

CONTENTS

Preface	ix
1 Introduction	1
2 The Einstein Equivalence Principle	11
2.1 The Dicke Framework	13
2.2 The Einstein Equivalence Principle	16
2.3 Experimental Tests of the Einstein Equivalence Principle	18
2.3.1 Tests of the Weak Equivalence Principle	18
2.3.2 Tests of Local Lorentz Invariance	23
2.3.3 Tests of Local Position Invariance	26
2.4 Schiff's Conjecture	33
2.5 The $TH\epsilon\mu$ formalism	38
2.5.1 Einstein Equivalence Principle in the $TH\epsilon\mu$ formalism	39
2.5.2 Proof of Schiff's conjecture	42
2.5.3 Energy conservation and anomalous mass tensors	45
2.5.4 Quantum systems and EEP	46
2.5.5 Local Lorentz Invariance and the c^2 formalism	49
2.5.6 Application to tests of EEP	49
2.6 The Standard Model Extension	52
2.7 EEP, particle physics, and the search for new interactions	54
3 Gravitation as a geometric phenomenon	57
3.1 Universal coupling	57
3.2 Nongravitational physics in curved spacetime	58
3.2.1 Charged point-particle dynamics	58
3.2.2 Geodesics	62
3.2.3 Actions and energy-momentum tensors	64
3.2.4 Hydrodynamics	66
3.3 Metric theories of gravity and the Strong Equivalence Principle	68
4 The parametrized post-Newtonian formalism	73
4.1 The post-Newtonian approximation	74
4.1.1 Newtonian gravity and the Newtonian limit	74
4.1.2 Post-Newtonian bookkeeping	75
4.1.3 Isolated post-Newtonian systems	77
4.2 Building the PPN formalism	79
4.2.1 Post-Newtonian potentials	79
4.2.2 The standard PPN gauge	82
4.3 Lorentz transformations and the PPN metric	84
4.4 Global conservation laws	89
4.5 Other post-Newtonian gauges	95

5	Metric theories of gravity and their post-Newtonian limits	99
5.1	Method of calculation	101
5.2	General relativity	103
5.2.1	Statement of the theory	103
5.2.2	The post-Newtonian limit and the PPN parameters	104
5.3	Scalar-tensor theories	106
5.3.1	Statement of the theory	106
5.3.2	Post-Newtonian limit and PPN parameters	107
5.3.3	General remarks	109
5.3.4	$f(R)$ theories	111
5.4	Vector-tensor theories	111
5.4.1	Constrained Theories	111
5.4.2	Unconstrained Theories	113
5.5	Tensor-vector-scalar (TeVeS) theories	114
5.6	Quadratic gravity and Chern-Simons theories	115
5.7	Massive gravity	116
5.8	The rise and fall of alternative theories of gravity	117
5.8.1	Whitehead's theory	117
5.8.2	Rosen's bimetric theory	119
5.8.3	Other theories	119
6	Equations of motion in the PPN formalism	121
6.1	Equations of motion for photons	122
6.2	PPN hydrodynamics	124
6.3	Equations of motion for massive bodies	125
6.4	Two-body systems	132
6.4.1	Effective one-body equation of motion	132
6.4.2	Perturbed two-body orbits and the Lagrange planetary equations	134
6.5	Semiconservative theories and N -body Lagrangians	136
6.6	The locally measured gravitational constant	138
6.7	Equations of motion for spinning bodies	141
7	The classical tests	147
7.1	Deflection of light	148
7.1.1	Tests of the deflection of light	151
7.1.2	Gravitational lenses	154
7.2	The Shapiro time delay	155
7.3	The perihelion advance of Mercury	156
8	Tests of the Strong Equivalence Principle	161
8.1	The Nordtvedt effect	162
8.2	Preferred frames and locations: Orbits	168
8.3	Preferred frames and locations: Structure of massive bodies	172
8.3.1	G_L and distortions of massive bodies	172
8.3.2	Precession of the spin axis of massive bodies	174
8.4	Preferred-frames and locations: Bounds on the PPN parameters	175
8.5	Constancy of Newton's gravitational constant	179
9	Other tests of post-Newtonian gravity	181
9.1	Testing the effects of spin	182
9.1.1	Effects of spin in two-body systems	182
9.1.2	Tests of spin precession	184
9.1.3	Tests of spin effects on orbits	189
9.2	De Sitter precession	191

