Black Hole Equilibrium States

Brandon Carter

Institute of Theoretical Astronomy,
Cambridge, Great Britain

This page intentionally left blank

Contents

This page intentionally left blank

PART I Analytic and Geometric Properties of the Kerr Solutions

1 Introduction

The Kerr solutions (Kerr 1963) and their electromagnetic generalizations (Newman *et al.* 1965) form a 4-parameter family of asymptotically flat solutions of the source-free Einstein-Maxwell equations, the parameters being most conveniently taken to be the asymptotically defined mass M, electric charge Q, and magnetic monopole charge P, together with a rotation parameter a, which is such that (in units of the form we shall use throughout, where the speed of light c and Newton's constant G are set equal to unity) the asymptotically defined angular momentum J is given by

$$J = Ma$$

The parameters may range over all real values subject to the restriction

$$M^2 \geqslant a^2 + P^2 + Q^2$$

which must be satisfied if the solution is to represent the exterior to a *hole* rather than *naked singularity.* It turns out (Carter 1968a) that the solutions all have the same gyromagnetic ratios as those predicted by the simple Dirac theory of the electron, as the asymptotic magnetic and electric dipole moments cannot be specified independently of the angular momentum but are given, in terms of the same rotation parameter, as Qa and Pa respectively.

The solutions are all geometrically unaltered by variations of P and Q provided that the sum $P^2 + Q^2$ remains constant, and since it is in any case widely believed that magnetic monopoles do not exist in nature, attention in most studies is restricted to the 3-parameter subfamily in which P is zero. This 3-parameter subfamily, and specially the 2-parameter pure vacuum subfamily in which Q is also zero, has come recently to be regarded as being at least potentially of great astrophysical interest, since the Kerr solutions do not merely represent the *only known* stationary source-free black hole exterior solutions: they are also widely believed (for reasons which will be presented in Part II of this course) to be the *only possible* such solutions.

We shall devote the whole of Part I of the present lecture course to the derivation and geometric investigation of these Kerr solutions. In a strictly logical approach, Part II of this course (which will consist of a general examination of stationary black hole states with or without external sources) should come first, but it is probably advisable for a reader who is not already familiar with the subject to start with Part I, since the significance of the reasoning to be presented in Part II will be more easily appreciated if one has in mind the concrete examples described in Part I. For the same reason many readers may find it easier to appreciate the accompanying lecture course of Hawking if they have first become familiar with the examples described here, although only the final stages of Hawking's course actually depend on the results presented here, the bulk of it being logically antecedent to both Part II and Part I of the present course. This present lecture course is intended to serve both logically and pedagogically as an introduction to the immediately following course of Bardeen and hence also (but less directly) to the subsequent courses in this volume.

2 Spheres and Pseudo-Spheres

A space time manifold $\mathscr{M}$ is said to be *spherically symmetric* if it is invariant under an action of the rotation group SO(3) whose surfaces of transitivity are 2-dimensional. The metric on any one of these 2-surfaces must have the form

$$ds_{\odot}^2 = r^2 \,(d\theta^2 + \sin^2\theta d\varphi^2) \tag{2.1}$$

in terms of a suitably chosen azimuthal co-ordinate θ running from 0 to π, and a periodic ignorable co-ordinate φ defined modulo 2π, where the scale factor r is the *radius of curvature* of the 2-sphere. In what follows we shall frequently find it convenient to use the equivalent alternative form

$$ds_{\odot}^2 = r^2 \left\{\frac{d\mu^2}{1-\mu^2} + (1-\mu^2)d\varphi^2\right\} \tag{2.2}$$

expressed in terms the customary alternative azimuthal co-ordinate

$$\mu = \cos\theta \tag{2.3}$$

running from -1 to $+1$. See Figure 2.1.

A space-time manifold $\mathscr{M}$ is said to be *pseudo-spherically* symmetric if it is invariant under an action of the 3-dimensional Lorentz group whose surfaces of transitivity are timelike and 2-dimensional. The metric on one of these 2-surfaces can be expressed in a form analogous to (2) as

$$ds_{\times}^2 = -s^2 \left\{\frac{dx^2}{1-x^2} - (1-x^2)dt^2\right\} \tag{2.4}$$

where s is the radius of curvature and where the co-ordinate t is ignorable.

The two simple and familiar metric forms $ds_{\odot}^2$ and $ds_{\times}^2$ illustrate a feature which will crop up repeatedly in the present course, that is to say they both have removable *co-ordinate singularities*. The spherical form (2.2) is obviously singular at $\mu = \pm 1$; moreover we cannot remove this singularity simply by returning to the form (2.1) in which the infinity is eliminated only at the price of introducing an equally undesirable vanishing determinant which will of course lead to an infinity in the inverse metric tensor. We can cure the singularity and show explicitly that the space is well behaved at the poles (as we know it must be by the homogeneity)

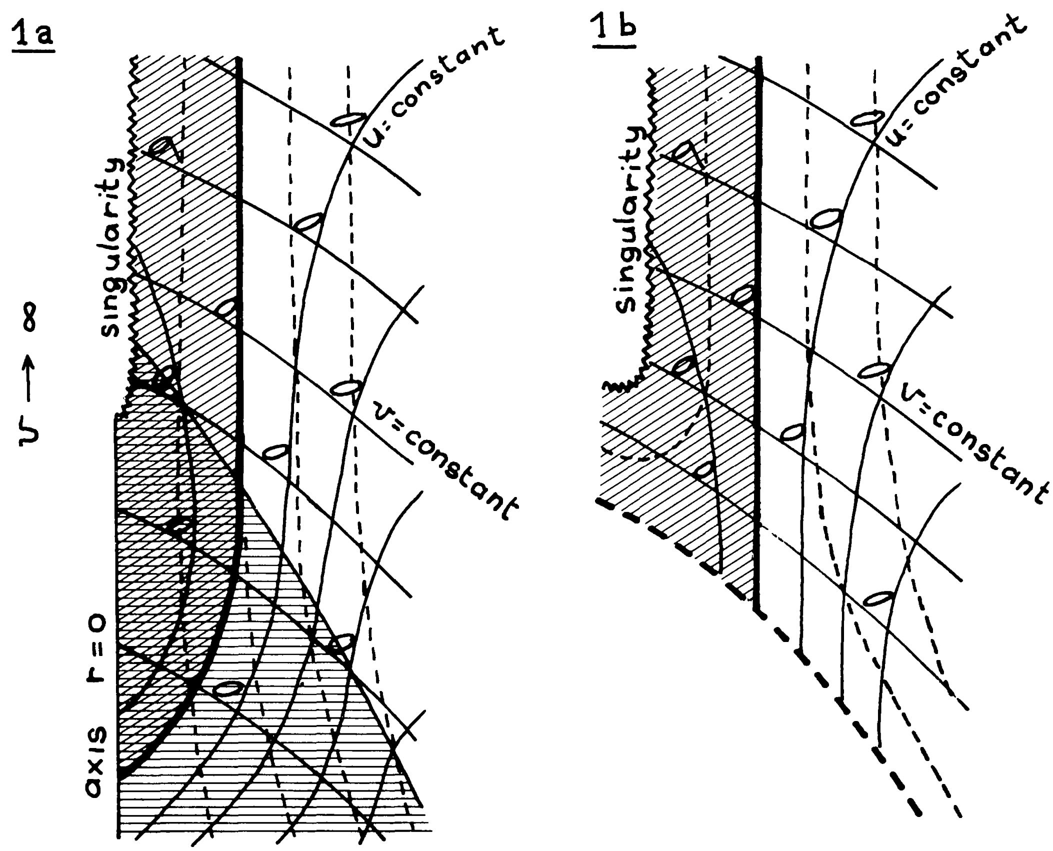


Figure 1. Figure 1a shows a plan of a timelike 2-section, in which the spherical co-ordinates θ, φ are held constant, of the space time manifold of a spherical collapsing star which occupies the part of space time marked by horizontal shading. Locuses on which r is constant are marked by dotted lines and null lines are marked by continuous lines. The horizon $\mathscr{H}^+$ is indicated by a heavy continuous line and the region outside the domain of outer communications is marked by diagonal shading. Figure 1b shows the extrapolation back under the group action of the pseudo-stationary empty outer region of Figure 1a. The past boundary marked by a heavy dotted line would become the horizon $\mathscr{H}^-$ in an extended manifold. (*See Part II, Section 1*)

only by giving up the use of the ignorable co-ordinate φ. Thus for example we can cure the singularity at the north pole $\mu = 1$ by introducing Cartesian type co-ordinates x, y defined by

$$x = 2r\frac{\sqrt{1-\mu^2}}{1+\mu}\sin\varphi \tag{2.5}$$

$$y = 2r\frac{\sqrt{1-\mu^2}}{1+\mu}\cos\varphi \tag{2.6}$$

to obtain the conformally flat form

$$ds_{\odot}^2 = \frac{dx^2 + dy^2}{1 + \dfrac{x^2+y^2}{4r^2}} \quad . \tag{2.7}$$

which is well behaved everywhere except at the south pole. It is of course impossible to obtain a form which is well behaved on the whole of the 2-sphere at the same time (see Figure 2.1.).

Let us now consider the algebraically analogous singularities at $x = \pm 1$ in the pseudo-spherical metric ds_x^2. These singularities are of a geometrically different nature since the metric is well behaved with the same signature, on *both sides* of the singularities i.e. both in the static regions with $|x| > 1$ and in the regions where the ignorable co-ordinate is spacelike $|x| < 1$. (In contrast with the spherical metric which has the right signature only for $|\mu| < 1$.) We can link the

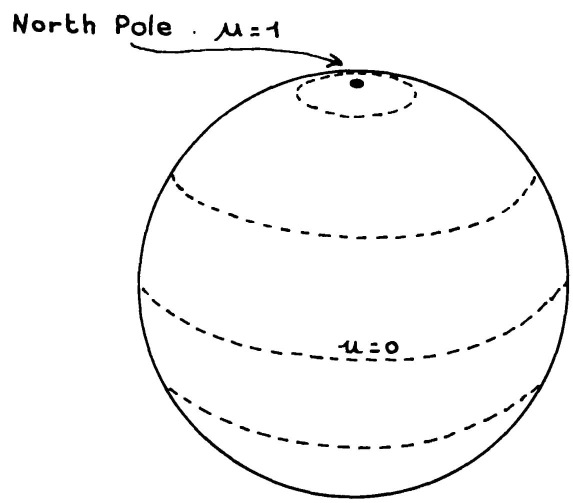


Figure 2.1. The 2-sphere. The light dotted lines represent trajectories of the Killing vector $\partial/\partial\varphi$.

well behaved domains $x > 1$ and $-1 < x < 1$ by a co-ordinate patch extending right across the divisions $x = \pm 1$, by introducing new co-ordinates

$$\tau = \frac{-x}{|1 - x^2|^{\frac{1}{2}}} e^{-t} \tag{2.8}$$

$$\lambda = |1 - x^2|^{\frac{1}{2}} e^t \tag{2.9}$$

which leads to the form

$$ds_x^2 = s^2 \left\{ \frac{d\lambda^2}{\lambda^2} - \lambda^2 d\tau^2 \right\} \tag{2.10}$$

in which the new co-ordinate τ is ignoreable. This form is well behaved over the whole of the region $\lambda > 0$ including the loci $\lambda\tau = \pm 1$ which correspond respectively to the divisions $-x = \pm 1$ in the original system (2.4). Moreover it can easily be seen that these loci are in fact *null lines*. The situation can be visualized most easily in terms of null-co-ordinates u, v which we introduce by the defining relations

$$u = \tau + \frac{1}{\lambda} \tag{2.11}$$

$$v = \tau - \frac{1}{\lambda} \tag{2.12}$$

which leads to the form

$$ds_x^2 = -4s^2 \frac{du dv}{(u - v)^2} \tag{2.13}$$

(whose u and v are restricted by the condition $u - v > 0$) in which the divisions $x = \pm 1$ are represented by the lines $u = 0$ and $v = 0$ respectively.

The relationship between the co-ordinate patches (2.13) (which is equivalent to (2.10) with $\lambda > 0$) and the co-ordinate patches of the form (2.7) is shown in the *conformal diagram* of Figure 2.2a. This is the first example of a technique, which we shall employ frequently, of representing the geometry of a timelike 2-surface on the plane of the paper, in a manner which takes advantage of the fact that any such metric is conformly flat and thus can be expressed in the canonical null form $ds^2 = C du dv$ where C is a conformal factor. The method consists simply of identifying the null co-ordinates u and v with ordinary Cartesian co-ordinates on the paper which conventionally are placed diagonally so that timelike directions lie in a cone of angles within 45 degrees of the vertical. There is nothing unique about such a representation since there is a wide choice of canonical null form preserving transformations in which the co-ordinates u, v are replaced respectively by new co-ordinates u^*, v^* of the form $u^* = u^*(u)$, $v^* = v^*(v)$ and in which the conformal factor C is replaced by $C^* = C(du^*/du)(dv^*/dv)$. This freedom can be used to arrange for the co-ordinate range u^*, v^* to be finite even if the original co-ordinate range is not (e.g. by taking $u^* = \tanh u$, $v^* = \tanh v$) thereby making it possible to represent an infinite timelike 2-manifold in its entirety on a finite piece of paper.

The co-ordinate patch of Figure 2a with the spacially homogeneous form (2.10) is in fact still incomplete, as would have been expected from the fact that starting from the time symmetric form (2.4) one could have extended into the past instead of the future by replacing t by $-t$ in the transformation equations (2.8) and (2.9). We can obtain a new form which covers both extensions by setting

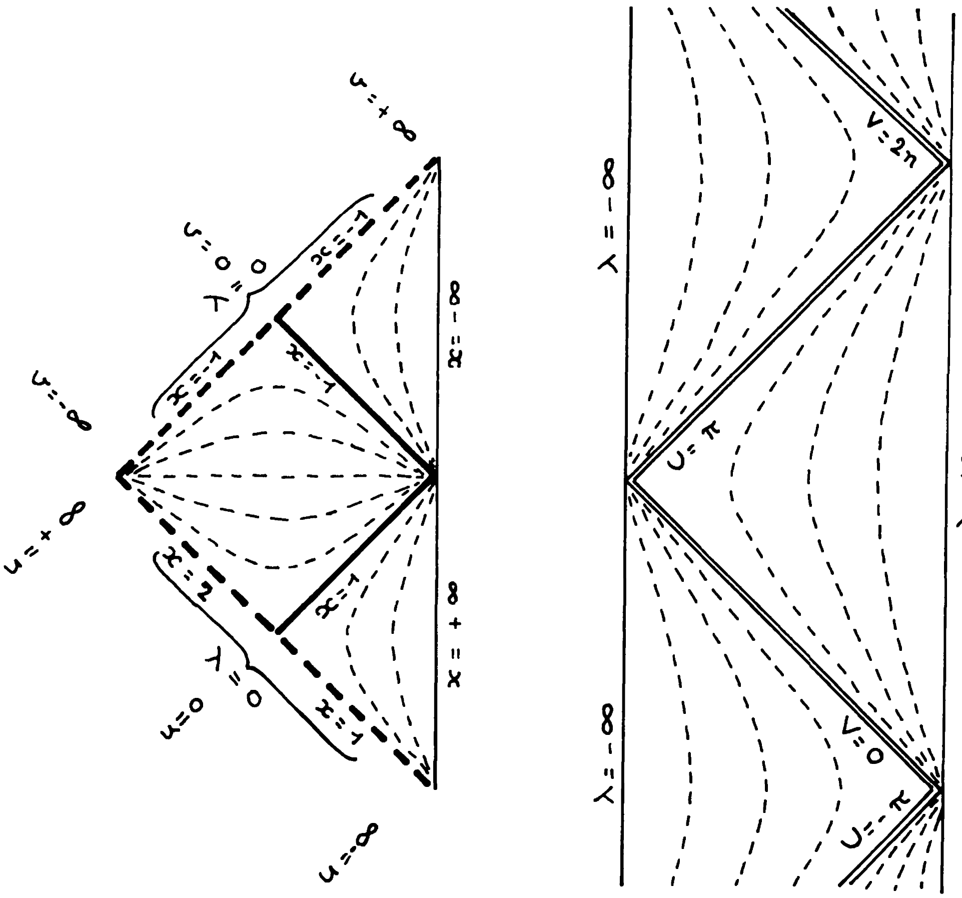
$$u = \tan \frac{U}{2} \tag{2.14}$$

$$v = -\cot \frac{V}{2} \tag{2.15}$$

which leads to the new null form

$$ds_x^2 = -s^2 \left\{1 + \tan^2 \frac{U - V}{2}\right\} dUdV \tag{2.16}$$

In terms of these co-ordinates (which can both range from $-\infty$ to ∞ subject to the restriction $-\pi < U - V < \pi$) the entire co-ordinate range of u and v subject to $u + v < 0$, i.e. the entire range of λ, τ subject to $\lambda > 0$, is covered by the range



2.2a 2.2b

Figure 2.2 Conformal diagram of the 2-dimensional pseudo-sphere. In Figure 2a the light dotted lines represent trajectories of the Killing vector $\partial/\partial t$ and the heavy lines represent the corresponding non-degenerate Killing horizons. In the extended diagram of Figure 2b the light dotted lines represent trajectories of the Killing vector $\partial/\partial\tau$, and the double lines represent the corresponding degenerate Killing horizons.

$0 < V < 2\pi$, and $-\pi < U < \pi$. The situation is illustrated in Figure 2b. The manifold in this figure is geodesically complete and hence maximal in the sense that no further extension can be made. The null form (2.16) covering this maximal extension can be converted to an equivalent static form by introducing co-ordinates X, T defined by

$$X = \tan \frac{U - V}{2} \tag{2.17}$$

$$T = \frac{U + V}{2} \tag{2.18}$$

which gives the maximally extended static form

$$ds_{\mathrm{x}}^2 = +s^2 \left\{ \frac{dX^2}{1 + X^2} - (1 + X^2) dT^2 \right\} \tag{2.19}$$

in which X and T can range from $-\infty$ to ∞ without restriction.

Looking back from the vantage point of this complete manifold we can see clearly what has been happening. The timelike Killing vector whose trajectories are the static curves x = constant, $|x| > 0$ in the system (2.4), becomes null on the surfaces on which U or V are multiples of π, these surfaces representing past and future *event horizons* for observers who are fixed in that x remains constant relative to this system. [The past and future event horizon (cf. Rindler 1966) of an observer being respectively the boundary of the past of his worldline, i.e. the set of events he will ultimately be able to know of, and the boundary of the future of his world line i.e. of the set of events which he could in principle have influenced.] Thus for the Killing vector whose trajectories are the static lines x = constant, $|x| > 0$ in the system (2.4), the corresponding event horizons are the lines on which U and V are multiples of π. In the extension (2.10) in which the ignorable co-ordinate t has been sacrificed, there is a new manifest symmetry generated by the Killing vector whose trajectories are the lines λ = constant, and for the corresponding static observers the event horizons coincide with alternate horizons of the previous set, specifically the lines on which $U + \pi$ and V are multiples of 2π. In the maximal extension in which the ignorable co-ordinate τ has in its turn been replaced by T, there is a third non-equivalent Killing vector field, whose static trajectories are the lines X = constant, and in this case the corresponding observers have no event horizons.

There is a fairly close analogy between the removable co-ordinate singularities associated with rotation axes, as exemplified by the case of the ordinary 2-sphere discussed earlier, and the removable co-ordinate singularities associated with Killing horizons. Both arise from the inevitable bad behaviour of an *ignorable* co-ordinate which is used to make manifest symmetry group action generated by a Killing vector. The former case arises in the case of a spacelike Killing vector generating a rotation group action when it becomes zero on a rotation axis, while the latter arises in the case of a static Killing vector (and also under appropriate conditions as we shall see later a stationary Killing vector) when it becomes null on a Killing horizon. Killing horizons can be classified as degenerate or non-degenerate according to whether the gradient of the square of the Killing vector is zero or not. In the case of a *non-degenerate Killing horizon* (as exemplified by the horizons on which the Killing vector whose trajectories are x = constant) the relevant Killing vector must change from being timelike on one side to being spacelike on the other. In the case of a *degenerate Killing horizon*, the relevant

Killing vector *may be* (and in the sample of the Killing vector whose trajectories are λ = constant, actually *is*) timelike on both sides of the horizon. We can see in the example of the pseudo-sphere another phenomenon that will later be shown to be true in general, namely that non-degenerate Killing horizons of a given Killing vector field cross each other at what we shall call a Boyer-Kruskal axis, on which the Killing vector is actually zero, this axis being very closely analogous to a rotation axis. On the other hand degenerate Killing horizons (which have no exact analogue in the case of axisymmetry) continue unimpeded over an unbounded range.

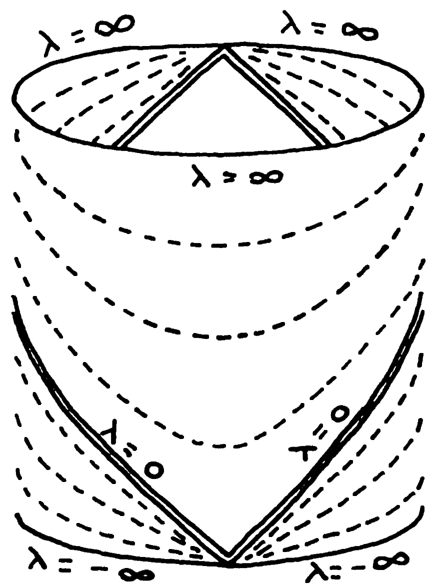


Figure 2.3. Sketch of conformal diagram of 2-dimensional pseudo-sphere with the canonical global topology obtained by identifying points with the same value of X but with values of T differing by 2π in the maximally extended covering manifold of Figure 2.2b.

Since T is ignorable, we could construct a new reduced manifold from the manifold of Figure 2.2b by identifying points with the same values of X but for which the values of T differ by any arbitrarily chosen fixed p period; the locus of points at unit spacelike distance from the origin in 3-dimensional Minkowski space is a manifold of the form obtained in this way where the period of T has the canonical value 2π (see Figure 2.3).

