SURVEYS IN DIFFERENTIAL GEOMETRY, 2000
Vol. VII ©2000, INTERNATIONAL PRESS
pp. 1-15

THE GEOMETRY OF CLASSICAL PARTICLES

MICHAEL ATIYAH

1. Introduction

In a recent paper [2] Berry and Robbins have described a classical approach to the spin-statistics theorem of quantum physics. In the course of their investigation they were led to a purely geometrical question in 3-dimensional Euclidean space. This paper grew out of an attempt to answer their question and to understand its significance.

The Berry-Robbins problem concerns two very well-known spaces:

(i) the configuration space $C_n(R^3)$, parametrizing n distinct ordered points in R^3;

(ii) the flag manifold $U(n)/T^n$, parametrizing 'flags' i.e., n ordered mutually orthogonal one-dimensional vector subspaces of C^n (here $U(n)$ is the unitary group, T^n the diagonal subgroup fixing a given flag).

The symmetric group Σ_n acts freely on both these spaces by permuting the points or the subspaces.

The problem is the following:

(1.1) Does there exist (for all n) a continuous map $f_n : C_n(R^3) \to U(n)/T^n$ which is compatible with the action of Σ_n?

In this paper I shall (in §6) give a positive answer to this problem using only elementary geometry. However, this solution has some unsatisfactory features and a much more elegant solution may exist. This

This research was supported in part by the Clay Mathematics Institute.

depends on a conjecture which is remarkably simple to state but appears to be difficult to settle. I shall explain this in §4 and I shall also describe a natural generalization to hyperbolic 3-space in §5.

There are clear indications that alternative approaches are possible which, though more complicated geometrically, have interesting physical interpretations. I hope to follow up on these in subsequent publications. There are also further generalizations in a number of directions which I hope to explore.

In §2 and §3 I make some preliminary comments on the topological aspects of the problem. These are helpful steps on the way to the solution.

2. Elementary comments

Before proceeding to study our problem seriously it is helpful to make a few preliminary remarks.

The case $n = 1$ is trivial, so the first interesting case arises from $n = 2$. By fixing the centre of mass we see that

$$C_2(R^3) = R^3 \times (R^3 - 0)$$

with Σ_2 being the antipodal map on the second factor. Since

$$U(2)/T^2 = P_1(C) = S^2$$

is the complex projective line, with Σ_2 as the antipodal map, it is clear how to define f_2. We simply use radial projection

$$(R^3 - 0) \to S^2.$$

This special case already shows two things. First it explains why the problem is a 3-dimensional one. We can define a configuration space $C_n(R^N)$ for all N, but only for $N = 3$ does the radial map

$$(R^N - 0) \to S^{N-1}$$

end up with S^2.

The second point, intimately related to the first, is that the complex numbers appear in our problem through the Riemann sphere "at ∞" in R^3. In fact the notable feature about the Berry-Robbins problems is that it relates **real** geometry in R^3 to linear independence of vectors over the **complex** numbers. Recalling that complex superposition is the hall-mark of quantum theory we see here the germ of a link between classical geometry and quantum theory. This point (in a Minkowski framework, to which we shall return in a later paper) has been persistently emphasized by Roger Penrose.

Mathematically this link between R^3 and the Riemann sphere is best understood by regarding R^3 as the space of imaginary quaternions. This is systematically exploited in the theory of hyperkähler manifolds and in the related Penrose twistor theory.

Note that our map f_2 has the following obvious properties.

(1) f_2 is translation invariant.

(2) f_2 is scale invariant.

(3) f_2 is compatible with the action of $SO(3)$ on both sides.

In fact these properties are easily seen to characterize f_2 uniquely (modulo composition with a fixed element of $0(3)$).

Our task, for general n, is to search for maps f_n which are generalizations of f_2. It would in particular be reasonable to ask for analogues of properties (1), (2), (3), but for (3) we have to specify an action of $SO(3)$ on the flag manifold. Any representation of $SU(2)$ on C^n will induce a representation of $SO(3)$ on the complex projective space $P(C^n)$ and more generally on the flag manifold $U(n)/T^n$ (which can be viewed as a subspace of $P(C^n)^n$). We shall see later that the correct choice of representation of $SO(3)$ is (for each n) given by the (unique) irreducible representation of dimension n.

