LIGO-P020018-00-E
Article for the proceedings of the
Eighth International Workshop on Advanced Computing and Analysis Techniques in Physics Research
Moscow, June 24-28, 2002
To be published in a special issue of Nuclear Instruments and Methods in Physics Research, Section A

LIGO Data Analysis

Peter S. Shawhan [a]*

[a]LIGO Laboratory, California Institute of Technology,
1200 E. California Blvd., Pasadena, CA 91125, USA

Gravitational waves promise to provide new information about massive astrophysical objects in the universe. Techological advances and engineering experience have finally made it feasible to construct detectors with sufficient sensitivity to detect the extremely weak waves which are believed to reach Earth. The Laser Interferometer Gravitational-Wave Observatory (LIGO) project has constructed two "observatories" in the United States which are poised to begin collecting scientifically interesting data. The LIGO Data Analysis System (LDAS) has been designed to support various types of scientific analysis to be performed using this data; its components include an interface to the raw data archive, a "Beowulf" cluster of PCs for parallel processing, and a database to store data analysis products. Attention has also been paid to the deployment of client interface programs and utility software to scientists throughout the LIGO Scientific Collaboration.

1. INTRODUCTION

LIGO, along with similar projects in Europe and Japan, represents a new branch in the field of observational astrophysics. These projects are designed to detect gravitational radiation, initially to confirm its existence and verify its properties, and ultimately to provide new information about particular astrophysical systems. By directly sampling the strong-field gravity which drives the most violent processes in the universe, gravitational-wave observations will complement observations made using electromagnetic waves and high-energy particles.

2. GRAVITATIONAL WAVES

According to Einstein's general theory of relativity, gravity is not a "force" like electromagnetism, but is a direct manifestation of the geometry of the four-dimensional space-time in which we live. Massive objects deform space-time, and gravitational attraction is simply the result of objects following locally "straight" trajectories in the inherently curved coordinate system. General relativity has been highly successful at explaining several small but measurable effects related to space and time, including the precession of the perihelion of Mercury, the bending of light rays by the sun, and the frequency shift of photons in a gravitational potential well [1].

In addition to the static deformation of space-time by massive objects, general relativity predicts that mass undergoing rapid acceleration can radiate *gravitational waves*, oscillations in the geometry of space which propagate away from the source at the speed of light. Unlike the dipole waves of electromagnetism, gravitational waves are quadrupolar, stretching space (and all objects in it) in one direction while shrinking it in the perpendicular direction, as shown in Figure 1. The

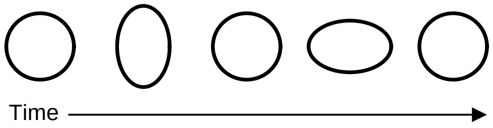

Figure 1. Distortions of a circular object caused by a gravitational wave propagating along the axis of the circle.

---

*Representing the LIGO Scientific Collaboration.

amplitude of a gravitational wave is expressed as a dimensionless strain, *i.e.* the amount of stretching is proportional to the size of the object, and is inversely proportional to the distance from the source.

3. SOURCES OF GRAVITATIONAL WAVES

Significant gravitational radiation can be produced only by systems containing a great deal of mass moving at a significant fraction of the speed of light. Specifically, gravitational waves will be emitted if the quadrupole moment of the mass distribution varies with time. All human activities, even nuclear explosions, produce utterly negligible gravitational radiation. However, certain astrophysical systems are expected to generate signals within the reach of realistic detectors.