3 Derivation of the Spherical Vacuum Solutions

In this section we shall run rapidly through the steps by which the spherically symmetric vacuum (both pure vacuum and source free Einstein Maxwell) solutions are derived. Thus we start from the condition that the space time manifold under consideration is invariant under an action of SO(3) which is transitive over 2-surfaces. Since the rotation group action is necessarily invertible in the sense that associated with any point there is a discrete subgroup action—a rotation by 180°—which reverses the senses of the tangent vectors to the surfaces of transitivity, it follows automatically (cf. e.g. Schmidt 1967, Carter 1969) that the group action is *orthogonally transitive* in the sense that the surfaces of transitivity must themselves be orthogonal to another family of 2-surfaces. Therefore by requiring that they be constant on these 2-surfaces, we can extend the polar co-ordinate θ, φ or equivalently μ, φ from the individual 2-surfaces to the 4-dimensional manifold $\mathscr{M}$. It follows further that apart from an overall conformal factor r^2 (where r is the local radius of the 2-spheres) the space-time will locally have the form of a direct product of the unit 2-sphere with a certain *time-like 2-space*, with metric dl^2 say, so that the overall metric will have the canonical form

$$ds^2 = r^2 \left\{ dl^2 + \frac{d\mu^2}{1-\mu^2} + (1-\mu^2)d\varphi^2 \right\}$$

Since we are only considering stationary spaces in the present course, we can introduce an ignorable co-ordinate t on the timelike 2-spaces (in fact by the well known theorem of Birkhoff, the stationary assumption involves no loss of generality at all in so far as spherically symmetric *pure vacuum* or *electromagnetic vacuum* solutions are concerned). Moreover except in the special case (which we shall consider later since it is not as irrelevant as is often assumed) when r is constant, we may take the radius r itself as a co-ordinate on the time-like 2-space. The t co-ordinate can then be fixed uniquely by the requirement that it be orthogonal to the r co-ordinate. This leads to the canonical spherical metric form

$$\boxed{ds^2 = r^2 \left\{ \frac{d\mu^2}{1-\mu^2} + (1-\mu^2)d\varphi^2 + \frac{dr^2}{\Delta_r} - \frac{\Delta_r dt^2}{Z_r^2} \right\}} \tag{3.2}$$

where Δ_r and Z_r are two arbitrary functions of r only, and where the factors Δ_r have been distributed in such a manner as to cancel out of the expression for the metric determinant g, so that the ubiquitous volume density weight factor $\sqrt{-g}$ which appears in so many expressions takes the very simple form

$$\sqrt{-g} = \frac{r^4}{Z_r} \tag{3.3}$$

The easiest way that I know of solving Einstein's equations in a case like this is the method described by Misner (1963), which is based on the application of Cartan type calculus to the differential forms of the canonical tetrad. In this method the metric is expressed in terms of 4 differential forms $\omega^{(i)} = \omega_a^{(i)} dx^a$, $i = 1, 2, 3, 0$, which are orthonormal, so that the metric takes the form

$$ds^2 = g_{(i)(j)}\omega^{(i)}\omega^{(j)} \tag{3.4}$$

where the $g_{(i)(j)}$ are scalar constants which are the components of the standard Minkowski metric tensor in a flat co-ordinate system. Instead of working with a large number of Christofel symbols one works with what will (specially if the original metric form is fairly simple) be a comparatively small number of connection forms, $\gamma^{(i)}{}_{(j)} = \gamma^{(i)}{}_{(j)}\, dx^a$ defined (but *not* in practice computed) by the equations

$$\omega^{(i)}{}_{a;b} = \gamma^{(i)}{}_{(j)b}\,\omega^{(j)}{}_a \tag{3.5}$$

These forms will automatically be symmetric in the sense that if the labelling indices are lowered by contractions with the Minkowski scalors i.e. setting $\gamma_{(i)(j)} = g_{(i)(k)}\gamma^{(k)}{}_{(j)}$, then we have they will satisfy

$$\gamma_{(i)(j)} = \gamma_{(j)(i)} \tag{3.6}$$

Moreover the antisymmetrized part of the equation (3.5) can be expressed in terms of Cartan form language as

$$d\omega^{(i)} = -\gamma^{(i)}{}_{(j)} \wedge \omega^{(j)} \tag{3.7}$$

and it can easily be seen that the two expression (3.6) and (3.7) together (which can easily be worked out without using covariant differentiation) can be used to determine the forms $\gamma^{(i)}{}_{(j)}$ instead of the computationally more awkward defining relation (3.4). Furthermore the tetrad components $R_{(i)(j)(k)(l)}$ of the Riemann tensor can also be read out without the use of covariant differentiation from the expression

$$\theta^{(i)}{}_{(j)} = \tfrac{1}{2}R^{(i)}{}_{(j)(k)(l)}\,\omega^{(k)} \wedge \omega^{(l)} \tag{3.8}$$

where the curvature two-forms $\theta^{(i)}{}_{(j)} = \theta^{(i)}{}_{(j)ab}\, dx^a \wedge dx^b$ are given (as can easily be checked by differentiating (1) and using the defining relation $\omega^{(i)}{}_{a[b;c]} = \frac{1}{2}R_{abcd}\,\omega^{(i)\alpha}$ bv

$$\theta^{(i)}{}_{(j)} = d\gamma^{(i)}{}_{(j)} + \gamma^{(i)}{}_{(k)} \wedge \gamma^{(k)}{}_{(j)} \tag{3.9}$$

The tetrad components of the Ricci tensor are obtained simply by contracting

$$R_{(i)(j)} = R^{(k)}{}_{(i)(j)(k)} \tag{3.10}$$

and if required the co-ordinate form is given by

$$R_{ab} = R_{(i)(j)}\,\omega^{(i)}{}_a\,\omega^{(j)}{}_b. \tag{3.11}$$

In the present case the obvious tetrad of forms consists of

$$\left.\begin{aligned} \omega^{(1)} &= \frac{r}{\sqrt{\Delta_r}}\, dr \\ \omega^{(2)} &= \frac{r}{\sqrt{1-\mu^2}}\, d\mu \\ \omega^{(3)} &= r\sqrt{1-\mu^2}\, d\varphi \\ \omega^{(0)} &= \frac{r\sqrt{\Delta_r}}{Z_r}\, dt \end{aligned}\right\} \tag{3.12}$$

and a straightforward calculation on the lines described above shows that the only solution of the pure Einstein vacuum equations

$$R_{ij} = 0 \tag{3.13}$$

are given (after use of co-ordinate scale change freedom to achieve a standard normalization) by

$$Z_r = r^2 \tag{3.14}$$

$$\Delta_r = r^2 - 2Mr \tag{3.15}$$

where M is a constant. The corresponding values of the curvature forms may be tabulated as

$$\left.\begin{array}{ll} \theta^{(1)}{}_{(2)} = -\frac{M}{r^3}\,\omega^{(1)} \wedge \omega^{(2)} & \theta^{(3)}{}_{(0)} = -\frac{M}{r^3}\,\omega^{(3)} \wedge \omega^{(0)} \\ \theta^{(3)}{}_{(1)} = \frac{M}{r^3}\,\omega^{(1)} \wedge \omega^{(3)} & \theta^{(3)}{}_{(2)} = 2\frac{M}{r^3}\,\omega^{(2)} \wedge \omega^{(3)} \\ \theta^{(0)}{}_{(1)} = 2\frac{M}{r^3}\,\omega^{(1)} \wedge \omega^{(0)} & \theta^{(0)}{}_{(2)} = \frac{M}{r^3}\,\omega^{(2)} \wedge \omega^{(0)} \end{array}\right\} \tag{3.16}$$

The comparative conciseness of this array, from which, if desired, all 20 of the ordinary Riemann tensor components can be read out, shows clearly the advantage of the Cartan formulation. If one is interested in the Petrof classification, it can be seen directly from the above array that the tetrad $\omega^{(i)}$ is in fact a canonical tetrad and that the Riemann tensor, which in the vacuum case is the same as the Weyl conformal tensor, is of type D. (See Pirani 1964, 1962; Ehlers and Kundt 1962.

The familiar Schwarzschild metric itself is given explicitly by

$$ds^2 = r^2 \left\{\frac{dr^2}{r^2 - 2Mr} + \frac{d\mu^2}{1 - \mu^2} - (1 - \mu^2)d\varphi^2\right\} - \frac{(r^2 - 2Mr)}{r^2}\,dt^2 \tag{3.17}$$

To obtain the *electromagnetic* vacuum solutions we must first find the most general forms of the electromagnetic field consistent with spherical symmetry. We start from the well known fact that there are no spherically symmetric vector fields on the 2-sphere, and only one unique (apart from a scale factor) spherically symmetric 2-form on the 2-sphere, which takes a very simple form in terms of the co-ordinate system (2.2), namely $d\mu \wedge d\varphi$. Since any cross components of the Maxwell field between directions orthogonal to and in the 2-spheres of transitivity would define vector fields in the surfaces of transitivity it follows that the most general spherical Maxwell field is a linear combination with co-efficients depending only on r, of $dr \wedge dt$ and $d\mu \wedge d\varphi$, and hence can be expressed in terms of the canonical tetrad in the form

$$F = 2E_r\omega^{(1)} \wedge \omega^{(0)} + 2B_r\omega^{(2)} \wedge \omega^{(3)} \tag{3.18}$$

where E_r and B_r are functions of r only. The dual field form $*F$ is then given simply by

$$*F = 2B_r\omega^{(1)} \wedge \omega^{(0)} + 2E_r\omega^{(2)} \wedge \omega^{(3)} \tag{3.19}$$

and in terms of these expressions the source free Maxwell equations simply take the form $dF = 0$, $d*F = 0$ and thus they too can be worked out by the Cartan method without recourse to covariant differentiation. The electromagnetic energy tensor will be given by

$$8\pi T_{ab} = (E_r^2 + B_r^2)\{\omega_a^{(0)}\omega_b^{(0)} + \omega_a^{(3)}\omega_b^{(3)} + \omega_a^{(2)}\omega_b^{(2)} - \omega_a^{(1)}\omega_b^{(1)}\} \tag{3.20}$$

and hence the Einstein-Maxwell equations

$$R_{ab} = 8\pi T_{ab} \tag{3.21}$$

can easily be worked out with the Ricci tensor evaluated in the way described above.

The solutions are given by

$$E_r = \frac{Q}{r^2} \qquad B_r = \frac{P}{r^2} \tag{3.22}$$

(where Q and P are constants which correspond respectively to electric and magnetic monopole charges) with

$$Z_r = r^2 \tag{3.23}$$

$$\Delta_r = r^2 - 2Mr + Q^2 + P^2 \tag{3.24}$$

In presenting the information giving the curvature, it is worthwhile to make a distinction now between the Riemann tensor, and the Weyl conformal tensor since the Ricci tensor which may be read out by substituting (3.20) in (3.21) and using (3.22), is no longer zero. The Weyl forms $\Omega_{(i)(j)} = \Omega_{(i)(j)ab}\,\omega_a^{(i)}\,\omega_b^{(j)}$ defined in terms of the tetrad components $C^{(i)}{}_{(j)(k)(l)}$ by

$$\Omega^{(i)}{}_{(j)} = \tfrac{1}{2}C^{(i)}{}_{(j)(k)(l)}\omega^{(k)} \wedge \omega^{(l)} \tag{3.25}$$

are related to the curvature forms $\theta^{(i)}{}_{(j)}$ by

$$\Omega^{(i)(j)} = \theta^{(i)(j)} + R^{[(i)}{}_{(k)}\omega^{(j)]} \wedge \omega^{(k)} + \tfrac{1}{6}R\omega^{(i)} \wedge \omega^{(j)} \tag{3.26}$$

where R is the Ricci scalar (which is of course zero in the present case). The Weyl form may be tabulated as

$$\left.\begin{aligned}
\Omega^{(1)}{}_{(2)} &= -\frac{Mr - Q^2 - P^2}{r^4}\,\omega^{(1)} \wedge \omega^{(2)} & \Omega^{(0)}{}_{(3)} &= -\frac{Mr - Q^2 - P^2}{r^4}\,\omega^{(3)} \wedge \omega^{(0)} \\
\Omega^{(3)}{}_{(1)} &= \frac{Mr - Q^2 - P^2}{r^4}\,\omega^{(3)} \wedge \omega^{(1)} & \Omega^{(3)}{}_{(2)} &= 2\frac{Mr - Q^2 - P^2}{r^4}\,\omega^{(3)} \wedge \omega^{(3)} \\
\Omega^{(0)}{}_{(1)} &= 2\,\frac{Mr - Q^2 - P^2}{r^4}\,\omega^{(0)} \wedge \omega^{(1)} & \Omega^{(0)}{}_{(2)} &= \frac{Mr - Q^2 - P^2}{r^4}\,\omega^{(0)} \wedge \omega^{(2)}
\end{aligned}\right\} \tag{3.27}$$

Again the form of this array makes it immediately clear to a connoisseur that the solution is of Petrov type D. The explicit form of the electromagnetic field is

$$F = \frac{2Q}{r^2}\, dr \wedge dt + 2P\, d\mu \wedge d\varphi \tag{3.28}$$

which may be derived via the relation

$$F = 2\, dA \tag{3.29}$$

from a vector potential A given by

$$A = \frac{Q}{r}\, dt - P\mu\, d\varphi \tag{3.30}$$

The explicit form of the metric itself is

$$ds^2 = r^2 \left\{ \frac{dr^2}{r^2 - 2Mr + Q^2 + P^2} + \frac{d\mu^2}{1-\mu^2} + (1-\mu^2)\, d\varphi^2 \right\} - \frac{r^2 - 2Mr + Q^2 + P^2}{r^2}\, dt \tag{3.31}$$

this being the solution of Riessner and Nordstrom, which of course includes the Schwarzschild solution in the limit when Q and P are set equal to zero.

Our search for spherical solutions is not quite complete at this point because by working with the spherical radius r as a co-ordinate, we have excluded the special case where r is a constant. Of course a solution with this property cannot be asymptotically flat, unlike the solutions which we have obtained so far, and therefore it might be thought not to have much physical interest. However as we shall see in the next section, it is impossible to have a full understanding of the global structures of the solutions we have obtained so far, without considering this special case, which arises naturally in a certain physically interesting limit.

To deal with this special case we must alter the canonical metric form (3.2) by replacing the co-ordinate r by a new co-ordinate, λ say, *except* in the conformal factor r^2 outside the large brackets which remains formally as before, and is now to be held constant.

Using the same methods as before, we find that there are no pure vacuum solutions of this form, but that there do exist source free electromagnetic solutions, which can be expressed as follows: The electromagnetic field takes the form

$$F = 2Q\, d\lambda \wedge dt + 2P\, d\mu \wedge d\varphi \tag{3.32}$$

which can be derived from a vector potential

$$A = Q\lambda\, dt - P\mu\, d\varphi \tag{3.33}$$

and the metric is given by

$$ds^2 = (Q^2 + P^2) \left\{ \frac{d\lambda^2}{\lambda^2} + \frac{d\mu^2}{1-\mu^2} + (1-\mu^2)\, d\varphi - \lambda^2\, dt^2 \right\} \tag{3.34}$$

This is the solution of Robinson and Bertotti. This metric is in fact almost as symmetric as possible, since it can easily be seen that it is the direct product of a 2-sphere whose radius is the square root of $Q^2 + P^2$ with a pseudo sphere (cf the form (2.10)) of the same radius. Its maximal symmetry group therefore has not four parameters as in the previous solutions (three for sphericity and one for stationarity) but six.

4 Maximal Extensions of the Spherical Solutions

It is relatively easy to analyse the global structures of the solutions which we have just derived, since, as indicated by (3.1), all of them are equivalent, modulo a conformal factor r^2, to the direct product of a 2-sphere with a timelike 2-surface. Thus the problem boils down to an analysis of the timelike 2-surfaces with metric $ds_\perp^2 = r^2 dl^2$, whose structures can be represented by the simple conformal diagrams whose use was described in section (1).

The simplest case is of course simply flat space to which the Schwarzschild form (3.17) reduces when the mass parameter is set equal to zero. The corresponding timelike 2-dimensional metric is simply $dr^2 - dt^2$ with the restriction $\pi > 0$, which, via the transformation $u = r - t$, $v = r + t$, is equivalent to the flat null form $2\, du\, dv$, with the restriction $u + v > 0$. The corresponding conformal diagram is given in Figure 4.1, in which the null boundaries $\mathscr{I}^+$ and $\mathscr{I}^-$ which play a key role in the Penrose definition of asymptotic flatness are marked. To qualify as asymptotically flat in the Penrose sense—which he refers to as the condition of weak asymptotic simplicity—a spacetime manifold must be conformally equivalent to an extended manifold-with-boundary with well behaved null boundary horizons $\mathscr{I}^+$ and $\mathscr{I}^-$ isomorphic to those of flat space. (Asymptotic flatness is discussed in more detail in the accompanying course of Hawking.)

The Schwarzschild solution is of course asymptotically flat in this sense, and indeed when M is negative the conformal diagram of the metric

$$ds_\perp^2 = \frac{r^2\, dr^2}{r^2 - 2Mr} - \frac{r^2 - 2Mr}{r^2}\, dt^2 \tag{4.1}$$

is topologically identical to the flat space diagram of Figure 4.1, although there is the important geometric difference that whereas in the former case the boundary $r = 0$ represented only a trivial co-ordinate degeneracy at the spherical centre, in the latter case, as can be seen from a glance at the array (3.16) it represents geometric singularity of the Riemann tensor. In the physically more interesting case where M is positive the Schwarzschild conformal diagram still agrees with the flat space diagram for large values of r, in accordance with the Penrose criterion, but it has an entirely different behaviour going in the other direction due to the fact that the metric form (3.17) becomes singular not only at $r = 0$ but also at $r = 2M$. The fact that the Riemann components, as exhibited in the array (3.16), are perfectly well behaved there suggests that this may not be a true geometric singularity but merely a removable co-ordinate singularity of the Killing horizon type which we have already come across in section 2.

It is not difficult to verify that this is indeed the case. By introducing just one null co-ordinate v, defined by

$$v = t + \{r + 2M \log |r - 2M|\} \tag{4.2}$$

we obtain the form

$$ds_{\perp}^2 = 2\, dv\, dr - \frac{r^2 - 2Mr}{r^2} dv^2 \tag{4.3}$$

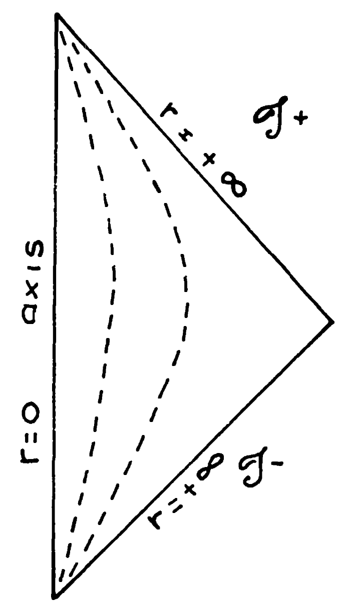


Figure 4.1. Conformal diagram of timelike 2-section with constant spherical co-ordinates θ, φ of Minkowski space.

This metric form is well behaved over the whole co-ordinate range $0 < r < \infty$ $-\infty < v < \infty$. The removability of the co-ordinate singularity at $r = 2M$ was first pointed out by Lemaitre (1933). The type of transformation we have used here in which the manifest stationary symmetry is preserved is due to Finkelstein (1958). The conformal diagram of this Finkelstein extension is shown in Figure 4.2.

The Finkelstein extension is not of course maximal since we could equally well have extended in the past direction, by introducing the alternative null co-ordinate

$$u = t - \{r + 2M \log |r - 2M|\} \tag{4.4}$$

and thereby obtaining a symmetric but distinct metric extension:

$$ds_{\perp}^2 = -2\, du\, dr - \frac{r^2 - 2Mr}{r^2} du^2 \tag{4.5}$$

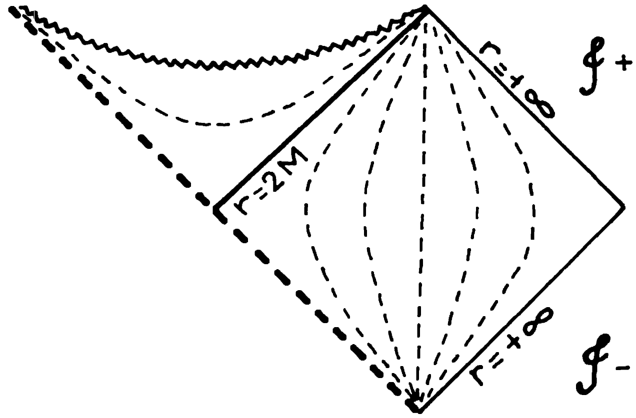


Figure 4.2. Conformal diagram of timelike 2-section with constant spherical co-ordinates θ, φ of Finkelstein extension of Schwarzschild manifold. In this figure and in all the conformal diagrams of section 2 and section 4 the convention is employed that the Killing vector trajectories (which in this case are the curves on which r is constant) are marked by light dotted lines, while Killing horizons are marked by heavy lines, except in that degenerate Killing horizons are marked by double lines. Curvature singularities are indicated by zig-zag lines.

In fact we can make a sequence of such extensions alternating between outgoing and ingoing null lines. After four such Finkelstein extensions, we can arrange to get back to our original starting point, in the manner shown in Figure 4.3. None of the four patches covers the central crossover point in the diagram, but our

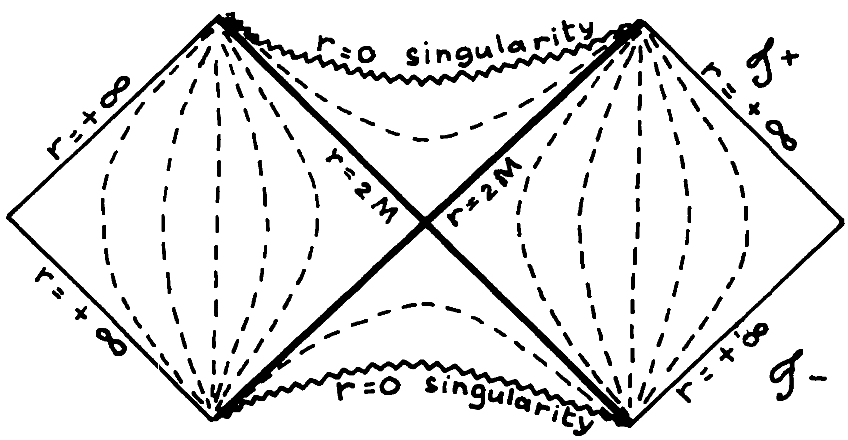


Figure 4.3. Conformal diagram of timelike 2-section with constant spherical co-ordinates θ, φ of Krushal's maximally extended Schwarzschild's solution.

experience with the analogous Killing horizon crossover points in the pseudo-sphere, suggests that just as the extension from (2.4) to (2.10) could be augmented to the maximal extension (2.19), so also here there should be a further extension to cover the crossover point. This is indeed the case, but unfortunately this time in dropping the manifest symmetry represented by the ignorability of the co-ordinate t we cannot expect to obtain any new compensating manifest symmetry as we did in the previous case, and so the new extended system will not be very elegant.

To construct the required extension we first introduce the null co-ordinates u, v simultaneously to obtain the null form

$$ds_\perp^2 = -\frac{r^2 - 2Mr}{r^2}\, du\, dv \tag{4.6}$$

where now r is given implicitly as the solution of

$$r + 2M \log |r - 2M| = \tfrac{1}{2}(u + v) \tag{4.7}$$

At this point we are even worse off than with the original Finkelstein extension, since we have restored the co-ordinate degeneracy at $r = 2M$. However we can now easily make an extension both ways at once by setting

$$u = -4M \ln |U| \tag{4.8}$$

$$v = 4M \ln |V| \tag{4.9}$$

which gives

$$ds_\perp^2 = \frac{e^{-r/2M}\, dU\, dV}{r} \tag{4.10}$$

where r is given implicitly as a function of U and V by

$$(r - 2M)\, e^{r/2M} = -UV \tag{4.11}$$

and where if required, the original time co-ordinate t can be recovered using the equation

$$e^{t/2} = -\frac{V}{U} \tag{4.12}$$

This form is well behaved in the whole of the UV plane. The original static patch is determined by $U < 0$, $V > 0$ and the original Finkelstein extension is determined by $V > 0$. It is fairly obvious that this extension is maximal since most geodesics either can be extended into the asymptotically flat region or to the geometrically singular limit. However to prove strictly that it is maximal we should check that there are no incomplete geodesics lying on or tending to the Killing horizons $U = 0$ and $V = 0$. Weshall cover this point later on when we discuss geodesics. [Checking that all geodesics are either complete or tend to a curvature singularity is a standard way of proving that an extension is maximal, but it is not the only way; for example a compact manifold must clearly be inextensible even though it may—as was pointed out by Misner (1963)—contain incomplete geodesics.] This maximal extension was first published—in a somewhat different co-ordinate system—by Kruskal (1960), surprisingly recently when one considers that the Schwarzschild solution, was discovered in 1916.