Clearly one would also like all the maps f_n, for different n, to be somehow related to each other. This can be made more precise by considering a "cluster decomposition" when the n points fall into two separate clusters of r and $n-r$ points far apart from each other. If we write $x = (x_1, ..., x_r)$ and $y = (y_1, ..., y_{n-r})$, we shall denote by $x * y$ the ordered n-tuple formed by putting x and y 'far apart'. We then expect an asymptotic formula.

(4) $$f_n(x * y) \sim f_r(x) \times f_{n-r}(y)$$

where on the right we use the obvious product map

$$U(r)/T^r \times U(n-r)/T^{n-r} \to U(n)/T^n.$$

While properties (1)-(4) put constraints on f_n they will not (for $n \neq 2$) uniquely determine f_n. The Berry-Robbins problem is, in the general case, a problem in equivariant homotopy theory and only the homotopy class will be unique. However we may hope for some natural representative map which has special geometric or physical significance.

3. Homology calculations

A standard test in homotopy theory for a putative map is to check its possible effect on homology or cohomology. It is frequently possible to disprove the existence of a map by showing inconsistency in its effects on cohomology. For a continuous map

$$f : X \to Y$$

we get an induced map

$$f^* : H^*(Y) \to H^*(X)$$

on the integral cohomology rings. If a group Σ acts on both X, Y and is compatible with f, then f^* must also be compatible with the Σ action on $H^*(X)$, $H^*(Y)$. This gives strong restrictions.

As a preliminary test let us consider the (integral) cohomology rings of our two spaces

$$X = C_n(R^3) \qquad Y = U(n)/T^n.$$

These have well-known descriptions [1] [3]. Each is multiplicatively generated by its 2-dimensional elements, and

$$H^2(X) \text{ is generated by } x_{ij} (i \neq j) \text{ with } x_{ij} = -x_{ji}$$

$$H^2(Y) \text{ is generated by } y_i \text{ with } \Sigma y_i = 0.$$

The action of Σ_n in each case is by permutation of the indices.

Any map $f : X \to Y$ will induce on H^2 a map f^* with

$$f^*(y_i) = \sum_{j \neq i} a_{ij} x_{ij}.$$

Compatibility with Σ_n implies that

$$a_{ij} = \lambda \text{ for all } i, j$$

and some integer λ.

For $n = 2$, and $f = f_2$, we have $\lambda = 1$ and the asymptotic requirement (4) will then imply $\lambda = 1$ for all n. Note that, since $x_{ij} = -x_{ji}$, we do indeed have

$$\sum_i f^*(y_i) = \sum_i \sum_{j \neq i} x_{ij} = 0$$

as required by $\Sigma y_i = 0$.

One might expect some inconsistency to appear in the ring structure, but in fact f^* turns out to be fully compatible with the multiplication. This follows from the following known facts [3] [4] (for the analogous case of R^2):

(i) $H^*(Y)$ is the polynomial ring in $y_1, ..., y_n$ modulo the symmetric functions (of positive degree).

(ii) Both $H^*(Y)$ and $H^*(X)$, after tensoring with C, give the regular representation of Σ_n.

From (ii) it follows that the only invariant element in $H^*(X)$ is in $H^0(X)$ and so all the symmetric functions of the y_i (of positive degree) get mapped to zero by f^*. In view of (i) this shows that f^* does indeed extend from $H^2(Y)$ to the whole of $H^*(Y)$.

Despite (ii), which gives an abstract Σ_n-isomorphism between $H^*(X) \otimes C$ and $H^*(Y) \otimes C$, the map f cannot possibly give such an isomorphism for $n > 2$. In fact f^* is not even an isomorphism on H^2 because the ranks of $H^2(X)$ and $H^2(Y)$ are different (for $n > 2$). The way the pieces of the regular representation of Σ_n get their dimensions differs in the two cases as is clear from the explicit formulae for the Poincaré polynomials:

$$
\begin{aligned}
P(X) &= (1+t^2)(1+2t^2) \dots (1+(n-1)t^2) \\
P(Y) &= (1+t^2)(1+t^2+t^4) \dots (1+t^2+t^4+\dots+t^{2n-2}).
\end{aligned}
$$

In both cases, putting $t = 1$, gives $n!$, the order of Σ_n. Only for $n = 2$ do the two Poincaré polynomials coincide.

