PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Low-Energy Scattering of Non-Abelian Magnetic Monopoles [and Discussion]

Michael Atiyah, N. J. Hitchin, J. T. Stuart and M. Tabor

Phil. Trans. R. Soc. Lond. A 1985 **315**, 459-469
doi: 10.1098/rsta.1985.0052

Phil. Trans. R. Soc. Lond. A **315**, 459–469 (1985)
Printed in Great Britain

Low-energy scattering of non-Abelian magnetic monopoles

BY SIR MICHAEL ATIYAH, F.R.S., AND N. J. HITCHIN
Mathematical Institute, University of Oxford, 24–29 St Giles, Oxford OX1 3LB

The Bogomolny equations represent static configurations of magnetic monopoles, for some non-Abelian group such as SU(2). Geodesic motion on this configuration space represents slowly moving interacting monopoles. Geometric information on this space can then be used to investigate the scattering of monopoles. The results show surprising features, including a 90° scattering and the conversion of angular momentum into electric charge.

1. INTRODUCTION

We begin by recalling the essential features of solitons. They are solutions of some nonlinear partial differential equations which are localized in space and preserve their identity under 'interaction'. In fact, essentially all the equations having soliton solutions arise as approximations of some more basic equation. The solitons are therefore interesting because they are approximate solutions to the more realistic equation. Finally we note that solitons have mainly arisen in situations of one space variable as in the K.d.V. equation (water in a channel) or in the nonlinear Schrödinger equation (pulses in optic fibres).

Here we shall describe a three-dimensional situation in which solitons arise. Physically the solitons represent 'magnetic monopoles' in the approximation when all velocities are small. We shall describe the interaction or scattering of these monopoles. As we shall show, this scattering exhibits surprising and interesting geometrical features that have no counterpart in one dimension.

The problem of monopole scattering in the form we present here was proposed by Manton (1982). Manton outlined the general way one should proceed, but at that time not enough was known to enable the project to be done. In the past two years, however, there has been significant progress in our understanding of the fundamental equations and this can now be exploited to solve the problem in considerable detail.

Before we introduce monopoles it may be helpful to make some very general and elementary remarks about dynamics. Suppose we have a classical mechanical system with a total energy $\mathscr{E}$ made up as usual of a kinetic part and a potential part. Then positions of stable equilibrium are given by minimizing the potential energy. These equilibrium positions can depend on continuous parameters and so will form some submanifold M of the total configuration space C. If the energy $\mathscr{E}$ is close to this minimum value of the potential energy then the dynamical flow of the system will stay close to M and will approximately be given by free flow on M. In geometric terms M will have a Riemannian metric (coming from the kinetic energy) and 'free' flow means geodesic flow for this metric.

The same principle applies (at least formally) for the dynamics of fields. Thus a nonlinear wave equation for a function $f(x, t)$, with $x \in R^3$, can be viewed as a flow on an infinite-dimensional configuration space C. Typically, minimizing the potential energy in this framework will involve an elliptic variational problem which, with appropriate boundary conditions, will lead to a *finite-dimensional* manifold M of solutions. Thus solutions of the nonlinear wave equation with energy close to the minimum should be approximated by the geodesic flow on M. There are, of course, analytical problems in this situation, due to the infinite-dimensionality of C, and these have to be investigated before rigorous statements about the approximation can be made. Following Manton (1982) we shall disregard these difficulties and concentrate simply on understanding the geodesic flow on M.

2. Magnetic monopoles

We turn now to a brief review of gauge theories and of the associated classical equations. The main idea in present theories of fundamental physics is that the electromagnetic force is only one component of a more complicated entity, which also involves the nuclear forces. The appropriate equations in this more complicated framework are essentially matrix generalizations of the Maxwell equations.