For studying the stationary exterior field of a collapsed star, that is to say the classical (if it is not premature to use such a word) black hole situation, the Finkelstein extension is sufficient, as is indicated by the conformal diagram given in Figure 4.4. The full Kruskal extension has the feature, which seems unlikely to be relevant except in a rather exotic situation, of possessing *two distinct asymptotically flat parts*, which are connected by wha' has come to be known as a bridge. The nature of the bridge can be under' ›od by considering the geometry of the three dimensional space sections, e.g. the locus $U = V$ (which coincides with $t = 0$ in the original co-ordinate system) whose geometry is suggested by Figure 4.5 which is meant to illustrate an imbedding in 3-dimensional

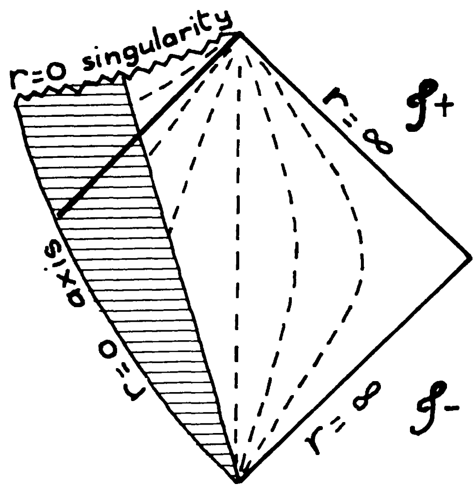


Figure 4.4. Conformal diagram of timelike 2-section with constant spherical co-ordinates θ, φ of maximally extended space-time of a collapsing spherically symmetric star.

flat space of the 2-dimensional circumferential section $\theta = \pi$ of the locus $U = V$. These three sections have spherical cross sections whose radius diminishes to a minimum value $r = 2M$ (at the throat of the bridge) and then increases again without bound. It can be seen from the conformal diagram of Figure 4.3 that an observer on a timelike trajectory cannot in fact cross this throat, since after crossing the horizon $U = 0$, the co-ordinate r must inevitably continue to decrease, until after a finite proper time the observer hits the geometric singularity at $r = 0$; this means in fact not only that such an observer is unable to reach the region of expanding r on the other side of the throat, but also that he is unable to return or even send a signal to the regions $r > 2M$ on the side from which he came. It is this latter phenomenon which is apparent even on the restricted Finkelstein extension (Figure 4.2), and in the dynamically realistic collapse diagram (Figure 4.4) which justifies the description of the null hypersurface at $r = 2M$ as a *horizon*. Thus technically, the hypersurface $U = 0$ is the past event horizon of $\mathscr{I}^+$, and together with the hypersurface $V = 0$ bounding the future of $\mathscr{I}^-$, it forms the boundary of the *domain of outer communications* which we shall denote by $\langle\!\langle \mathscr{I} \rangle\!\rangle$ meaning the region which can both receive signals from, and send them back to, asymptotically large distances which in the present case is the connected region with $r > 2M$ specified by $U < 0$, $V > 0$. The horizon $U = 0$ has the *infinite red shift* property (which is quite generally characteristic of a past event horizon of $\mathscr{I}^+$) that the light emitted by any physical object which crosses it is spread out over an infinite time as seen by a stationary observer in the asymptotically flat region, and thus not only gets infinitely red shifted as the body approaches the horizon (which of course it can reach in a finite proper time from its own point of view) but also progressively fades out. Thus the body, which might for example be the collapsing star represented on Figure 4.4, will

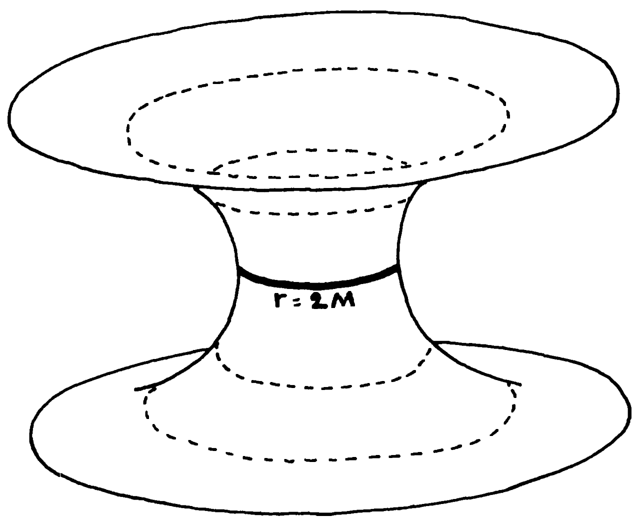


Figure 4.5. Sketch of part of a space-like equatorial 2-section ($\cos\theta = 0$) of one of the t = constant hypersurfaces through the throat of Kruskal's maximally extended Schwarzschild manifold. Trajectories of the Killing vector $\partial/\partial\varphi$ are indicated by light dotted lines.

appear to become not only redder but also blacker until with a characteristic time of the order of that required for light to cross the Schwarzschild radius ($\sim 10^{-4}$ seconds in the case of a collapsing star of typical mass) it effectively fades out of view altogether. It is for this reason that the region inside the horizon is referred to as a black hole.

Of course in the more exotic situation when the full Kruskal extension is present, the region of the past of $V = 0$ is by no means invisible from outside, and it would in fact be possible to see right back to the singularity (unless of course, like the big bang of cosmology theory, it were hidden in some opaque cloud of particles). The region to the past of $V = 0$ is often referred to as a *white* hole, although with less justification than the application of the description black to the future of $U = 0$. Without worrying about the question of colour, it is obviously reasonable in all cases to describe the regions outside the domain of outer communications $\langle\!\langle \mathscr{I} \rangle\!\rangle$ as *holes.*

Let us now move on to see how the situation is modified in the electromagnetic case. In the Reissner-Nordstrom solutions the timelike 2-sections have the metric form

$$ds_\perp^2 = \left(\frac{r^2}{r^2 - 2Mr + Q^2 + P^2}\right) dr^2 - \left(\frac{r^2 - 2Mr + Q^2 + P^2}{r^2}\right) dt^2 \qquad (4.13)$$

When the charge is large, more precisely when $Q^2 + P^2 > M^2$, this metric is well behaved over the whole range from the singularity at $r = 0$ (which from the array (3.26) can be seen to be a genuine geometric curvature singularity) to the asymptotic limit $r \to \infty$, and the conformal diagram is the same as in the negative M

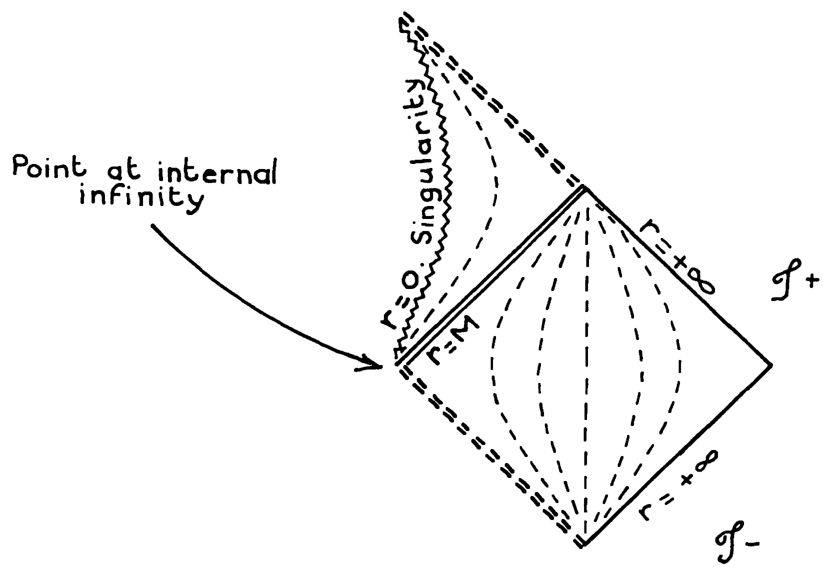


Figure 4.7. Conformal diagram of timelike 2-section with constant spherical co-ordinates θ, φ of Finkelstein type extension of Riessner-Nordstrom solution with $M^2 = P^2 + Q^2$.

Schwarzschild case and is given by Figure 4.6 (This also applies to the Reissner-Nordstrom case with negative M.) However when $M^2 > Q^2 + P^2$ there is a singularity of the metric form (4.13) at two different values of r, which we shall label r_+ and r_-, and which are given by

$$r_\pm = M \pm \sqrt{M^2 - Q^2 - P^2} \qquad (4.14)$$

We can use this notation to write the metric in the more convenient form

$$ds_\perp^2 = \frac{r^2}{(r - r_+)(r - r_-)} dr^2 - \frac{(r - r_+)(r - r_-)}{r^2} dt^2 \qquad (4.15)$$

As is suggested by the regularity of the Weyl curvature array (3.26), these singularities are removable in the manner with which we are beginning to become familiar, as can easily be seen by making the obvious Finkelstein type transformation

$$v = t + r + \frac{r_+^2}{2M} \ln |r - r_+| - \frac{r_-^2}{2M} \ln |r - r_-| \tag{4.16}$$

which leads to the form

$$ds_\perp^2 = 2\, dv\, dr - \frac{(r - r_-)(r - r_+)}{r^2}\, dv^2 \tag{4.17}$$

which is clearly well behaved in the whole range $0 < r < \infty$. In this system there is an inner static domain $0 < r < r_-$ in addition to the outer static domain as is shown in the conformal diagrams of Figure 4.7. In the limit as the charges Q and P tend to zero the inner horizon $r = r_-$ collapses down on to the curvature singularity $r = 0$ and the inner domain is squeezed out of existence, so that the manifold goes over continuously to the Schwarzschild Finkelstein extension. The way in which this limiting process takes place is illustrated in the conformal diagram of Figure 4.8 in which the freedom to adjust the conformal factor is used to represent the Schwarzschild Finkelstein extension in a manner which approaches the zero charge limit of the Reissner-Nordstrom Finkelstein extension.

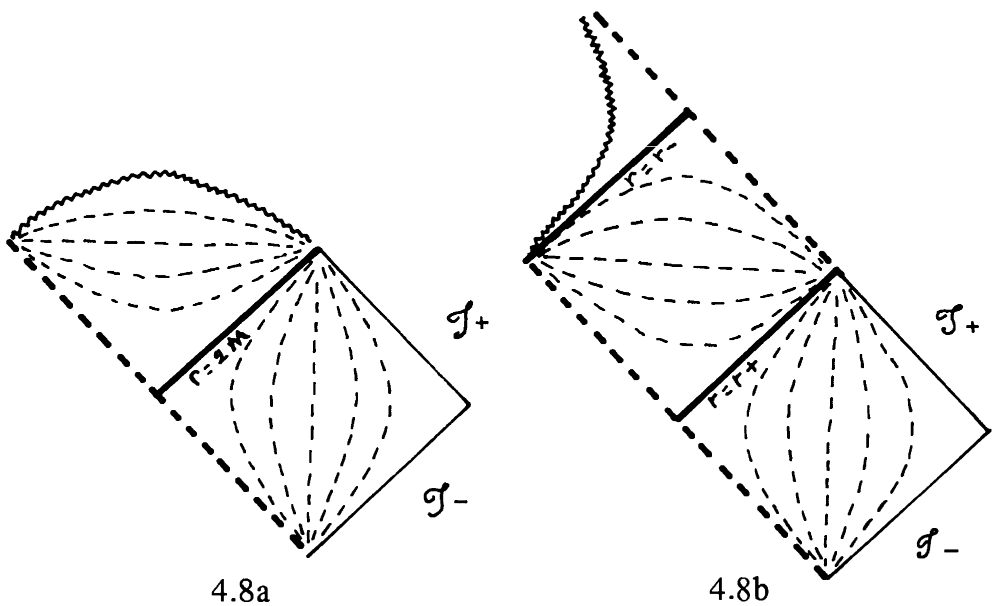


Figure 4.8. Figure 4.8a is a modified version of Figure 4.2 and Figure 4.8b is a modified version of Figure 4.7 showing how the Reissner-Nordstrom solution approaches the Schwarzschild solution as $P^2 + Q^2 \to 0$.

As in the Schwarzschild case, we can make a Finkelstein type extension not only in the forward direction but also in the backward, by introducing the outgoing null co-ordinate

$$u = t - r - \frac{r_+^2}{2M} \ln |r - r_+| + \frac{r_-^2}{2M} \ln |r - r_-| \tag{4.18}$$

which leads to the form

$$ds_{\perp}^2 = -2\,du\,dr - \frac{(r - r_+)(r - r_-)\,du^2}{r^2} \tag{4.19}$$

By patching together forward and backward going extensions of this kind we can build up what is in fact a maximal extension in the manner illustrated in the

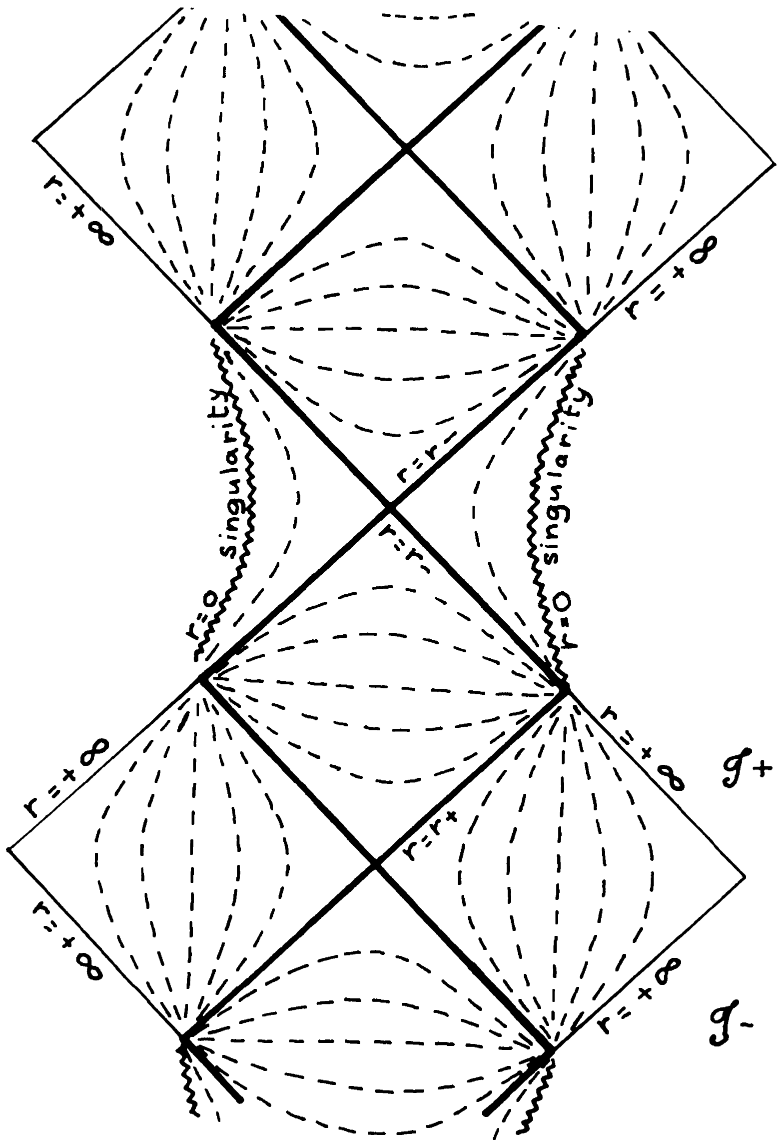


Figure 4.9. Conformal diagram of timelike 2-section with constant spherical co-ordinates θ, φ of maximally extended Riessner-Nordstrom solution with $M^2 > P^2 + Q^2$.

conformal diagram of Figure 4.9 to obtain an infinite double chain of asymptotically flat universes linked together in pairs by wormholes. This manifold was first described by Graves and Brill (1960). As in the Schwarzschild case it is possible to construct Kruskal type co-ordinate extensions to cover the crossover points of the horizons. It is possible to construct a co-ordinate system covering the whole manifold at once, but it is simpler to construct local patches adapted to the kinds of horizon $r = r_+$ and $r = r_-$ separately. The notation we are using enables us to describe both kinds of patch at once as follows. Starting as in the previous case from the double null form.

$$ds_\perp^2 = -\frac{(r - r_+)(r - r_-)}{r^2}\, du\, dv \tag{4.20}$$

where r is given implicitly as a function of u and v by

$$r + \frac{r_+^2}{2M} \ln |r - r_+| - \frac{r_-^2}{2M} \ln |r - r_-| = \tfrac{1}{2}(v - u) \tag{4.21}$$

we introduce new co-ordinates U^+, V^+ or U^-, V^- depending on whether we want to remove the singularity at $r = r_+$ or $r = r_-$, defined by

$$\left.\begin{aligned} u &= \mp \frac{r_\pm^2}{M} \ln |-U^\pm| \\ v &= \pm \frac{r_\pm^2}{M} \ln |+V^\pm| \end{aligned}\right\} \tag{4.22}$$

which leads to the form

$$ds_\perp^2 = -\left(\frac{r_\pm}{M}\right)^2 \exp\left(-\frac{2Mr}{r_\pm^2}\right) |r - r_\mp|^{(r_\pm^2 + r_\mp^2)/r_\pm^2} \frac{dU^\pm\, dV^\pm}{r^2} \tag{4.23}$$

where r is given implicitly as a function of $U^\pm$, $V^\pm$ by

$$|r - r_\pm| \exp\left(\pm\frac{2Mr}{r_\pm^2}\right) |r - r_\mp|^{-r_\mp^2/r_\pm^2} = -U^\pm V^\pm \tag{4.24}$$

and the co-ordinate t, if required, is given by

$$\exp\left(\frac{2M}{r_\pm^2} t\right) = -\frac{V^\pm}{U^\pm} \tag{4.25}$$

It can be seen that these forms are well behaved in the neighbourhood of $r = r_\pm$ respectively, (although not in the neighbourhood of $r = r_\mp$).

In the Graves and Brill manifold the time symmetric 3-dimensional space sections (t = constant) fall into two classes, as illustrated in Figure 4.10 which shows imbeddings in 3-dimensional flat space of 2-dimensional sections. The first kind, exemplified by the locuses $U^+ + V^+ = 0$, are topologically similar to the Kruskal bridges connecting two asympototically flat regions except that the throat is somewhat narrower, and can be (if the charges are larger) very much longer. The second kind, as exemplified by the locus $U^- + V^- = 0$, represents a sort of tube connecting two curvature singularities $r = 0$; only part of the tube is illustrated since the flat space imbedding breaks down before the curvature singularities are reached.

Let us pause for a moment to consider the implications of what we have seen so far for a gravitational collapse situation. We notice that a material sphere can only undergo catastrophic gravitational collapse if the condition $Q^2 + P^2 < M^2$ (or simply $Q^2 < M^2$ provided we take for granted the physical condition that there are no material magnetic monopoles) is satisfied since otherwise electromagnetic repulsion would predominate over gravitational attraction, and that it

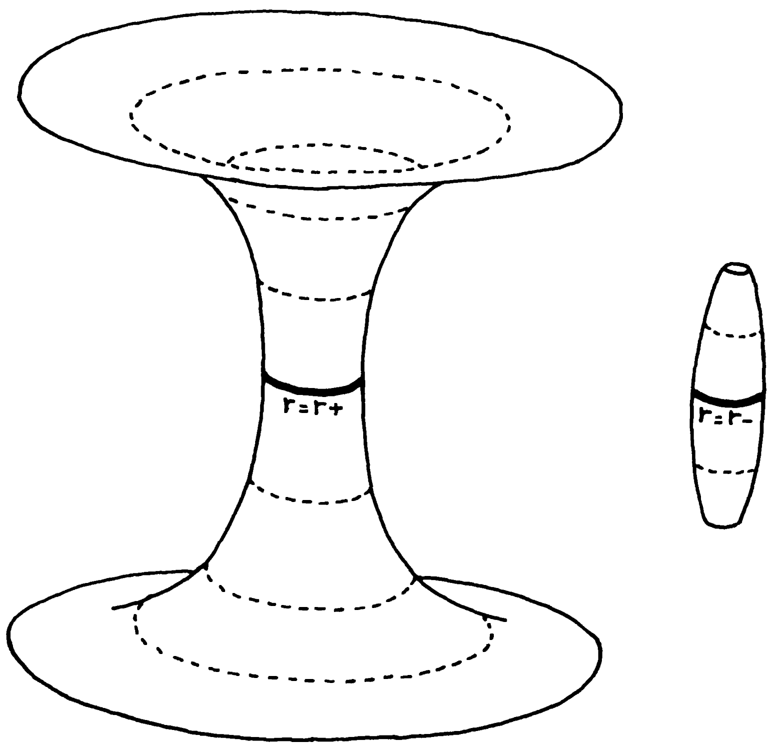


Figure 4.10. Sketches of spacelike equatorial 2-sections ($\cos\theta = 0$) of t = constant hypersurfaces in the maximally extended Reissner-Nordstrom when $M^2 > P^2 + Q^2$. The sketches are intended loosely to suggest imbeddings in 3-dimensional flat space; strictly, however, the section passing through the Killing horizon crossover point $r = r_-$ ceases to be symmetrically imbeddable when it gets too near to the curvature singularity $r = 0$.

is *precisely the same in equality* $Q^2 + P^2 < M^2$ which determines whether a horizon will be formed. This suggests what Penrose (1969) has termed the *cosmic censorship hypothesis* according to which, in any situation arising from the gravitational collapse of an astrophysical object (such as the central part of a star after a supernova explosion) starting from a well behaved initial situation, the singularities which result are hidden from outside by an event horizon i.e. that *naked singularities* which can both be approached from and seen from outside (as in the case of the singularity at $r = 0$ in the Reissner-Nordstrom solutions with $Q^2 + P^2 > M^2$) *cannot arise naturally* from a well behaved initial situation. This hypothesis does not exclude the visibility of *preexisting* singularities (such as that of the big bang in cosmology theory not to mention the so called white holes) which cannot be reached in the future by a timelike trajectory and in this sense are not completely naked.

Let us now move on to consider the special case of the Reissner-Nordstrom solutions for which the critical condition $Q^2 + P^2 = M^2$ is satisfied, i.e. which are poised between the normal hole situation $Q^2 + P^2 < M^2$ and the naked singularity situation $Q^2 + P^2 > M^2$. According to the cosmic censorship hypothesis a critical case such as this should represent a physically unattainable limit, but one which is approachable and therefore of considerable interest in the same way as for example the extreme relativistic limit is of great interest in particle scattering theory.

The metric in this limiting case simply takes the form

$$ds_{\perp}^2 = \frac{(r-M)^2}{r^2}dr^2 - \frac{r^2}{(r-M)^2}dt^2 \tag{4.26}$$

This metric form is static *both* for $r > M$ *and* for $r < M$ but it is none the less singular at $r = m$. Its extension was first discussed by myself (1966). The co-ordinate singularity is in fact the first example in the present section of a Killing horizon which is *degenerate* in the sense described in section 2, as would be expected when it is thought of as the limiting case in which the two non-degenerate Killing horizons $r = r_+$ and $r = r_-$ have coalesced.

As usual the Finkelstein type extension can be carried out without difficulty by introducing the null co-ordinate

$$v = t + (r - M) + 2M \ln |r - M| - \frac{M^2}{r - M} \tag{4.27}$$

which leads to the metric form

$$ds_{\perp}^2 = 2\,dr\,dv - \frac{(r-M)^2}{r^2}dv^2 \tag{4.28}$$

The corresponding co-ordinate patch which is well behaved over the whole range $0 < r < \infty$, $-\infty < v < \infty$, is illustrated in Figure 4.10, in which we have also shown the way in which this limit is approached by the Finkelstein extensions of the ordinary Riessner-Nordstrom metrics as $Q^2 + P^2$ approaches M^2 from below.

As usual we can also make the symmetric past extension, by introducing a co-ordinate

$$u = t - (r - M) - 2M \ln |r - M| + \frac{M^2}{r - M} \tag{4.29}$$

thus obtaining the form

$$ds_{\perp}^2 = -2\,du\,dr - \frac{(r-M)^2}{r^2}dr^2 \tag{4.30}$$

This time however there is no way of carrying out a further Kruskal type extension, because the transformation (4.27) contains not only the by now familiar logarithmic singularity but also a first order pole singularity. However it turns out that a Kruskal type extension is quite unnecessary, since the Finkelstein extensions can be fitted together in the manner illustrated in Figure 4.1.2 to form an extended manifold which is in fact maximal since (as we shall verify later) all geodesics either intersect the curvature singularity or can be completed. The situation is rather different from those which we have come across so far, because

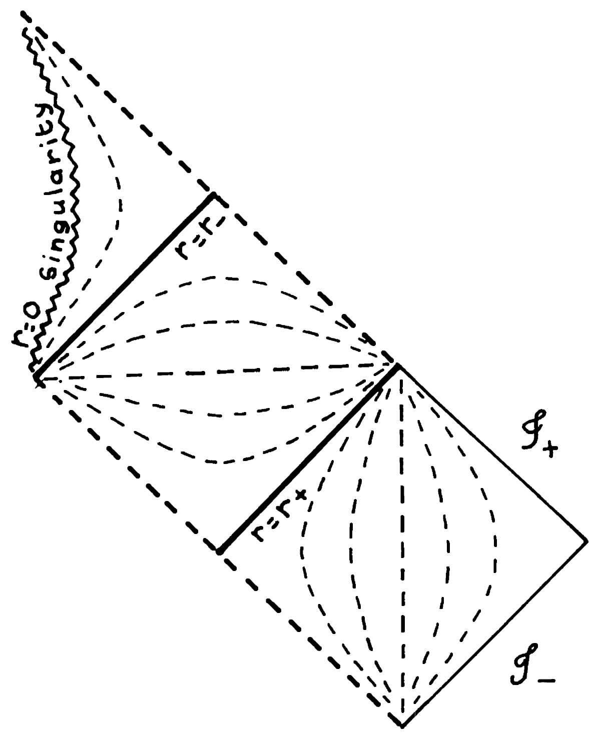


Figure 4.11. Conformal diagram of timelike 2-section with constant spherical co-ordinates θ, φ of Finkelstein-type extension of Reissner-Nordstrom solution in degenerate case when $M^2 = P^2 + Q^2$.

the completed geodesics can not only approach the outer parts of the diagram labelled $\mathscr{I}^+$ and $\mathscr{I}^-$, but can also approach the *inner* boundary points labelled x, each of which therefore represents (in a manner which is disguised by the conformal factor) infinitely distant limit in the inward direction.