9.3	Tests of post-Newtonian conservation laws	192
10	Structure and motion of compact objects	195
10.1	Structure of neutron stars	196
10.2	Structure of black holes	201
10.3	The motion of compact objects	204
10.3.1	A modified EIH formalism	207
10.3.2	Scalar-tensor theories	211
10.3.3	Constrained vector-tensor theories	215
11	Gravitational radiation	219
11.1	The problem of motion and radiation	219
11.2	Gravitational wave detectors	223
11.3	Speed of gravitational waves	224
11.4	Polarization of gravitational waves	227
11.4.1	The $E(2)$ framework	227
11.4.2	$E(2)$ class of specific metric theories	230
11.4.3	Measurement of gravitational-wave polarizations	232
11.5	Generation of gravitational waves	236
11.5.1	Calculating gravitational waves in alternative theories	238
11.5.2	General relativity	246
11.5.3	Scalar-tensor theories	250
11.5.4	Other theories	253
12	Strong-field and dynamical tests of relativistic gravity	255
12.1	Binary pulsars	255
12.1.1	Arrival-time analysis for binary pulsars	258
12.1.2	A zoo of binary pulsars	264
12.1.3	Tests of alternative theories	269
12.2	Inspiralling compact binaries and gravitational waves	272
12.2.1	Compact binary inspiral in general relativity	273
12.2.2	Compact binary inspiral in scalar-tensor theories	276
12.2.3	Compact binary inspiral in other theories	278
12.3	Exploring spacetime near compact objects	281
12.3.1	SgrA*: A black hole in the galactic center	281
12.3.2	Neutron stars and black holes	285
12.4	Cosmological tests	285
	Index	327

PREFACE

The year 2015 marked the 100th anniversary of the publication of Einstein’s general theory of relativity, and relativists worldwide celebrated this historic occasion. As if this were not enough, on September 14, 2015, scientists at the LIGO gravitational-wave observatories in the US detected, for the first time, gravitational waves passing the Earth, emitted by a pair of merging black holes over a billion light years away. This event provided a kind of fairy-tale capstone to a remarkable century.

Indeed, some popular accounts of the history of general relativity read like a fairy tale, going something like this: in 1905, Einstein discovered special relativity. He then turned his attention to general relativity and after 10 years of hard work, he got general relativity in November 1915. In 1919, Eddington verified the theory by measuring the bending of starlight during a solar eclipse. Einstein became famous. And everybody lived happily ever after.

The real history of general relativity is rather more complex. At the time of Eddington’s measurements of light bending, there was considerable skepticism about the results. There were major conceptual difficulties with the theory; it was very hard to understand what this new theory was and what it really predicted. And finally, there was an abiding sense that the theory mainly predicted some very tiny corrections to Newtonian gravity, and that it really wasn’t all that important for physics.

As a result, within about 10 years of its development, general relativity entered a period of decline, dubbed the “low-water mark” by Jean Eisenstaedt (2006), so that by the end of the 1950s, general relativity was considered to be in the backwaters of physics and astronomy, not a fit subject for a serious scientist to pursue.

But during the 1960s there began a remarkable renaissance for the theory. This was driven in part by the discovery of quasars, pulsars and the cosmic background radiation, systems where it became clear that general relativity would play a central role. It was also fueled by the beginnings of a world-wide effort to put the theory to the test using new precision tools such as atomic clocks and radio telescopes, together with the emerging space program. And gravitational theorists developed a variety of tools that allowed them to clarify what the theory and its competitors predicted, analyze the new experimental results and devise new tests.

Today, general relativity is fully integrated into the mainstream of physics, and in fact is central to some of the key scientific questions of today, such as How did the universe begin and what is its future? What governs physics at the shortest distances and the longest distances? Do black holes really exist and how do they affect their surroundings? How can we reconcile gravity and quantum mechanics?

Yet, is it the correct theory of gravity? Was Einstein really right?

By the time of the centenary of general relativity, Einstein’s theory had been tested in many ways and to high precision, and had passed every test. So far, no experiment has been able to put an unambiguous dent in the armor of general relativity. And yet “experimental gravity” is as active and exciting a field as it was in 1981, when the first edition of this book came out. This is motivated in part by the ongoing mysteries and conundra associated with the acceleration of the universe, the apparent existence of dark matter, and the difficulty of marrying general relativity with quantum mechanics.

But it is also motivated by our newfound ability to explore regimes for testing Einstein's theory far beyond the relatively weak and benign realms of the laboratory and the solar system. This exploration began, of course, with the discovery of the binary pulsar in 1974, leading to the first tests involving neutron stars, but in recent years it has accelerated dramatically.

As we look toward the second century of general relativity, two important themes are going to be (i) testing general relativity in the strong-gravity regime near black holes and neutron stars, going beyond the weak-gravity conditions of the solar system; and (ii) testing general relativity in the highly dynamical regime, where gravitational radiation is both a phenomenon to be scrutinized and a tool for studying dynamical, strong-gravity sources.

There is no better illustration of this new era for testing general relativity than the outpouring of papers following the first detections of gravitational waves by the LIGO-Virgo collaboration, showing how the data place new and compelling constraints on a wide range of alternative theories, in ways that would not have been possible using solar-system measurements. It is my hope that this book will serve not only as an update of the 1981 edition, but also as a foundation for students and researchers who wish to join in this new effort to check whether Einstein was right.