The starting point of such a theory is a compact Lie group G, called the gauge group. We shall concentrate on the simplest non-Abelian case when $G = \mathrm{SU}(2)$ so that its Lie aglebra $L(G)$ consists of 2×2 complex matrices T with $T^* = -T$ and trace $T = 0$. We use the natural norm on $L(G)$

$$|T|^2 = -\tfrac{1}{2}\,\mathrm{trace}\,T^2.$$

The basic fields are now the gauge potential $A_\mu(x,t)$ and the Higgs field $\phi(x,t)$. Both of these are functions of $x \in R^3$ and time t, and take their values in $L(G)$. The index μ takes the four values, 0, 1, 2, 3 corresponding to the time and space coordinates respectively. The covariant derivative D_μ is defined, acting on ϕ, by

$$\mathrm{D}_\mu \phi = \partial_\mu \phi + [A_\mu, \phi],$$

where $\partial_\mu = \partial/\partial x_\mu$ (and $x_0 = t$). The gauge field $F_{\mu\nu}$ is defined as the commutator $[\mathrm{D}_\mu, \mathrm{D}_\nu]$, or more explicitly

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu].$$

In analogy with the usual electromagnetic field one can then define the magnetic and electric parts of F by:

$$\text{magnetic} \quad H_1 = -F_{23}, \text{ etc.}; \quad \text{electric} \quad E_1 = -F_{01}, \text{ etc.}$$

The total energy $\mathscr{E}$ of such a Yang–Mills–Higgs field is the sum of a kinetic energy

$$\frac{1}{2}\int_{R^3} (\sum_i |E_i|^2 + |\mathrm{D}_0\phi|^2)\,\mathrm{d}x$$

and a potential energy

$$\frac{1}{2}\int_{R^3} (\sum_i |H_i|^2 + \sum_i |\mathrm{D}_i\phi|^2)\,\mathrm{d}x,$$

where the spatial index i takes the values, 1, 2, 3.

The corresponding equations of motion (evolution equations) are called the Yang–Mills–Higgs equations. Actually, in the usual model there is an additional term $\lambda(|\phi|^2 - C^2)^2$ in the potential energy and we have taken the Prasad–Sommerfield limit in which $\lambda = 0$. This can be viewed as an approximation, for small λ, to the more physical model.

If we look for solutions of finite energy $\mathscr{E}$ one can show that $|\phi|$ tends to a constant value C at infinity. If $C \neq 0$ we can, by rescaling, take $C = 1$. Now consider the map

$$\phi: S^2_\rho \rightarrow S^2$$

given by restricting ϕ to a sphere of large radius ρ in R^3: its values lie (approximately) on the unit sphere in $L(G)$. Clearly we get a well-defined degree k, which is a topological invariant. This is the magnetic charge and it determines a lower bound for the potential energy: one shows that

$$\text{potential energy} \geqslant 4\pi k.$$

Moreover equality is attained if and only if the Bogomolny equations

$$\mathrm{D}_i \phi = H_i$$

are satisfied. Solutions of these equations describe static magnetic monopoles.

The Bogomolny equations are indeed elliptic provided we work modulo gauge transformations. Recall that if $g(x)$ is a function on R^3 with values in the group SU(2) it defines a gauge transformation of A_μ and ϕ by the formulae

$$A_\mu \rightarrow g^{-1} A_\mu g + g^{-1} \partial_\mu g, \quad \phi \rightarrow g^{-1} \phi g.$$

Physically (or geometrically) the true configuration space C is the space of fields A_μ, ϕ modulo gauge transformations. Thus we expect the manifold $M_k \subset C$ representing solutions of the Bogomolny equations with magnetic charge k to be finite-dimensional. Moreover, M_k will have a natural Riemannian metric arising from the kinetic energy.

As explained in §1 one now argues that *geodesic flow on* M_k *gives the approximate evolution of the Yang–Mills–Higgs equations for* $\mathscr{E}$ *close to* $4\pi k$. This is Manton's (1982) observation. To make use of it we now have to solve the following problems:

(1) find the manifold M_k;
(2) find the metric on M_k;
(3) find the geodesics on M_k;
(4) for each point $m \in M_k$ find the associated solution of the Bogomolny equations;
(5) use (3) and (4) to describe the evolution of the solutions.

In the next sections we shall describe at least partial solutions to all these problems. For the moment we would just like to point out the difference between (1) and (4). This is best illustrated by considering the situation for a linear elliptic problem. There the analogue of M_k is a finite-dimensional linear space and (1) is only a question of finding the dimension of this space, whereas the analogue of (4) would involve exhibiting a basis of explicit solutions. In the present nonlinear case M_k is itself nonlinear and (1) involves identifying M_k as a manifold, not just finding its dimension, but (4) requires us to get the explicit solutions as functions on R^3.

