General introduction

General Relativity is a theory proposed by Einstein in 1915 as a unified theory of space, time and gravitation. The theory's roots extend over almost the entire previous history of physics and mathematics.

Its immediate predecessor, Special Relativity, established in its final form by Minkowski in 1908, accomplished the unification of space and time in the geometry of a 4-dimensional affine manifold, a geometry of simplicity and perfection on par with that of the Euclidean geometry of space. The root of Special Relativity is Electromagnetic Theory, in particular Maxwell's incorporation of Optics, the theory of light, into Electrodynamics.

General Relativity is based on and extends Newton's theory of Gravitation as well as Newton's equations of motion. It is thus fundamentally rooted in Classical Mechanics.

Perhaps the most fundamental aspect of General Relativity however, is its geometric nature. The theory can be seen as a development of Riemannian geometry, itself an extension of Gauss' intrinsic theory of curved surfaces in Euclidean space.

The connection between gravitation and Riemannian geometry arose in Einstein's mind in his effort to uncover the meaning of what in Newtonian theory is the fortuitous equality of the inertial and the gravitational mass. Identification, via the equivalence principle, of the gravitational tidal force with spacetime curvature at once gave a physical interpretation of curvature of the spacetime manifold and also revealed the geometrical meaning of gravitation.

One sees here that descent to a deeper level of understanding of physical reality is connected with ascent to a higher level of mathematics. General Relativity constitutes a triumph of the geometric approach to physical science.

But there is more to General Relativity than merely a physical interpretation of a variant of Riemannian Geometry. For, the theory contains physical laws in the form of equations – Einstein's equations – imposed on the geometric structure. This gives a tightness which makes the resulting mathematical structure one of surpassing subtlety and beauty. An analogous situation is found by comparing the theory of differentiable functions of two real variables with the theory of differentiable functions of one complex variable. The latter gains, by the imposition of the Cauchy–Riemann equations, a tighter structure which leads to a greater richness of results.

The domain of application of General Relativity, beyond that of Newtonian theory, is astronomical systems, stellar or galactic, where the gravitational field is so strong that it implies the potential presence of velocities which are not negligible in comparison with the velocity of light. The ultimate domain of application is the study of the structure and evolution of the universe as a whole.

General Relativity has perhaps the most satisfying structure of all physical theories from the mathematical point of view. It is a wonderful research field for a mathematician. Here, results obtained by purely mathematical means have direct physical consequences.

One example of this is the incompleteness theorem of R. Penrose and its extensions due to Hawking and Penrose known as the “singularity theorems". This result is relevant to the study of the phenomenon of gravitational collapse. It shall be covered in the second volume of the present work. The methods used to establish the result are purely geometrical – the theory of conjugate points. In fact, part of the main argument is already present in the theory of focal points in the Euclidean framework, a theory developed in antiquity.

Another example is the positive energy theorem, the first proof of which, due to R. Schoen and S. T. Yau, is based on the theory of minimal surfaces and is covered in the the present volume. In this example a combination of geometric and analytic methods are employed.

A last example is the theory of gravitational radiation, a main theme for both volumes of this work. Here also we have a combination of geometric and analytic methods. A particular result in the theory of gravitational radiation is the so-called memory effect [11], which is due to the non-linear character of the asymptotic laws at future null infinity and has direct bearing on experiments planned for the near future. This result will also be covered in our second volume.

The laws of General Relativity, Einstein’s equations, constitute, when written in any system of local coordinates, a non-linear system of partial differential equations for the metric components. Because of the compatibility conditions of the metric with the underlying manifold, when piecing together local solutions to obtain the global picture, it is the geometric manifold, namely the pair consisting of the manifold itself together with its metric, which is the real unknown in General Relativity.

The Einstein equations are of hyperbolic character, as is explained in detail in this first volume. As a consequence, the initial value problem is the natural mathematical problem for these equations. This conclusion, reached mathematically, agrees with what one expects physically. For, the initial value problem is the problem of determining the evolution of a system from given initial conditions, as in the prototype example of Newton’s equations of motion. The initial conditions for Einstein’s equations, the analogues of initial position and velocity of Newtonian mechanics, are the intrinsic geometry of the initial spacelike hypersurface and its rate of change under a virtual normal displacement, the second fundamental form. In contrast to the case of Newtonian mechanics however, these initial conditions are, by virtue of the Einstein equations themselves, subject to constraints, and it is part of the initial value problem in General Relativity – a preliminary part – to analyze these constraints. Important results can be obtained on the basis of this analysis alone and the positive energy theorem is an example of such a result.