One type of source which has received much attention is a compact binary system consisting of two neutron stars, two black holes, or one of each. Such a system emits gravitational waves with a frequency twice the orbital frequency, and these waves carry away angular momentum, causing the separation distance and orbital period to shrink. The rate of orbital decay is very slow at first, but accelerates as the two bodies come together and finally merge into a single object. This "inspiral" process produces a distinctive gravitational waveform that may be described as a "chirp", with frequency and amplitude rising over time to the point of coalescence. The exact waveform is accurately known for binary stars and low-mass black holes, and somewhat less well-known for higher-mass systems in which higher-order relativistic effects become significant. Although the density of compact binary systems in the universe is not accurately known, we are fortunate to have direct evidence of the existence of a few binary neutron star systems from radio observations. In particular, the gradual orbital decay of the binary pulsar PSR 1913+16 has been tracked over many years and exactly matches the expected rate due to gravitational radiation [2], providing excellent additional evidence for the correctness of general relativity.

Supernovae are another potential source of detectable gravitational waves. Due to our limited knowledge of supernova dynamics, the degree of asymmetry of the explosions—and thus the strength of the emitted gravitational waves—is unknown, as are the waveforms. Nevertheless, a sufficiently close supernova is likely to produce a detectable transient signal. There is also the possibility of establishing a coincidence with a gamma-ray burst or an event in an underground neutrino detector.

A rapidly-spinning neutron star will emit gravitational waves if it is slightly non-axisymmetric, and a weak signal may be detected by integrating over a period of weeks or months (with appropriate Doppler and antenna-pattern corrections due to the motion of the Earth). Whether a neutron star can have a significant non-axisymmetry is currently an open question, since it depends on the properties of the crust material. It is straightforward to analyze the data to search for signals from pulsars whose positions and rotation frequencies are known from radio observations; on the other hand, an all-sky, all-frequency search poses a formidable computational challenge.

Other potential sources of detectable gravitational waves include stellar collapse, "ringing" oscillations of a newly-formed black hole, and stochastic gravitational radiation from the motion of matter in the early universe (which would be detected as correlated "noise" in independent detectors). Perhaps more importantly, there may be unanticipated sources waiting to be discovered by the new science of gravitational-wave astronomy.

4. THE DETECTION CHALLENGE

Gravitational waves have not yet been directly detected, but the success of the general theory of relativity and the data from PSR 1913+16 give us great confidence that they do, in fact, exist. Although our understanding of the various potential sources is far from complete, it seems that all of them are either rare (*e.g.* in the case of binary coalescences and supernovae) or intrinsically weak (in the case of spinning neutron stars and stochastic radiation from the early universe, in

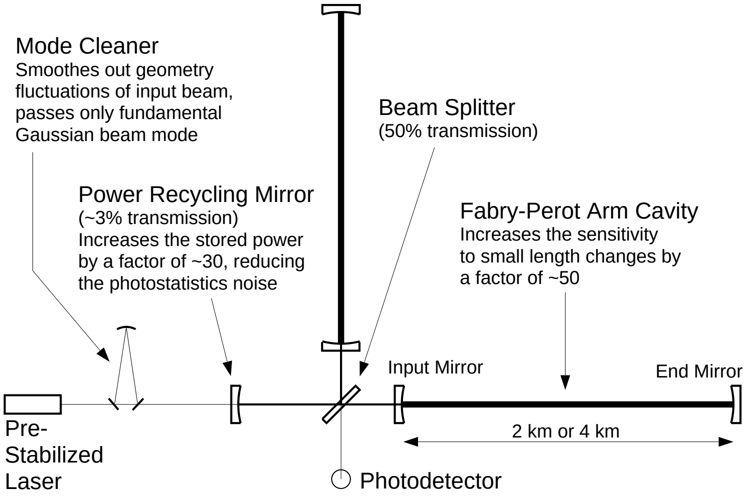

Figure 2. Simplified optical layout of a LIGO interferometer.

most models). Thus, a highly sensitive detector is needed to reach far into the universe and to be able to pick weak signals out of the noise over the finite lifetime of the experiment. In fact, a typical value of the strain amplitude of gravitational waves reaching the Earth is 10^{-21} ! Amazingly, there is reason to believe that such a small strain will be detectable in the near future.