Acknowledgments

Over my 50-year long adventure in the world of testing general relativity, it has been my great fortune to have had three inspiring mentors: Kip Thorne, Subrahmanyan Chandrasekhar and Bob Wagoner. Kip assigned me the PhD project of figuring out what had to be done to test general relativity better. He also taught by example the ways of research, teaching, scientific writing and presenting science to broad audiences. From Chandra I learned a style of meticulous attention to detail in calculations that has been a hallmark of my own research. He also gave me a deep appreciation for science history. Bob's infectious enthusiasm for physics and astronomy taught me that, if you aren't having fun doing this work, you're in the wrong business. I have always had fun!

I also want to acknowledge some of the colleagues, collaborators and students who had an impact directly or indirectly on my career in experimental gravity, and who thus, in one way or another, left a mark on this book: Eric Adelberger, John Anderson, K. G. Arun, Emanuele Berti, Bruno Bertotti, Luc Blanchet, Alessandra Buonanno, Thibault Damour, Robert Dicke, Douglas Eardley, Francis Everitt, William Fairbank, Mark Haugan, Timothy Krisher, Bala Iyer, Larry Kidder, Ryan Lang, David Lee, Alan Lightman, Craig Lincoln, Saeed Mirshekari, Wei-Tou Ni, Anna Nobili, Ken Nordtvedt, Ho-Jung Paik, Michael Pati, Eric Poisson, Bob Reasenberg, Peter Saulson, Bernard Schutz, Irwin Shapiro, Adam Stavridis, Joe Taylor, Robert Vessot, Martin Walker, Joseph Weber, Alan Wiseman, Nicolás Yunes, and Helmut Zaglauer.

I am grateful to the Institut d'Astrophysique de Paris for its hospitality during annual stays in 2015, 2016 and 2017, where large parts of this book were written. This work was supported in part by the US National Science Foundation under Grants No. PHY 13-06069 and PHY 16-00188.

1

INTRODUCTION

On September 14, 1959, 12 days after passing through her point of closest approach to the Earth, the planet Venus was bombarded by pulses of radio waves sent from Earth. A handful of anxious scientists at Lincoln Laboratories in Massachusetts waited to detect the echo of the reflected waves. To their initial disappointment, neither the data from this day, nor from any of the days during that month-long observation, showed any detectable echo near inferior conjunction of Venus. However, a later, improved reanalysis of the data showed a bona fide echo in the data from one day: September 14. Thus occurred the first recorded radar echo from a planet.

Exactly 56 years later, on September 14, 2015, a rather different signal was received by scientists, this time in Hanford, Washington and Livingston, Louisiana. The signal was not electromagnetic, but instead was a wave in the fabric of space-time itself. It was the final burst of gravitational waves from two black holes that merged to form a single black hole somewhere in the southern sky some 1.3 billion years ago. The signal was recognized within minutes by automated data processing software. This time, the scientists, numbering over 1000, were anxious lest the signal be an unlucky artifact of instrumental noise. But on February 11, 2016, after an intensive and secretive five months of detailed analysis, checks and cross-checks, they announced at a Washington DC press conference that LIGO, the Laser Interferometer Gravitational-wave Observatory had made the first direct detection of gravitational waves.

As if the 100th anniversary of the general theory of relativity during 2015 was not already something to celebrate, the detection of gravitational waves was icing on the cake.¹ It was also the capstone of a half-century period during which general relativity experienced a remarkable renaissance, from a subject relegated to the backwaters of physics and astronomy, to one that is regarded as central to the major scientific questions of the day, from the nature of the fundamental particles to the fate of the universe. It was a perfect illustration of how the field was transformed from one that was once called “a theorist’s paradise and an experimentalist’s purgatory”, to one in which experimentalists and theorists work hand in hand. It highlighted the field’s evolution from the world of “small science”, where individuals

¹And a scoop of ice cream on top of the icing was the award of the 2017 Nobel Prize in Physics to three of LIGO’s founders, Rainer Weiss, Barry Barish and Kip Thorne. The other pioneer of the project, Ron Drever, had already died in March of that year.

or small groups scratched out mathematical formulas in their tiny offices, to that of “big science”, where world-wide collaborations of scientists conduct their affairs via telecon and Skype, spend budgets measured in units of megabucks or megaeuros, and require project managers to keep matters on track.

The origins of this remarkable transformation of general relativity from an obscure niche of mathematics and physics to a major subfield of physics and astronomy, today called “Gravitational Physics”, can be found in a set of events of the academic year 1959 - 60, beginning with that first radar echo from Venus. Four key events followed.

