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

Configurations of points

M. Atiyah

Phil. Trans. R. Soc. Lond. A 2001 **359**, 1375-1387
doi: 10.1098/rsta.2001.0840

10.1098/rsta.2001.0840

THE ROYAL SOCIETY

Configurations of points

By Michael Atiyah

Department of Mathematics and Statistics, University of Edinburgh, James Clerk Maxwell Buildings, King's Buildings, Edinburgh EH9 3JZ, UK

Berry & Robbins, in their discussion of the spin-statistics theorem in quantum mechanics, were led to ask the following question. Can one construct a continuous map from the configuration space of n distinct particles in 3-space to the flag manifold of the unitary group $U(n)$? I shall discuss this problem and various generalizations of it. In particular, there is a version in which $U(n)$ is replaced by an arbitrary compact Lie group. It turns out that this can be treated using Nahm's equations, which are an integrable system of ordinary differential equations arising from the self-dual Yang–Mills equations. Our topological problem is therefore connected with physics in two quite different ways, once at its origin and once at its solution.

Keywords: configurations; flag manifolds; symmetric group

1. Introduction

In this paper I will discuss a problem in elementary geometry that has arisen from the investigations of Berry & Robbins (1997) on the spin-statistics theorem of quantum mechanics. The question concerns n distinct particles in Euclidean 3-space, idealized as points, and it aims to bridge the gap to the complex wave-functions of quantum mechanics.

Let us first recall two well-known manifolds. The first, denoted by $C_n(\mathbb{R}^3)$, is the configuration space of n distinct ordered points in $\mathbb{R}^3$. It is an open set of $\mathbb{R}^{3n}$, obtained by removing the linear subspaces of codimension 3 where any two of the points coincide. The second manifold is the famous flag manifold $U(n)/T^n$, which represents n orthonormal vectors in $\mathbb{C}^n$, each ambiguous up to a phase.

Clearly, the symmetric group Σ_n acts freely on each of these manifolds by permutation of the points and vectors, respectively. The question posed by Berry & Robbins is now simply as follows.

Does there exist, for each n, a continuous map

$$f_n : C_n(\mathbb{R}^3) \to U(n)/T^n, \tag{1.1}$$

compatible with the action of the symmetric group?

The first non-trivial case is for $n = 2$, then

$$(x_1, x_2) \to (\tfrac{1}{2}(x_1 + x_2), \tfrac{1}{2}(x_2 - x_1)), \quad x_i \in \mathbb{R}^3,$$

identifies

$$C_3(\mathbb{R}^2) \cong \mathbb{R}^3 \times (\mathbb{R}^3 - 0),$$

while

$$U(2)/T^2 = P_1(\mathbb{C}) = S^2$$

is the complex projective line or Riemann sphere. Observing that Σ_2 reverses the sign of $x_2 - x_1$, and is the antipodal map on S^2, there is therefore an obvious solution

$$f_2(x_1, x_2) = \frac{x_2 - x_1}{|x_2 - x_1|}.$$

Remarks 1.1

(1) *The case $n = 2$ already shows that the complex numbers enter the problem through the natural complex structure of $S^2 \subset \mathbb{R}^3$.*

(2) *A map f_n as in (1.1) can be viewed as assigning to n point-particles in their classical states (positions in $\mathbb{R}^3$) n quantum states (vectors in $\mathbb{C}^n$). This is the 'bridge' referred to above.*

(3) *The map f_2, in addition to its compatibility with Σ_2, is also geometrically defined so that it is*

  (i) *translation invariant,*

  (ii) *compatible with rotations in $\mathbb{R}^3$,*

  (iii) *scale invariant.*

We could ask for similar natural properties for all f_n. The only point to note is that we should require SO(3) to act on $U(n)/T^n$ via some (projective) representation on $\mathbb{C}^n$). As will emerge later, the natural choice is the irreducible representation of dimension n.

This question has already been considered (see Atiyah 2001), when a positive answer was given by constructing an explicit map f_n. However, this construction has some unsatisfactory features; in particular, it involves a choice of origin and so does not have translation invariance. One could fix the centre of mass as the origin, thus preserving translation invariance, but one pays the price elsewhere when comparing f_n for different values of n.

An alternative, more sophisticated, solution can be derived using Nahm's differential equation

$$\frac{\mathrm{d}T_i}{\mathrm{d}t} = [T_j, T_k],$$

where the T_i $(i = 1, 2, 3)$ are $n \times n$ matrix-valued functions of the real variable t and (i, j, k) is a cyclic ordering of $(1, 2, 3)$. This approach will be described in a subsequent publication with Roger Bielawski. It also has the advantage of generalizing $U(n)$ to other Lie groups, and it fits naturally into Lie theory. It is intriguing that Nahm's equation also occurs in a physical context as a method of constructing non-abelian magnetic monopoles.

