first order in order to give enough amplitude. But this is not currently thought to be likely. Another possibility is associated with suggested effects of extra dimensions, which may have become very small at a time which would produce a peak in the gravitational-wave spectrum near the
LISA frequency band. In any case, LISA will improve the limits on a possible cosmic background in its frequency band, but the chances of seeing anything appear to be very uncertain.

10.2.6 Future prospects

Provided that signals involving massive black holes are indeed seen by LISA, it is likely that there will be an opportunity for a later advanced gravitational-wave mission. Depending on what is seen, there are several directions in which major instrumental improvements could be made. If improvements in sensitivity at frequencies above 3 mHz are of most interest, this can be achieved by increasing the telescope size and the laser power, provided that comparable reductions in other noise sources such as beam pointing jitter and phase measurement errors also can be made. Shortening the antenna arm length also would be desirable. If improvements below about 0.1 mHz are the main objective, then improved free mass sensors and longer arm lengths would be needed.

As an example of a mission to achieve strongly increased high-frequency sensitivity, a goal of using 10 W lasers and 1 m diameter telescopes might be chosen, along with a reduction in the antenna arm length to 50 000 km. The resulting factor of about $10^7$ increase in the received power would reduce the shot noise contribution to measuring changes in the arm length difference from about $1 \times 10^{-11}$ m/rHz to $3 \times 10^{-15}$ m/rHz, but the signal strength for frequencies below about 1 Hz would be reduced also because of the reduced arm lengths. If other errors such as from beam pointing jitter and from the phase measurements only double the distance measurement error, the level of the LISA threshold sensitivity curve can be reduced by a factor ranging from about 20 at 10 mHz to 3000 at 1 Hz, and then retaining that value at higher frequencies. However, with the same assumptions about the confusion noise level due to extragalactic CWDBs as used for figures 10.8–10.11, and plausible estimates for extragalactic neutron star and black hole binaries, the overall sensitivity will be limited more by the confusion noise than the instrumental noise up to at least 100 mHz.

One strategy that is not planned for LISA but probably would be used in a later high-frequency mission with shorter arm lengths is to apply enough force to each test mass to keep the rates of change of the arm lengths constant. This is because the force required is about a factor 6000 less for 50 000 km arms, or roughly $4 \times 10^{-10}$ m s$^{-2}$. Also, the main emphasis is on the noise level above 3 mHz, where keeping fluctuations in the applied voltages low is easier than at lower frequencies. This would make the phase measurements very much simpler than would be the case with substantial Doppler shifts.

Quite a few types of measurements could be improved considerably with the above sensitivity. For example, gradual coalescences of $10M_\odot$ black holes with MBHs with masses down to $10^3 M_\odot$ could be observed virtually anywhere in the universe, if they occur. And tests of general relativity for coalescences with highly unequal masses would have considerably higher S/N ratios, as well as large numbers of cycles. In addition, sensitivity to events involved in the initial formation of intermediate mass black holes would be improved.

For improved measurements at low frequencies, below about 0.1 mHz, the most important goals are reduced noise levels from the free mass sensors and longer arm lengths. Moderate goals might be a factor ten reduction in the free mass sensor noise level plus a factor three increase in the arm length. Basically the same geometry could be used as for LISA, but the antenna probably would have to be considerably farther from the Earth to keep the arm length changes from being too large. This antenna would make it possible to observe MBH–MBH binaries with both masses above about $10^5 M_\odot$ much longer before coalescence, and thus increase the number of such events observed by a large factor.

A much more challenging goal would be to try for roughly 1 AU arm lengths and a factor 100 reduction in the free mass sensor noise. This would permit improved measurements mainly at frequencies below about 0.02 mHz, assuming that the confusion noise level is nearly constant below this frequency, as given in Hils, Bender and Webbink [39]. This would further push back the time before coalescence when MBH–MBH binaries could be observed, and probably tell us considerably more about such coalescences after mergers during the process of galaxy formation.

One possible geometry would be to locate spacecraft near the $L_4$ and $L_5$ points of the Earth–Sun system, 60° in front of and behind the Earth, and near either the $L_1$ or $L_2$ point, within about 1500 000 km of the Earth. In this case the third arm would be about 30.5 times the length of the other two. Because the $L_1$ and $L_2$ points are unstable, putting the third spacecraft near the $L_3$ point on the opposite side of the Sun from the Earth and various other possibilities would be considered. One disadvantage of such geometries is that the antenna is in the ecliptic plane, and the sensitivity to the ecliptic latitude of sources at low latitudes is reduced. However, the extra impulse required to go considerably out of the ecliptic is large.