On March 9, 1960, the editorial office of *Physical Review Letters* received a paper by Robert Pound and Glen Rebka, Jr., entitled “Apparent Weight of Photons.” The paper reported the first successful laboratory measurement of the gravitational redshift of light. The paper was accepted and published in the April 1 issue.

In June, 1960, there appeared in volume 10 of the *Annals of Physics* a paper on “A Spinor Approach to General Relativity” by Roger Penrose. It outlined a streamlined calculus for general relativity based upon “spinors” rather than upon tensors.

Later that summer, Carl Brans, a young Princeton graduate student working with Robert Dicke, began putting the finishing touches on his Ph.D. thesis, entitled “Mach’s Principle and a Varying Gravitational Constant.” Part of that thesis was devoted to the development of a “scalar-tensor” alternative to the general theory of relativity. Although its authors never referred to it this way, it came to be known as the Brans-Dicke theory.

On September 26, 1960, just over a year after the recorded Venus radar echo, astronomers Thomas Matthews and Allan Sandage and co-workers at Mount Palomar used the 200-inch telescope to make a photographic plate of the star field around the location of the radio source 3C48. Although they expected to find a cluster of galaxies, what they saw at the precise location of the radio source was an object that had a decidedly stellar appearance, an unusual spectrum, and a luminosity that varied on a timescale as short as 15 minutes. The name quasistellar radio source or “quasar” was soon applied to this object and to others like it.

These disparate and seemingly unrelated events of the academic year 1959-60, in fields ranging from experimental physics to abstract theory to astronomy, signaled the beginning of a new era for general relativity. This era was to be one in which general relativity not only would become an important theoretical tool of the astrophysicist, but would have its validity challenged as never before. Yet it was also to be a time in which experimental tools would become available to test the theory in unheard-of ways and to unheard-of levels of precision.

The optical identification of 3C48 (Matthews and Sandage, 1963) and the subsequent discovery of the large redshifts in its spectral lines and in those of 3C273 (Schmidt, 1963; Greenstein and Matthews, 1963) presented theorists with the problem of understanding the enormous output of energy from a region of space compact enough to permit the luminosity to vary systematically over timescales as short as days or hours. Many theorists turned to general relativity and to the strong relativistic gravitational fields it predicts, to provide the mechanism underlying such violent events. This was the first use of the theory’s strong-field aspect, in an attempt to interpret and understand observations. The subsequent discovery of the cosmic microwave background (CMB) radiation in 1964, of pulsars in 1967, and of the first black hole candidate in 1971 showed that it would not be the last. However, the use of relativistic gravitation in astrophysical model building forced theorists and experimentalists to address the question: Is general relativity the correct relativistic theory of gravitation? It would be difficult to place much confidence in models for such phenomena as quasars and pulsars if there were serious doubt about one of the basic underlying physical theories. Thus, the growth of “relativis-

tic astrophysics” intensified the need to strengthen the empirical evidence for or against general relativity.

The publication of Penrose’s spinor approach to general relativity (Penrose, 1960) was one of the products of a new school of relativity theorists that came to the fore in the late 1950s. These relativists applied the elegant, abstract techniques of pure mathematics to physical problems in general relativity, and demonstrated that these techniques could also aid in the work of their more astrophysically oriented colleagues. The bridging of the gaps between mathematics and physics and mathematics and astrophysics by such workers as Bondi, Dicke, Sciama, Pirani, Penrose, Sachs, Ehlers, Misner, and others changed the way that research (and teaching) in relativity was carried out, and helped make it an active and exciting field of physics. Yet again the question had to be addressed: Is general relativity the correct basis for this research?

The other three events of 1959-60 contributed to the rebirth of a program to answer that question, a program of experimental gravitation that had been semi-dormant for 40 years.

The Pound-Rebka (1960) experiment, besides verifying the principle of equivalence and the gravitational redshift, demonstrated the powerful use of quantum technology in gravitational experiments of high precision. The next two decades would see further uses of quantum technology in such tools as atomic clocks, laser ranging, superconducting gravimeters and gravitational-wave detectors, to name only a few. Recording radar echos from Venus (Smith, 1963) opened up the solar system as a laboratory for testing relativistic gravity. The rapid development of the interplanetary space program during the early 1960s made radar ranging to both planets and artificial satellites a vital new tool for probing relativistic gravitational effects. Coupled with the theoretical discovery in 1964 of the relativistic time-delay effect (Shapiro, 1964), it provided new and accurate tests of general relativity. For the next decade and a half, until the summer of 1974, the solar system would be the primary arena for high-precision tests of general relativity. Finally, the development of the Brans-Dicke (1961) theory provided a viable alternative to general relativity. Its very existence and agreement with the experimental results of the day demonstrated that general relativity was not a unique theory of gravity. Some even preferred it over general relativity on aesthetic and theoretical grounds. At the very least, it showed that discussions of experimental tests of relativistic gravitational effects should be carried on using a broader theoretical framework than that provided by general relativity alone. It also heightened the need for high-precision experiments because it showed that the mere *detection* of a small general relativistic effect was not enough. What was now required was measurements of these effects to accuracies of 10 percent, one percent, or fractions of a percent and better, to distinguish among competing theories of gravitation.