The nature of the limits represented by the points x, (which take the place of the missing Kruskal crossover axes) can best be understood by considering the time symmetric space sections, t = constant. In this case the analogue of Figure 4.10 is given by Figure 4.13 which shows that instead of having a minimum (the throat) or a maximum of r in the respective cases $r > M$, $r < M$, the space sections extend indefinitely, both approaching asymptotically the same infinite spherical cylinder.

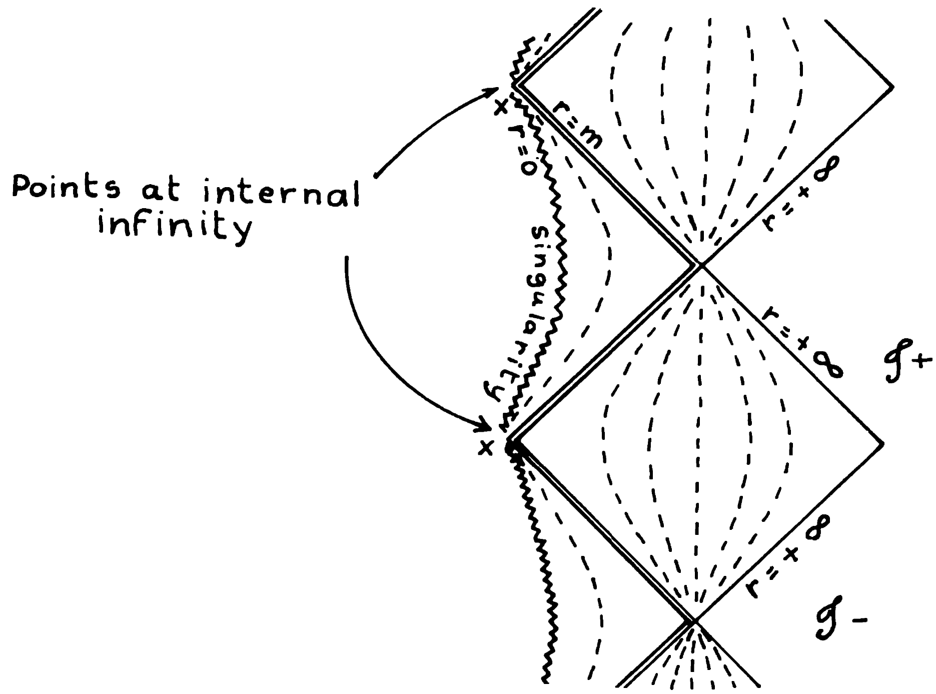


Figure 4.12. Conformal diagram of timelike 2-section with constant spherical co-ordinates θ, φ of maximal extension of Reissner-Nordstrom solution in degenerate case when $M^2 = P^2 + Q^2$.

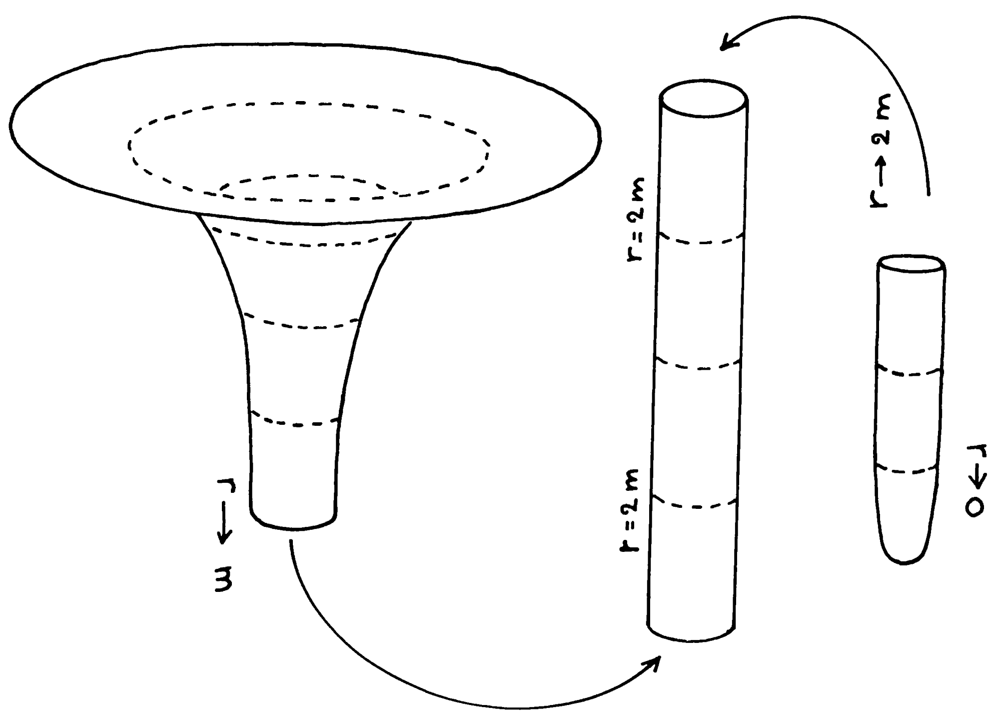
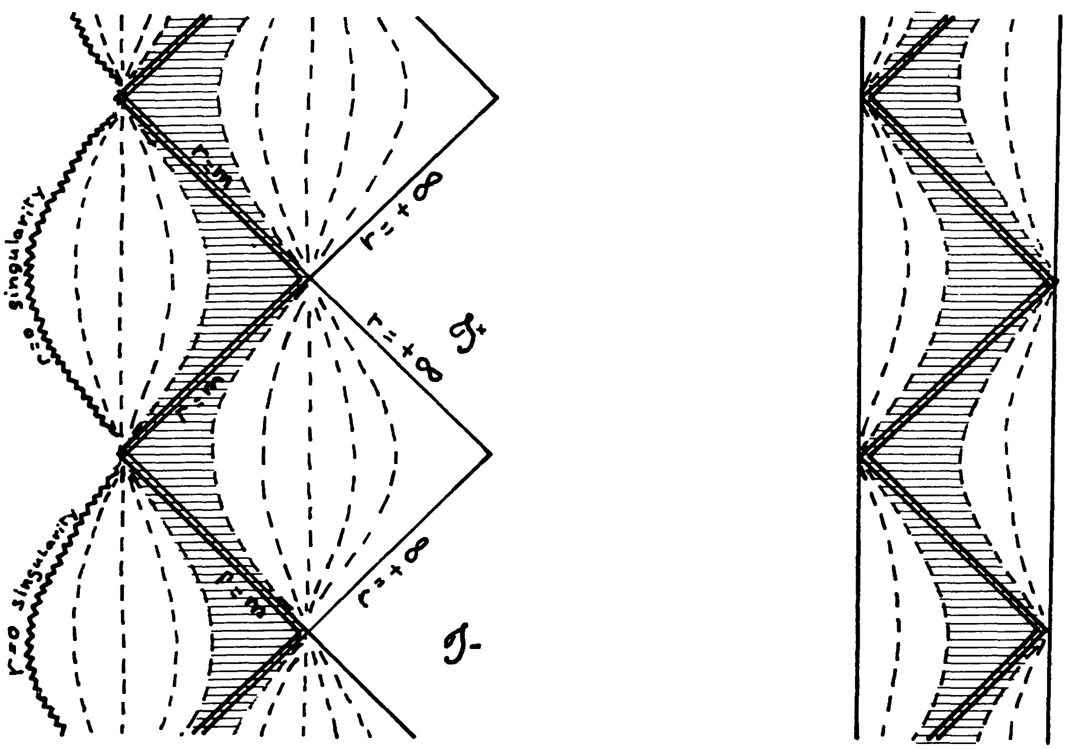


Figure 4.13. Sketches of spacelike equatorial 2-sections ($\cos\theta = 0$) of t = constant hyper surfaces in the maximally extended Reissner-Nordstrom solution in the case $M^2 = P^2 + Q^2$ and also in a Robertson-Bertotti universe, indicating how the latter represents an asymptotic limit of the former.

It is at this stage that the Robinson-Bertotti solution (3.34) enters into the discussion. We do not need to do any further work to find its global structure, since its time sections have the metric

$$ds_{\perp}^2 = (Q^2 + P^2)\left(\frac{d\lambda^2}{\lambda^2} - \lambda^2\, d\tau^2\right) \tag{4.31}$$

which, as we have already remarked is the metric of a pseudo-sphere, so that the



4.14a 4.14b

Figure 4.14. Figure 4.14a represents a modified version of Figure 4.12 and Figure 4.14b represents a modified version of Figure 2.2b. The shaded regions of the two diagrams can be made to approximate each other arbitrarily closely in the neighbourhood of the horizons, showing that the maximally extended Reissner-Nordstrom solution effectively contains a Robertson-Bertotti universe within it in the bottomless hole case when $M^2 = P^2 + Q^2$.

required extension past the singularity $\lambda = 0$ is given by (2.19) and the corresponding conformal diagram is that of Figure 2.2b. Now it can easily be seen by setting $r = \lambda + M$, $t = M^2\tau$, that the Reissner-Nordstrom solution with $Q^2 + P^2 = M^2$ approaches the Robertson-Bertotti solution in the asymptotic limit as $r \to \infty$. The limiting spherical cylinder illustrated in Figure 4.13 is in fact the same as a τ = constant cross section of a Robinson-Bertotti universe. The conformal diagrams of the $Q^2 + P^2 = M^2$ Reissner-Nordstrom solution, and of the Robinson-Bertotti solution are shown side by side for comparison in Figure 4.14. The shaded regions

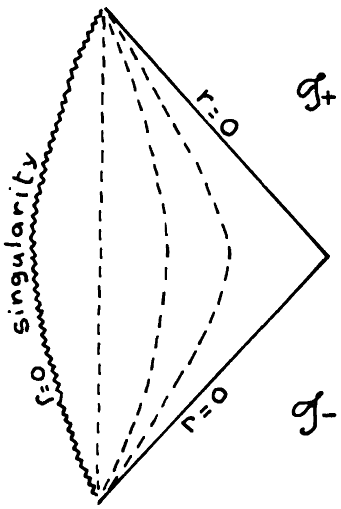


Figure 4.6. Conformal diagram of timelike 2-section with constant spherical co-ordinates θ, φ in the maximally extended Reissner-Nordstrom solution in the naked singularity case when $M^2 < P^2 + Q^2$.

can be made to coincide as closely as one pleases by adjusting the conformal factor so that they correspond to a sufficiently small range of the co-ordinate $\lambda = r - M$.

5 Derivation of the Kerr Solution and its Generalizations

Having now looked fairly thoroughly at the spherical vacuum solutions, we have clearly arrived at a stage where it would be interesting to see how the various phenomena—horizons, naked singularities etc.—which we have come across would be modified in more general non-spherical situations, particularly when angular momentum, the most obvious source of deviation from spherical symmetry, is present.

Since, as we have seen, the derivation of the spherical Schwarzschild solution is very easy (it was achieved in 1916 within a year of the publication of Einstein's Theory in 1915) one might have guessed that it would be comparatively not too difficult to derive a rotating (and hence non-static but still stationary) generalization in a fairly straightforward manner, starting from some suitably simple and natural canonical metric form. Moreover, as I plan to make clear in this section, one would have guessed correctly. Nevertheless, after forty years, repeated attempts to find a canonical form leading to a natural vacuum generalization of the Schwarzschild solution had turned up nothing (or more precisely nothing which was asymptotically flat) except the Weyl solutions, which are unfortunately static, and therefore useless in so far as showing the effect of angular momentum is concerned. Moreover the first (and so far the only) non-static pure vacuum generalization of the Schwarzschild solution was found at last by Kerr in 1963 using a method which is by no means straightforward, and which arose as a bi-product of the sophisticated Petrov-Pirani approach to gravitational radiation theory which was developed during the nineteen fifties. In consequence of this history it is still widely believed that the Kerr solution can only be derived using advanced modern techniques. There is however an elementary approach of the old fashioned kind which was rather surprisingly overlooked by the searchers in the nineteen twenties and thirties, and which I actually found myself, with the aid of hindsight, in 1967. This approach now seems so obvious that I am sure it will be clear to anyone who follows it that despite the apparent messiness of the form in which it is customarily presented, no non-static generalization of the Schwarzschild solution which may be discovered in the future can possibly be simpler in its algebraic structure than that of Kerr.

This approach starts from the observation that for practical computational purposes one of the most useful, indeed almost certainly *the* most useful, algebraic property which the Schwarzschild solution possesses as a consequence of spherical symmetry, and which one might hope to return in a simple non-spherical generalization, is that of separability of its Dalembertian wave equation and the associated integrability of its geodesic equations.

Now it is physically evident from the correspondence principle of quantum mechanics, (and it follows mathematically from standard Hamilton Jacobi theory) that integrability of the geodesics as well, obviously, as separability of the Dalembertian wave equation $\psi^{;a}{}_{;a} = 0$, will follow if the slightly more general Klein-Gordon wave equation $\psi^{;a}{}_{;a} - m^2 \psi = 0$ (where m^2 is a freely chosen constant which may be interpreted as a squared test particle mass) is separable.

Separability is of course something which depends not only on the geometry but on a particular choice of co-ordinates x^2 $(a = 0, 1, 2, 3)$, and in terms of these co-ordinates it depends less directly on the form of the ordinary covariant tensor g_{ab} defined by

$$ds^2 = g_{ab}\, dx^a\, dx^b \tag{5.1}$$

than on the form of the contravariant metric tensor g^{ab} defined in terms of the inverse co-form

$$\left(\frac{\partial}{\partial s}\right)^2 = g^{ab} \frac{\partial}{\partial x^a} \frac{\partial}{\partial x^b} \tag{5.2}$$

In terms of the contravariant metric components and of the determinant

$$g = \det(g_{ab}) = \{\det(g^{ab})\}^{-1} \tag{5.3}$$

the Klein Gordon equation can be expressed in terms of simple partial derivatives in the form

$$\psi^{-1} \frac{\partial}{\partial x^a} \sqrt{-g}\, g^{ab} \frac{\partial}{\partial x^b} \psi - m^2 \sqrt{-g} = 0 \tag{5.4}$$

The standard kind of separability takes place if substitution of the product expression

$$\psi = \Pi_i' \psi_i \tag{5.5}$$

where each function $\psi_i (i = 0, 1, 2, 3)$ is a function of just the single variable x^i, causes the left hand side of (5.4) to split up into four independent single variable ordinary differential equations, expressed in terms of four independent freely chosen constants of which $\bar{m}^2$ is one.

To see how this works out in the spherical case, we note that the inverse metric corresponding to our general spherical canonical form (3.2) is

$$\left(\frac{\partial}{\partial s}\right)^2 = \frac{1}{r^2}\left\{(1-\mu^2)\left(\frac{\partial}{\partial \mu}\right)^2 + \frac{1}{1-\mu^2}\left(\frac{\partial}{\partial \varphi}\right)^2 + \Delta_r\left(\frac{\partial}{\partial r}\right)^2 - \frac{Z_r^2}{\Delta_r}\left(\frac{\partial}{\partial t}\right)^2\right\} \tag{5.6}$$

and hence using (3.3), the Klein Gordon equation takes the form

$$\frac{r^2}{Z_r}\left\{\psi^{-1}\frac{\partial}{\partial \mu}(1-\mu^2)\frac{\partial \psi}{\partial \mu} + \psi^{-1}\frac{1}{1-\mu^2}\frac{\partial^2}{\partial \varphi^2}\psi - m^2 r^2\right\}$$
$$+ \psi^{-1}\frac{\partial}{\partial r}\frac{r^2\Delta_r}{Z_r}\frac{\partial \psi}{\partial r} - \psi^{-1}\frac{rZ_r}{\Delta_r}\frac{\partial^2 \psi}{\partial t^2} = 0 \tag{5.7}$$

which has solutions of the form

$$\psi = R_r P_n^l(\mu) e^{in\varphi} e^{i\omega t} \tag{5.8}$$

where l, n, ω are separation constants (of which l and n must be integers if the solution is to be regular) and $P_n^l(\mu)$ is a solution of the (l, n) associated Legendre equation (and is thus an associated Legendre-polynomial in the regular case) and where R_r is a solution of the equation

$$R_r^{-1}\frac{d}{dr}\frac{r^2\Delta_r}{Z_r}\frac{dR_r}{dr} + \frac{\omega^2 r^2 Z_r}{\Delta_r} + \frac{r^2}{Z_r}[l(l+1) - m^2 r^2] = 0 \tag{5.9}$$

What we want to do now is find the simplest possible non-static generalization of the canonical coform (5.6), including in our criteria for simplicity not only the most obvious requirement of all, namely that the manifest symmetry property of stationarity and axisymmetry represented by the ignorability of the co-ordinates t and φ be retained but also the requirement that the separability property of the corresponding Klein Gordon equation be retained. There is a third obvious simplicity property of the coform (5.6) which can be retained without prejudice to the other two, (and which greatly simplifies the computation of the Riemann tensor etc.) namely the fact that it determines a natural canonical orthonormal tetrad, such that two of the tetrad vectors contribute in the separation only to the terms independent of r, while the other two contribute only to the terms independent of μ.

This leads us to try the canonical coform

$$\left(\frac{\partial}{\partial s}\right)^2 = \frac{1}{Z}\left\{\Delta_\mu\left(\frac{\partial}{\partial\mu}\right)^2 + \frac{1}{\Delta_\mu}\left[Z_\mu\frac{\partial}{\partial t} + Q_\mu\frac{\partial}{\partial\varphi}\right]^2\right\} + \frac{1}{Z}\left\{\Delta_r\left(\frac{\partial}{\partial r}\right)^2 - \frac{1}{\Delta_r}\left[Z_r\frac{\partial}{\partial t} + Q_r\frac{\partial}{\partial\varphi}\right]^2\right\} \tag{5.10}$$

where Δ_μ, Z_μ, Q_μ are functions of μ only, and where Δ_r, Z_r, Q_r are functions of r only, and the form of the conformal factor Z remains to be determined.

It is to be observed that the factors Δ_μ are redundant, since the one in the first term could be eliminated by renormalizing μ as a function of itself while the one in the second term could be eliminated by a proportional readjustment of Z_μ and Δ_μ. The same applies to the factor Δ_r. These factors have been included explicitly however firstly because they are suggested by the canonical spherical coform (5.6) (to which (5.10) reduces when one sets $\Delta_\mu = 1 - \mu^2$ with $\mu = 1$, $Q_\mu = Q_r = 0$ and $Z = r^2$) but also for the more compelling reason that the freedom of adjustment of Δ_μ and Δ_r can be used to achieve considerable simplification of the form of Z required to achieve separability. In any case they are arranged so as to cancel out of the determinant which is simply given by

$$\sqrt{-g} = \frac{Z^2}{|Z_r Q_\mu - Z_\mu Q_r|} \tag{5.11}$$

Let us now investigate the conditions which must be imposed on Z to achieve separability. Using (5.10) and (5.11) we obtain the Klein Gordon equation in the form

$$\psi^{-1}\left\{\frac{\partial}{\partial\mu}\frac{\sqrt{-g}}{Z}\Delta_\mu\frac{\partial}{\partial\mu} + \frac{\partial}{\partial r}\frac{\sqrt{-g}}{Z}\Delta_r\frac{\partial}{\partial r} - \frac{\sqrt{-g}}{Z}m^2 Z\right\}\psi + \psi^{-1}\left\{\frac{\partial}{\partial\varphi}Q_\mu + \frac{\partial}{\partial t}Z_\mu\right\}\frac{1}{\Delta_\mu}\left(\frac{\sqrt{-g}}{Z}\right)\left(Q_\mu\frac{\partial}{\partial\varphi} + Z_\mu\frac{\partial}{\partial t}\right)\psi - \psi^{-1}\left(\frac{\partial}{\partial\varphi}Q_r + \frac{\partial}{\partial t}Z_r\right)\frac{1}{\Delta_r}\left(\frac{\sqrt{-g}}{Z}\right)\left(Q_r\frac{\partial}{\partial\varphi} + Z_r\frac{\partial}{\partial t}\right)\psi = 0 \tag{5.12}$$

It is clear that if these terms are to separate, the factor $Z^{-1}\sqrt{-g}$ which occurs in each one must depend on r and μ only as a product of single variable functions which can be absorbed into Δ_r and Δ_μ (using our freedom to rescale these functions) so as to reduce the factor $Z^{-1}\sqrt{-g}$ to unity. Thus we are led to choose the definition

$$Z = Z_r Q_\mu - Z_\mu Q_r \tag{5.13}$$

for the conformal factor. This is still not quite sufficient for separability except in the case of the pure Dalembertian wave equation, since there remains the mass term which now takes the form $m^2 Z$. In this expression we have not only made no provision for cross terms between the non-ignorable co-ordinates, i.e. terms proportional to $(\partial/\partial r)(\partial/\partial\mu)$, which would obviously destroy the separability, but we have also, in accordance with our principle of maximum simplicity, excluded all other cross terms *except* those directly between the ignorable co-ordinates, i.e. those proportional to $(\partial/\partial t)(\partial/\partial\varphi)$ whose presence is essential if we are to have non-zero angular momentum. To achieve complete separability this term also must split up into the sum of two parts each depending on only one variable, i.e. Z must have the algebraic form

$$Z = U_\mu + U_\lambda \tag{5.14}$$

where U_μ depends only on μ and U_λ only on λ. From the expression (5.13) we see that this requirement will be satisfied if and only if

$$\frac{dZ_r}{dr}\frac{dQ_\mu}{d\mu} - \frac{dZ_\mu}{d\mu}\frac{dQ_r}{dr} = 0 \tag{5.15}$$

There are basically two ways in which this can be satisfied: a more general case in which either both Z_r and Z_μ or both Q_r and Q_μ are constants, and a more special case in which at least one of the four functions is zero, or can be reduced to zero by a form preserving co-ordinate transformation in which φ and t are replaced by linear combinations of themselves. In the more general case we can take it without loss of algebraic generality that it is the Q's which are constants. Thus replacing Q_r, Q_μ by constant C_r, C_μ respectively we obtain the basic separable canonical form

$$\left(\frac{\partial}{\partial s}\right)^2 = \frac{1}{[C_\mu Z_r - C_r Z_\mu]}\left\{\Delta_\mu\left(\frac{\partial}{\partial\mu}\right)^2 + \frac{1}{\Delta_\mu}\left[C_\mu\frac{\partial}{\partial\varphi} + Z_\mu\frac{\partial}{\partial t}\right]^2\right\}$$
$$+\frac{1}{[C_\mu Z_r - C_r Z_\mu]}\left\{\Delta_r\frac{\partial}{\partial r} - \frac{1}{\Delta_r}\left[C_r\frac{\partial}{\partial\varphi} + Z_r\frac{\partial}{\partial t}\right]^2\right\} \tag{5.16}$$

The original expression (5.10) from which we started had algebraic symmetry not only between r and μ (apart from a sign change) but also between t and φ. In the coform (5.16), in which the symmetry between t and φ has been lost, we have chosen to set Q_r and Q_μ rather than Z_r and Z_μ constant in order that it should include the spherical coform (5.6), in which t and φ have their usual quite distinct interpretations in the appropriate limit (that is to say when $C_\mu = 1$, $C_r = 0, Z_\mu = 0, Z_r = r^2, \Delta_\mu = 1 - \mu^2$).

In the alternative way of satisfying (5.15) it can be arranged either that Z_μ is zero, and Q_r a constant or that Q_μ is zero and Z_r a constant. I have studied the resulting canonical forms (Carter 1968) and found that there are no vacuum solutions except those which are in fact special cases of (5.16) and therefore we shall not consider these alternative possibilities any further here.