Here, however, I prefer to follow the elementary approach, indicated in Atiyah (2001), and discuss various aspects of this.

2. A candidate map

Since any set of n linearly independent vectors in $\mathbb{C}^n$ can (in various ways) be orthogonalized, we can relax the unitarity condition in (1.1) and simply ask for a map

$$f_n : C_n(\mathbb{R}^3) \to GL(n, \mathbb{C})/(\mathbb{C}^*)^n. \tag{2.1}$$

An explicit reduction from (2.1) to (1.1) is given in Atiyah (2001). The only point to note is that we must choose our orthogonalization procedure to be compatible with Σ_n, i.e. not to depend on an ordering.

Equation (2.1) is equivalent to defining n points $f_n^i(x_1, \dots, x_n)$ in the complex projective space $P_{n-1}(\mathbb{C})$ that are linearly independent (i.e. do not lie in a proper linear subspace).

We shall think of $P_{n-1}(\mathbb{C})$ as the space of polynomials of degree less than or equal to $n-1$ in a complex variable $t \in S^2 = P_1(\mathbb{C})$. More formally, let (t_0, t_1) be homogeneous coordinates for $P_1(\mathbb{C})$. Then $P_{n-1}(\mathbb{C})$ is the space of homogeneous polynomials of degree $n-1$ in (t_0, t_1),

$$p(t) = a_0 t_1^{n-1} + a_1 t_1^{n-2} t_0 + \cdots + a_{n-1} t_0^{n-1}.$$

Here we assume that $p(t)$ is not identically zero and we consider it up to a scalar factor.

Considering $S^2 \subset \mathbb{R}^3$ as acted on by SO(3), the variables (t_0, t_1) are in the spin representation and p is in the (projective) irreducible representation.

For convenience of calculation, we shall usually work with the inhomogeneous coordinate

$$t = t_1/t_0,$$

with the understanding that $t = \infty$ (i.e. $t_0 = 0$) is included. We can then think of t as the variable in the complex plane given by stereographically projecting S^2 from the north pole ($t = \infty$).

With these preliminaries out of the way, we now proceed as follows. For each pair $i \neq j$, we define

$$t_{ij} = \frac{x_j - x_i}{|x_j - x_i|},$$

where $(x_1, \dots, x_n) \in \mathbb{C}_n(\mathbb{R}^3)$. Note that this is just using the map f_2 for the pair (x_i, x_j). For each i, we then define the polynomial $p_i(t)$ to be the one with roots t_{ij} $(j \neq i)$

$$p_i(t) = \prod_{j \neq i} (t - t_{ij}). \tag{2.2}$$

We now make the following conjecture.

Conjecture 2.1. *For any $(x_1, \dots, x_n) \in C_n(\mathbb{R}^3)$, the polynomials $p_1, \dots, p_n$ defined by (2.2) are linearly independent.*

If this can be proved, then putting

$$f_n^i(x_1, \dots, x_n) = p_i, \quad i = 1, \dots, n,$$

we get the desired solution of (2.1), the compatibility with Σ_n being clear from the construction. The geometric character of our definition then shows that f_n has all the same invariance properties as f_2, relative to the natural action of SO(3) on the space of polynomials.

The first crucial case that must be checked is for collinear points. But this is easy. Taking the line joining the x_i to correspond to $t=\infty$, and ordering the x_i by increasing magnitude, we see that

$$p_1 = 1,\ p_2 = t,\ p_3 = t^2, \dots,\ p_n = t^{n-1},$$

which are clearly independent.

Similar reasoning (see Atiyah 2001) shows that if the conjecture holds for $n-1$ and if we add x_n 'very far away from' $x_1, \dots, x_{n-1}$, linear independence will still hold. But this fails to provide an inductive proof, because the argument breaks down as x_n moves closer to the other points (see Atiyah (2001) for an ingenious if inelegant way around the problem).

The first non-trivial case is for $n=3$, and since three points lie in a plane in $\mathbb{R}^3$, one can give a simple geometric proof (see Atiyah 2001). Here, I shall give an alternative algebraic computation that has proved fruitful. But first, I shall digress to show how to define a normalized determinant.

3. The normalized determinant

The aim now is to define a natural determinant function

$$D(x_1, \dots, x_n)$$

whose non-vanishing is equivalent to the required independence of $p_1, \dots, p_n$. Since the p_i are only defined up to scalar factors, we have to find some normalization procedure to define D. It is easy to define the absolute value $|D|$. All we have to do is to choose each p_i to have norm 1. To preserve SO(3)-invariance, we should use the natural invariant norm on the space of polynomials. Since the representation is irreducible, this norm is unique, up to an overall factor (which can be fixed by taking $\|1\| = 1$).