To appreciate more fully the regenerative effect that these events had on gravitational theory and its experimental tests, it is useful to review briefly the history of general relativity in the 45 years following Einstein’s publication of the theory.

In deriving general relativity, Einstein was not particularly motivated by a desire to account for unexplained experimental or observational results. Instead, he was driven by theoretical criteria of elegance and simplicity. His primary goal was to produce a gravitation theory that incorporated the principle of equivalence and special relativity in a natural way. In the end, however, he had to confront the theory with experiment. This confrontation was based on what came to be known as the “three classical tests.”

One of these tests was an immediate success - the ability of the theory to account for the anomalous perihelion shift of Mercury. This had been an unsolved problem in celestial mechanics for over half a century, since the announcement by Urbain Jean Joseph Le Verrier in 1859 that, after the perturbing effects of the planets on

Mercury's orbit had been accounted for, there remained in the data an unexplained advance in the perihelion of Mercury. The modern value for this discrepancy is about 43 arcseconds per century. A number of ad hoc proposals were made in an attempt to account for this excess, including the existence of a new planet, dubbed "Vulcan", near the Sun, and a deviation from the inverse-square law of gravitation. A half century of astronomical searches for Vulcan yielded numerous claimed sightings, but in the end, no solid evidence for the planet was found. And while a change in the Newtonian inverse-square law proposed by Simon Newcombe, from the power 2 to the power 2.000000157, could account for the perihelion advance of Mercury, it ultimately conflicted with data on the motion of the Moon.

Einstein was well aware of the problem of Mercury, and in fact, he used it as a way to test his early attempts at a theory of gravity; for example, he finally rejected the 1912 "Entwurf" or "draft" theory that he had developed with Marcel Grossmann in part because it gave the wrong perihelion advance. But when he thought he had obtained the final theory in November 1915, the fact that it gave the correct advance convinced him that he had succeeded.

The next classical test, the deflection of light by the Sun, was not only a success, it was a sensation. Shortly after the end of World War I, two expeditions organized by Arthur Stanley Eddington set out from England: one for Sobral, in Brazil; and one for the island of Principe off the coast of Africa to observe the solar eclipse of May 29, 1919. Their goal was to measure the deflection of light as predicted by general relativity: 1.75 arcseconds for a ray that grazes the Sun. The observations had to be made in the path of totality of a solar eclipse, during which the Moon would block the light from the Sun and reveal the field of stars around it. Photographic plates taken of the star field during the eclipse were compared with plates of the same field taken when the Sun was not present, and the angular displacement of each star was determined. The results were 1.13 ± 0.07 times the Einstein prediction for the Sobral expedition, and 0.92 ± 0.17 for the Principe expedition (Dyson et al., 1920). The November 1919 announcement of these results confirming the theory caught the attention of a war-weary public and helped make Einstein a celebrity. Nevertheless, the experiments were plagued by systematic errors, and subsequent eclipse expeditions did little to improve the situation.

The third classical test was actually the first proposed by Einstein (1908): the gravitational redshift of light. But by contrast with the other two tests, there was no reliable confirmation of it until the 1960 Pound-Rebka experiment. One possible test involved the red shift of spectral lines from the sun. A 1917 measurement by astronomer Charles St. John (1917) failed to detect the effect, sowing considerable doubt about the validity of the theory. Thirty years of such measurements revealed mainly that the observed shifts in solar spectral lines are dominated by Doppler shifts due to radial mass motions in the solar photosphere, and by line shifts due to the high pressures in the solar atmosphere, making detection of the Einstein shift very difficult. It would be 1962 before a reliable solar redshift measurement would be made. Similarly inconclusive were attempts to measure the gravitational redshift of spectral lines from white dwarfs, primarily from Sirius B and 40 Eridani B, both members of binary systems. Because of uncertainties in the determination of the masses and radii of these stars, and because of possible complications in their spectra due to scattered light from their companions, reliable, precise measurements were not possible. Furthermore, by the late 1950s, it was being suggested that the gravitational red shift was not a true test of general relativity after all. According to Leonard Schiff and Robert Dicke, the gravitational red shift was a consequence purely of the principle of equivalence, and did not test the specific field equations of gravitational theory.