The steps by which we have arrived at the canonical separable coform 5.16 are so simple that it is really quite surprising that it was not found by any researcher of the nineteen thirties, such as for example Eisenhart or Robertson who both worked on separability of wave equations in curved spacetime (cf. Robertson 1927, Eisenhart 1933). The corresponding metric determinant is given by

$$\sqrt{-g} = |C_\mu Z_r - C_r Z_\mu| \tag{5.17}$$

and the covariant metric form is given by

$$\boxed{ds^2 = [C_\mu Z_r - C_r Z_\mu]\left\{\frac{dr^2}{\Delta_r} + \frac{d\mu^2}{\Delta_\mu}\right\} + \frac{\Delta_\mu[C_r dt - Z_r d\varphi]^2 - \Delta_r[C_\mu dt - Z_\mu d\varphi]^2}{[C_\mu Z_r - C_r Z_\mu]}} \tag{5.18}$$

It is very easy to derive the vacuum solutions corresponding to this canonical form, using the same method as in the spherical case, working in terms of the natural canonical orthogonal tetrad of forms,

$$\omega^{(1)} = \left\{\frac{C_\mu Z_r - C_r Z_\mu}{\Delta_r}\right\}^{1/2} dr \tag{5.19}$$

$$\omega^{(2)} = \left\{\frac{C_\mu Z_r - C_r Z_\mu}{\Delta_\mu}\right\} d\mu \tag{5.20}$$

$$\omega^{(3)} = \left\{\frac{\Delta_\mu}{C_\mu Z_r - C_r Z_\mu}\right\}^{1/2} [C_\mu\, dt - Z_\mu\, d\varphi] \tag{5.21}$$

$$\omega^{(0)} = \left\{\frac{\Delta_r}{C_\mu Z_r - C_r Z_\mu}\right\} [C_\mu\, dt - Z_\mu' d\varphi] \tag{5.22}$$

The complete set of solutions for the four unknown variable functions Z_r, Z_μ, Δ_r, Δ_μ in terms of the constants C_μ, C_r turns out to be remarkably simple: the variables functions are all required to be quadratic polynomials, whose coefficients are subject to a few linear restraints. These quadratic functions can be stated in a very compact form if we temporarily exclude the special cases which arrive when either C_μ or C_r is zero (in the same way that it was convenient to give the general spherical solutions in a form in which r was assumed to be a variable, dealing separately with the special Robinson Bertotti case). When C_r and C_μ are non-zero they may both be normalized to unity by a change of scale of ψ and t. We shall impose this normalization only for C_μ which is unity in the spherical canonical form, but in order to be able to go over smoothly to the spherical limit we shall retain C_r as a freely renormalizable parameter which—with a certain prescience—we shall relable a. Thus setting

$$C_\mu = 1 \tag{5.23}$$

$$C_r = a \tag{5.24}$$

we can express the general solution (excluding special limiting cases where C_μ or C_r is zero) in the concise form

$$Z_r = r^2, \qquad Z_\mu = -a\mu^2 \tag{5.25}$$

$$\Delta_r = hr^2 - 2Mr + pa^2 \tag{5.26}$$

$$\Delta_\mu = -h\mu^2 - 2q\mu + p \tag{5.27}$$

where h, M, q and p are arbitrary parameters. (This would look more symmetric between r and μ if we chose to use the available co-ordinate freedom to adjust a to unity.) The complete set of solutions, including the special limiting cases, can be obtained from these by first making linear co-ordinate transformations in which r and $a\mu$ are replaced respectively by $cr + d$ and $a\mu + f$, and then allowing the parameters c, d, a, f to vary freely over *all* values in the resulting somewhat more complicated expressions including the values $c = 0$ and $a = 0$ for which the original transformation would have been singular. The solutions, (5.25), (5.26), (5.27) represent a vast class of vacuum metrics most of which are unacceptable for our present purposes for global geometric reasons, examples being the Taub-N.U.T. space (cf. Misner, 1963) which is included in the family as one of the special limiting cases and the more general family discovered by Newman and Demianski (1966).

In order to understand the global geometry of these solutions, we note that the canonical metric form (5.18) has singularities whenever Δ_r and Δ_μ are zero which are very similar respectively to the Killing horizon type and symmetry axis type co-ordinate singularities with which we are already familiar in the spherical case. Moreover while we can conceive that r should be able to vary across a region where Δ_r vanishes in an extended metric, it is clear that if the metric is to retain the correct signature μ must be *absolutely restricted to the range in which* Δ_μ *is non-negative.* Now since we wish the co-ordinate to represent an azimuthal variation between extending from a south polar to a north polar symmetry axis, it is clear that we must require that the quadratic function Δ_μ be positive in a restricted range, whose limits will we hope—if things work out—turn out to be the symmetry south and north polar symmetry axes. To achieve these conditions it is clear that we need $h > 0$ and also $hp + q^2 > 0$. Now we know that if the symmetry axis is to be well behaved the overall coefficient of $d\varphi^2$ in the metric form (5.18) must be zero that is to say Z_μ must vanish for the same values of μ as Δ_μ. Now the values of Z_μ and Z_r may be adjusted to the extent of the addition of *the same* arbitrary constant to both of them (by replacing t by a constant coefficient linear combination of t and φ) but it is clear that even with these adjustments the zeros of Z_μ will always occur for equal and opposite values of μ, and can therefore match the zeros of Δ_μ only if q is zero. Thus we see that the restrictions

$$q = 0 \tag{5.28}$$

$$h > 0 \tag{5.29}$$

$$p > 0 \tag{5.30}$$

are necessary for regular angular behaviour of the solutions. When the two latter conditions are satisfied we can make scale changes of μ and r so as to obtain $h = p = 1$ thereby ensuring that μ varies over the conventional co-ordinate range $-1 < \mu < 1$. In doing so we use up our co-ordinate freedom to adjust a which thereafter becomes a geometrically well determined parameter). The adjustment of t and φ necessary to ensure that the zeros of Z_μ coincide with those of Δ_μ leads us to replace the forms (5.25) of the solution (which were previously adjusted for maximum algebraic simplicity) by

$$Z_r = r^2 + a^2 \tag{5.31}$$

$$Z_\mu = a(1 - \mu^2) \tag{5.32}$$

while the other conditions we have imposed cause the expressions (5.26), (5.27) to reduce to the standard expressions

$$\Delta_r = r^2 - 2Mr + a^2 \tag{5.33}$$

$$\Delta_\mu = 1 - \mu^2 \tag{5.34}$$

On substituting these expressions together with (5.23) and (5.24) back into the canonical form (5.18) and making the substitution

$$\mu = \cos\theta \tag{5.35}$$

we obtain the solution of Kerr (1963) in the standard co-ordinate system introduced by Boyer and Lindquist (1966), which takes the explicit form

$$ds^2 = (r^2 + a^2\cos^2\theta)\left\{\frac{dr^2}{r^2 - 2Mr + a^2} + d\theta^2\right\}$$
$$+ \frac{\sin^2\theta\,[a\,dt - (r^2 + a^2)\,d\varphi]^2 - (r^2 - 2Mr + a^2)[dt - a(1 - \mu^2)\,d\varphi]^2}{r^2 + a^2\cos^2\theta} \tag{5.36}$$

where we have retained the grouping of the terms to make manifest the canonical tetrad with respect to which the separability (which of course is not affected by the replacement of μ by a function of itself) takes place. In terms of this tetrad as given explicitly by (5.19), (5.20), (5.21), (5.22) and of the curvature forms defined in section (3), the Weyl tensor, which in this case is equal to the Riemann tensor, may be presented in the form

$$\begin{aligned}
\Omega^{(1)}{}_{(2)} &= -I_1\,\omega^{(1)} \wedge \omega^{(2)} - I_2\,\omega^{(0)} \wedge \omega^{(3)} \\
\Omega^{(0)}{}_{(3)} &= -I_1\,\omega^{(0)} \wedge \omega^{(3)} + I_2\,\omega^{(1)} \wedge \omega^{(2)} \\
\Omega^{(0)}{}_{(1)} &= 2I_1\,\omega^{(1)} \wedge \omega^{(0)} - 2I_2\,\omega^{(2)} \wedge \omega^{(3)} \\
\Omega^{(3)}{}_{(2)} &= 2I_1\,\omega^{(2)} \wedge \omega^{(3)} + 2I_2\,\omega^{(1)} \wedge \omega^{(0)} \\
\Omega^{(0)}{}_{(2)} &= I_1\,\omega^{(2)} \wedge \omega^{(0)} - I_2\,\omega^{(1)} \wedge \omega^{(3)} \\
\Omega^{(3)}{}_{(1)} &= I_1\,\omega^{(1)} \wedge \omega^{(3)} + I_2\,\omega^{(2)} \wedge \omega^{(0)}
\end{aligned} \tag{5.37}$$

where

$$I_1 = Mr\,\frac{(r^2 - 3a^2\mu^2)}{(r^2 + a^2\mu^2)^3} \tag{5.38}$$

$$I_2 = Ma\mu\,\frac{(3r^2 - a^2\mu^2)}{(r^2 + a^2\mu^2)^3} \tag{5.39}$$

It is obvious (to an expert) from this array that the canonical separation tetrad is also a canonical Petrov tetrad, and that the Weyl tensor is of Petrov type D. It was by searching for vacuum solutions with type D Weyl tensors that Kerr originally found this metric.

Let us now move on to consider the electromagnetic generalization of the method we have just applied. The obvious thing to do is to seek a generalization of the spherical canonical form (3.32) to canonical form of the electromagnetic potential A in a separable background metric of the canonical form (5.18) which will be such that not only the ordinary Klein Gordon equation but also its electromagnetic generalization is separable, thus ensuring (as a consequence of Hamilton-Jacobi theory, or from a physical point of view by the correspondence principle) that not only geodesics but also charged particle orbits will be integrable. In terms of an electromagnetic field potential

$$A = A_a\,dx^a \tag{5.40}$$

the electromagnetic Klein-Gordon equation takes the form

$$\psi^{-1}\left(\frac{\partial}{\partial x^a} - ieA_a\right)\sqrt{-g}\,g^{ab}\left(\frac{\partial}{\partial x^a} - ieA_a\right)\psi - m^2\psi = 0 \tag{5.41}$$

In order not to introduce unnecessary cross terms we shall start off by requiring that A contain only the same two components as were necessary in the spherical case, i.e. that it has the form

$$A = A_3\, d\varphi + A_0\, dt \tag{5.42}$$

where A_3 and A_0 are scalars, independent of φ and t, whose functional dependence on μ and r remains to be determined. On substituting this expression into (3.41), using the canonical co-form (5.18) we obtain

$$\begin{aligned}
&\psi^{-1}\left\{\frac{\partial}{\partial\mu}\Delta_\mu\frac{\partial}{\partial\mu}+\frac{\partial}{\partial r}\Delta_r\frac{\partial}{\partial r}-m^2(C_\mu Z_r - C_r Z_\mu)\right\}\psi \\
&+\psi^{-1}\left\{\frac{\partial}{\partial t}Z_\mu+\frac{\partial}{\partial\varphi}C_\mu+ieX_\mu\right\}\frac{1}{\Delta_\mu}\left\{\frac{\partial}{\partial t}Z_\mu+\frac{\partial}{\partial\varphi}C_\mu+ieX_\mu\right\}\psi \\
&-\psi^{-1}\left\{\frac{\partial}{\partial t}Z_r+\frac{\partial}{\partial\varphi}C_r-ieX_r\right\}\frac{1}{\Delta_r}\left\{\frac{\partial}{\partial t}Z_r+\frac{\partial}{\partial\varphi}C_r-ieX_r\right\}\psi = 0
\end{aligned} \tag{5.43}$$

where we have set

$$\left.\begin{aligned} A_0 Z_r + A_3 C_r &= X_r \\ A_0 Z_\mu + A_3 C_\mu &= -X_\mu \end{aligned}\right\} \tag{5.44}$$

In order for the equation to separate we clearly need to choose A_0 and A_3 in such a way that X_μ and X_r are respectively functions of μ and r only. This can be done simply by taking arbitrary functions X_r, X_μ of r and μ respectively and solving (5.44) for A_0 and A_3. Thus we obtain the expression

$$\boxed{A = \frac{X_r(C_\mu\, dt - Z_\mu\, d\varphi) + X_\mu(C_r\, dt - Z_r\, d\varphi)}{C_\mu Z_r - C_r Z_\mu}} \tag{5.45}$$

for the canonical separable vector potential associated with the form (5.18). We can now go on to solve the source free Einstein-Maxwell equations for the forms (5.45) and (5.18) in conjunction, using the method described in section 3. As in the pure vacuum case the functions Δ_μ, Δ_r, Z_r, Z_μ turn out to just quadratic polynomials, and the new functions X_r, X_μ are even simpler—they are linear. Again there is a large class of solutions most of which have undesirable global behaviour, and from which a small subclass can be selected by requiring that μ should be a well behaved azimuthal angle co-ordinate. We shall not repeat this selection procedure but merely present the final form of the well behaved subclass, using the same normalization conditions as in our presentation of the vacuum Kerr solutions. Thus the solutions for which C_r or C_μ are non-zero are given directly, setting

$$C_\mu = 1 \tag{5.46}$$

$$C_r = a \tag{5.47}$$

by

$$X_r = Qr \tag{5.48}$$

$$X_\mu = P\mu \tag{5.49}$$

$$Z_r = r^2 + a^2 \tag{5.50}$$

$$Z_\mu = a(1 - \mu^2) \tag{5.51}$$

$$\Delta_r = r^2 - 2Mr + a^2 + Q^2 + P^2 \tag{5.52}$$

$$\Delta_\mu = 1 - \mu^2 \tag{5.53}$$

where in addition to the parameters M and a which by comparison with the Lenz-Thirring solution can be seen to represent mass and angular momentum per unit respectively, we now have two more parameters Q and P which represent electric and magnetic monopole moments. This solution was first obtained by Newman and his co-workers (1965) using a guessing method based on an algebraic trick. Written out explicitly, and again setting $\mu = \cos\theta$, this Kerr-Newman solution takes the standard form

$$ds^2 = (r^2 + a^2\cos^2\theta)\left\{\frac{dr^2}{r^2 - 2Mr + a^2 + Q^2 + P^2} + d\theta^2\right\}$$
$$+ \frac{\sin^2\theta[a\,dt - (r^2 + a^2)\,d\varphi]^2 - (r^2 - 2Mr + a^2 + Q^2 + P^2)[dt - a\sin^2\theta\,d\varphi]^2}{r^2 + a^2\cos^2\theta} \tag{5.54}$$

with

$$A = \frac{Qr[dt - a\sin^2\theta\,d\varphi] + P\cos\theta\,[a\,dt - (r^2 + a^2)\,d\varphi]}{r^2 + a^2\cos^2\theta} \tag{5.55}$$

[These solutions can be obtained from the general solution [A] (Carter 1968) by making the restrictions $q = 0$, $h > 0$, $p > 0$ and setting $r = \lambda$, $a\cos\theta = \mu$, $\varphi = a\psi$, $t = \chi + a^2\psi$ where $a^2 = p$.] The metric (but not of course the field) is invariant under a duality rotation in which a and b are altered in such a way as to preserve the value of the sum of their squares. The Ricci tensor for this metric can be expressed in terms of the canonical tetrad as

$$R_{ab} = \frac{Q^2 + P^2}{r^2 + a^2\mu^2}\,[\omega_a^{(0)}\omega_b^{(0)} + \omega_a^{(3)}\omega_b^{(3)} + \omega_a^{(2)}\omega_b^{(2)} - \omega_a^{(1)}\omega_b^{(1)}] \tag{5.56}$$

and the Weyl tensor has the same basic form (5.37) as in the vacuum case, but this time with the more general expressions

$$I_1 = \frac{Mr[r^2 - 3a^2\mu^2] - (Q^2 + P^2)(r^2 - a^2\mu^2)}{(r^2 + a^2\mu^2)^3} \tag{5.57}$$

$$I_2 = \frac{Ma\mu[3r^2 - a^2\mu^2] - 2(Q^2 + P^2)a\mu r}{(r^2 + a^2\mu^2)^3} \tag{5.58}$$

for the coefficients.

We now come to the special case—the generalization of the Robinson-Bertotti solution—which has been left out so far. These special cases can be obtained from the general form by making the linear co-ordinate transformation $r \to cr + k$ and then taking the singular limit when c tends to zero. Again we select a subset from a much wider class by the requirement that μ should behave as a regular azimuthal co-ordinate. The solutions could be expressed algebraically in terms of the general canonical form (5.18) but it would be geometrically misleading to do so since t and φ would come the wrong way round. Instead we express the solution directly in an alternative and in this case geometrically more enlightening canonical form as follows. [These solutions can be obtained from the form $[\hat{B}(-)]$ (Carter 1968) in the subcase $q = 0$, $h > 0$, $k^2 > Q^2 + P^2$ by adjusting m to zero, n to unity, setting $a^2 = k^2 - Q^2 - P^2$ and replacing μ by $a\mu$, ψ by t and $a\chi$ by $(a^2 + k^2)\varphi$]

$$ds^2 = (a^2\mu^2 + k^2)\left\{\frac{d\lambda^2}{\Delta_\lambda} + \frac{d\mu^2}{\Delta_\mu} - \Delta_\lambda\, d\tau^2\right\}$$
$$+ \frac{\Delta_\mu}{a^2\mu^2 + k^2}\,[(a^2 + k^2)\,d\varphi - 2ak\lambda\, d\tau]^2 \tag{5.59}$$

with

$$A = Q\lambda\, d\tau + P_\mu\,\frac{[(a^2 + k^2)\,d\varphi - 2ak\lambda\, d\tau]}{a^2\mu^2 + k^2} \tag{5.60}$$

where

$$\left.\begin{aligned}\Delta_\mu &= 1 - \mu^2\\ \Delta_\lambda &= \lambda^2 + n\end{aligned}\right\} \tag{5.61}$$

where Q, P, a, n are independent parameters, of which the last, n, can be adjusted to zero by suitable but non-trivial form preserving co-ordinate transformations using the fact that the solution has a 4-parameter isometry group (of which only 2 degrees of freedom—those corresponding to the ignorability of t and φ—are manifest) under which the hypersurfaces on which μ is constant are homogeneous and partially isotropic. The parameter k is not independent of the others, but must satisfy

$$k^2 = a^2 + Q^2 + P^2 \tag{5.62}$$

[These solutions are not fundamentally new since they could have been obtained from Taub-N.U.T. space (in the case where the "mass" term is zero) by analytic continuation in complex co-ordinate planes—the parameter k turning up as the analogue of the standard N.U.T. parameter l.]

In a more explicit form, having chosen $n = 0$, and having set $\mu = 0$, these solutions can be expressed as

$$ds^2 = [a^2(1 + \cos^2\theta) + Q^2 + P^2] \left\{\frac{d\lambda^2}{\lambda^2} + d\theta^2 - \lambda^2\, d\tau^2\right\}$$
$$+ \frac{\sin^2\theta\, [(2a^2 + Q^2 + P^2)\, d\varphi - 2a(a^2 + Q^2 + P^2)^{1/2}\, \lambda\, d\tau]^2}{[a^2(1 + \cos^2\theta) + Q^2 + P^2]} \tag{5.63}$$

with

$$A = Q\lambda\, d\tau + P\cos\theta \frac{[(2a^2 + Q^2 + P^2)\, d\varphi - 2a(a^2 + Q^2 + P^2)^{1/2}\, \lambda\, d\tau]}{[a^2(1 + \cos^2\theta) + Q^2 + P^2]} \tag{5.64}$$

which clearly reduces to the Robinson-Bertotti metric when a is set equal to zero.

The derivation of the source-free solutions of the separable form (5.18) with (5.45) worked out so well, and the separability property is so valuable for subsequent applications, that I was tempted to go on and search for more general solutions, in which for example, a perfect fluid is present. Unfortunately the separability conditions are in fact very restrictive—so much so that it is rather remarkable that there are any vacuum solutions at all—and nothing that was of any obvious physical interest turned up (not that I would by any means claim to have exhausted the subject). However, there was one generalization which came out at once, namely to solutions with no material sources but in which a cosmological Λ term is present. Although I don't think there is much physical justification for believing a non-zero Λ term, it is perhaps worth quoting the result as a geometrical curiosity. When a Λ term is present the solutions are still polynomials, X_r and X_μ being linear and Z_r and Z_μ being quadratic, as before, but now Δ_r and Δ_μ are no longer quadratic but quartic. As before, the solutions which I quote below are selected out of a much wider class by the requirement that μ should behave as a regular and azimuthal angle co-ordinate, which ensures that the solutions tend asymptotically this time not to Minkowski space as before but to de-Sitter space. The pure vacuum solutions are given by

$$ds^2 = (r^2 + a^2\cos\theta)\left\{\frac{dr^2}{\Delta_r} + \frac{d\theta^2}{1 - \frac{1}{3}a^2\Lambda\cos^2\theta}\right\}$$
$$+ \sin^2\theta\left(\frac{1 - \frac{1}{3}a^2\Lambda\cos^2\theta}{r^2 + a^2\cos^2\theta}\right)\left[\frac{a\, dt - (r^2 + a^2)\, d\varphi}{1 - \frac{1}{3}a^2\Lambda}\right]^2$$
$$- \frac{\Delta_r}{r^2 + a^2\cos^2\theta}\left[\frac{dt - a\sin^2\theta\, d\varphi}{1 - \frac{1}{3}a^2\Lambda}\right]^2 \tag{5.65}$$

where

$$\Delta_r = \tfrac{1}{3}\Lambda(r^4 + a^2r^2) + r^2 - 2Mr + a^2 \tag{5.66}$$

A straightforward electromagnetic generalization exists, but I shall not bother to write it out here. [This solution can be obtained from the general solution [A] (Carter 1968) by setting $q = 0, p = a^2, h = 1 + \frac{1}{3}a^2\Lambda$ with the restriction $1 - \frac{1}{3}a^2\Lambda > 0$, and setting $r = \lambda$, $a \cos\theta = \mu$, $\varphi = a(1 - \frac{1}{3}a^2\Lambda)\psi$ and $t = (1 - \frac{1}{3}a^2\Lambda)(\chi + a^2\psi)$].

This solution has the property that it can be expressed in the form

$$ds^2 = ds_0^2 + 2Mr\left[\frac{d\tilde{t} - a\sin^2\theta\, d\tilde{\varphi}}{(1 - \frac{1}{3}a^2\Lambda)(r^2 + a^2\cos^2\theta)} - \frac{dr}{(r^2 + a^2)(1 + \frac{1}{3}\Lambda r^2)}\right]^2 \tag{5.67}$$

where

$$ds_0^2 = (r^2 + a^2\cos^2\theta)\left\{\frac{dr^2}{(a^2 + r^2)(1 + \frac{1}{3}\Lambda r^2)} + \frac{d\theta^2}{1 - \frac{1}{3}a^2\Lambda\cos^2\theta}\right\} + \sin^2\theta\left(\frac{1 - \frac{1}{3}a^2\Lambda\cos^2\theta}{r^2 + a^2\cos^2\theta}\right)\left[\frac{a\,dt - (r^2 + a^2)\,d\varphi}{(1 - \frac{1}{3}\Lambda a^2)}\right]^2 \tag{5.68}$$

where the new time and angle co-ordinates $\tilde{t}$ and $\tilde{\varphi}$ are defined by

$$d\tilde{t} = dt + \frac{2Mr\,dr}{(1 + \frac{1}{3}\Lambda r^2)\,\Delta_r} \tag{5.69}$$

$$d\tilde{\varphi} = d\varphi + \left(\frac{a}{r^2 + a^2}\right)\frac{2Mr\,dr}{(1 + \frac{1}{3}\Lambda r^2)\,\Delta_r} \tag{5.70}$$

The form whose square appears in the second term of (5.67) is in fact a null form, either with respect to the full metric ds^2 or with respect to the metric ds_0^2 to which ds^2 reduces when m is set equal to zero. The existence of the expression (5.67) establishes the claim that the solution is asymptotically de-Sitter (or asymptotically flat when $\Lambda = 0$) since the metric ds_0^2 to which ds^2 tends in the limit is in fact exactly de-Sitter space (or exactly Minkowski space when $\Lambda = 0$) albeit in a somewhat twisted co-ordinate system. The co-ordinate system may be untwisted, and (5.68) reduced to a familiar expression for de-Sitter space (or flat space when $\Lambda = 0$) by introducing new co-ordinates $\hat{r}, \hat{\theta}, \hat{\varphi}, \hat{t}$ defined by

$$(1 - \tfrac{1}{3}a^2\Lambda)\hat{r}^2 = r^2 + a^2\sin^2\theta - \tfrac{1}{3}a^2\Lambda r^2\cos^2\theta \tag{5.71}$$

$$\hat{r}\cos\hat{\theta} = r\cos\theta \tag{5.72}$$

$$\hat{\varphi} = \tilde{\varphi} + \frac{\frac{1}{3}a\Lambda t}{1 - \frac{1}{3}a^2\Lambda^2} \tag{5.73}$$

$$\hat{t} = \frac{\tilde{t}}{1 - \frac{1}{3}a^2\Lambda} \tag{5.74}$$

which leads to

$$ds_0^2 + \frac{d\hat{r}^2}{1 + \frac{1}{3}\Lambda\hat{r}^2} + \hat{r}\,(d\hat{\theta}^2 + \sin^2\hat{\theta}\; d\hat{\varphi}^2) - (1 + \tfrac{1}{3}\Lambda\hat{r}^2)\; d\hat{t}^2 \tag{5.75}$$

The fact that the original Kerr metric (with $\Lambda = 0$) can be expressed as flat-space-plus-squared-null-form was discovered by Kerr and Schild (1965). This property, with its cosmological term generalization to de-Sitter-space-plus-squared-null-form (which also applies to the electromagnetic generalizations) distinguishes the asymptotically well behaved solutions selected above from the much wider class of separable solutions (including Taub-N.U.T. space) which have been rejected.