If

$$p(t) = a_0 t^{n-1} + a_1 t^{n-2} + \dots + a_{n-1},$$

then

$$\|p\|^2 = \sum_{j=0}^{n} C_j^{-1} |a_j|^2, \quad \text{where } C_j = \binom{n-1}{j}. \tag{3.1}$$

This normalization of the absolute value of the determinant can be done for all sets of polynomials. However, the phase is more delicate. In fact, if $g_1, \dots, g_n$ are any polynomials, then

$$g_1 \wedge g_2 \wedge \dots \wedge g_n$$

lies naturally in a complex line-bundle over the product of n copies of $P_{n-1}(\mathbb{C})$. While this line-bundle has a natural norm, it is topologically non-trivial and so one cannot define the phase of the determinant. However, in our case, the polynomials $p_1, \dots, p_n$ have an additional property, namely that the root t_{ij} of p_i and the root t_{ji} of p_j are related by the antipodal map, i.e.

$$t_{ji} = t_{ij}^* = -(\bar{t}_{ij})^{-1}.$$

This additional property will enable us to define the phase of D. We proceed as follows.

It will be convenient to use the quaternions H, an element $u \in H$ being written as

$$u = a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d, \quad a, b, c, d \in \mathbb{R}.$$

We can also identify H with $\mathbb{C}^2$, writing

$$u = t_0 + \mathrm{j}t_1, \quad t_0, t_1 \in \mathbb{C},$$

with the complex numbers $a + \mathrm{i}b$ acting by right multiplication.

The quaternions of unit norm give a 3-sphere and the right action of $U(1)$ gives the standard Hopf fibration

$$S^1 \to S^3 \to S^2,$$

where $S^2 = P(\mathbb{C}^2)$ is the projective line of $\mathbb{C}^2$. The group $SU(2)$ acts on $\mathbb{C}^2$ (the spin representation) and induces the SO(3) action on S^2. Multiplication on the right by j defines an anti-linear map on $\mathbb{C}^2$, which induces a 'real structure' σ on $P(\mathbb{C}^2)$; this is just the antipodal map of S^2, since

$$(t_0 + \mathrm{j}t_1)\mathrm{j} = -\bar{t}_1 + \mathrm{j}\bar{t}_0.$$

Given a point $t \in P(\mathbb{C}^2)$, we can lift it to a non-zero vector $u \in \mathbb{C}^2$ with norm 1, determined up to a complex number λ of modulus 1. For the antipodal point $\sigma(t)$, we will then choose the representative $u\mathrm{j}$ as our lift. Note that this procedure is skew-symmetric between t and $\sigma(t)$. If we start from $s = \sigma(t)$ and choose a representative $v \in \mathbb{C}^2$, then $v\mathrm{j}$ becomes our choice over $\sigma(s) = t$, but since $\mathrm{j}^2 = -1$, $(v\mathrm{j})\mathrm{j} = -v$ gives the opposite sign. We will return to this point later.

Notice that a difference choice $u\lambda$ over t leads to

$$u\lambda\mathrm{j} = u\mathrm{j}\bar{\lambda}$$

over $\sigma(t)$. It is this ambiguity in phase factors we must handle.

Now† the roots t_{rs} of the polynomials $p_1, \dots, p_n$ occur in pairs (rs) and (sr), which are antipodes. We now use the natural ordering and start with t_{rs} for $r < s$ (the positive roots in Lie theory). Choosing a lift $u_{rs} \in \mathbb{C}^2$, and then the lift $u_{sr} = (u_{rs})\mathrm{j}$ over $t_{sr} = \sigma(t_{rs})$, we have definite choices for all the vectors $u_{rs} \in \mathbb{C}^2$, which we think of as linear forms in the variable $t = t_0/t_1$. More precisely, using the canonical skew form on $\mathbb{C}^2$ (invariant under $SU(2)$), we identify $(a_0, a_1) \in \mathbb{C}^2$, with the linear form $a_0 t_1 - a_1 t_0$, so that if $\alpha = a_1/a_0$ and $t = t_1/t_0$, then $t - \alpha$ is the polynomial with root α. The product

$$p_r = \prod_{r \neq s} u_{rs}$$

is then a definite choice for our polynomial p_r. At present, we have been concentrating on giving it a well-defined phase. It has been normalized as an element of the tensor product and this norm differs from that of the polynomials (e.g. for $n = 2$, the $\otimes$ norm squared is the sum of the squares of the symmetric and anti-symmetric parts). For the present, we will stick with this normalization. It is SO(3)-invariant but does not come from a Hermitian metric on the space of polynomials (it is a Banach space norm). Later we will correct to obtain the right geometric norm.