Cosmology was one area where general relativity could conceivably be confronted with observation. Initially the theory met with success in its ability to account for

the observed expansion of the universe, yet by the 1940s there was considerable doubt about its applicability. According to pure general relativity, the expansion of the universe originated in a dense primordial explosion called the “big bang.” However, at that time, the measured value of the expansion rate was so high that working backward in time using the cosmological solutions of general relativity led to the conclusion that the age of the universe was less than that of the Earth! One result of this doubt was the rise in popularity during the 1950s of the steady-state cosmology of Herman Bondi, Thomas Gold, and Fred Hoyle. This model avoided the big bang altogether, and allowed for the expansion of the universe by the continuous creation of matter. But by the late 1950s, revisions in the cosmic distance scale had reduced the expansion rate by a factor of five, and had thereby increased the age of the universe in the big bang model to a more acceptable level. Nevertheless, cosmology was still in its infancy, hardly suitable as an arena for testing theories of gravity. The era of “precision cosmology” would not begin until the launch of the COBE satellite in 1989 followed by its precise measurements of the spectrum and fluctuations of the cosmic background radiation.

Meanwhile, a small “cottage industry” had sprung up, devoted to the construction of alternative theories of gravitation. Some of these theories were produced by such luminaries as Henri Poincaré, Alfred North Whitehead, Edward Arthur Milne, George Birkhoff, Nathan Rosen and Frederick Belinfante. Many of these authors expressed an uneasiness with the notions of general covariance and curved spacetime, which were built into general relativity, and responded by producing “special relativistic” theories of gravitation. Many of these theories considered spacetime itself to be governed by special relativity, and treated gravitation as a field on that background. As of 1960, it was possible to enumerate at least 25 such alternative theories, as found in the primary research literature between 1905 and 1960; for a partial list, see Whitrow and Morduch (1965).

Thus, by 1960, it could be argued that the validity of general relativity rested on the following empirical foundation: one test of moderate precision (the perihelion shift, approximately 10 percent), one test of low precision (the deflection of light, approximately 25 percent), one inconclusive test that was not a real test anyway (the gravitational redshift), and cosmological observations that could not distinguish between general relativity and the steady-state theory. Furthermore, a variety of alternative theories laid claim to viability.

In addition, the attitude toward the theory seemed to be that, whereas it was undoubtedly important as a fundamental theory of nature, its observational contacts were limited. This view was present for example in the standard textbooks on general relativity of this period, such as those by Møller (1952), Synge (1960) and Landau and Lifshitz (1962). As a consequence, general relativity was cut off from the mainstream of physics. It was during this period that one newly minted graduate of the California Institute of Technology was advised not to pursue this subject for his graduate work, because general relativity “had so little connection with the rest of physics and astronomy” (his name: Kip Thorne).

However, the events of 1959-60 changed all that. The pace of research in general relativity and relativistic astrophysics began to quicken and, associated with this renewed effort, the systematic high-precision testing of gravitational theory became an active and challenging field, with many new experimental and theoretical possibilities. These included new versions of old tests, such as the gravitational red shift and deflection of light, with accuracies that were unthinkable before 1960. They also included brand new tests of gravitational theory, such as the gyroscope precession, the time delay of light, and the “Nordtvedt effect” in lunar motion, all discovered theoretically after 1959.

Because many of the experiments involved the resources of programs for interplanetary space exploration and observational astronomy, their cost in terms of

money and manpower was high and their dependence upon increasingly constrained government funding agencies was strong. Thus, it became crucial to have as good a theoretical framework as possible for comparing the relative merits of various experiments, and for proposing new ones that might have been overlooked. Another reason that such a theoretical framework was necessary was to make some sense of the large (and still growing) number of alternative theories of gravitation. Such a framework could be used to classify theories, elucidate their similarities and differences, and compare their predictions with the results of experiments in a systematic way. It would have to be powerful enough to be used to design and assess experimental tests in detail, yet general enough not to be biased in favor of general relativity.

A leading exponent of this viewpoint was Dicke (1964). It led him and others to perform several high-precision null experiments which greatly strengthened the empirical support for the foundations of gravitation theory. Within this viewpoint one asks general questions about the nature of gravity and devises experiments to test them. The most important dividend of the Dicke framework is the understanding that gravitational experiments can be divided into two classes. The first consists of experiments that test the foundations of gravitation theory, one of these foundations being the principle of equivalence. These experiments (Eötvös experiment, Hughes-Drever experiment, gravitational redshift experiment, and others) accurately verify that gravitation is a phenomenon of curved spacetime, that is, it must be described by a “metric theory” of gravity, at least to a high level of precision. General relativity and Brans-Dicke theory are examples of metric theories of gravity.