The conclusion of this little excursion into cohomology is that there appears to be no obvious obstruction to the existence of the required map f. On the other hand, algebraic topology alone has difficulty in giving a positive solution. With a bit more effort one can construct a map f having the desired cohomological properties but making it genuinely compatible with Σ_n (beyond its cohomological action) is far from easy. For this a direct geometrical construction must be sought. The cohomology calculations do provide a clue which will be followed up in the next section.

4. A candidate map

We shall now, by direct and elementary construction exhibit a map f_n which will be our first candidate for a solution. Its success depends however on the non-vanishing of a certain determinant and this, though highly probable, has not yet been established. First let us recall the polar decomposition

$$GL(n, C) = P \times U(n)$$

where P is the space of positive self-adjoint matrices.

If $g = pk$ is the decomposition of $g \in GL(n, C)$ then

$$\begin{aligned} gg^* &= pkk^*p^* \\ &= pp^* \text{ since } k \in U(n) \\ &= p^2 \text{ since } p = p^*. \end{aligned}$$

Hence p is the **positive square root** of the self-adjoint matrix gg^* and then $k = p^{-1}g$ gives the explicit retraction map

$$\phi : GL(n, C) \to U(n).$$

Note that this is compatible with the action of $U(n)$ on both left and right:

if $g = pk$ and $u \in U(n)$, then

$$\phi(gu) = \phi(pku) = ku = \phi(g)u$$

and $\phi(ug) = \phi(upk) = \phi(upu^{-1}.uk) = uk = u\phi(g)$.

In particular, for any permutation σ,

$$\phi(g\sigma) = \phi(g)\sigma$$

so that ϕ is equivariant with respect to the action of the symmetric group Σ_n on the column vectors of the matrices. Moreover, factoring on the right by the action of the maximal torus T^n of $U(n)$, we get an induced Σ_n-equivariant retraction (still denoted by ϕ)

$$\phi : GL(n, C)/T^n \to U(n)/T^n.$$

What we shall try to do is to construct a Σ_n-equivariant map

$$F_n : C_n(R^3) \to GL(n, C)/T^n$$

and then follow this by the retraction ϕ, so that $f_n = \phi F_n$.

We now describe our construction. Given a configuration $x = (x_1, ..., x_n) \in C_n(R^3)$ we consider the points on the unit 2-sphere S^2 given by

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

In other words t_{ij} is the **direction** of the line $x_i x_j$ (equivalently $t_{ij} = f_2(x_i, x_j)$). Fixing i and taking all values $j \neq i$ we get $n-1$ points on S^2 and hence a point

$$p_i \in S^{n-1}(S^2) = P_{n-1}(C)$$

in the symmetric product. If we identify S^2 with the Riemann sphere we can think of each t_{ij} as a complex number (or ∞), and then p_i is simply the polynomial whose roots are the t_{ij} for $j \neq i$. The coefficients of the polynomial are the homogeneous coordinates of $P_{n-1}(C)$. The polynomial p_i is only determined up to a non-zero scalar, but we can normalize so that $||p_i|| = 1$ in C^n and then p_i is only ambiguous up to a phase factor. The metric we use in C^n is the natural metric induced by a metric in C^2. This means it is invariant under $SU(2)$, the double cover of $SO(3)$. Since the representation is irreducible the metric is uniquely determined up to an overall scale. We now have n (normalized) polynomials, associated to $x \in C_n(R^3)$, namely

$$p_1, p_2, ..., p_n.$$

Let us assume that these are **linearly independent** (we shall discuss this question later). Then the matrix g, whose columns are $p_1, ..., p_n$, is well-defined in $GL(n, C)/T^n$, the factor T^n corresponding to the ambiguous phase factors. Clearly, from their construction, permuting the points $x_1, ..., x_n$ just leads to the corresponding permutation of $p_1, ..., p_n$. Hence

$$x \to (p_1, ..., p_n)$$

gives a Σ_n-equivariant map

$$F_n : C_n(R^3) \to GL(n, C)/T^n.$$

Following this by the retraction ϕ would then give us our map

$$f_n : C_n(R^3) \to U(n)/T^n.$$

This is compatible with Σ_n and also with the action of $SU(2)$ (or $SO(3)$). Note that $SU(2)$ acts on the left on $U(n)/T^n$.

Instead of working with matrices a more invariant way of thinking is to say that the p_i define a linear map $C^n \to V$, where V is the $(n-1)^{\text{th}}$ symmetric power of C^2. The symmetric group Σ_n acts by permuting the basis of C^n, while $SU(2)$ acts on V.