6 Maximal Extensions of the Generalized Kerr Solutions

It turns out that despite their much greater complexity, the generalized Kerr solutions have maximally extended manifolds which are closely analogous to those of the spherical special cases as described in section 4. The construction of the appropriate maximally extended manifolds has been discussed in detail by Boyer and Lindquist (1967) and Carter (1968). The present section will consist primarily of a somewhat abbreviated description of these extension constructions.

The metric forms (5.36) and (5.54) are of course singular on the locus $r = 0$, $\mu \equiv \cos\theta = 0$ where the factor Z is zero, and where the curvature array (5.37) is clearly singular itself. However, when $M^2 > Q^2 + P^2 + a^2$, these forms are also singular on the loci where $r = r_+$ or $r = r_-$ with the definitions

$$r_\pm = M \pm \sqrt{M^2 - a^2 - Q^2 - P^2} \tag{6.1}$$

i.e. where Δ_r is zero. Since the curvature array (5.36) is well behaved where Δ_r is zero, one might have guessed that the metric form singularities at $r = r_+$ and $r = r_-$ should be removable by co-ordinate transformations, as is in fact shown to be the case by the existence of the transformation (5.69), (5.70) to the form (5.67) which is perfectly well behaved where $r = r_\pm$. It will be convenient to make a slightly different transformation, exactly analogous to the Finkelstein transformation described in section 4, introducing an ingoing null co-ordinate v, and a corresponding angle co-ordinate $\tilde{\varphi}$ by

$$\left.\begin{aligned} dv &= dt + \frac{r^2 + a^2}{\Delta_r}\, dr \\ d\tilde{\varphi} &= d\varphi + \frac{a}{\Delta_r}\, dr \end{aligned}\right\} \tag{6.2}$$

so as to obtain from (5.54) the form

$$\begin{aligned} ds^2 &= 2\,dr\,dv - 2a\sin^2\theta\,dr\,d\tilde{\varphi} + Z\,d\theta^2 \\ &\quad + Z^{-1}[(r^2+a^2)^2 - \Delta_r a^2\sin^2\theta]\sin^2\theta\,d\tilde{\varphi}^2 \\ &\quad - 2aZ^{-1}(2Mr - Q^2 - P^2)\sin^2\theta\,d\tilde{\varphi}\,dv \\ &\quad - [1 - Z^{-1}(2Mr - Q^2 - P^2)]\,dv^2 \end{aligned} \tag{6.3}$$

with

$$Z = r^2 + a^2\cos^2\theta \tag{6.4}$$

the corresponding electromagnetic potential form being

$$\tilde{A} = Z^{-1}[Qr + Pa\cos\theta)\,dv - Z^{-1}[Qra\sin^2\theta + P(r^2+a^2)\cos\theta]\,d\tilde{\varphi} \tag{6.5}$$

where the new potential $\tilde{A}$ differs from A by a gauge transformation. It was in this form that both the pure vacuum solution (Kerr 1963) and its electromagnetic generalization (Newman *et al.* 1965) were originally discovered (the curves on which the co-ordinates v, $\tilde{\varphi}$ and θ are all held constant being the integral trajectories of one of the two degenerate Debever (1958) eigenvector fields, namely $\omega^{(0)} - \omega^{(1)}$, of the Weyl tensor). The transformation (6.2), from this original Kerr-Newman form (6.3) to the standard symmetric form (5.36) was first discovered by Boyer and Lindquist (1967). The Kerr-Newman form (6.3) is well behaved over the whole of a co-ordinate patch whose topology is that of the product of a 2-plane and a 2-sphere, with co-ordinates v, r running from $-\infty$ to ∞ and with spherical type co-ordinates θ running from 0 to π and φ periodic with period 2π, except for the usual rotation axis co-ordinate singularity at $\theta = 0$ and $\theta = \pi$ where φ ceases to be well defined, and except also for a *ring singularity* on the locus $r = 0$, $\theta = \pi/2$ where the function Z is zero. The rotation axis singularity is easily removable, for for example by introducing Cartesian-type co-ordinates x, y, z defined by

$$x + iy = (r + ia)\,e^{i\varphi}\sin\theta \qquad z = r\cos\theta \tag{6.6}$$

in terms of which the form (5.67) is transformed (in the case $\Lambda = 0$ under consideration here) to the Cartesian Kerr-Schildt form

$$\begin{aligned} ds^2 &= dx^2 + dy^2 + dz^2 - dt^2 \\ &\quad + \frac{2Mr - Q^2 - P^2}{r^4 + a^2z^2}\left\{\frac{r(x\,dx + y\,dy) - a(x\,dy - y\,dx)}{r^2+a^2} + \frac{z\,dz}{r} + dt\right\}^2 \end{aligned} \tag{6.7}$$

The ring singularity is irremovable, since as we have already remarked the curvature array (5.37) is singular when Z is zero.

Two dimensional cross sections at constant values of v and $\tilde{\varphi}$ are illustrated in Figures 6.1 to 6.5. The ring singularity (which is irremovable since as we have already remarked, the curvature components in the array (5.37) are unbounded in the limit as Z tends to zero) can be thought of as representing the frame of an

Alice-type looking glass separating the positive r part of the patch (which is asymptotically flat in the limit $r \to \infty$) from the negative r part of the patch (which can easily be seen to be asymptotically flat also in the limit $r \to -\infty$).

It can be seen directly from the form (6.3) that when they exist (i.e. when the inequality $M^2 > Q^2 + P^2 + a^2$ is satisfied) the locuses $r = r_+$ and $r = r_-$are null

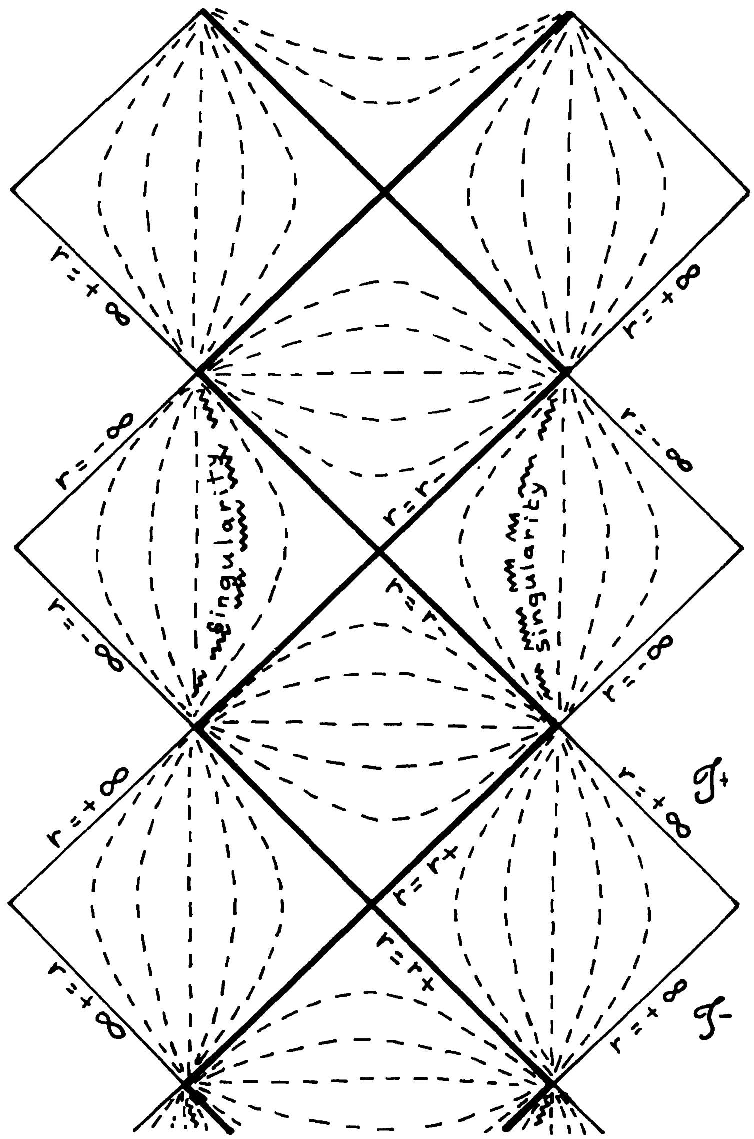


Figure 6.1. Conformal diagram of symmetry axis $\theta = 0$ of maximally extended Kerr or Kerr-Newman solution when $M^2 > a^2 + P^2 + Q^2$. In all the diagrams of this section the locus $\theta = 0$, where the axis passes through (without intersecting) the ring singularity, is marked by a broken zig-zag line.

hypersurfaces which can be crossed in the ingoing (decreasing r) direction only by future directed timelike lines. The surfaces on which r has a constant value between these limits (i.e. $r_- < r < r_+$) are spacelike and can also be crossed only in the sense of decreasing r by a future directed timelike line. Of course starting from the t, φ reversal symmetric standard form 5.54, we could also have made an analogous extension by introducing outgoing null co-ordinates $u, \vec{\varphi}$ defined by

$$\left.\begin{aligned} du &= dt - \frac{r^2 + a^2}{\Delta_r}\, dr \\ d\vec{\varphi} &= d\varphi - \frac{a}{\Delta_r}\, dr \end{aligned}\right\} \qquad (6.8)$$

The metric form in the resulting co-ordinate system is identical to (6.4) except that u is replaced by $-v$ and $\overleftarrow{\varphi}$ by $-\vec{\varphi}$. This metric form is well behaved on a co-

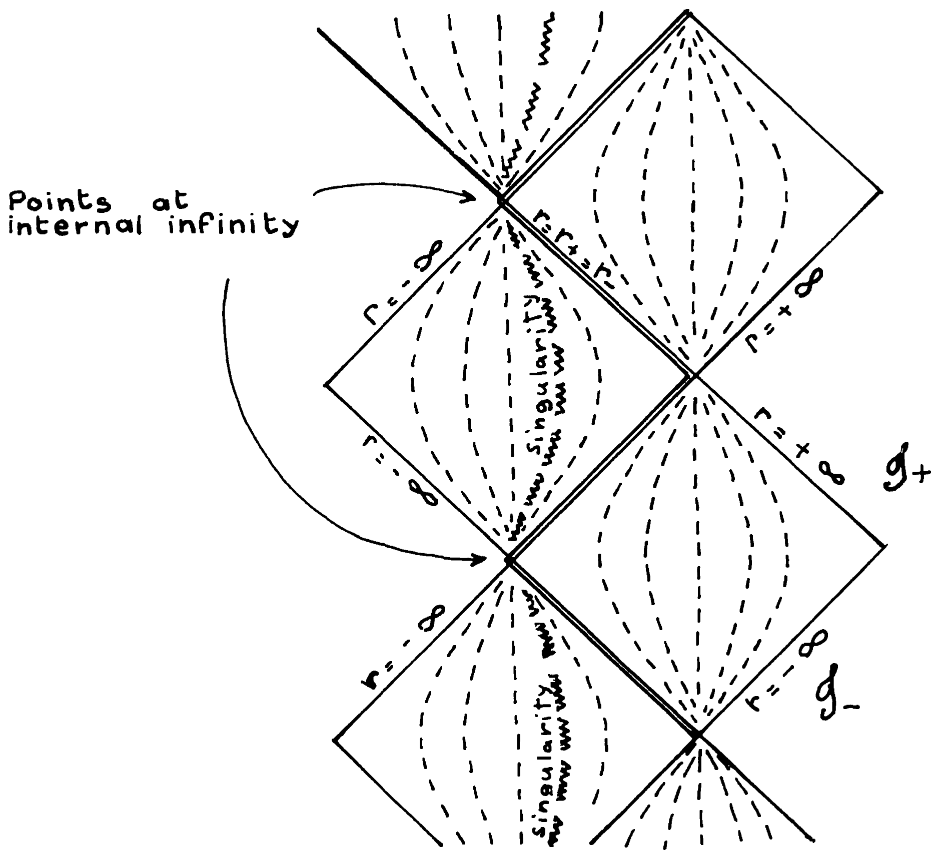
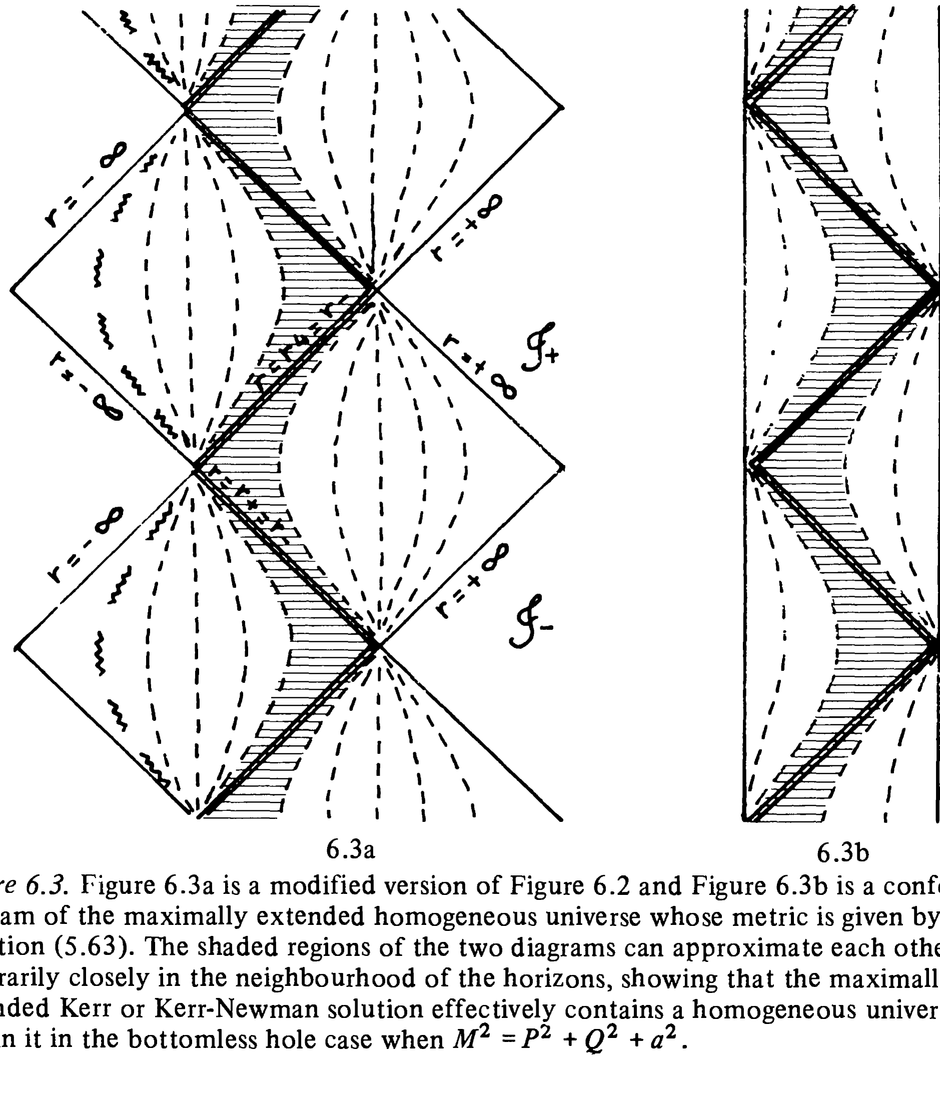


Figure 6.2. Conformal diagram of symmetry axis $\theta = 0$ of maximally extended Kerr or Kerr-Newman solution when $M^2 = a^2 + P^2 + Q^2$ ($a^2 > 0$).

ordinate patch with the topology of the product of a 2-plane and a 2-sphere, with co-ordinates u, r running from $-\infty$ to ∞ and with spherical type co-ordinates θ running from 0 to π and φ periodic with periodic 2π, except as before for the standard rotation axis co-ordinate singularity at $\theta = 0, \theta = \pi$, and the ring singularity at $r = 0, \theta = \pi/2$. Just as was done in the special case of the Reissner-Nordstrom solutions in section 4, so also in this more general case, it is possible to build a maximally extended manifold by combining ingoing null co-ordinate patches $(r, \theta, \overleftarrow{\varphi}, v)$ and outgoing null co-ordinate patches $(r, \theta, \vec{\varphi}, u)$ in the criss-cross pattern illustrated in Figures 6.1, 6.2 and 6.4 (which are analogous to 4.9,



6.3a 6.3b

Figure 6.3. Figure 6.3a is a modified version of Figure 6.2 and Figure 6.3b is a conformal diagram of the maximally extended homogeneous universe whose metric is given by equation (5.63). The shaded regions of the two diagrams can approximate each other arbitrarily closely in the neighbourhood of the horizons, showing that the maximally extended Kerr or Kerr-Newman solution effectively contains a homogeneous universe within it in the bottomless hole case when $M^2 = P^2 + Q^2 + a^2$.

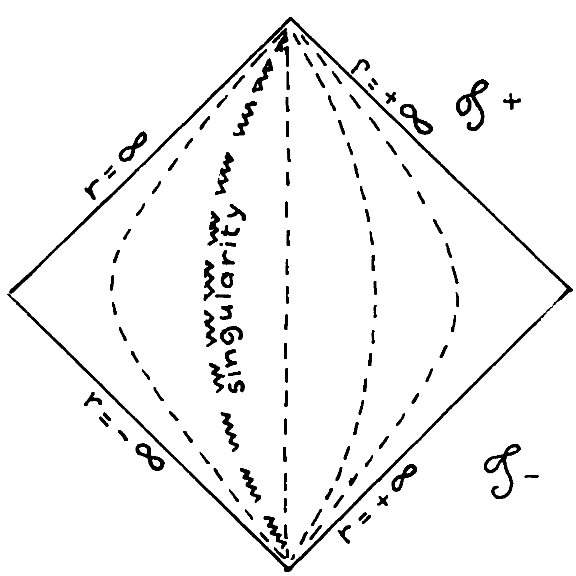


Figure 6.4. Conformal diagram of symmetry axis $\theta = 0$ of maximally extended Kerr or Kerr-Newman solution in the naked singularity case when $M^2 < a^2 + P^2 + Q^2$.

4.12 and 4.6 respectively). These figures can be thought of as showing the basic pattern which the 4-dimensional ingoing and outgoing Finkelstein-type extensions fit together to form a maximally extended manifold, and they can also be given a more specific interpretation (as described in more detail by Carter 1966) as exact conformal diagrams of the symmetry axis $\theta = 0$ of the maximally extended manifold. As in the Reissner-Nordstrom case, when there are no horizons, i.e. when $M^2 < Q^2 + P^2 + a^2$, a single patch suffices as illustrated in Figure 6.4; in this case the singularity where Z tends to zero is naked in the sense that it can both receive light signals from and send light signals to the same original asymptotically flat limit $\mathscr{I}$ where $r \to \infty$. In the general non-degenerate case when $M^2 > Q^2 + P^2 + a^2$, i.e. when the horizons $r = r_+$ and $r = r_-$ both exist and are distinct, it is necessary to construct Kruskal type co-ordinate patches analogous to those of 4.9 to cover the crossover points in Figure 6.1, which represent 2-dimensional spacelike surfaces in the 4-dimensional extended manifold, where the horizons $r = r_-$ and $r = r_+$ intersect themselves. As in the spherical cases this construction can be carried out by first introducing the co-ordinates u and v simultaneously in place of r and t. We also take advantage of the fact noticed by Boyer and Lindquist (1968) who first carried out this construction, that there are two particular Killing vector fields corresponding to the two operators $(r_\pm^2 + a^2)\,\partial/\partial t + a\,\partial/\partial\varphi$ in the standard co-ordinates of the form (5.64), which coincide everywhere with the null generator of the horizons $r = r_\pm$ respectively. This makes it possible to define new ignorable angle co-ordinates $\varphi^\pm$ given by

$$2\,d\varphi^\pm = d\vec{\varphi} - d\overleftarrow{\varphi} + a(r_\pm^2 + a^2)^{-1}\,(du - dv) \tag{6.9}$$

which are constant on the null generators of the horizons at $r = r_\pm$ respectively. We thus obtain two symmetric double null co-ordinate systems, analogous to (4.20) (and adapted respectively to the horizons $r = r_\pm$) given by

$$\begin{aligned} ds^2 = {} & \frac{\Delta_r}{Z}\left(\frac{Z}{r^2 + a^2} + \frac{Z_\pm}{r_\pm^2 + a^2}\right)\frac{(r^2 - r_\pm^2)a^2 \sin^2\theta}{(r^2 + a^2)(r_\pm^2 + a^2)}\,\frac{(du^2 + dv^2)}{4} \\ & + \frac{\Delta_r}{Z}\left[\frac{Z^2}{(r^2 + a^2)^2} + \frac{Z_\pm^2}{(r^2 + a^2)^2}\right]\frac{du\,dv}{2} + Z\,d\theta^2 \\ & - \frac{\Delta_r\, a \sin^2\theta}{Z}\left[a \sin^2\theta\, d\varphi^\pm - \frac{Z_\pm}{r_\pm^2 + a^2}(du - dv)\,d\varphi^\pm\right]^2 \\ & + \frac{\sin^2\theta}{Z}\left[a\,\frac{(r_\pm^2 - r^2)}{r_\pm^2 + a^2}\frac{(du - dv)}{2} - (r^2 + a^2)\,d\varphi^\pm\right]^2 \end{aligned} \tag{6.10}$$

with the obvious abbreviation $Z_\pm = r_\pm^2 + a^2 \cos^2\theta$ and where r is defined implicitly as a function of u and v by

$$F(r) = u + v \tag{6.11}$$

where

$$F(r) = 2r + \kappa_+^{-1} \ln |r - r_+| + \kappa_-^{-1} \ln |r - r_-| \tag{6.12}$$

with the constants κ_+ and κ_- defined by

$$\kappa_\pm = \tfrac{1}{2}(r_\pm^2 + a^2)^{-1}(r_\pm - r_\mp) \tag{6.13}$$

As in section 4, we now introduce new co-ordinates U^+ and V^+ or U^- and V^- (depending on whether we wish to remove the co-ordinate singularity at $r = r_+$ or $r = r_-$) defined by

$$\left.\begin{aligned} u &= -\kappa_\pm^{-1} \ln |U^\pm| \\ v &= +\kappa_\pm^{-1} \ln |V^\pm| \end{aligned}\right\} \tag{6.14}$$

which leads directly to the forms

$$\begin{aligned} ds^2 = {} & Z^{-1}\left(\frac{Z}{r^2 + a^2} + \frac{Z_\pm}{r_\pm^2 + a^2}\right)\frac{(r - r_\mp)(r + r_\pm)a \sin^2\theta}{(r^2 + a^2)(r_\pm^2 + a^2)}\kappa_\pm^{-2}G_\pm^2(r) \\ & \times \frac{(U^{\pm 2}dV^{\pm 2} + V^{\pm 2}dU^{\pm 2})}{4} + Z^{-1}\left(\frac{Z^2}{(r^2 + a^2)^2} + \frac{Z_\pm^2}{(r_\pm^2 + a^2)^2}\right) \\ & \times (r - r_\mp)\kappa_\pm^{-2}G_\pm(r)\frac{dU^\pm dV^\pm}{2} + Z\, d\theta^2 - \frac{a^2 \sin^2\theta}{Z} \\ & \times \left[\Delta_r a \sin^2\theta\, d\varphi^\pm - \frac{Z_\pm^2}{r_\pm^2 + a^2}(r - r_\mp)G_\pm(r)\kappa_\pm^{-1}(V^\pm dU^\pm - U^\pm dV^\pm)\right]d\varphi^\pm \\ & + \frac{\sin^2\theta}{Z}\left[(r^2 + a^2)\, d\varphi^\pm + a\frac{(r + r_\pm)}{r_\pm^2 + a^2}\kappa_\pm^{-1}G_\pm(r)\frac{(V^\pm dU^\pm - U^\pm dV^\pm)}{2}\right]^2 \end{aligned} \tag{6.15}$$

[which is the generalization of (4.23)] where r is now determined implicitly as a function of $U^\pm$ and $V^\pm$ by

$$U^\pm V^\pm = (r - r_\pm)G_\pm^{-1}(r) \tag{6.16}$$

where $G_\pm(r)$ is defined by

$$G_\pm(r) = e^{-2\kappa_\pm r}|r - r_\mp|^{|\kappa_\pm/\kappa_\mp|} \tag{6.17}$$

The corresponding electromagnetic field potential forms are

$$\begin{aligned} A^\pm = {} & \frac{[Q(rr_\pm - a^2\cos^2\theta) - Pa\cos\theta(r + r_\pm)]}{Z(r_\pm^2 + a^2)}\frac{(V^\pm dU^\pm - U^\pm dV^\pm)}{2} \\ & + \frac{[Qra\sin^2\theta + P\cos\theta(r^2 + a^2)]}{Z}d\varphi^\pm \end{aligned} \tag{6.18}$$

where the new potentials $A^{\pm}$ differ from A by gauge transformations. Since the relations (6.16) and (6.17) ensure that r is an analytic function of U^+ and V^+ in the neighbourhood of the horizons $r = r_+$ and an analytic function of U^-, V^- in the neighbourhood of the horizons $r = r_-$, it can easily be checked that the forms (6.15) and (6.18) are well behaved in the neighbourhood of the horizons $r = r_{\pm}$ respectively except for the usual co-ordinate singularities on the rotation axes $\theta = 0$, $\theta = \pi$ which could easily be removed by changing to a Cartesian type system. Since it is rather complicated the form (6.15) is of very little practical use, but its *existence* is important in order to establish that the extended manifold includes the *intersections* of the two null hypersurfaces $U^{\pm} = 0$ and $V^{\pm} = 0$ which make up the locuses $r = r_{\pm}$. The fitting together of the co-ordinate patches covering the maximally extended manifold of Figure 6.1 is described in more explicit detail by Boyer and Lindquist (1967) and Carter (1968).