The second class of experiments consists of those that test metric theories of gravity. Here another theoretical framework was developed that takes up where the Dicke framework leaves off. Known as the “Parametrized post-Newtonian” or PPN formalism, it was pioneered by Kenneth Nordtvedt, Jr. (1968b), and later extended and improved by Will (1971c), Will and Nordtvedt (1972), and Will (1973). The PPN framework takes the slow motion, weak field, or post-Newtonian limit of metric theories of gravity, and characterizes that limit by a set of 10 real-valued parameters. Each metric theory of gravity has particular values for the PPN parameters. The PPN framework was ideally suited to the analysis of solar system gravitational experiments, whose task then became one of measuring the values of the PPN parameters and thereby delineating which theory of gravity is correct. A second powerful use of the PPN framework was in the discovery and analysis of new tests of gravitation theory, examples being the Nordtvedt effect (Nordtvedt, 1968a), preferred-frame effects (Will, 1971b) and preferred-location effects (Will, 1971b, 1973). The Nordtvedt effect, for instance, is a violation of the equality of acceleration of massive bodies, such as the Earth and Moon, in an external field; the effect is absent in general relativity but present in many alternative theories, including the Brans-Dicke theory. The third use of the PPN formalism was in the analysis and classification of alternative metric theories of gravitation. After 1960, the invention of alternative gravitation theories did not abate, but changed character. The crude attempts to derive Lorentz-invariant field theories described previously were mostly abandoned in favor of metric theories of gravity, whose development and motivation were often patterned after that of the Brans-Dicke theory. A “theory of gravitation theories” was developed around the PPN formalism to aid in their systematic study. The PPN formalism thus became the standard theoretical tool for analyzing solar system experiments, looking for new tests, and studying alternative metric theories of gravity.

But by the middle 1970s it became apparent that the solar system could no longer be the sole testing ground for gravitation theories. The reason was that many alternative theories of gravity agreed with general relativity in their weak-

field, slow-motion limits closely enough to pass all solar system tests. But they did not necessarily agree in other predictions, such as neutron stars, black holes, gravitational radiation, or cosmology, phenomena that involved strong or dynamical gravity.

This was confirmed in the summer of 1974 with the discovery by Joseph Taylor and Russell Hulse of the binary pulsar (Hulse and Taylor, 1975). Here was a system that featured, in addition to significant post-Newtonian gravitational effects, highly relativistic gravitational fields associated with the pulsar (and possibly its companion) and the possibility of the emission of gravitational radiation by the binary system. The role of the binary pulsar as a new arena for testing relativistic gravity was confirmed four years later with the announcement (Taylor et al., 1979) that the rate of change of the orbital period of the system had been measured. The result agreed with the prediction of general relativity for the rate of orbital energy loss due to the emission of gravitational radiation. But it disagreed strongly with the predictions of many alternative theories, even some with post-Newtonian limits identical to that of general relativity.

By 1981, when the first edition of this book was published, it was not uncommon to describe the period 1960 – 1980 as a “golden era” for experimental gravity. Many of the events of that period were described for a lay audience in my 1986 book “Was Einstein Right?” (Will, 1986). But the phrase “golden era” suggests that it was downhill from that time forward. Quite the opposite was true.

Solar-system tests of relativistic gravity continued, with highlights including dramatically improved measurements of light deflection and the Shapiro time delay, measurements of “frame-dragging” by the Gravity Probe B and LAGEOS experiments, and steadily improving lunar laser ranging. Binary pulsar tests continued, aided by remarkable discoveries, including the famous “double pulsar” and a pulsar in a triple system.

At the same time, the central thrust of testing gravity began to shift away from the weak-field limit. Two themes began to emerge as the key themes for the future.

The first theme is Dynamical Gravity. This involves phenomena in which the variation with time of the spacetime geometry plays an important role. In the solar system, velocities are small compared to the speed of light and the masses of the planets are small compared to the mass of the sun, so the underlying spacetime geometry can be viewed either as being stationary or as evolving in a quasistationary manner. But in the binary pulsar, for example, the two bodies have almost the same mass and are orbiting each other ten times faster than planets in the solar system, and consequently the varying spacetime geometry that they both generate devolves into gravitational waves propagating away from the system, causing it to lose energy. A more dramatic example is the final inspiral of the two black holes whose gravitational signal was detected by LIGO in 2015. The black holes are made of pure curved spacetime, and the manner in which that geometry evolved during the final fractions of a second of the inspiral and merger left its imprint on the gravitational waves that were detected. The final black hole that was left over even oscillated a few times, emitting a specific kind of gravitational radiation called ringdown waves. This is the regime of dynamical gravity. Dynamical gravity often goes hand in hand with gravitational radiation.

The second theme is Strong Gravity. Much like modern art, the term “strong” means different things to different people. To some one steeped in general relativity, the principal figure of merit that distinguishes strong from weak gravity is the quantity $\epsilon \sim Gm/c^2r$, where m is the characteristic mass scale of the phenomenon, r is the characteristic distance scale, and G and c are the Newtonian gravitational constant and the speed of light, respectively. Near the event horizon of a non-rotating black hole, or for the expanding observable universe, $\epsilon \sim 1$; for neutron stars, $\epsilon \sim 0.2$. These are the regimes of strong gravity. For the solar system,

$\epsilon < 10^{-5}$; this is the regime of weak gravity.