† We adopt a different index notation now to avoid confusion with the quaternions i, j, k.

Consider now the element

$$D = p_1 \wedge p_2 \wedge \cdots \wedge p_n \tag{3.2}$$

in the nth exterior power of the space $\mathbb{C}^n$ of polynomials of degree $n-1$. Since there is a canonical isomorphism

$$\Lambda^n(\mathbb{C}^n) \cong \mathbb{C}, \tag{3.3}$$

we can regard D as a complex number. In fact, there is a sign convention involved in the isomorphism (3.3). We fix this by taking the generating element of the left-hand side of (3.3) to be

$$e_1 \wedge e_2 \wedge \cdots \wedge e_n,$$

where e_r is the properly normalized monomial

$$e_r = (C_{r-1})^{-1/2} t^{r-1}, \qquad C_r = \binom{n-1}{r}.$$

For $n = 1$, this agrees with our choice of skew-form on $\mathbb{C}^2$.

If we write each normalized polynomial p_r as

$$p_r = \sum_{s=0}^{n-1} a_{rs} t^s,$$

then

$$D = \mu(n) \det A, \tag{3.4}$$

where A is the matrix of coefficients (a_{rs}) and

$$\mu(n) = \left\{ \prod_s \binom{n-1}{s} \right\}^{-1/2}. \tag{3.5}$$

We must now check that D is well defined independently of our choice of lifts u_{rs}. But (for $r < s$) changing u_{rs} to $u_{rs}\lambda$, with $|\lambda| = 1$, changes u_{sr} to $u_{sr}\bar{\lambda}$, while all other linear factors are unchanged. Thus p_r gets multiplied by λ and p_s by $\bar{\lambda}$, so that the element D defined by (3.2) is unchanged. Thus D is a well-defined function

$$D : C_n(\mathbb{R}^3) \to \mathbb{C}.$$

Although we have used a basis (t_0, t_1) of $\mathbb{C}^2$, our construction is compatible with the action of SO(2) and since this acts trivially on $\mathbb{C}$ it follows that D is invariant under SO(3). It is also more trivially invariant under translation and scale change in $\mathbb{R}^3$.

Finally, consider the action of the permutation group Σ_n. It is sufficient to consider a transposition of consecutive indices rs with $s = r+1$. Because of the way we defined our lifts to $\mathbb{C}^2$, we see that we pick up a sign change (coming from $\mathrm{j}^2 = -1$). But this cancels the sign change in the determinant. *Thus D is invariant under Σ_n.*

It is also interesting to consider the effect of reflection $x \to -x$ on D. This corresponds to the antipodal map $t \to t^*$ on S^2. This is induced by multiplication by the quaternion j on $\mathbb{C}^2$ and on linear forms gives

$$\alpha + t\beta \to -\bar{\beta} + t\bar{\alpha}. \tag{3.6}$$

For each lifting u_{rs} of t_{rs}, we can then choose the lifting

$$u^*_{rs} = u_{rs}\mathrm{j}$$

of t^*_{rs}. Note that this is consistent with our conventions on the relation between rs and sr, namely (for $r < s$)

$$u^*_{sr} = u_{sr}\mathrm{j} = (u_{rs}\mathrm{j})\mathrm{j} = -u_{rs} = u^*_{rs}\mathrm{j}.$$

Hence if

$$p_r(t) = \prod_{s \neq r} u_{rs} = \sum_{s=0}^{n-1} a_{rs} t^s,$$

then, using (3.6), our reflected polynomial is

$$p^*_r(t) = \prod_{s \neq r} u^*_{rs} = \sum_{s=0}^{n-1} a^*_{rs} t^s,$$

where

$$a^*_{rs} = (-1)^{n-1-s} \bar{a}_{r(n-1-s)}. \tag{3.7}$$

Reversing the order of the columns of the matrix A (indexed by s) produces a factor

$$(-1)^{n(n-1)/2},$$

and this precisely cancels out the signs occurring in (3.7). Hence we conclude that

$$\det A^* = \det \bar{A},$$

and so

$$D(-x) = \overline{D(x)}. \tag{3.8}$$

Since $D(x)$ is invariant under SO(3), it follows from (3.8) that $D(x)$ gets conjugated under any reflection. This implies that

$$\textit{if } x_1, \ldots, x_n \textit{ are coplanar, then } D(x) \textit{ is real.} \tag{3.9}$$

In particular, this applies to $n = 3$, since any three points are coplanar. In the next section we shall give an explicit formula for $D(x)$ when $n = 3$.