Clearly our map f_n would have the invariance properties (1)(2)(3) of §2.

So our problem is solved, provided we can establish the following.

(4.1) Conjecture. For any configuration of points in $C_n(R^3)$ the complex polynomials $p_1, \dots, p_n$ are linearly independent.

We now turn to examine this point. At first glance one might expect the p_i would become dependent for certain degenerate configurations of points. In particular the worst case would seem to be when the points $x_1, \dots, x_n$ are all **collinear**. Let us consider this case, and choose our complex coordinate t on S^2 so that the line of the x_i gives $t = 0$ and $t = \infty$. We can assume that $x_1, \dots, x_n$ appear on the line in that order and let $x_1 x_n$ be the direction $t = \infty$. Then we see that our polynomials p_i are:

$$\begin{aligned} p_1 &= 1 \\ p_2 &= t \\ &\vdots \\ p_n &= t^{n-1} \end{aligned}$$

and these are clearly linearly independent. This is very encouraging!

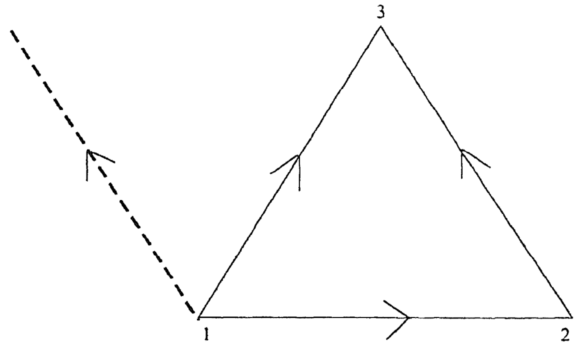
Suppose next that our n points fall into 2 clusters x and y far apart, as discussed in §2. We again choose coordinates on S^2 so that $t = \infty$ is the direction connecting the clusters (from x to y). We then find that

$$\begin{aligned} p_i(x * y) &= p_i(x) & 1 \leq r \\ &= t^r p_{i-r}(y) & 1 > r. \end{aligned}$$

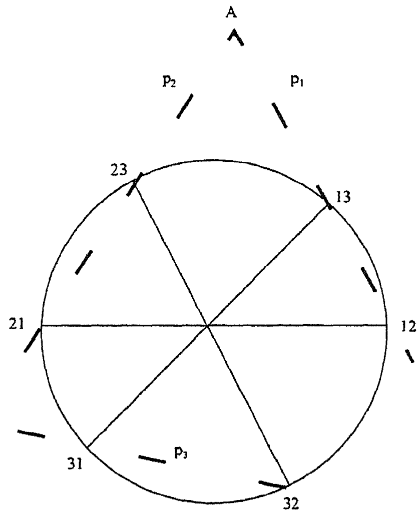
Since the $p_i(x)$ have degree $r - 1$, linear independence for $x * y$ would follow from linear independence for x and y separately. This also would establish property (4) of §2.

We have therefore reduced our problem to that of establishing the linear independence of the polynomials $p_i, \dots, p_n$ for all configurations $x_i, \dots, x_n$ in R^3. We have seen that linear independence holds in the collinear case (which includes the trivial case $n = 2$), so the first significant case is for $n = 3$. We shall now give a direct geometric proof for $n = 3$.

We can assume that the three points x_1, x_2, x_3 are not collinear (since this case is already covered) and so they lie in a definite plane. Starting with the triangle $1, 2, 3$, take 1 (i.e., x_1) as origin and draw the parallel to 23 through 1.



Taking intersections of the extended lines (both directions) with the unit circle we get the following picture, where t_{ij} is just the point denoted by ij on the circle:



The essential point is the order in which the points i (which are all distinct) appear on the circle.

The three polynomials p_1, p_2, p_3 can be represented (linearly) by the three dotted lines in the picture. Linear independence means that the three lines should not be concurrent. But this is clear: p_1 and p_2 meet in a point A in the upper-half plane (above the diameter joining 1 and 2), and every line through A meets the circle either in no points or in a pair, one in each half-plane, while p_3 meets the circle in 2 points in the lower half-plane.

The case $n = 4$ is already much more complicated and I know of simple geometric proof. However computer calculations by Robbins appear to provide convincing evidence.