3. Monopole scattering

The basic monopole occurs when $k = 1$. In this case there is an explicit solution due to Prasad & Sommerfield (1975), which is spherically symmetric about a given origin (it is referred to as the B.P.S. monopole, where B stands for Bogomolny). Moreover the energy density has a maximum at the origin so that this may be reasonably viewed as the 'location' of the monopole, regarded as some kind of particle. The Higgs field vanishes only at the origin so that elsewhere it defines a one-dimensional subspace of the Lie algebra of SU(2). The projection of $F_{\mu\nu}$ on this subspace can be interpreted as an ordinary electromagnetic field and asymptotically the B.P.S. monopole then looks like a Dirac monopole. However, near the origin the two types of monopole are different. The Dirac monopole has a point singularity whereas the B.P.S. monopole is everywhere regular: it is our soliton.

In addition to its location a B.P.S. monopole has a 'phase' angle and, although this can be removed by a gauge transformation, it is best not to do this so that the manifold M_1 is then

$$M_1 = R^3 \times S^1,$$

where S^1 is the circle parametrizing the phase. Moreover the natural metric on M_1 coincides (up to a scale factor) with the standard metric on $R^3 \times S^1$.

If our monopole now starts to move then motion in R^3 gives rise to linear and angular momentum while motion in S^1, i.e. phase change, is interpreted as electric charge. Such a 'particle' with both magnetic and electric charge is called a dyon.

For $k > 1$ the manifolds M_k have been extensively studied, culminating in the work of Donaldson (1984). In particular, one knows that M_k is non-singular and has dimension $4k$. Moreover, asymptotically M_k is a product of k (unordered) copies of M_1. For a point $m \in M_k$ in this asymptotic region the corresponding solution of the Bogomolny equations looks approximately like a superposition of k B.P.S. monopoles: this is a k-soliton.

Let us now consider the dynamics of slowly moving monopoles as proposed by Manton. Suppose we start with a point $m \in M_k$ in the asymptotic region and that we give it a small initial velocity v. Back in space this means that our collection of k far-separated B.P.S. monopoles have been given initial velocities u_i and electric charges e_i. Following the evolution of this system of monopoles corresponds, by Manton's argument, to following the geodesic on M_k starting at m in the direction of v.

Since M_k is asymptotically a product of copies of M_1 and since the natural metric on M_1 is flat it follows that M_k is asymptotically flat. Thus geodesics are asymptotically straight lines, whereas in the interior of M_k we expect curvature corresponding to the interaction of the monopoles. Instead, therefore, of attempting to follow the geodesic flow in detail we can simply ask for its scattering behaviour. This means that given (m, v) as an initial 'in-state' we ask for the 'out-state' (m', v') describing the straight line on M_k to which the geodesic is asymptotic as $t \to +\infty$. In space this means that we want the asymptotic trajectories and electric charges of the k monopoles when they have emerged from their 'collisions'. This will give the scattering of the monopoles (i.e. of the solitons).

In the remainder of the section we shall describe the results of the scattering of two monopoles for various initial data. In the next section we shall indicate the basis on which these results are derived. We shall restrict ourselves to the case when the initial monopoles are pure monopoles with no electric charge. Moreover we shall fix our origin at the centre of mass of the two monopoles.

We begin with the simplest case of a linear collision. This means that the initial motion of the two monopoles is directed towards the origin. The result in this case is that the monopoles scatter at 90° to their original motion, as indicated in the diagram:

$$\begin{array}{ccc} & \downarrow \text{in} & \\ \text{out} \leftarrow & \cdot & \rightarrow \text{out}. \\ & \uparrow \text{in} & \end{array}$$

This result is quite surprising since one might have expected the monopoles simply to stay on the same straight line (bouncing back or passing through). Moreover the scattering takes place in a certain plane, breaking the apparent symmetry of the situation, and it is not clear what determines this choice of plane.