An alternative view of “strong” gravity comes from the world of particle physics. Here the figure of merit is $Gm/c^2r^3 \sim \ell^{-2}$, where the curvature of spacetime associated with the phenomenon, represented by the left-hand-side, is comparable to the inverse square of a favorite length scale ℓ . If ℓ is the Planck length $(\hbar G/c^3)^{1/2} \sim 10^{-35}$ m, this would correspond to the regime where one expects conventional quantum gravity effects to come into play. Another possible scale for ℓ is the TeV scale associated with many models for unification of the forces, or models with extra spacetime dimensions. From this viewpoint, strong gravity is where the radius of curvature of spacetime is comparable to the fundamental length. Weak gravity is where the radius of curvature is much larger than this. The universe at the Planck time is strong gravity. Just outside the event horizon of an astrophysical black hole is weak gravity.

We will adopt the relativist’s view of strong gravity.

The boundary between dynamical gravity and strong gravity is somewhat fuzzy. One can explore strong gravity alone by studying the motion of a star around a static supermassive black hole or of gas around a neutron star. Gravitational waves can be emitted by a binary system of white dwarfs, well characterized by weak gravity. However, the strongest waves tend to come from systems with compact, strongly gravitating bodies, because only such bodies can get close enough together to reach the relativistic speeds required to generate strong gravitational waves. And the universe as a whole can be thought of as both “strong gravity” and dynamical, yet because of the high degree of symmetry, gravitational waves do not play a major role in its evolution. On the other hand primordial gravitational waves could be detectable, in fluctuations of the cosmic background radiation, for example. Regardless of the specific context, testing general relativity in the strong-field and dynamical regimes will dominate this field for some time to come.

As a young student of 17 at the Polytechnical Institute of Zürich, Einstein studied the work of Helmholtz, Maxwell, and Hertz, and ultimately used his deep understanding of electromagnetic theory as a foundation for special and general relativity. He appears to have been especially impressed by Hertz’s confirmation that light and electromagnetic waves are one and the same (Schilpp, 1949). The electromagnetic waves that Hertz studied were in the radio part of the spectrum, at 30 MHz. It is amusing to note that, 60 years later, the “golden age” for testing relativistic gravity began with radio waves, the 440 MHz waves reflected from Venus, and ended with radio waves, the pulsed signals from the binary pulsar, observed at 430 MHz. We are now in a new era for testing general relativity, an era in which we can exploit and study an entirely new kind of wave, a wave in the fabric of spacetime itself.

During the half-century that closed on the centenary of Einstein’s formulation of general relativity, the empirical foundations of his great theory were strengthened as never before. The question then arises, why bother to continue to test it? One reason is that gravity is a fundamental interaction of nature, and as such requires the most solid empirical underpinning we can provide. Another is that all attempts to quantize gravity and to unify it with the other forces suggest that the standard general relativity of Einstein may not be the last word. Furthermore, the predictions of general relativity are fixed; the pure theory contains no adjustable constants, so nothing can be changed. Thus every test of the theory is either a potentially deadly test or a possible probe for new physics. Although it is remarkable that this theory, born 100 years ago out of almost pure thought, has managed to survive every test, the possibility of finding a discrepancy will continue to drive experiments for years to come. These experiments will search for new physics beyond Einstein in many different directions: the large distance scales of the cosmological realm; scales of very

short distances or high energy; and the realms of strong and dynamical gravity.

Box 1.1 Units and conventions

Throughout this book, we will adopt the units and conventions of standard textbooks such as Misner, Thorne and Wheeler (1973) (hereafter referred to as MTW) or Schutz (2009). For a pedagogical development of many of the topics presented here, such as Newtonian gravity, post-Newtonian theory and gravitational radiation, we will refer readers to Poisson and Will (2014) (hereafter referred to as PW). Although we have attempted to produce a reasonably self-contained account of gravitation theory and gravitational experiments, the reader's path will be greatly smoothed by a familiarity with general relativity at the level of one of these texts.

We will use “geometrized units”, in which $G = c = 1$ (except in Chapter 2) and in which mass and time have the same units as distance. Greek indices on vectors and tensors will run over the four spacetime dimensions, while Latin indices will run only over spatial dimensions. We will use the Einstein summation convention, in which one sums repeated indices over their range. Multi-index objects, such as products $x^j x^k x^l \dots$ will be denoted using capital superscripts, eg. x^N , where N is the number of indices. Partial derivatives and covariant derivatives will be denoted by commas and semicolons preceding indices, respectively. Parentheses enclosing indices will denote symmetrization, while square brackets will denote antisymmetrization.