It is interesting to note that the requirement on the accuracy of distance measurements between the test masses for frequencies below 0.1 mHz are much less severe than those for LISA, even for roughly 1 AU arm lengths. The distance measurement error would have to be about $1 \times 10^{-8}$ m/rHz to equal the limitation due to the confusion noise at 0.1 mHz, and more than a factor $(0.1 \text{ mHz}/f)^2$ larger at a lower frequency $f$. Thus the required combination of laser power and telescope size will be affected more by considerations such as staying sufficiently above possible noise in the scattered light level rather than the desire to keep the shot noise level low. In any case, 10 W of laser power and 1 m diameter telescopes would at least be adequate, since it would give the same received power as for LISA.

Achieving a factor 100 lower free mass sensor noise than for LISA would be a major challenge. It might be desirable, for example, to consider having each free mass sensor in a small separate slave spacecraft containing as little other equipment as possible in order to minimize disturbances. Several such slave
spacecraft could be included if they could be made much smaller than the main
craft. With the increased tolerance on the distance measurement uncertainty
below 0.1 mHz, a small local diode laser system could track changes in the
internal geometry of each combination of a spacecraft plus its slaves. However,
this approach would help only if the main noise sources in each free mass sensor
were independent. This would not be the case if common fluctuations in the solar
intensity contributed substantially to the noise.

In view of the apparent advantages of improving the sensitivity at both high
and low frequencies after the LISA mission, it may be desirable to consider flying
two separate antennae, with one at least somewhat like the high frequency antenna
discussed earlier and the other like the moderately improved low frequency
antenna. However, questions such as this can be addressed much more easily
after some data from LISA have been received.

Acknowledgments

It is a pleasure to thank all of the people who have been involved in the various
studies of LISA, including in particular, Karsten Danzmann, Jim Hough, Bernard
Schutz, Stefano Vitale and the rest of our European colleagues. In the USA, Rai
Weiss and Ron Drever and later Mark Vincent were very helpful in getting the
discussions started; Kip Thorne and Jan Hall have provided constant information
and encouragement throughout; and Bill Folkner, Sterl Phinney, Doug Richstone,
Ron Webbink, John Armstrong, Ron Hellings and other members of the LISA
Mission Definition Team have been essential to the progress that has been
achieved. The 1999–2000 ESA Industrial Study has added a number of other
people who have made important contributions to future progress. At JILA,
essential ideas and long-term creative efforts from Tuck Stebbins, Dieter Hils
and Jim Faller have been responsible for much of our present understanding of
the opportunity presented by the LISA mission.

References

Kilometric Optical Arrays in Space, Cargese (Corsica), ESA Report SP-226
pp 157–63
Max-Planck Institute for Quantum Optics) pp 1–191
(AIP Conf. Proc. 456) ed W M Folkner (Woodbury, NY: American Institute of
Physics) pp 11–16
(AIP Conf. Proc. 456) ed W M Folkner (Woodbury, NY: American Institute of
Physics) pp 17–23
(Woodbury, NY: American Institute of Physics) pp 31–38
[16] Staff of Space Science Department, Johns Hopkins University Applied Physics
Laboratory and Staff of Guidance and Control Laboratory, Stanford University
1974 AIAA J. Spacecraft 11 637–44
[17] Lange B 1964 AIAA J. 2 1500–606
See also section 3.3 of LISA 1996 Pre-Phase A Report MPQ-208, 1st edn, ed
K V Danzmann (Garching: Max-Planck Institute for Quantum Optics) pp 60–75
[19] Speake C C 1996 Class. Quantum Grav. 13 A291
Institute of Physics) pp 172–7
Proc. 2nd Int. LISA Symp. (AIP Conf. Proc. 456) ed W M Folkner (Woodbury,
NY: American Institute of Physics) pp 236–9
University of California, San Diego, CA
Proc. 2nd Int. LISA Symp. (AIP Conf. Proc. 456) ed W M Folkner (Woodbury,
NY: American Institute of Physics) pp 207–13
[24] Thorne K 1987 300 Years of Gravitation ed S W Hawking and W Israel (Cambridge:
[28] Iben I 1984 Paper Presented at Workshop on Grav. Waves and Axions (Santa Barbara,
CA: Institute for Theoretical Physics)
References