5. DETECTORS

Up to now, the most sensitive gravitational-wave detectors have been based on the resonant "bar" design pioneered by Joseph Weber in the 1960s. A gravitational wave with the proper frequency and polarization induces the bar to ring at its resonant frequency; a transducer at one end of the bar converts this to an electrical signal. Modern bar detectors operate at cryogenic temperatures to minimize thermal noise and achieve high sensitivity over a very narrow frequency band.

The idea of a detector based on a large *interferometer*, which would be sensitive over a wide frequency range, was developed beginning in the 1970s. The basic design is that of a Michelson interferometer, with a beam splitter and two long perpendicular arms. A gravitational wave lengthens one arm while shortening the other, resulting in a phase difference as the light returns to the beam splitter, and thus a signal at the photodetector. A number of prototypes were built and operated over the following decades, allowing many of the technical challenges to be discovered and addressed, while industrial progress has resulted in the availability of stable, high-power lasers and low-loss optical components and coatings. These advances and experience finally made it feasible to build detectors capable of detecting the weak gravitational waves expected to reach Earth, and have led to the large-scale detectors now being commissioned and operated in the U.S., Europe and Japan. Figure 2 shows the major optical components of the design used by LIGO, which augments the Michelson layout with three additional semi-transparent mirrors, forming coupled optical "cavities" which increase the power stored in the interferometer and enhance its sensitivity to displacements of its end mirrors.

The new large-scale detectors have required sophisticated design and engineering to deal with fundamental noise sources such as ground vibrations, and complex servo control systems necessary to operate at maximum sensitivity without introducing electronic noise. It is anticipated that a few years of commissioning and tuning will be needed before the detectors can be operated at their design sensitivity.

6. THE LIGO PROJECT

The Laser Interferometer Gravitational-wave Observatory (LIGO) project [4] is funded by the U.S. National Science Foundation and headed by the California Institute of Technology (Caltech) and the Massachusetts Institute of Technology (MIT), which make up the "LIGO Laboratory" [3]. Two observatory facilities have been constructed, one in Washington state on the Hanford Nuclear Reservation and the other in Livingston Parish, Louisiana. Each observatory has two perpendicular evacuated "beam tubes", 4 km long, containing the light in the interferometer arms. At Hanford, two independent detectors coexist in the beam tubes, one with 4-km arms and the other with 2-km arms. The LIGO detectors are sensitive to gravitational waves with frequencies between about 50 Hz and 4 kHz.

After several years of facility construction, detector installation, and commissioning, LIGO is now making a transition to scientific operations. The first "Science Run" collected data from August 23 to September 9, 2002; this data is expected to yield interesting scientific results, even though the sensitivities of the interferometers were a few orders of magnitude away from the ultimate design goal.

Analysis of the data collected by LIGO is the responsibility of the LIGO Scientific Collaboration, a group of over 300 scientists from over 30 institutions. At present, working groups have been organized for four of the most promising types of sources, each requiring distinct data analysis algorithms and procedures. An additional working group focuses on characterizing the performance of the detectors.

7. LIGO DATA ANALYSIS

A gravitational wave shows up in a LIGO detector as a difference in arm lengths. The servo signal which measures this is sampled at 16384 Hz, synchonized to GPS time signals. The data stream also includes hundreds of auxiliary channels to record various signals from the interferometer control systems as well as environmental sensors (seismometers, magnetometers, etc.), sampled at various rates. All of the channels are recorded continuously; there is no "trigger". The total data rate from each detector is roughly 3 MB/s, and the central data archive at Caltech is designed to accept 100−200 TB/year. Data is written using the VIRGO "frame" format, which has been adopted by all gravitational-wave experiments to facilitate the exchange of data.

The different scientific topics to be addressed by LIGO involve different data access models: searches for transient signals (of known or unknown waveform) involve optimal filtering or time-frequency analysis, while searches for periodic signals require integrating over very long time periods, and searches for stochastic signals require cross-correlating data from different detectors. A customized computing environment called the LIGO Data Analysis System (LDAS) has been created to support these and other usage models.