The explanation of this surprising behaviour lies in the fact that our monopoles are not just

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

'point-particles', even asymptotically: they also have an internal structure represented by their phases. Thus the plane of the scattering has to be determined by the initial relative phases of the two monopoles. This in turn shows that the internal phase variables are linked to the ordinary space variables. This linkage is produced by the interaction since, for a single monopole, the phase and space variables are quite independent.

As we shall see later, the relative phase produces initially only exponentially small dynamic effects, which would be experimentally negligible. The result of the scattering, involving a symmetry-breaking, would therefore appear as a random effect. Nevertheless the process is entirely deterministic. The point is that however far apart the two monopoles are initially they cannot be regarded as independent particles: their phases are linked in a way that determines the outcome of the collision.

We move on now to consider a planar interaction, obtained by displacing the initial motions in the plane of the scattering just described. We shall refer to this as type I scattering, and it is illustrated in figure 1. To simplify the picture we have indicated only the motion of one of the two monopoles: the motion of the other is obtained by symmetry.

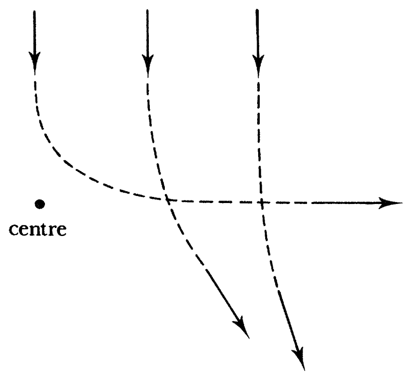


FIGURE 1. Type I scattering.

The distance apart, μ, of the initial directions of motion is the angular momentum of the monopoles and is conserved. The deflection or scattering angle S is a function of μ, which increases steadily from 0° to 90° as μ decreases from ∞ to 0. The effective forces are therefore 'repulsive'. For monopole trajectories, which are far apart (μ large), the forces and the deflection are small, while $\mu = 0$ corresponding to linear collision gives the maximum 90° scattering described earlier.

Perhaps we should emphasize that the broken lines in the diagram are merely meant to correlate the various in–out states for the various values of μ. Far out they do in fact indicate approximate trajectories of the monopoles, but in the interaction region the monopoles lose their particle-like identity and so the trajectories should not be taken seriously. In particular the diagram appears to suggest that monopoles 'turn left' as they approach. However, this is misleading, as one sees by letting μ become negative. A more accurate description of the direct collision ($\mu = 0$) is that one half of each monopole turns left while the other half turns right. This is borne out by the fact that, at the moment of collision (explained in the next section), the energy density is concentrated in a ring around the origin. If one considers the region of maximal energy density as representing the physical particle then the various stages of the direct collision process can be represented schematically by figure 2.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

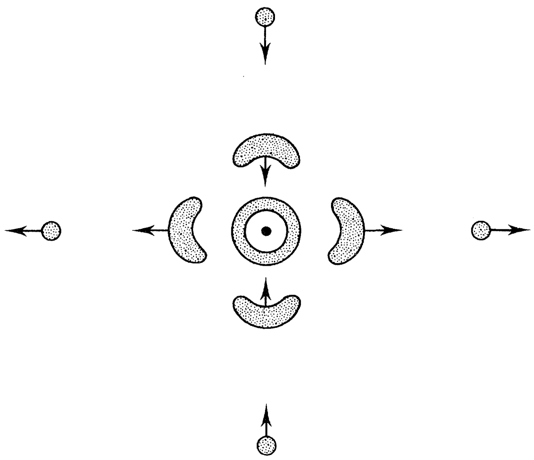
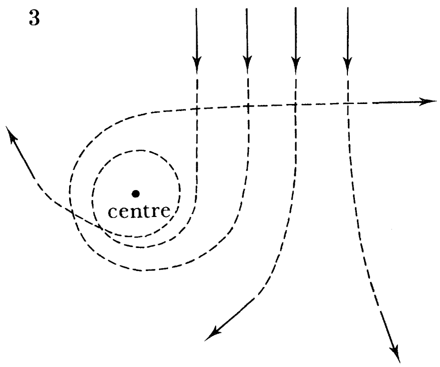


FIGURE 2. Schematic diagram of the direct collision process.

We next describe type II planar interactions, in which the initial motions are this time displaced perpendicularly to the scattering plane of the linear collision. Figures 3 and 4 indicate the scattering as a function of the initial angular momentum μ. Note finally that figure 4 is three-dimensional.