In the naked-singularity case, i.e. when $M^2 < Q^2 + P^2 + a^2$, the patches covered by the original standard metric form (5.54) and by the Kerr–Newman form (6.3) are equivalent, and so a single patch gives the maximally extended manifold as illustrated in Figure 6.4. In the limiting intermediate case $M^2 = Q^2 + P^2 + a^2$ where the horizons $r = r_+$ and $r = r_-$ have coalesced to form a single degenerate horizon at $r = M$, the Kerr–Newman patches of the form (6.3) are sufficient to build the whole maximally extended manifold illustrated in Figure 6.2, as is described in more explicit detail by Carter (1968). As in the degenerate spherical case of Figure 4.12, so also in Figure 6.2 there are *internal* boundary points which represent limits which can only be reached by curves of infinite affine length (in addition to the boundary points representing the ordinary external asymptotically flat limit, i.e. the boundary hypersurfaces congruent to $\mathscr{I}$). As in the spherical case, so also in this more general case, it can easily be seen that the degenerate limit manifold illustrated by Figure 6.2 approximates a homogeneous universe whose metric is given by the special solution (5.63), in the limit as the degenerate horizon $r = M$ is approached, in the manner shown by Figure 6.3 (which is analogous to Figure 4.14). This limiting case will be given special attention in the following course of Bardeen.

The rigorous proof that these manifolds (as illustrated by Figures 6.1 to 6.4) really are maximal (in the sense that they are not submanifolds of more extended solution manifolds) depends (as in the spherical case) on the demonstration that all geodesics can either be extended to arbitrary affine length or else approach the curvature singularities, i.e. they are such that Z tends to zero along them. The method of integrating the geodesics is described in the following section. The explicit demonstration that the only incomplete geodesics are those which approach the curvature singularity is decribed by Carter (1968) and will not be repeated in this course.

We conclude this section by remarking that the procedures described here can also be applied in a straightforward manner to the asymptotically de Sitter Λ-term generalization (5.65) of the Kerr solutions. In this case when the rotation parameter a is small compared with M, and when M is small compared with the radius parameter $\Lambda^{-1/2}$ of the asymptotic de Sitter universe, there will be four zeros of Δ_r, of which two, which may be denoted by $r = r_+, r = r_-$, will be analogues of the zeros $r = r_\pm$ in the ordinary Kerr solutions, while the other two, which may be denoted by $r = r_{++} \gg r_+$ and $r = r_{--} \ll r_-$ would exist even in the ordinary de Sitter universe. The conformal diagram for the symmetry axis of the ordinary de Sitter universe has already been given by Figure 2.3. It is easy to guess that the

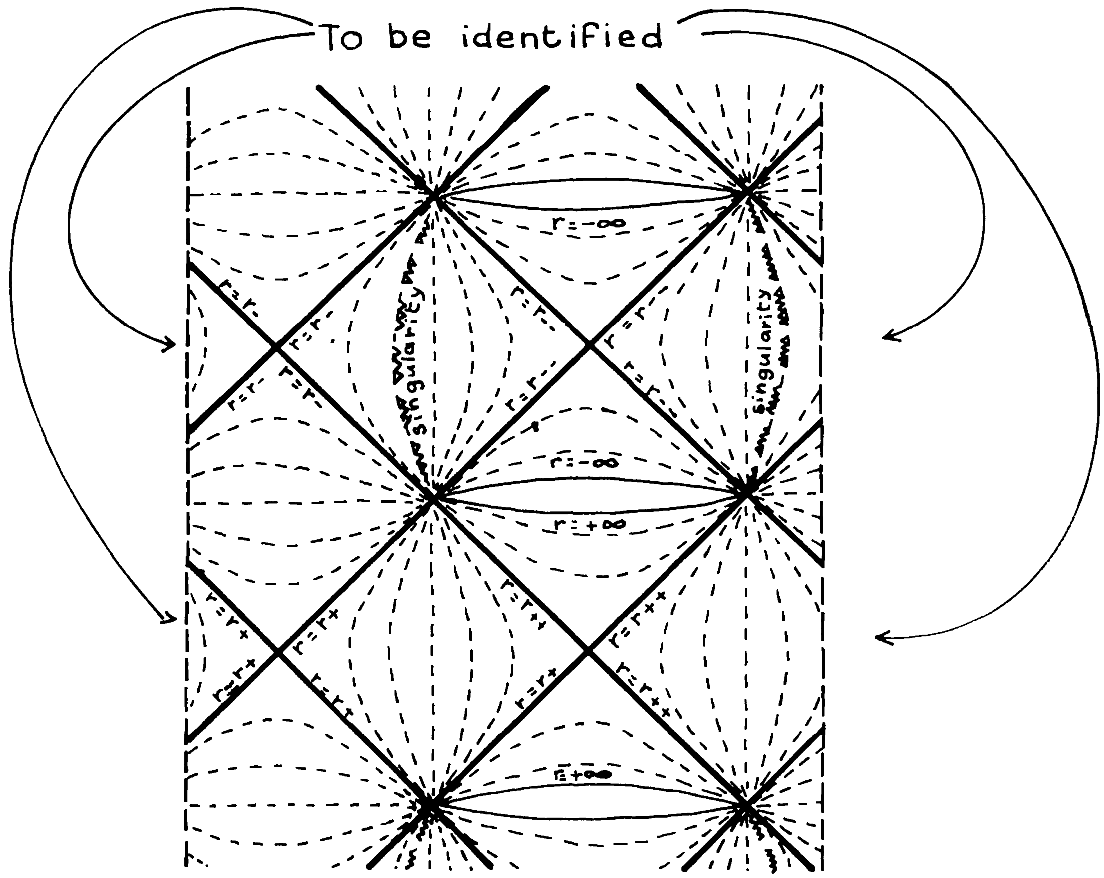


Figure 6.5. Conformal diagram of the symmetry axis $\theta = 0$ of the Kerr-de Sitter solution when $\Lambda^{-4} \gg M^2 \gg a^2$. The opposite sides of the diagram are to be identified [compare Figure 2.3 which is also the conformal diagram of the symmetry axis $\theta = 0$ for the pure de Sitter universe].

corresponding conformal diagram of the symmetry axis of the simplest maximally extended Kerr-de Sitter space will be as shown by Figure 6.5 which shows a sequence of de Sitter universes in which antipodal points are connected by Kruskal type throats. It is a straightforward (albeit lengthy) exercise, using the techniques described in this section, to construct appropriate Finkelstein type and Kruskal type co-ordinate patches to cover the manifold whose structure is shown by Figure 6.5 and then to integrate the geodesic equations and use them to prove that this manifold is indeed maximal. Since the Kerr-de Sitter solution is of more geometrical than physical interest, we shall not discuss it further in this course.

7 The Domains of Outer Communication

The only parts of the manifolds described in the previous sections which are strictly relevant to studies of black holes are the *domains of outer communications* (i.e. the parts which can be connected to one of the asymptotically flat regions by both future and past directed timelike lines), more particularly the domains of communications which are non-singular in the sense of having geodesically complete closures in the extended manifolds, since (as is discussed in the introduction to Part II of this course) it is these domains which are believed to represent the possible asymptotic equilibrium states of a source-free exterior field of a black hole. It is fairly obvious (and follows directly from the lemma 2.1 to be given in Part II of this course) that in the cases when $M^2 < Q^2 + P^2 + a^2$ the domains of outer communications consist of the entire manifold, whereas when $M^2 > Q^2 + P^2 + a^2$ the domains of outer communications consist only of the regions $r > r_+$ bounded by the outer Killing horizons at $r = r_+$. In the former case —the case of naked singularities—the domain of outer communications is itself singular, since there are always some geodesics which approach the curvature singularities. (Nevertheless there are fewer incomplete geodesics than one might have expected since one can show (Carter 1968) that only geodesics confined to the equatorial symmetry plane $\mu = 0$ can actually approach the curvature singularities.) In the latter case, when the horizons at $r = r_+$ exist, the singularities will always be hidden outside the observable region $r > r_+$ (since we always have $r_+ \geqslant r_- > 0$); therefore, since the only incomplete geodesics in the maximally extended manifolds are those which approach the curvature singularities, it follows that closures of the domains $r > r_+$ are indeed complete in the sense that any geodesic in one of these domains must either have infinite affine length or else have an end point within the domain $r > r_+$ or on the boundary $r = r_+$ in the extended manifold.

The remainder of our discussion will therefore be limited to these black hole exterior domains $r > r_+$ in the cases when the inequality $M^2 \geqslant a^2 + P^2 + Q^2$ is satisfied. An important property of these domains (without which they could not be taken seriously as the basis of a physical theory) is that they satisfy the causality condition that there are no compact (topologically circular) causal (i.e. timelike or null) curves within them. This follows from the fact that the hypersurfaces on which t is constant form a well-behaved congruence of *spacelike* hypersurfaces within the domains $r > r_+$ [which implies that t must increase monotonically along any timelike or null curve which remains in one of these domains]. The fact that these hypersurfaces are spacelike when $r > r_+$ follows from the fact that the two-dimensional metric on the surfaces on which t and φ are *both* constant is positive definite whenever Δ_r is positive, (i.e. both when $r > r_+$ and when $r < r_-$) and from the fact that the coefficient given by

$$X = \frac{(r^2 + a^2)^2 - \Delta_r a^2 \sin^2 \theta}{r^2 + a^2 \cos^2 \theta} \sin^2 \theta \tag{7.1}$$

(which represents the squared magnitude of the Killing vector corresponding to the operator $\partial/\partial\varphi$ in r, θ, t, φ co-ordinates) turns out to be strictly positive also throughout the domains $r = r_-$ except of course on the rotation axis where it is zero. The coefficient X does become negative in a subregion [indicated by double shading in Figures 7.1 to 7.7] in the neighbourhood of the curvature singularity, thereby giving rise to causality violation in the inner regions of the extended manifold. In the naked singularity case, when $M^2 < a^2 + P^2 + Q^2$, it can easily be shown (Carter 1968, Carter 1972) that the entire extended manifold is a single *vicious set* in the sense that any event can be connected to any other by both a

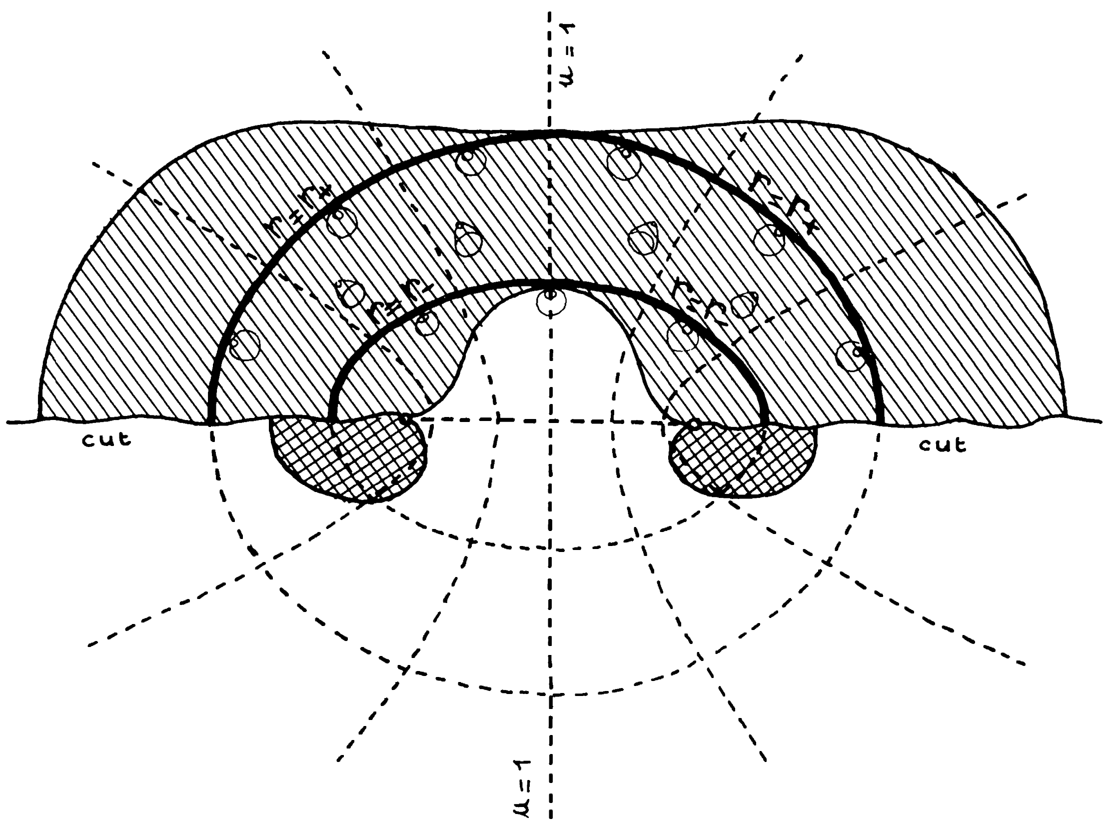


Figure 7.1. Plan of a polar 2-section on which v and $\tilde{\varphi}$ are constant through maximally extended Kerr solution with $M^2 > a^2$. The ring singularity is treated as a branch point and only half of the 2-section (corresponding roughly to $\cos\theta > 0$) bounded by cuts is shown—the other half should be regarded as being superimposed on the first half in the plane of the paper. The same comments apply to Figures 7.2 and 7.3. In all the diagrams of this section dotted lines are used to represent locuses on which r or θ is constant, and the positions of the Killing horizons are marked by a heavy line except for degenerate horizons which are marked by a double line. The regions in which V is negative are indicated by single shading and the regions where X is negative are marked by double shading. Some projected null cones are marked.

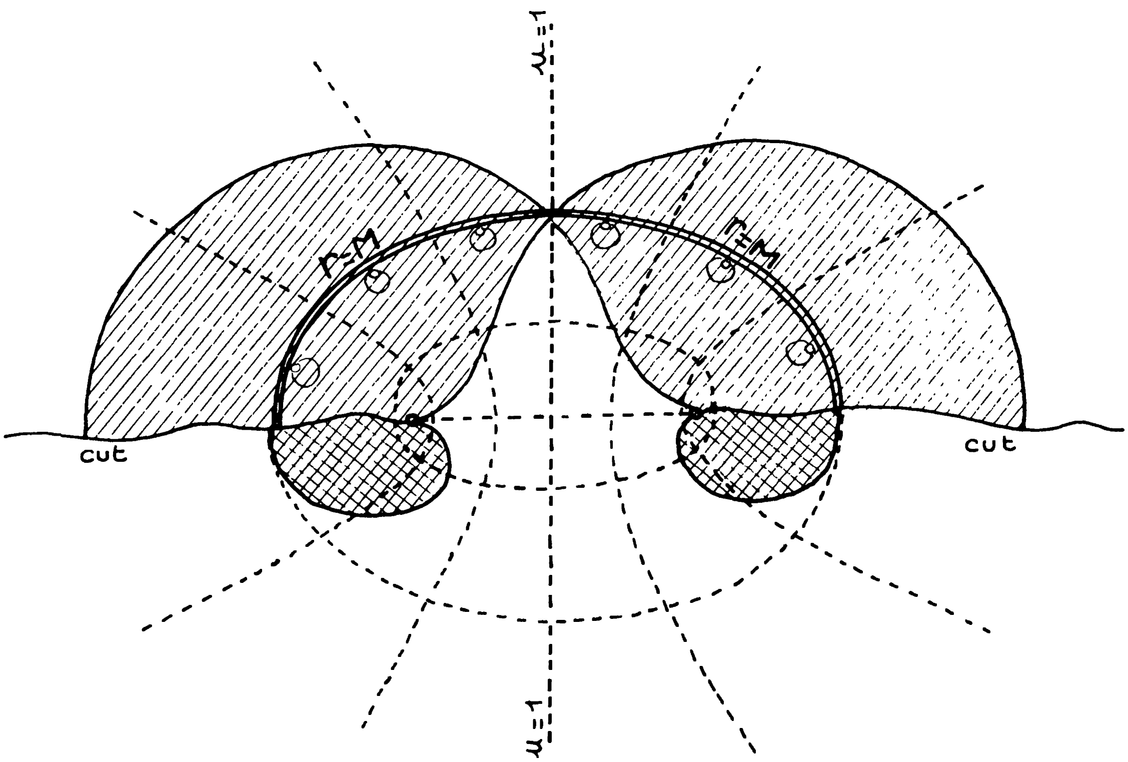


Figure 7.2. Plan of polar 2-section in which v and $\tilde{\varphi}$ are constant through a maximally extended Kerr solution in the degenerate case when $M^2 = a^2$.

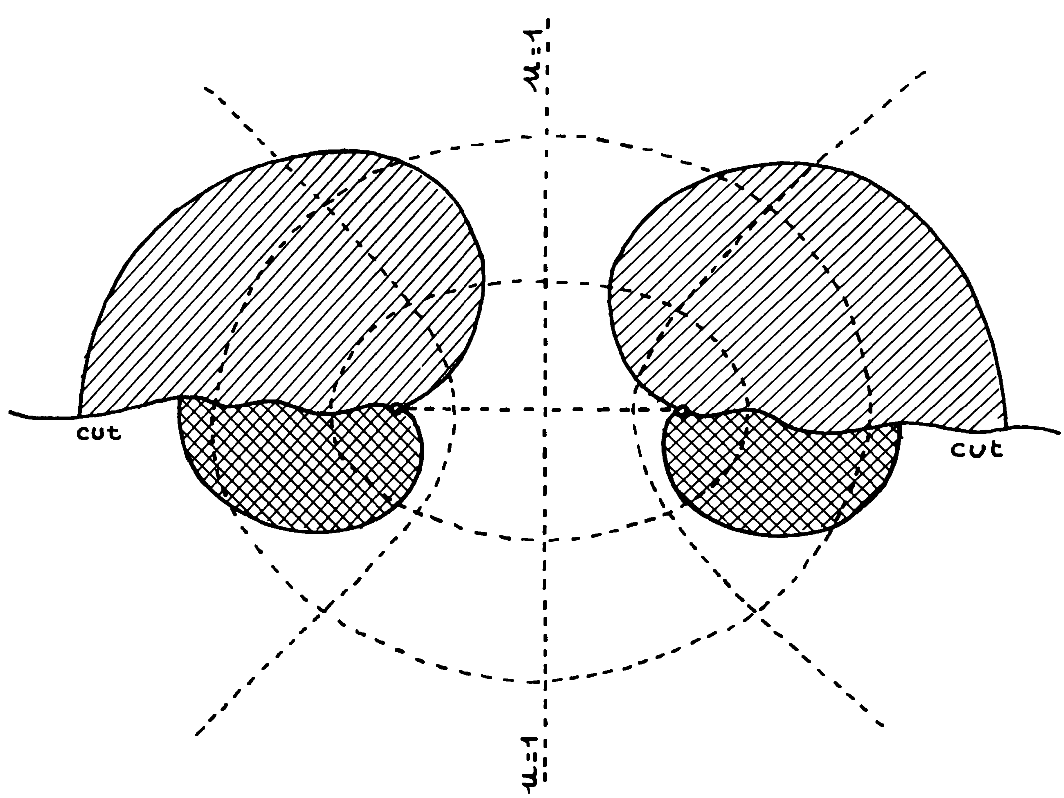


Figure 7.3. Plan of a polar 2-section in which v and $\tilde{\varphi}$ are constant through a maximally extended Kerr solution when $M^2 < a^2$

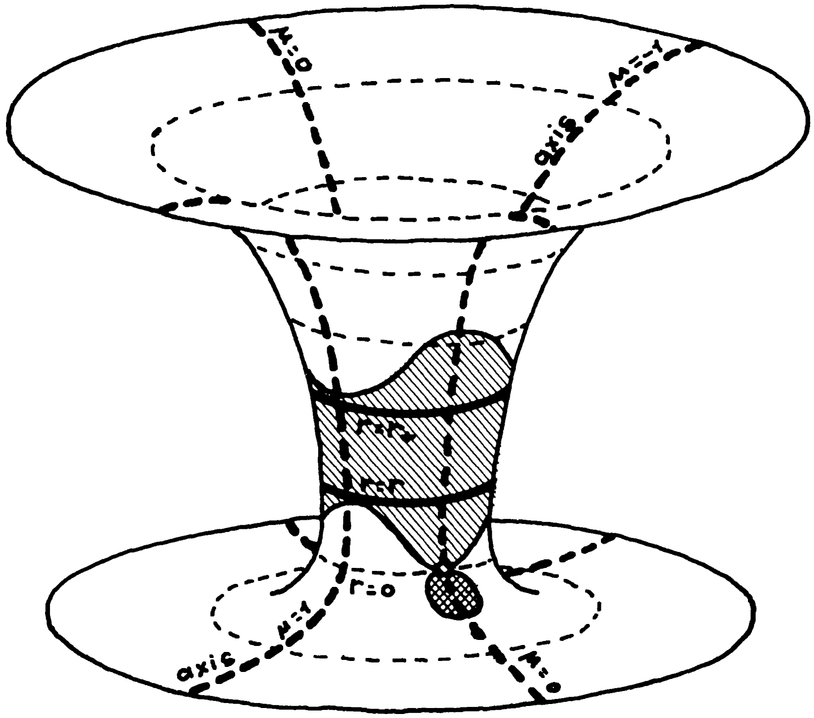


Figure 7.4. Rough sketch in perspective of a polar 2-section in which v and $\tilde{\varphi}$ are constant through a maximally extended Kerr solution when $M^2 > a^2$ (this is an alternative representation of the same section as is shown in Figure 7.1).

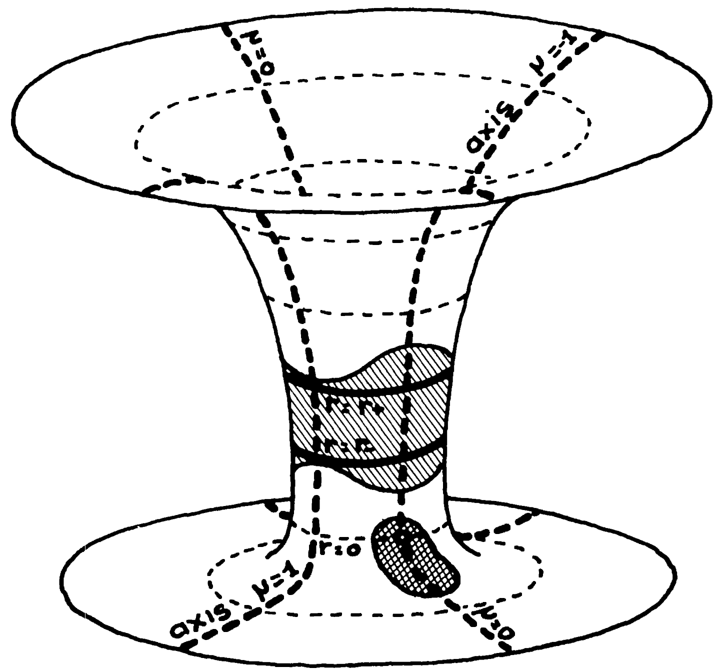


Figure 7.5. Rough sketch in perspective of a polar 2-section in which v and $\tilde{\varphi}$ are constant through a maximally extended Kerr-Newman solution when $M^2 > a^2 + P^2 + Q^2$.