An important notion in physics is that of an isolated system. In the context of the theory of gravitation, examples of such systems are a planet with its moons, a star with its planetary system, a binary or multiple star, a cluster of stars, a galaxy, a pair or multiplet of interacting galaxies, or, as an extreme example, a cluster of galaxies – but not the universe as a whole. What is common in these examples is that each of these systems can be thought of as having an asymptotic region in which conditions are trivial. Within General Relativity the trivial case is the flat Minkowski spacetime of Special Relativity. Thus the desire to describe isolated gravitating systems in General Relativity leads us to consider spacetimes with asymptotically Minkowskian regions. However it is important to remember at this point the point of view of the initial value problem: a spacetime is determined as a solution of the Einstein equations from its initial data. Consequently, we are not free to impose our own requirements on a spacetime. We are only free to impose requirements on the initial data – to the extent that the requirements are consistent with the constraint equations. Thus the correct notion of an isolated system in the context of General Relativity is a spacetime arising from asymptotically flat initial conditions, namely an intrinsic geometry which is asymptotically Euclidean and is a second fundamental form which tends to zero at infinity in an appropriate way. This is discussed in detail in this volume.

Trivial initial data for the Einstein equations consists of Euclidean intrinsic geometry and a vanishing second fundamental form. Trivial initial data gives rise to the trivial solution, namely the Minkowski spacetime. A natural question in the context of the initial value problem for the vacuum Einstein equations is whether or not every asymptotically flat initial data which is globally close to the trivial data gives rise to a solution which is a complete spacetime tending to the Minkowski spacetime at infinity along any geodesic. This question was answered in the affirmative in the joint work of the present author with Sergiu Klainerman, which appeared in the monograph [14]. One of the aims of the present work is to present the methods which went into that work in a more general context, so that the reader may more fully understand their origin and development as well as be able to apply them to other problems. In fact, problems coming from fields other than General Relativity are also treated in the present work. These fields are Continuum Mechanics, Electrodynamics of Continuous Media and Classical Gauge Theories (such as arise in the mesoscopic description of superfluidity and superconductivity). What is common to all these problems from our perspective is the mathematical methods involved.

One of the main mathematical methods analyzed and exploited in the present work is the general method of constructing a set of quantities whose growth can be controlled in terms of the quantities themselves. This method is an extension of the celebrated theorem of Noether, a theorem in the framework of the action principle, which associates a conserved quantity to each 1-parameter group of symmetries of the action (see [12]). This extension is involved at a most elementary level in the very definition of the notion of hyperbolicity for an Euler–Lagrange system of partial differential equations, as discussed in detail in this first volume. In fact we may say that such a system is hyperbolic at a particular background solution if linear perturbations about this solution possess positive energy in the high frequency limit.

The application of Noether's Principle to General Relativity requires the introduction of a background vacuum solution possessing a non-trivial isometry group, as is explained in this first volume. Taking Minkowski spacetime as the background, we have the symmetries of time translations, space translations, rotations and boosts, which give rise to the conservation laws of energy, linear momentum, angular momentum and center of mass integrals, respectively. However, as is explained in this first volume, these quantities have geometric significance only for spacetimes which are asymptotic at infinity to the background Minkowski spacetime, so that the symmetries are in fact asymptotic symmetries of the actual spacetime.

The other main mathematical method analyzed and exploited in the present work is the systematic use of characteristic (null) hypersurfaces. The geometry of null hypersurfaces has already been employed by R. Penrose in his incompleteness theorem mentioned above. What is involved in that theorem is the study of a neighborhood of a given null geodesic generator of such a hypersurface. On the other hand, in the work on the stability of Minkowski spacetime, the global geometry of a characteristic hypersurface comes into play. In addition, the properties of a foliation of spacetime by such hypersurfaces, also come into play. This method is used in conjunction with the first method, for, such characteristic foliations are used to define the actions of groups in spacetime which may be called quasi-conformal isometries, as they are globally as close as possible to conformal isometries and tend as rapidly as possible to conformal isometries at infinity. The method is introduced in this first volume and will be treated much more fully in the second volume. It has applications beyond General Relativity to problems in Fluid Mechanics and, more generally, to the Mechanics and Electrodynamics of Continuous Media.

This book is based on Nachdiplom Lectures held at the Eidgenössische Technische Hochschule Zurich during the Winter Semester 2002/2003. The author wishes to thank his former student Lydia Bieri for taking the notes of this lecture, from which a first draft was written, and for making the illustrations.