We now return to correct our normalizations. For each p_r, let $\|p_r\|$ be its invariant norm. Note that, since p_r has norm 1 in the tensor product, we certainly have

$$\|p_r\| \leqslant 1. \tag{3.10}$$

Finally, therefore, the natural geometric determinant Δ is given by the formula

$$\Delta(x) = \frac{D(x)}{\prod_1^n \|p_r\|}. \tag{3.11}$$

As far as our conjecture is concerned, we can work equally with either D or Δ; the non-vanishing of either is equivalent to the conjecture. The geometric significance as a volume shows that

$$|\Delta(x)| \leqslant 1, \tag{3.12}$$

and equality holds when the x_r are collinear, with the p_r being orthonormal. From (3.10), it follows trivially that

$$|D(x)| \leqslant 1$$

but, as is easily seen, equality is never achieved.

4. The case $n = 3$

Denote the three points x_1, x_2, x_3 simply by the symbols 1, 2, 3, as in figure 1, and let α, β, γ be the unit vectors in the directions shown, so that

$$\alpha = t_{23}, \qquad \beta = t_{31}, \qquad \gamma = t_{12}.$$

These are points on the unit circle of the complex plane and their antipodes are now just their negatives

$$t_{32} = t^*_{23} = -1/\bar{\alpha} = -\alpha \qquad (\text{since } |\alpha| = 1).$$

Note that we have chosen coordinates so that $t = 0, \infty$ are the directions perpendicular to the plane. Denote by A, B, C the angles of the triangle $\alpha\beta\gamma$ (see figure 2), so that

$$\beta\bar{\alpha} = \mathrm{e}^{2\mathrm{i}C}, \qquad \alpha\bar{\gamma} = \mathrm{e}^{2\mathrm{i}B}, \qquad \gamma\bar{\beta} = \mathrm{e}^{2\mathrm{i}A}. \tag{4.1}$$

In terms of the angles X, Y, Z of the original triangle 1, 2, 3, we have

$$2A = Y + Z, \qquad 2B = Z + X, \qquad 2C = X + Y. \tag{4.2}$$

The polynomial p_1 has roots γ and $-\beta$. If we pick

$$\frac{t-\gamma}{\sqrt{2}}$$

as the normalized representative of γ, then we must pick

$$\frac{\bar{\gamma}t + 1}{\sqrt{2}} = \frac{\bar{\gamma}(t+\gamma)}{\sqrt{2}}$$

as the normalized representative over $t = -\gamma$.

Proceeding cyclically (but remembering that, in the cyclic ordering 123, point 31 is 'negative'), we find

$$\left.\begin{aligned} p_1 &= \tfrac{1}{2}(t-\gamma)(t+\beta)\bar{\beta}, \\ -p_2 &= \tfrac{1}{2}(t-\alpha)(t+\gamma)\bar{\gamma}, \\ p_3 &= \tfrac{1}{2}(t-\beta)(t+\alpha)\bar{\alpha}. \end{aligned}\right\} \tag{4.3}$$

Using (3.4) and noting that $\mu(3) = 1/\sqrt{2}$, we see that

$$\begin{aligned} D &= \frac{-1}{\sqrt{2}}\frac{\Pi\bar{\alpha}}{8}\det\begin{pmatrix} 1 & \beta-\gamma & -\beta\gamma \\ 1 & \gamma-\alpha & -\gamma\alpha \\ 1 & \alpha-\beta & -\alpha\beta \end{pmatrix} \\ &= -\frac{\overline{\alpha\beta\gamma}}{8\sqrt{2}}(\Sigma\alpha^2\beta - 6\alpha\beta\gamma) \\ &= \frac{1}{8\sqrt{2}}(6 - \Sigma\alpha\bar{\beta}). \end{aligned}$$

Using (4.1), this can be written in terms of the angles A, B, C,

$$\begin{aligned} D &= \frac{1}{4\sqrt{2}}\{3 - \Sigma\cos 2A\} \\ &= \frac{1}{2\sqrt{2}}(\Sigma\sin^2 A). \end{aligned} \tag{4.4}$$

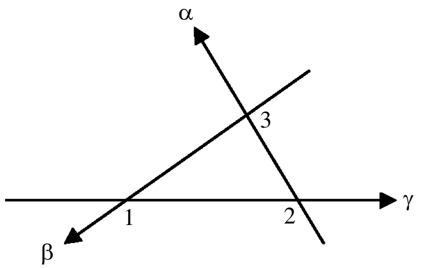

Figure 1.

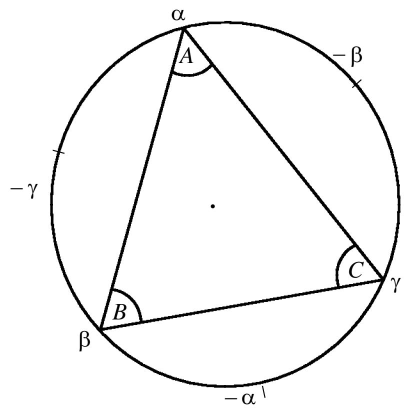

Figure 2.