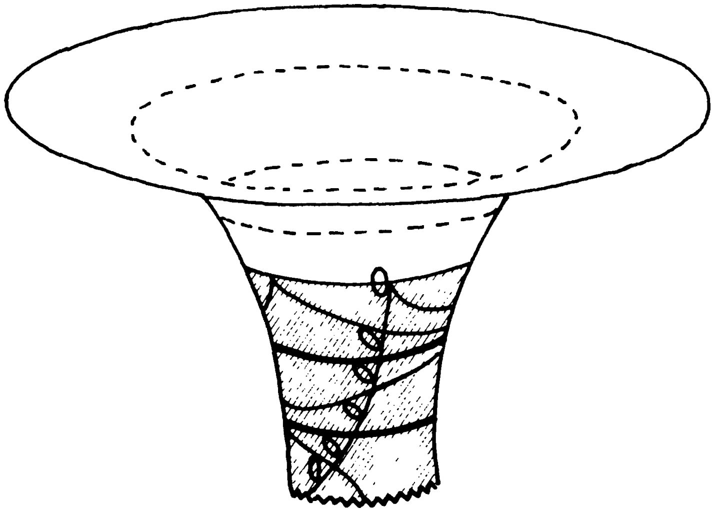


Figure 7.6. Rough sketch in perspective of an equatorial 2-section in which v is constant and $\cos\theta = 0$ through a maximally extended Kerr or Kerr-Newman solution when $M^2 > a^2 + P^2 + Q^2$. The continuous lines represent envelopes of projected null cones.

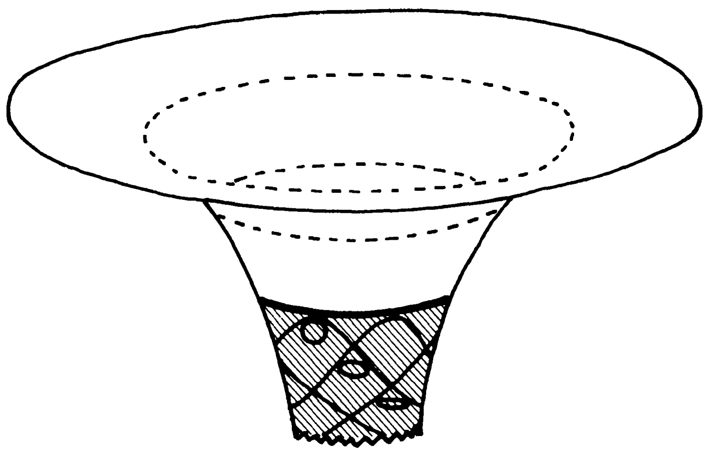


Figure 7.7. Rough sketch in perspective of equatorial 2-section in which v is constant and $\cos\theta = 0$ through maximally extended Schwarzschild solution. This represents the static limit of Figure 7.6.

future and a past directed timelike curve. However, in the black hole cases, i.e. when $M^2 \leqslant a^2 + P^2 + Q^2$, the causality violation is restricted to the domains $r < r_-$, and so does not affect the domains of outer communications.

Another noteworthy feature of the domains of outer communications (although its significance is often exaggerated) is that except in the spherical cases they overlap with the regions [indicated by single shading in Figures 7.1 to 7.6] in which the coefficient V given by

$$V = 1 - \frac{2Mr - P^2 - Q^2}{r^2 + a^2\cos^2\theta} \tag{7.2}$$

is negative. This coefficient V is the negative squared magnitude of the time Killing vector corresponding to the operator $\partial/\partial t$ in r, θ, φ, t co-ordinates; it is the same as the metric component g_{00} in a notation system where the co-ordinates x^1, x^2, x^3, x^0 are taken to be r, θ, φ, t. We shall refer to the outer boundary of this region as the *outer ergosurface* since it bounds the region within which the energy E of an uncharged particle orbit (as defined by equation (8.22) of the next section) can be negative. It is sometimes misleadingly described as an "infinite red shift surface" on the grounds that the light emitted by a particle which is static (in the sense that its co-ordinates remain fixed) will be redshifted at large asymptotic distances by an amount which tends to infinity as the position of the static particle approaches the ergosurface. However, the real significance of the region $V < 0$ bounded by the ergosurface is that physical particles which are static in this sense may not exist at all within it (if their world lines are to be timelike). In practice, moreover, one would not expect static particles to exist even in the neighbourhood of the ergosurface, and therefore the infinite red shift phenomenon would never be observed. On the other hand a genuine physically observable infinite red shift will take place whenever a particle crosses the actual horizon at $r = r_+$, since it is obvious that the light emitted by such a particle in the finite proper time-interval preceding the moment at which it crosses the horizon will emerge at large distances spread during an unbounded time interval.

8 Integration of the Geodesic Equations

The geodesic equation, and also the orbit equations

$$mv^a{}_{;b}v^b = eF^a{}_b v^b \tag{8.1}$$

for a particle of mass m charge e moving in a free orbit with unit tangent vector v^a can both be obtained from the same simple Lagrangian

$$L = \tfrac{1}{2}g_{ab}\dot{x}^a\dot{x}^b - eA_a\dot{x}^a \tag{8.2}$$

where a dot denotes differentiation with respect to an affine parameter, λ say. In the case of a particle orbit, λ must be related to the proper time τ along the orbit by

$$\tau = m\lambda \tag{8.3}$$

which is equivalent to imposing the normalization condition

$$g_{ab}\dot{x}^a\dot{x}^b = -m^2 \tag{8.4}$$

which means that we shall have

$$\dot{x}^a = mv^a \tag{8.5}$$

The geodesic equations are obtained from the Lagrangian (8.2) when e is set equal to zero.

Introducing the momentum co-vector

$$p_a = g_{ab}\dot{x}^b - eA_a \tag{8.6}$$

we see that the Hamiltonian corresponding to the Lagrangian (8.2) has the very simple form

$$H = \tfrac{1}{2}g^{ab}(p_a + eA_a)(p_b + eA_b) \tag{8.7}$$

Since the affine parameter λ does not appear explicitly in the Hamiltonian, the Hamiltonian itself will automatically be a constant of the motion, and by the normalization condition (8.4) it is clear that this constant is given simply by

$$H = -\tfrac{1}{2}m^2 \tag{8.8}$$

When one of the co-ordinates, x^0 say, is ignorable (i.e. does not appear explicitly in the Hamiltonian) the corresponding momentum p_0 will commute with the Hamiltonian (in the sense that its Poisson bracket with the Hamiltonian will be zero) and hence it will also be a constant of the motion. In the case of the metric form (5.18) with the electromagnetic field potential (5.45) there are two independent ignorable co-ordinates, namely t and φ and hence the corresponding momenta p_t and p_φ will both be conserved, giving two constants of the motion in addition to the constant given by (8.7) which would always exist.

To obtain a complete set of explicit first integrals of the motion we need to have a *fourth* constant of the motion. In a general stationary-axisymmetric space-time an independent fourth integral would not exist, but the condition that the charged Klein Gordon equation be separable (which we imposed in our derivation of the Kerr solutions) automatically guarantees the weaker condition that the Hamilton-Jacobi equation corresponding to the Hamiltonian (8.6) will be separable in the present case (Carter 1968a, Carter 1968b), thus ensuring the existence of the required fourth integral. We shall not use the Hamilton-Jacobi formalism in the present section, but instead shall show how the fourth constant may be obtained directly by inspection of the Hamiltonian using the following lemma:

LEMMA Let the Hamiltonian have the form

$$H = \frac{1}{2}\frac{H_r + H_\mu}{U_r + U_\mu} \tag{8.9}$$

where U_r and U_μ are single variable functions of the co-ordinates r and μ respectively and where H_r is independent of the momentum p_μ and of all the co-ordinate functions other than r, and H_μ is independent of the momentum p_r and of all the co-ordinate functions other than μ. Then the quantity

$$K = \frac{U_r H_\mu - U_\mu H_r}{U_r + U_\mu} \tag{8.10}$$

commutes with H and hence is a constant of the motion.

Proof Since the stipulated conditions clearly ensure that H_r commutes with H_μ, and since H_μ naturally commutes with itself, we obtain

$$[H_r, H] = \tfrac{1}{2}(H_r + H_\mu)\left[H_r, \frac{1}{U_r + U_\mu}\right] \tag{8.11}$$

Similarly since U_r clearly commutes with H_μ and with $U_r + U_\mu$ we obtain

$$[U_r, H] = \frac{1}{2(U_r + U_\mu)}[U_r, H_r] \tag{8.12}$$

Working out the Poisson brackets on the right hand sides we find

$$\begin{aligned}\left[H_r, \frac{1}{U_r + U_\mu}\right] &= -\frac{1}{(U_r + U_\mu)^2}\frac{\partial H_r}{\partial p_r}\frac{dU_r}{dr} \\ &= \frac{1}{(U_r + U_\mu)^2}[U_r, H_r]\end{aligned} \tag{8.13}$$

Thus we can eliminate the right hand sides of (8.11) and (8.12) to obtain

$$[H_r, H] = 2H[U_r, H] \tag{8.14}$$

It follows immediately that the quantity

$$K = 2U_r H - H_r \tag{8.15}$$

commutes with H and hence is a constant of the motion. It is clear from (9.8) that this quantity K can also be expressed in the form

$$K = H_\mu - 2U_\mu H \tag{8.16}$$

and also in the more symmetric form (8.10). This completes the proof.

Now the Hamiltonian corresponding to our canonical separable metric form (5.18) with the canonical vector potential (5.45) can be read off directly from the inverse metric (5.16) as

$$\begin{aligned}H = &\frac{\Delta_\mu p_\mu^2 + \Delta_\mu^{-1}[C_\mu p_\varphi + Z_\mu p_t - eX_\mu]^2}{2(C_\mu Z_r - C_r Z_\mu)} \\ &+ \frac{\Delta_r p_\mu^2 - \Delta_r^{-1}[C_r p_\varphi + Z_\mu p_t + eX_r]^2}{2(C_\mu Z_r - C_r Z_\mu)}\end{aligned} \tag{8.17}$$

which clearly has the form (8.9) with

$$H_r = \Delta_r p_r^2 - \Delta_r^{-1}[C_r p_\varphi + Z_r p_t + eX_r]^2 \tag{8.18}$$

$$H_\mu = \Delta_\mu p_\mu^2 + \Delta_\mu^{-1}[C_\mu p_\mu + Z_\mu p_t - eX_\mu]^2 \tag{8.19}$$

$$U_r = C_\mu Z_r \tag{8.20}$$

$$U_\mu = -C_r Z_\mu \tag{8.21}$$

It follows at once from the lemma that there will indeed be a fourth constant of the motion K given equivalently by (8.10) or (8.16) in terms of the expressions (8.18), (8.19), (8.20), (8.21), in addition to the constant H given by (8.17) and the energy and polar angular momentum constants E and L_z given by

$$E = -p_t \tag{8.22}$$

$$L_z = p_\varphi \tag{8.23}$$

In terms of the actual source free solutions (5.46) to (5.53) we shall have

$$H = \frac{(1-\mu^2)p_\mu^2 + (1-\mu^2)^{-1}[p_\varphi + a(1-\mu^2)p_t - eP\mu]^2}{2(r^2+a^2\mu^2)} + \frac{\Delta_r p_r^2 - \Delta_r^{-1}[ap_\varphi + (r^2+a^2)p_t + eQr]^2}{2(r^2+a^2\mu^2)} \tag{8.24}$$

and the corresponding expression for K will be

$$\mathrm{K} = \frac{a^2(1-\mu^2)\{\Delta_r p_r^2 - \Delta_r^{-1}[ap_\varphi + (r^2+a^2)p_t + eQr]^2\}}{r^2+a^2\mu^2} + \frac{(r^2+a^2)\{(1-\mu^2)p_\mu^2 + (1-\mu^2)^{-1}[p_\varphi + a(1-\mu^2)p_t - eP\mu]^2\}}{r^2+a^2\mu^2} \tag{8.25}$$

where Δ_r is given by (5.52).

Transforming from momentum to velocity co-ordinates and setting $\mu = \cos\theta$ we obtain the four constants of the motion explicitly as two linear combinations

$$E = \frac{(r^2 - 2Mr + a^2\cos^2\theta + P^2 + Q^2)\dot{t}}{r^2 + a^2\cos^2\theta} + \frac{(2Mr - P^2 - Q^2)a\sin^2\theta\dot{\varphi} + e(Qr + aP\mu)}{r^2+a^2\cos^2\theta} \tag{8.26}$$

$$L_z = \frac{[(r^2+a^2)^2 - \Delta_r a^2\sin^2\theta]\sin^2\theta\dot{\varphi}}{r^2+a^2\cos^2\theta} + \frac{-(2Mr - P^2 - Q^2)a\sin^2\theta\,\dot{t} + e[aQr\sin^2\theta + P(r^2+a^2)\cos\theta]}{r^2+a^2\cos^2\theta} \tag{8.27}$$

and two quadratic combinations

$$m^2 = -(r^2 + a^2\cos^2\theta)\left(\frac{\dot{r}^2}{\Delta_r} + \dot{\theta}^2\right) - \frac{\sin^2\theta}{r^2 + a^2\cos^2\theta}[(r^2 + a^2)\dot{\varphi} - a\dot{t}]^2 + \frac{\Delta_r}{r^2 + a^2\cos^2\theta}[a\sin^2\theta\dot{\varphi} - \dot{t}]^2 \tag{8.28}$$

$$\mathrm{K} = a^2\sin^2\theta\left\{\frac{r^2 + a^2\cos^2\theta}{\Delta_r}\dot{r}^2 - \frac{\Delta_r}{r^2 + a^2\cos^2\theta}[a\sin^2\theta\dot{\varphi} - \dot{t}\,]^2\right\} + (r^2 + a^2)\left\{(r^2 + a^2\cos^2\theta)\dot{\theta}^2 + \frac{\sin^2\theta}{r^2 + a^2\cos^2\theta}[(r^2 + a^2)\dot{\varphi} - a\dot{t}]^2\right\} \tag{8.29}$$

Introducing the functions

$$R(r) = [E(r^2 + a^2) - aL_z - eQr]^2 - \Delta_r(m^2r^2 + \mathrm{K}) \tag{8.30}$$

$$\Theta(\theta) = \mathrm{K} - m^2a^2\cos^2\theta - \sin^{-2}\theta\,[Ea\sin^2\theta - L_z + eP\cos\theta\,]^2 \tag{8.31}$$

$$= C + 2aePE\cos\theta + [a^2(E^2 - m^2) - (L_z^2 + e^2P^2)\sin^{-2}\theta\,]\cos^2\theta \tag{8.32}$$

where C is a constant given in terms of K by

$$C = \mathrm{K} - (L_z - aE)^2 \tag{8.33}$$

we can express the non-ignorable co-ordinate velocities by the equations

$$Z^2\dot{r}^2 = R(r) \tag{8.34}$$

$$Z^2\dot{\theta}^2 = \Theta(\theta) \tag{8.35}$$

It is clear from the form of (8.35) and (8.31) that K must always be positive if $\dot{\theta}$ is to be real. The ignorable co-ordinate velocities can be expressed directly as

$$Z\dot{\varphi} = L_z\left[\frac{1}{\sin^2\theta} - \frac{a^2}{\Delta_r}\right] - \frac{aE(2Mr - P^2 - Q^2)}{\Delta_r} \tag{8.36}$$

$$Z\dot{t} = E\left[\frac{(r^2 + a^2)^2}{\Delta_r} - a^2\sin^2\theta\right] - aL_z\frac{(2Mr - P^2 - Q^2)}{\Delta_r} \tag{8.37}$$

Hence we can obtain the final fully integrated form of the orbit equations as

$$\int^{\theta}\frac{d\theta}{\sqrt{\Theta}} = \int^{r}\frac{dr}{\sqrt{R}} \tag{8.38}$$

$$\lambda = \int^{\theta}\frac{a^2\cos^2\theta\,d\theta}{\sqrt{\Theta}} + \int^{r}\frac{r^2\,dr}{\sqrt{R}} \tag{8.39}$$

$$\varphi = +\int^{r} \frac{[L_z + aE(2Mr - P^2 - Q^2)]\,dr}{\Delta_r\sqrt{R}} + \int^{\theta} \frac{L_z\,d\theta}{\sin^2\theta\sqrt{\Theta}} \tag{8.40}$$

$$t = +\int^{r} \frac{[E(r^2 + a^2)^2 - aL_z(2Mr - P^2 - Q^2)]\,dr}{\Delta_r\sqrt{R}} - \int^{\theta} \frac{a^2 E\sin^2\theta\,d\theta}{\sqrt{\Theta}} \tag{8.41}$$

Applications of these equations are discussed in the following course by Bardeen.

We conclude by remarking that the linear constants of the motion E and L_z can be expressed in the form

$$E = -k^a p_a \tag{8.42}$$

and

$$L_z = m^a p_a \tag{8.43}$$

in terms of the Killing vector components k^a and m^a defined by

$$\frac{\partial}{\partial t} = k^a \frac{\partial}{\partial x^a} \tag{8.44}$$

$$\frac{\partial}{\partial \varphi} = m^a \frac{\partial}{\partial x^a} \tag{8.45}$$

It can easily be seen that the Killing equations

$$k_{(a;b)} = 0 \tag{8.46}$$

$$m_{(\alpha;b)} = 0 \tag{8.47}$$

are both necessary and sufficient for E and L_z as defined by (8.42) and (8.43) to be constant along all geodesics. In order for them to be constant along *charged* particle orbits as well it is further necessary and sufficient that the invariance conditions

$$A^a k_{a;b} + A_{b;a} k^a = 0 \tag{8.48}$$

$$A^a m_{a;b} + A_{b;a} m^a = 0 \tag{8.49}$$

be satisfied, as they are in the present case.

Now the quadratic constant of the motion (8.29) can analogously be expressed in the form

$$\mathrm{K} = a_{ab}\dot{x}^a\dot{x}^b \tag{8.50}$$

where a_{ab} is a symmetric tensor given by

$$a_{ab}\, dx^a\, dx^b = (r^2 + a^2 \cos^2\theta)\left[\frac{a^2 \sin^2\theta\, dr^2}{\Delta_r} + (r^2 + a^2)\, d\theta^2\right]$$
$$+\frac{(r^2 + a^2)\sin^2\theta\,[(r^2 + a^2)\, d\varphi - a\, dt]^2}{r^2 + a^2\cos^2\theta}$$
$$-\frac{\Delta_r a^2 \sin^2\theta\,[a\sin^2\theta\, d\varphi - dt]^2}{r^2 + a^2\cos^2\theta} \tag{8.51}$$

[This expression (8.50) for K having a similar form to the expression (8.4) for the constant m^2.] It is a standard result (see Eisenhart 1926) that the equations

$$a_{(ab;c)} = 0 \tag{8.52}$$

are both necessary and sufficient for an expression of the form (8.50) to be constant along all geodesics. Since the equation (8.52) is analogous to the Killing equations (8.46) which are necessary and sufficient for the analogous linear expression to be constant along geodesics, Penrose and Walker (1969) in a recent discussion have introduced the term Killing tensor to describe a symmetric tensor satisfying the equations (8.52). Personally I would prefer to call it a Stackel tensor since Stackel (a contemporary of Killing) actually made the first studies of the kind of Hamilton-Jacobi separability which gives rise to the existence of such a tensor (Stackel 1893).

It is to be noted that the existence of the Stackel-Killing tensor is not in itself sufficient to ensure that the quadratic expression (8.50) is constant along charged particle orbits as well as geodesics. It is easy to see that in general a necessary and sufficient condition for the quadratic expression $a_{ab}\dot{x}^a\dot{x}^b$ to be constant along charged particle orbits is that in addition to (8.52) the electromagnetic field should be such that

$$a_{a(b}F_{c)}{}^a = 0 \tag{8.53}$$

It is clear that in any space the metric tensor g_{ab} will satisfy the Stackel-Killing equations (8.52) and also that it will satisfy the condition (8.53) for *arbitrary* F_{ab}, thus giving rise to the constant of motion (8.4). Also when there are Killing vectors present their symmetrized products also form Stackel-Killing tensors, so that in the Kerr metrics there is actually a five parameter family of Stackel-Killing tensors of which five linearly independent members may be taken to be g_{ab}, the tensor a_{ab} defined by (8.51), together with $k_{(a}k_{b)}$, $k_{(a}m_{b)}$ and $m_{(a}m_{b)}$. These give rise to five linearly independent constants of which of course only four are *algebraically* independent.

The approach to the Kerr metrics which we have adopted here is to *start* with the separability properties as a fundamental postulate, and then to derive other basic algebraic properties such that the fact that the Weyl tensor is of Petrov type D. The more traditional approach has been to start from the Petrov type D property and then to *derive* the separability properties. The investigations initiated by Penrose and Walker (1969) are aimed at deriving the separability properties directly from the Petrov type D property without actually going through the explicit solution of the field equations. [It is to be emphasized however that the direct derivation of the existence of the Stackel-Killing tensor can only be an intermediate stage in this programme, and is by no means sufficient to establish the full separability properties of the Kerr metrics: there is at present no known way of deducing the separability of the *wave equation*—as opposed merely to the separability of the Hamilton-Jacobi equation—from the Stackel-Killing tensor.] Some important recent work by Teukolsky (1972) has shown that the Kerr metrics possess even stronger separability properties than those we have used in the present course, by which *higher spin* wave equations (as well as the ordinary scalar wave equation) can be at least partially separated by combining the separation methods described at the beginning of section 5 with Petrov-type analysis. This discovery is likely to prove to be of very great value in future perturbation analyses of the Kerr solutions.

Bibliography

Bertotti, B. (1959) *Phys. Rev.* **116,** 1331.
Boyer, R. H. and Price, T. G. (1965) *Proc. Camb. Phil. Soc.* **61,** 531.
Boyer, R. H. and Lindquist, R. W. (1967) *J. Math. Phys.* **8,** 265.
Carter, B. (1966a) *Phys. Rev.* **141,** 1242.
Carter, B. (1966b) *Physics Letters* **21,** 423.
Carter, B. (1968a) *Phys. Rev.* **174,** 1559.
Carter, B. (1968b) *Commun. Math. Phys.* **10,** 280.
Carter, B. (1969) *J. Math. Phys.* **10,** 70.
Debever, R. (1969) *Bull. Akad. Roy. de Belgique* **55,** 8.
Eisenhart, L. P. (1926) *Riemmannian Geometry* (Princeton)
Finhelstein, D. (1958) *Phys. Rev.* **110,** 965.
Graves, J. C. and Brill, D. R. (1960) *Phys. Rev.* **120,** 1507.
Goldberg, J. N. and Sachs, R. K. (1962) *Acta Phys. Polon.* **13.**
Hughston, L. P., Penrose, R., Sommers, P. and Walker, M. (1972) *Commun Math. Phys.* **27,** 303.
Kerr, R. P. (1963) *Phys. Rev. Letters* **11,** 238.
Kerr, R. P. and Schild, A. (1964) Am. Math. Soc. Symposium.
Kundt, W. and Trumper, M. (1962) *Akad. Wiss. Lit. Mainz* **12.**
Lemaitre, G. E. (1933) *Am. Soc. Scient. Bruxelles,* **A53,** 51.
Misner, C. (1963) *J. Math. Phys.* **4,** 924.
Robinson, I. (1959) *Bull. Akad. Pol.* 7, 351.
Schwarzschild, K. (1916) *Berl. Ber.* **189.**
Stachel (1893) *Compt. Rendus Acad. Sci.* **116,** 1284.
Teukolski, S. A. (1972) preprint Caltech.
Walker, M. and Penrose, R. (1970) *Commun. Math. Phys.* **18,** 265.