3



4

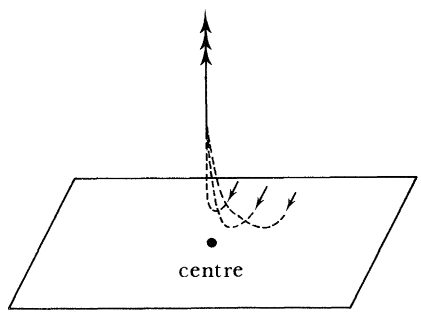


FIGURE 3. Planar diagram of the scattering for $\mu > 1$.
FIGURE 4. Three-dimensional diagram of the scattering for $\mu \leqslant 1$.

For large μ the deflection is small and repulsive, but as μ decreases the deflection turns round and becomes attractive. The attractive deflection angle $S(\mu)$ increases without bound as μ decreases to the critical value 1. *For* $\mu < 1$ *the monopoles leave the plane* along the perpendicular line through the origin. Their orbital angular momentum is therefore destroyed but, in compensation, they now acquire equal and opposite electric charges. Thus *they emerge from the collision as dyons*. This possibility was envisaged by Manton (1982) and it emphasizes once again the linkage between spatial and internal phase variables.

As P. Goddard has pointed out to us, although *orbital* angular momentum is lost, *total* angular momentum is still conserved provided one takes into account the angular momentum of the electromagnetic field (cf. Goddard & Olive (1978), §2.2). The type II collision process, by producing dyons, has converted *orbital* angular momentum into *field* angular momentum.

The fact that $\mu = 1$ is the critical value arises from our normalization of $|\phi|$ at ∞. Effectively this means that we have normalized the value of the magnetic charge, and this becomes apparent when one computes the angular momentum of the electromagnetic field as indicated above.

For values of μ slightly larger than 1, one can describe the interaction by saying that the monopoles, after collision, leave the plane for a time (of the order of $(\mu-1)^{-\frac{3}{2}}$), but they then fall back into the plane and separate.

4. The geometry of M_2^0

We shall now explain how the scattering results in the previous section are obtained. First of all, by discarding translations and an overall phase factor we can reduce the manifold M_2 (of dimension 8) to a manifold M_2^0 (of dimension 4). Asymptotically the four parameters of M_2^0 represent the relative locations of the two monopoles and their relative phase.

The geometry of M_2^0 has been studied in detail by Hurtubise (1983) and the more general case of M_k^0 has been treated by Donaldson (1984). From these we can give several descriptions of M_2^0, of which the simplest is perhaps the following. A point m of M_2^0 is represented parametrically by a pair of (unoriented) vectors $\pm \boldsymbol{x}$ and $\pm \boldsymbol{y}$ in R^3 subject to the conditions:

$$\boldsymbol{x}\cdot\boldsymbol{y} = 0, \quad \boldsymbol{y}^2 = 1.$$

Thus $\boldsymbol{y}$ is a unit vector perpendicular to $\boldsymbol{x}$ and changing the sign of $\boldsymbol{y}$ or that of $\boldsymbol{x}$, or both, gives the same point m of M_2^0.

The points $\pm\boldsymbol{x}$ are related to the spectral lines of the monopole introduced by Hitchin (1982). In fact the two spectral lines parallel to $\boldsymbol{y}$ intersect the plane $\boldsymbol{y}\cdot\boldsymbol{z} = 0$ in the two points $\boldsymbol{z} = \pm\boldsymbol{x}$. For $|\boldsymbol{x}| \to \infty$ the points $\pm\boldsymbol{x}$ give the locations of the two monopoles and $\boldsymbol{y}$ gives their relative phase. Note that this exhibits the geometrical linkage between the phase variables and the space variables referred to earlier. For $\boldsymbol{x} = 0$, the vector $\boldsymbol{y}$ has any direction and the 2-monopole is axially symmetric with axis y. Moreover the two zeros of the Higgs field now coincide at the origin. We shall therefore refer to these 2-monopoles as the 'collision states'. They give rise to a copy of the projective plane P_2 in M_2^0. Topologically we can retract M_2^0 onto P_2 (by just shrinking $\boldsymbol{x}$ to 0) so that M_2^0 is not simply connected: its fundamental group is of order 2.