Since $A + B + C = \pi$ and (from (4.1)) $A, B, C \leqslant \frac{1}{2}\pi$, we see that $D > 0$ and so *the conjecture is established for* $n = 3$.

In fact, the simple form of (4.4) enables us to be more precise. Differentiating we see that, for a critical point of D, we have

$$2\Sigma \sin A \cos A \, \mathrm{d}A = 0$$

or

$$\sin 2A \, \mathrm{d}A + \sin 2B \, \mathrm{d}B - \sin 2(A + B)[\mathrm{d}A + \mathrm{d}B] = 0.$$

This implies

$$\sin 2A = \sin 2B = \sin 2C.$$

In the allowed region of values this implies either

(i) $A = B = \frac{1}{2}\pi$, $C = 0$ (or cyclic permutation), or

(ii) $A = B = C = \frac{1}{3}\pi$.

Case (i) is the collinear case, with $D = 1/\sqrt{2}$, while case (ii) is the equilateral case with

$$D = \frac{9}{8\sqrt{2}}. \tag{4.5}$$

This shows that D has a *maximum for equilateral triangles and a minimum for collinear triples* (and no other critical points).

Now let us return to the geometric determinant $\Delta(x)$. This is given, in terms of $D(x)$, by (3.11). In our case,

$$\|p_1\|^2 = \tfrac{1}{4}(1 + \tfrac{1}{2}|\beta - \gamma|^2 + |\beta\gamma|^2) \\ = \tfrac{1}{4}(2 + 2\sin^2 A),$$

so

$$\|p_1\| = (\tfrac{1}{2}(1 + \sin^2 A))^{1/2}.$$

Thus, from (4.4),

$$\Delta = \frac{\Sigma \sin^2 A}{\Pi(1 + \sin^2 A)^{1/2}}. \tag{4.6}$$

Note that, for the equilateral triangle, this gives

$$\Delta = \frac{18}{7\sqrt{7}}.$$

5. Numerical computations†

Because of the simple behaviour of $D(x)$ for $n = 3$ (see (4.4)), it was reasonable to suppose that $\Delta(x)$, which has a maximum value 1 for collinear points, would have a minimum value for an equilateral triangle. However, Paul Sutcliffe has done computer calculations which show that the behaviour of $\Delta(x)$ is more complicated. He finds the minimum is the value $\frac{2}{3}\sqrt{2}$, and this arises in the limit where x_1, x_2 are fixed and x_3 tends to ∞ along the perpendicular bisector of x_1x_2. Note that

$$\frac{2\sqrt{2}}{3} \approx 0.9428, \qquad \frac{18}{7\sqrt{7}} \approx 0.9719,$$

the latter coming from the equilateral triangle. In fact, among isosceles triangles, the equilateral triangle is a local maximum and there is one further local minimum. In the notation of the previous section, this occurs when $B = C$ and

$$\cos A = \tfrac{1}{2}(\sqrt{5} - 1).$$

The value of Δ is then *ca.* 0.9717, just slightly less than the value for the equilateral case.

This peculiar behaviour arises from the fact that the invariant norm on polynomials changes when we alter the degree (e.g. by multiplying a power of t). This is clear from (3.1). By contrast, the tensor product norm is stable as we increase n. Adding a new point far away from a given configuration leaves D unchanged (up to an overall constant factor), whereas it decreases Δ. This suggests that, although Δ is geometrically more natural, we should consider working with D.

The case $n = 3$ gives rise to the inequality (the reverse of that for Δ)

$$|D(x)| \geqslant \mu(n). \tag{5.1}$$

This suggested, perhaps optimistically, that the lower bound (5.1) (for the absolute value) might hold for all n. This would, of course, establish our conjecture. Very

† The numerical results reported in this section were mainly obtained after the Discussion Meeting and arose directly out of that occasion.

recent computations by Sutcliffe have indicated *that (5.1) does indeed hold for* $|D|$ *for all* $n \leqslant 20$. This verifies the validity of our conjecture for $n \leqslant 20$ and is a large improvement on earlier calculations, which had only gone as far as $n \leqslant 4$.

One might also ask for configurations that give a maximum of $|D(x)|$, generalizing the equilateral case. Sutcliffe finds extremely interesting results, which will be reported on elsewhere (Atiyah & Sutcliffe 2001).