Needless to say, since $n = 4$ remains unproved, the same is true of the general case. It is possible that, with sufficient ingenuity, an elementary proof can be constructed for all n but despite my publicizing the problem on several occasions no solution has yet emerged.

Since a direct assault along these lines has reached an impasse we can try to think of ways round the problem. One well known mathematical procedure, when faced by an intractable problem, is to generalize it in the hope that new insight might follow. We shall follow this strategy in the next section and see that it does in fact lead to a solution of the Berry-Robbins problem. This solution, although explicit and elementary, does suffer from some aesthetic drawbacks and it should not be regarded as the end of the story, as I shall explain in subsequent papers.

5. The hyperbolic analogue

In this section I shall replace Euclidean space R^3 by hyperbolic space H^3. The curvature will play no role and it can be normalized to -1, though at a later stage we may wish to allow it to vary. We can therefore introduce the space $C_n(H^3)$ representing ordered configurations of n distinct points in H^3, and we can again ask for an analogue of the Berry-Robbins conjecture. **Is there a continuous map**

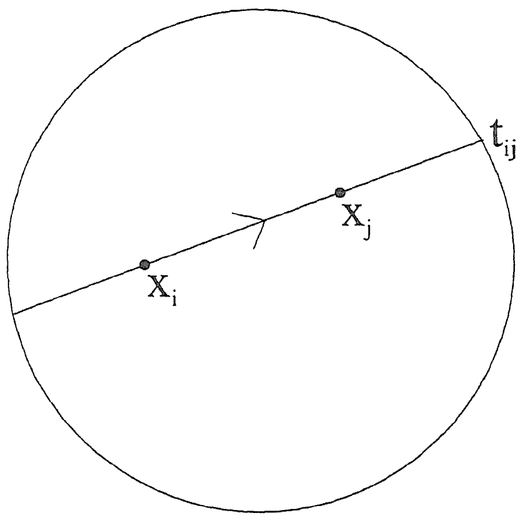
$$f : C_n(H^3) \to U(n)/T^n$$

compatible with the action of Σ_n? Since H^3 and R^3 are topologically equivalent this question, as it stands, is equivalent to the original conjecture concerning $C_n(R^3)$. The problem however becomes more natural and interesting if we ask for the existence of a continuous map

$$f : C_n(H^3) \to GL(n, C)/(C^*)^n \tag{5.1}$$

which is compatible with Σ_n **and with the action of** $SL(2, C)$. Here $SL(2, C)$ acts on H^3 as its group of isometries and on $GL(n, C)$ via the irreducible n-dimensional representation. Note that there is no invariant metric on C^n, so we use the full diagonal $(C^*)^n$, not just T^n.

The construction of §4, based on the polynomials, $p_1, ..., p_n$, can be repeated here in essentially the same way. Given two distinct points x_i, x_j of H^3 we define t_{ij} to be the "point at ∞" along the (oriented) geodesic $x_i x_j$. If we take the projective model of H^3, as the interior of the unit ball in R^3, the geodesics are just the usual straight lines, and t_{ij} is just the point where the (oriented) line $x_i x_j$ meets the unit sphere of R^3



As before p_i is the polynomial with roots t_{ij} $(j \neq i)$, and is determined up to a scalar. Again we may

(5.2) Conjecture. For any configuration in $C_n(H^3)$ the complex polynomials $p_1, \ldots, p_n$ are linearly independent.

If this conjecture is true then the polynomials p_i define the required map (5.1).

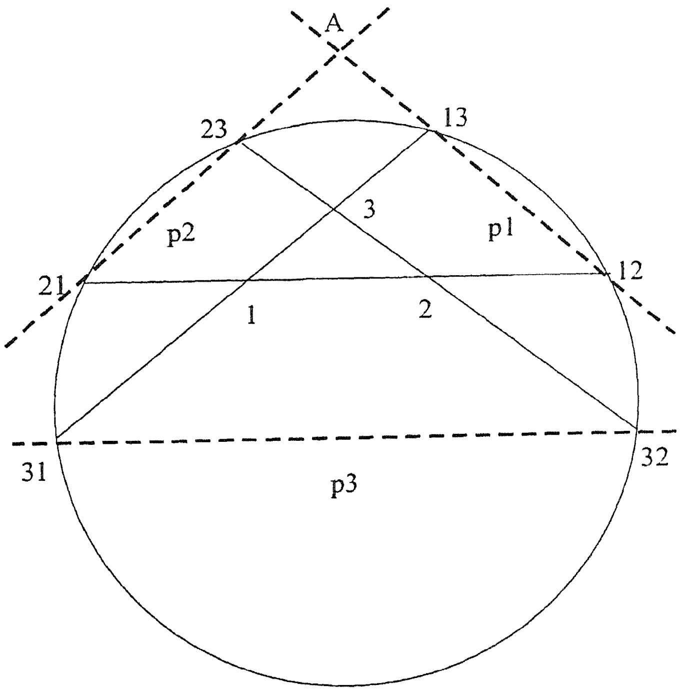
Evidence[1] for the conjecture parallels that for the Euclidean case. The collinear case follows by the same argument as before, and the case $n = 3$ can be proved explicitly (see below). The cluster decomposition property also holds as before.