If $\boldsymbol{x} \neq 0$ the 2-monopole parametrized by $(\pm\boldsymbol{x}, \pm\boldsymbol{y})$ has only a finite number of symmetries given by reflection in the three axes $\boldsymbol{x}, \boldsymbol{y}$ and $\boldsymbol{x} \wedge \boldsymbol{y}$. Suppose now we fix a direction through the origin, call it the z-axis, and consider reflection in this axis. This induces an isometry on M_2^0 whose fixed-point set consists of two totally geodesic surfaces. One surface (type I) consists of pairs $(\pm\boldsymbol{x}, \pm\boldsymbol{y})$ with $\pm\boldsymbol{y} = \boldsymbol{z}$, while the other (type II) consists of pairs with $\boldsymbol{y}\cdot\boldsymbol{z} = 0$. A type I surface intersects P_2 in just one point and is topologically a plane, while a type II surface intersects P_2 in a P_1 (i.e. a circle) and is topologically a punctured plane. Both types are surfaces of revolution and so geodesics on them are easily described (once the metric is known). This leads to the type I and type II scattering diagrams described in the previous section.

A type I surface, with $\boldsymbol{y}$ fixed, is parametrized by the pair $\pm\boldsymbol{x}$ in the orthogonal plane. Viewing $\boldsymbol{x}$ as a complex variable in this plane and putting $\xi = \boldsymbol{x}^2$ we get a natural complex parameter for the type I surface on M_2^0. Fixing the argument of ξ gives a geodesic through the origin and motion along this geodesic clearly converts the pair $\pm\boldsymbol{x}$ (at $+\infty$) to $\pm \mathrm{i}\boldsymbol{x}$ (at $-\infty$). This explains the 90° scattering of a direct collision.

To understand the geometry of type II surfaces it is helpful to consider figure 5. The line

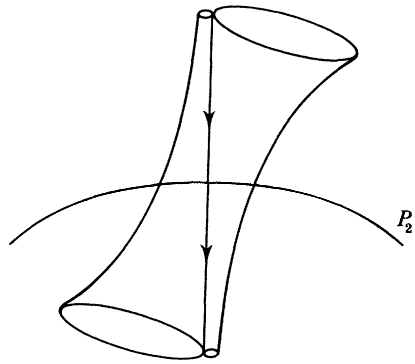


FIGURE 5. The geometry of type II surfaces.

with arrows indicates the geodesic of a direct collision. The large circles correspond to a rotation in R^3, while the small circles are phase circles. Each of the two surfaces is of type II and is funnel-shaped, approaching a cone at the large end and a cylinder at the small end. A geodesic on such a surface, starting at the cone end, will fall through the cylindrical hole if its angular momentum is small but will fall and then rise again if its angular momentum is large. This explains the complicated behaviour of the type II scattering.

5. THE METRIC ON M_2^0

The geometry of M_2^0 as explained in the previous section gives qualitative results on the monopole scattering, but quantitative results depend on further knowledge concerning the Riemannian metric. We shall now explain how one finds the explicit metric, which turns out to have remarkable features and to be of interest in its own right.

The key observation is that, for quite general reasons, all the manifolds M_k and M_k^0 are Hamiltonian (or hyper-Kähler). This means that their holonomy group reduces from $SO(4n)$ to the symplectic group $Sp(n)$, where $n = k$ or $k-1$ respectively. Alternatively it is equivalent to say that there are three covariant constant tensors I, J, K acting on the tangent space and satisfying the quaternion algebra identities

$$I^2 = J^2 = K^2 = -1, \quad IJ = -JI = K, \text{ etc.}$$

Each of I, J, K (and any combination $aI+bj+cK$ with $a^2+b^2+c^2 = 1$) gives a complex Kähler structure. The standard example is, of course, the Euclidean 4-space R^4 or alternatively $R^3 \times S^1$, i.e. the manifold M_1 parametrizing 1-monopoles. The I, J, K of $R^3 \times S^1$ in turn induce corresponding covariant constant tensors acting on all the M_k and M_k^0. The associated complex structures occur explicitly in the approach of Donaldson (1984) and can be understood in a general framework developed by Hitchin *et al.* (1985).