Equation (5.1) suggests that the most natural quantity to consider is $\mu(n)^{-1}D(x)$, i.e. the 'naive determinant' (without the normalization constant $\mu(n)$ inserted in (3.4)) of the coefficients of the polynomials $p_r(t)$, each given the tensor product norm. It would be interesting to understand the significance of this. One merit of this normalization is that it makes our determinant multiplicative for separated clusters.

In studying the minima and maxima of functions such as $D(x)$ or $\Delta(x)$, it is clear that it would be useful to make an appropriate compactification of the configuration space $C_n(\mathbb{R}^3)$. In fact, there is one that is very suitable for our purposes, and has already been used elsewhere, e.g. in connection with knot invariants.

The non-compactness of $C_n(\mathbb{R}^3)$ arises from two sources. In the first place, two points x_i and x_j can come together and tend to coincidence. Secondly, points can tend to infinity in $\mathbb{R}^3$. In fact, because our functions are scale invariant, these two types of non-compactness are related. We could, for instance, scale any configuration so that it lies in a ball of radius 1. To deal with points coalescing, we proceed as follows.

In the region where $x_j \to x_i$, we add one point for each limit direction. For example, when $n = 2$ and we discard the centre of mass, this would add a sphere as an internal boundary round the origin, making our space the product of S^2 with the closed half-line $r \geqslant 0$.

The process is akin to 'blowing up' in algebraic geometry, except that here we use oriented directions and get a manifold with boundary. Repeating this process, allowing several points to coalesce along fixed directions (by local rescaling), we end up with a partial compactification of $\mathcal{C}_n(\mathbb{R}^3)$ in which a polyhedral boundary has been added. Finally, if we factor out by translation and rescaling, we end up with a compact space, which we might denote by $\mathcal{C}_n(\mathbb{R}^3)$. For $n = 2$, it is just S^2. As this example shows, in factoring out the scale we must take the closures of the orbits under scale multiplication. Our definition of the partial compactification ensures that the closed orbits are all disjoint.

Our polynomials p_i are defined by the directions t_{ij} and these are preserved under this compactification. Hence the p_i and our determinant functions extend to give maps defined on $\mathcal{C}_n(\mathbb{R}^3)$.

6. Some generalizations

As pointed out in Atiyah (2001), our conjecture about the linear independence of the polynomials $p_1, \ldots, p_n$ has a natural generalization to $C_n(H^3)$, the configuration of ordered distinct points in *hyperbolic* 3-space. Given two points $x_i, x_j \in H^3$, we define t_{ij} to be the point on the 2-sphere 'at ∞' along the oriented geodesic $x_i x_j$. This S^2 has a natural complex structure, since the group $\mathrm{SL}(2, \mathbb{C})$ is (up to ± 1) the (oriented) isometry group of H^3. This means we can define the polynomials p_i as

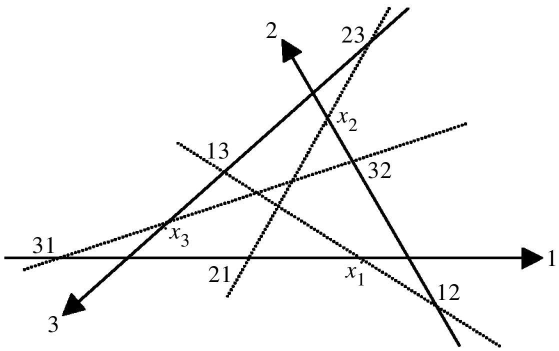

Figure 3.

before and conjecture their linear independence. A geometric proof of this conjecture for $n = 3$ is given in Atiyah (2001).

Since there is no metric on the space of polynomials invariant under $\mathrm{SL}(2, \mathbb{C})$, we cannot define a fully invariant determinant†. However, if we fix an origin in H^3, the symmetry gets reduced to $\mathrm{SO}(3)$ and we can then define the analogues of $|D(x)|$ and $|\Delta(x)|$. There seems to be no obvious way to define the complex phase because the points t_{ij} and t_{ji} are no longer antipodal.

H^3 has a constant scalar curvature. This plays no role in the definition of the polynomials p_i, but it will enter in the computation of the normalized determinants. Essentially, there is an intrinsic scale in H^3 and, in an appropriate sense, we can study the dependence of our determinants on this scale, i.e. on the curvature. As the curvature tends to zero, we recover the Euclidean case.

Because the isometry group of H^3 is $\mathrm{SL}(2, \mathbb{C})$, which is also the Lorentz group, this suggests that it might be possible to formulate a further generalization involving Minkowski space. We can start as follows. Let $\xi_1, \ldots, \xi_n$ be the n (non-intersecting) world-lines of n moving particles (or 'stars'). On each world-line ξ_i, pick an event x_i. Imagine an observer at this point of space-time. He looks out into the sky and sees $n-1$ other stars on his 'celestial sphere', i.e. on the base of his backward light-cone. These positions describe the light-rays emitted by the other stars, at some time in their past, which happen to arrive at star i at the time (or event) x_i.