The LDAS architecture may be likened to a "computing center", with dedicated hardware and a software infrastructure providing specific services, into which individual users plug their analysis-specific code. The main services include: a data repository for raw and reduced data sets; a "data conditioning" server which performs various operations appropriate for gravitational-wave data (filtering, resampling, removal of coherent "line" noise, calculation of power spectra, etc.); a "Beowulf" cluster of many PCs running Linux for parallel processing; a relational database to store results of analyses; data input/output functions; system control and monitoring; and a "manager" which controls all of the other components. These components are independent unix processes running on several different machines, with socket-based flow control and data transmission. They are written in a combination of Tcl (for job control, interprocess communication, and high-level operations) and C++ functions (for CPU-intensive operations) called from the Tcl layer.

Parallel processing on the Beowulf cluster uses the LAM implementation of the MPI protocol. The parallelization scheme is under the complete control of the scientific user, who provides analysis-specific code in the form of a shared-object library, with a standardized set of entry points, which is loaded at run time. Typically, the data (which LDAS delivers to the master node at the beginning of the job) is broadcast to many other nodes, each of which processes it differently—for example, using optimal filters for different waveforms. Results are collected on the master node, then passed back to the LDAS for insertion into the relational database or other post-processing.

LDAS operates as a batch system. To submit a job, a user communicates directly only with the manager, sending the job specification via a socket-based protocol which requires a username and password but does not provide access to a unix shell. Output from the job, if any, is retrieved from a web server. All other LDAS components generally run on machines on a private network, inaccessible from the Internet. At present, LDAS clusters have been set up at Caltech, MIT, the LIGO observatory sites, and a few LSC institutions.

Design and implementation of LDAS has involved balancing functionality and sophistication against practical issues of schedule and cost, in order to be ready to support the analysis of the data which is now being collected. Essentially all of the basic functionality is in place, but the software is expected to evolve substantially in response to the emerging needs of scientific users.

8. CLIENT AND UTILITY SOFTWARE

The LIGO Scientific Collaboration is a large and diverse group, and the wide geographical distribution presents a special challenge for sharing software tools and expertise. To address this problem, we have developed a software suite and software deployment system called "LIGOtools" which avoids the usual burdens of software installation and maintenance. LIGOtools can be set up in any directory and automatically installs precompiled binaries for the main platforms used by LIGO scientists. Thereafter, a single command, "`ligotools_update`", checks for new or updated software packages and automatically installs them. There is also a central web site with documentation, FAQs, and useful links related to the software.

Currently, LIGOtools includes several packages supporting interaction with LDAS and with LIGO data products, including graphical, command-line, and library interfaces. It also includes the LIGO Algorithm Library and various other software tools that have been found to be useful, in particular for the post-analysis that follows the execution of batch jobs using LDAS.

9. CONCLUSION

LIGO and similar gravitational-wave projects are now entering a phase of scientific operations. A customized computing environment has been developed to serve the needs of scientific analyses using LIGO data and has already had some success with the first scientific data set, but it surely will evolve over time as data analysis procedures mature and experience reveals new requirements. We are attempting to address the higher-level challenges of communication among a large group of scientific users, to ensure the efficient extraction of scientific results from the gravitational-wave data.

10. ACKNOWLEDGEMENTS

LIGO is funded by the U.S. National Science Foundation under Cooperative Agreements PHY-9210038 and PHY-0107417.

REFERENCES

1. C. M. Will, *Was Einstein Right? Putting General Relativity to the Test*, 2nd ed., Basic Books, New York, 1993.
2. G. Taylor and J. M. Weisberg, Astrophys. J. 345 (1989) 434.
3. `http://www.ligo.caltech.edu`
4. B. C. Barish and R. Weiss, Physics Today (Oct. 1999) 44.