In dimension 4 a Hamiltonian manifold is just a self-dual Einstein manifold. Thus our manifold M_2^0 is a self-dual Einstein manifold admitting $SO(3)$ as a symmetry group: this is what remains of the Euclidean symmetries of R^3 once we have eliminated translations. Note that $SO(3)$ does not preserve the complex structures. Moreover, as we saw in the last section, $SO(3)$ acts with only finite isotropy groups (of order 4) in general. The only exceptional orbit is the P_2 representing collision states where the isotropy group is a circle. Thus the self-dual Einstein equations can be reduced to a system of ordinary differential equations. This was in fact done

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

by Gibbons & Pope (1979) and the result is the following. One first puts the metric in the form

$$ds^2 = (abc)^2\,d\eta^2 + a^2\sigma_1^2 + b^2\sigma_2^2 + c^2\sigma_3^2, \tag{5.1}$$

where σ_1, σ_2, σ_3 are a standard orthonormal base for the Lie algebra of SO(3). Here a, b, c are functions of the independent variable η and they satisfy the system of first-order differential equations

$$\frac{2}{a}\frac{da}{d\eta} = (b-c)^2 - a^2, \text{ etc.}, \tag{5.2}$$

where etc. means we permute a, b, c cyclically.

It turns out that there are essentially only two solutions of these differential equations which yield complete non-singular 4-manifolds. One solution, known for a long time, is the Taub–NUT solution, for which two of a, b, c are equal, say $a = b$. Then a, b, c are all positive and $c \to$ constant at ∞ while $a \to 0$ at the unique degenerate SO(3)-orbit. The fact that $a = b$ means that, besides the SO(3)-symmetry, there is an additional U(1)-symmetry.

The second solution is the one that gives the metric on M_2^0. For this solution a, b, c are all unequal and one, say c, is negative. At ∞ the difference $a-b$ is exponentially small and the metric looks asymptotically like the Taub–NUT solution, except that the parameter associated with c has the opposite sign. The degenerate orbit, i.e. P_2, corresponds to $a = 0$.

It turns out that (5.2) can, in a sense, be linearized. More precisely, consider the linear differential equation for a function $u(\theta)$:

$$u'' + \tfrac{1}{4}u \operatorname{cosec}^2\theta = 0. \tag{5.3}$$

For any solution $u(\theta)$ of (5.3) define $\eta(\theta)$ by

$$d\eta/d\theta = u^{-2}. \tag{5.4}$$

Finally, define a, b, c as functions of θ, and hence implicitly of η, by the three equations

$$\left.\begin{aligned} bc &= -uu' - \tfrac{1}{2}u^2 \operatorname{cosec}\theta, \\ ca &= -uu' + \tfrac{1}{2}u^2 \cot\theta, \\ ab &= -uu' + \tfrac{1}{2}u^2 \operatorname{cosec}\theta. \end{aligned}\right\} \tag{5.5}$$

One can then verify by direct substitution that a, b, c satisfy the original equation (5.2). Since (5.3) has two independent solutions while (5.4) yields a further constant of integration we see that we have a 3-parameter family of solutions of (5.2). This procedure should therefore give the general solution. Now we pick a distinguished solution to give our metric on M_2^0 by taking for u the solution of (5.3), which satisfies

$$u(\theta) \sim \theta^{\frac{1}{2}} \quad \text{as} \quad \theta \to 0. \tag{5.6}$$

This is unique up to a multiplicative constant and an explicit solution is given by

$$u(\theta) = (2\sin\theta)^{\frac{1}{2}} K(\sin\tfrac{1}{2}\theta) \quad 0 \leqslant \theta < \pi, \tag{5.7}$$

where $K(k)$ is the complete elliptic integral

$$K(k) = \int_0^{\pi/2} \frac{d\phi}{(1-k^2\sin^2\phi)^{\frac{1}{2}}}. \tag{5.8}$$

If $\theta = 0$ one gets the degenerate orbit P_2, while $\theta \to \pi$ gives the asymptotic region on M_2^0.