In this way, we can again define points $t_{ij} \in S^2$, where we identify all celestial spheres by parallel translation in Minkowski space. This gives us our polynomials $p_i, \ldots, p_n$, and we can again ask if they are linearly independent.

It is not hard to see that the cases we have previously studied for $\mathbb{R}^3$ and H^3 are indeed special cases of this Minkowski situation. The Euclidean case just corresponds to n static stars relative to a definite space-time decomposition of Minkowski space. The world-lines are just parallel to the time-axis and the points t_{ij} are essentially the same as the ones we had before.

To get the hyperbolic space situation, we take our n stars to have arisen from a 'big bang' and to have exploded from this past event at uniform (but not necessarily the same) velocities. Again, since ξ_i, ξ_j now lie in a common plane, it is easy to see that t_{ij} is the same as before.

† In fact there is an alternative approach that gives a fully invariant determinant function in the hyperbolic case (see a forthcoming joint paper (Atiyah & Sutcliffe 2001)).

Since the polynomials p_i can be defined for any world-lines, not necessarily straight lines, one might optimistically wonder whether linear independence held in full generality. It does so for $n = 2$, as a little thought will show, but it fails for $n = 3$, even for world-lines that lie in a three-dimensional linear subspace $\mathbb{R}^{2,1}$ of Minkowski space $\mathbb{R}^{3,1}$. This will be the dynamic version of the $n = 3$ coplanar case studied in § 4. To avoid three-dimensional pictures, we shall just look at the spatial paths in a plane and consider our stars to be moving along these paths. We shall give a counterexample to the optimistic conjecture, even when the spatial paths are straight lines (but the velocities are not uniform). We consider the figure 3 (which has cyclic symmetry). The sides 1, 2, 3 of the equilateral triangle are the paths ξ_1, ξ_2, ξ_3, and, at time 0, our stars are located at the points x_1, x_2, x_3, one-quarter of the distance from each vertex. The dotted lines joining the x_i to the mid-points of the opposite sides of the triangle represent the paths of the light-rays from the past of the other stars. Note that the point 12 on line 2 is indeed in the past ($t < 0$) of the trajectory ξ_2 of x_2, and similarly for all the others. Thus this diagram represents a possible scenario of our three stars. However, the directions 12 and 13 (the light-rays reaching x_1) are in the antipodal directions. Since this holds by cyclic symmetry for the others, we see that the three lines joining the points $t_{ij}t_{ik}$ are concurrent (note that t_{ij} is not the point ij in the diagram, but the direction from x_i to ij). But this shows that the three polynomials p_i (which represent these lines) are linearly dependent. This geometrical reasoning is similar to that used in the proof of the Euclidean conjecture for $n = 3$ in Atiyah (2001).

A careful inspection of the diagram will show, however, that the velocities of the stars required to produce it cannot be uniform. Consider, for instance, star x_1. In its past history light emitted at 31 and 21 reaches x_3 and x_2, respectively, at time $t = 0$. If the velocity of x_1 was uniform, say k times the velocity of light (with $k < 1$), then we should have the following relations for distances:

$$d(31, x_1) = kd(31, x_3), \qquad d(21, x_1) = kd(21, x_2).$$

While the second of these is consistent with $k < 1$, the first is clearly not.

It might be possible to modify the geometry of this example to satisfy the relations above (for all three points), and so be consistent with straight lines in Minkowski space (i.e. uniform velocity). Alternatively, we might hope to prove the general conjecture about linear independence of the polynomials whenever $\xi_1, \dots, \xi_n$ are straight lines in Minkowski space. This would be a satisfactory generalization of the $\mathbb{R}^3$ and H^3 cases. It would also bring us back to physics in an interesting way, since it combines relativity with spin, both ingredients of the standard proof of the spin-statistics theorem. It is also very much in the spirit of Roger Penrose's ideas, in which the complex structure of the celestial sphere should tie in (or explain) the role of complex numbers in quantum mechanics. Recall that we have interpreted $p_1, \dots, p_n$ as 'quantum states' associated to the classical point states $x_1, \dots, x_n$.

References

Atiyah, M. F. 2001 The geometry of classical particles. In *Surveys in differential geometry,* vol. 7. Cambridge, MA: International Press.

Atiyah, M. F. & Sutcliffe, P. 2001 The geometry of point particles. hep-th/0105179 (preprint).

Berry, M. V. & Robbins, J. M. 1997 Indistinguishability for quantum particles: spin, statistics and the geometric phase. *Proc. R. Soc. Lond.* A **453**, 1771–1790.