The fact that the p_i are constructed purely geometrically means, both for R^3 and for H^3, that the construction is compatible with the appropriate isometry group. In this respect the hyperbolic case is more interesting and more natural, because the whole group $SL(2, C)$ (modulo ± 1), acts effectively, whereas in the Euclidean case the translations act trivially and only the orthogonal group acts effectively.

As promised we now prove the conjecture for $n = 3$. The picture is very similar to the Euclidean case.

---

[1]**Added in proof:** Further computer calculations by P. Sutcliffe have now extended the evidence for the conjecture up to $n = 20$



As before the lines p_1, p_2, p_3 are not concurrent for essentially the same reason as before. Lines through A which cut the circle do so at points on either side of the line joining 1 and 2, whereas p_3 cuts the circle at two points (31) and (32), both of which are below the line.

Any fixed configuration of R^3 lies inside some ball and, by rescaling, this can be taken as the unit ball, so that our configuration can now be thought of as in H^3. Varying the size of the ball corresponds to changing the curvature of H^3 and, if the size of the ball tends to ∞ the curvature tends to zero, so that we recover flat space. Moreover, taking a minimal ball, i.e., one which has a point of the configuration on its boundary corresponds to a configuration in H^3 with a point "at ∞". This case of the conjecture reduces easily to that for $(n-1)$ points, so this suggests a possible inductive proof of the conjecture for R^3 by using balls of different sizes, i.e., copies of H^3 with different curvatures. In fact, and perhaps surprisingly, such a naive procedure can actually be made to work and will be explained in detail in the next section. This will justify our belief in the advantage of generalizing difficult conjectures!

6. An explicit map

In this section we shall construct an explicit map

$$f_n : C_n(R^3) \to U(n)/T^n$$

compatible with the action of Σ_n and $SO(3)$.

We shall break the translation symmetry of the problem by fixing an origin. We shall also identify, by radial projection, all spheres with this origin as centre. Given a configuration $x_1, ..., x_n$ of R^3 we shall define polynomials, $p_1, ..., p_n$ with roots t_{ij} $(j \neq i)$ by a slight variant of the construction in §5. To define the roots of p_i we shall distinguish between the values of j:

(a) if $|x_j| \geq |x_i|$ we take $t_{ij} = x_j$(on the sphere of radius $|x_j|$).

(b) if $|x_j| \leq |x_i|$ we take t_{ij} to be the second intersection of the line $x_i x_j$ with the sphere of radius $|x_i|$.

Roughly speaking we treat points inside the ball of radius $|x_i|$ as in hyperbolic space while leaving alone the external points. Note that when $|x_i| = |x_j|$ (a) and (b) agree, so that t_{ij} and hence p_i is a continuous function of $(x_1, ..., x_n)$.

Although the points are treated differently in (a) and (b) our construction is still compatible with Σ_n. The action of Σ_n should be viewed as just altering the labels (suffixes) of the points, and the dichotomy leading to (a) or (b) does not depend on the labelling, but on the intrinsic geometry of the configuration.

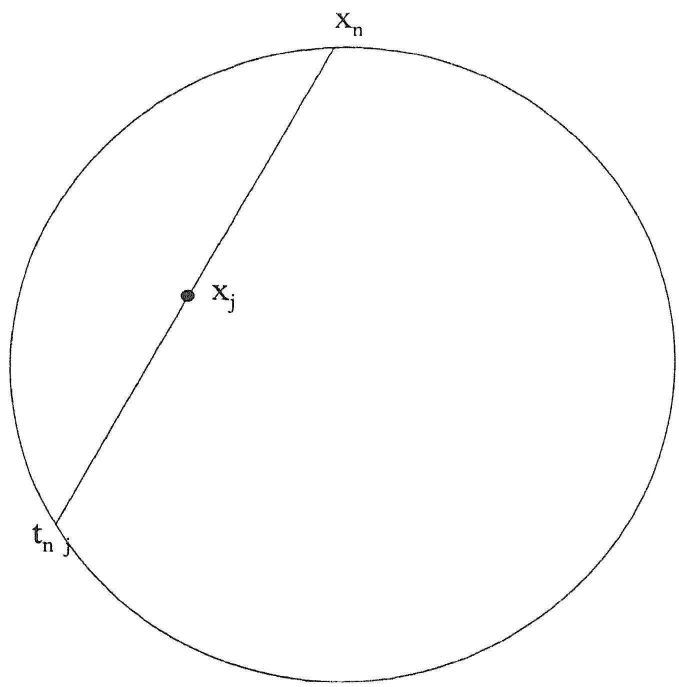
The key claim is that **the polynomials** $p_1, ..., p_n$ **given by this new construction are linearly independent**. This is easily proved by induction on n. Given $(x_1, ..., x_n)$, choose an index j for which $|x_j|$ is maximal. For simplicity of notation we may take $j = n$. Now let $q_1, ..., q_{n-1}$ be the polynomials defined by the smaller configuration $(x_1, ..., x_{n-1})$. By the inductive hypothesis these are independent polynomials (of degree $n-2$). Now choose our complex parameter t on the 2-sphere so that x_n is $t = \infty$. Then $p_i = q_i$ for $i \leq n-1$, while p_n is a polynomial of genuine degree $n-1$ (i.e., none of its roots t_{nj} is ∞)

Thus adding p_n to the set $p_1, ..., p_{n-1}$ we still have linear independence, establishing the induction (which starts trivially with $n = 2$).

Just as in §4 we can normalize the p_i and then use the polar decomposition to end up with the required map

$$f_n : C_n(R^3) \rightarrow U(n)/T^n.$$

We have thus settled the original question posed by Berry and Robbins. As a bonus our map has $SO(3)$-invariance relative to our chosen origin (and also dilation-invariance). Unfortunately, and this is certainly a drawback, our map is definitely not translation invariant. We could make the origin depend on the configuration by choosing the "centre of mass", and this would restore translation invariance. However we would then lose the "cluster decomposition" property (4) of §2, which our construction with a fixed origin does satisfy.



There is another variant of our construction which uses the upper half space model of H^3. We identify H^3 with $C \times R^+$, so that a point x is represented by a point (t, u) with $t \in C$ and $u > 0$. The geodesics are now circles orthogonal to $u = 0$. The dichotomy (a) versus (b) now depends on the u-component (with (a) corresponding to $u_j \leq u_i$).

The role of expanding concentric spheres exhausting R^3 is here played by parallel planes u = constant, and these are all identified with C through their t-component (we now take $R^3 = C \times R$).

This alternative construction is compatible with the subgroup of $SL(2, C)$ keeping a point ($t = \infty$) fixed. This consists of transformations $t \to at + b$. It is also compatible with u-translations (the analogue of dilations).

All these constructions, though continuous, are not actually differentiable because of the sharp transition from (a) to (b). This can be overcome by a Σ_n-equivariant smoothing, but this is a little cumbersome.

To sum up, while the Berry-Robbins problem has been settled by constructing an explicit map there are some unsatisfactory features of this solution. One might hope for a more elegant geometric solution, for example by settling the conjecture of §4 (and its hyperbolic analogue in

§5). One might also ask for a solution which has some physical meaning. I hope to return to these questions in future publications.

References

[1] M. F. Atiyah & J. D. S. Jones, *Topological aspects of Yang-Mills theory*, Commun. Math. Phys. **61** (1978) 97–118.

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

[3] A. Borel & F. Hirzebruch, *Characteristic classes and homogeneous spaces. I,* Amer. J. Math. **80** (1958) 458–538.

[4] G. I. Lehrer, *On the Poincáre series associated with Coxeter group actions on complements of hyperphones*, J. London Math. Soc. (2) (1987) 275–294.

Department of Mathematics & Statistics
University of Edinburgh