Using these explicit formulae one deduces, for example, that a type I surface has positive curvature and this leads to the monotonic behaviour of the scattering angle $S(\mu)$. On the other hand a type II surface has positive curvature at ∞ but then, at a certain critical distance, the curvature changes sign and then remains negative. This explains the type II scattering, where an initial repulsion changes eventually into an attraction.

It should perhaps be pointed out that the Taub–NUT solution arises from a 'special' solution of (5.2): it cannot be obtained from the general 3-parameter family of solutions described above. Presumably it is a limiting case of this general solution.

6. Further problems

While we have described the geodesic flow, and hence the monopole scattering, for some special surfaces in M_2^0 the general problem of geodesic flow on M_2^0 remains to be investigated. In particular, it would be interesting to see if there are any closed geodesics, since these would represent 'bound states' of monopoles. Probably they do not exist. One might also conjecture that the geodesic scattering is complete in the sense that every geodesic coming in from 'infinity' eventually returns to 'infinity'. Given our explicit knowledge of the metric, together with its SO(3)-symmetry these problems should prove tractable.

Another direction for investigation would be the study of M_k^0 for $k > 2$. Although SO(3)-symmetry is no longer sufficient to tie down the metric, other techniques, based on twistor theory, should permit one to find the metric explicitly. Moreover, the work of Donaldson (1984) gives a simple parametrization of M_k^0 for all k.

References

Donaldson, S. K. 1984 Nahm's equations and the classification of monopoles. *Communs math. Phys.* **96**, 387–407.

Gibbons, G. W. & Pope, C. N. 1979 The positive action conjecture and asymptotically Euclidan metrics in quantum gravity. *Communs math. Phys.* **66**, 267–290.

Goddard, D. & Olive, D. I. 1978 Magnetic monopoles in gauge field theories. *Rep. Prog. Phys.* **41**, 1357–1437.

Hitchin, N. J. 1982 Monopoles and geodesics. *Communs math. Phys.* **83**, 579–602.

Hitchin, N. J., Karlhede, A., Lindström, U. & Rocek, M. 1985 Algebraic constructions of hyperkähler manifolds. (In the press.)

Hurtubise, J. 1983 SU(2) monopoles of charge 2. *Communs math. Phys.* **92**, 137–162.

Manton, N. 1982 Multimonopole dynamics. In *Monopoles in quantum field theory*. Singapore: World Scientific.

Prasad, M. K. & Sommerfield, C. M. 1975 *Phys. Rev. Lett.* **35**, 760–762.

Discussion

J. T. Stuart, F.R.S. (*Department of Mathematics, Imperial College, London, U.K.*). Many important evolution equations come from dissipative systems and then the coefficients, as, for example, in the relevant cubic Schrödinger equation, are all complex numbers. Thus, we do not have the beauty of the algebraic and geometric structures underlying integrable systems. Even so, the following points are worth making in relation to the complex cubic Schrödinger equation: (i) for some initial conditions and coefficients, the solution may develop a singularity in a finite time, as shown by Hocking, Stewartson and Stuart; (ii) a side-band mechanism, analogous to the Benjamin–Feir instability, is important for stability of, for example, convection cells and Taylor vortices, and is due to Eckhaus; (iii) the solution may remain bounded, but not periodic (chaotic?).

M. Tabor (*Department of Applied Physics, Columbia University, New York, U.S.A.*). We have seen that completely integrable systems have both very strong group properties (for example Kac–Moody algebras) and analytic properties (for example meromorphicity). It would clearly be of interest to further develop this connection (and if possible determine which property follows from which!).

Sir Michael Atiyah (*Mathematical Institute, University of Oxford, U.K.*). In searching for a common theme or explanation of solitons one might consider, in addition to symmetry, the notion of *duality*. This may take different forms. The duality between electricity and magnetism has led Dr Olive and his collaborators to suggest that there may be some analogue in the context of non-Abelian gauge theories. Magnetic monopoles figure prominently in these ideas. There is also the wave–particle duality in quantum theory and the 'field–space democracy' by A. Schwartz, in which points (of space–time) and fields are put on an equal footing. In view of the important role of the self-dual Yang–Mills equations, as emphasized in the talk by Dr Ward, it seems that a deeper understanding of duality in all its aspects might shed light on the soliton phenomenon.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES