THE N-BODY PROBLEM, THE BRAID GROUP,

AND ACTION-MINIMIZING PERIODIC

SOLUTIONS.

Richard Montgomery

Dept. of Mathematics,

Univ. of California, Santa Cruz

Santa Cruz, CA 95064, USA

email: rmont@cats.ucsc.edu

June 26, 2003

Abstract: A reduced periodic orbit is one which is periodic modulo a rigid motion. If such an orbit for the planar N-body problem is collision-free then it represents a conjugacy class in the projective colored braid group. Under a 'strong force' assumption which excludes the original 1/r Newtonian potential we prove that in *most* conjugacy classes there is a collision-free reduced periodic solution to Newton's N-body equations. These are the classes that are "tied" in the sense of Gordon [?]. We give explicit homological conditions which insure that a class tied. The method of proof is the direct method of the calculus of variations. For the three-body problem we obtain qualitative information regarding the shape of our solutions which leads to a partial symbolic dynamics.

1 Introduction and Results.

Poincare emphasized the importance of periodic orbits in his famous work ([?], esp. §36 and §39-48 and p. I 42 of Goroff's introduction). The orbits he investigates there are not periodic in the standard sense. Rather, the mutual distances between bodies are periodic functions of time, but the placement and orientation of the triangle formed by the bodies is allowed to change in one period. In center of mass coordinates this means that the entire system may suffer a rigid rotation after one period.

We will call such orbits 'reduced periodic'. To be precise, a curve or motion $\gamma(t)$ will be called "reduced periodic" with period T if there is a rigid motion R such that $\gamma(t+T)=R\gamma(t)$. The geometric arena for studying reduced periodic solutions to the **planar** N-body problem is the space C of proper congruence classes of planar N-gons. This is the quotient space Q/G of the usual configuration space Q for the problem, by the group G of rigid motions

of the plane. ($Q = \mathbb{R}^2 \times \mathbb{R}^2 \times ... \times \mathbb{R}^2$ is the N-fold product of the plane.) A curve in Q is reduced periodic if and only if its projection to C is periodic in the usual sense. We call C the "shape space" since its points represent shapes (congruence classes) of N-gons.

The standard kinetic energy induces a very simple Riemannian metric on C. Namely, $C = C(\mathbf{CIP^{N-2}})$ is the CONE over complex projective N-2-space with its standard unitary-invariant (Fubini-Study) metric. The complex projective space represents similarity classes of oriented N-gons. For this reason we will sometimes refer to it as S, so that $S = \mathbf{CIP^{N-2}}$. (We recall the construction of the cone over S momentarily.) The remarkable simplicity of this metric space is central to our investigations. It was known to Iwai [?], and almost certainly before Iwai.

In general, the cone C(X) over a Riemannian manifold X with metric ds^2 is a metric space of one higher dimension constructed as follows. Topologically, $C(X) = (X \times [0, \infty))/\sim$, where the " \sim " means that all points of the form (x, 0) are "squashed", or identified to a single point called the cone point, and denoted 0. The metric on C(X) is $dr^2 + r^2 ds^2$ where r parameterizes the $[0, \infty)$ factor, and equals the distance from 0. This metric is smooth everywhere except at 0. The homeomorphisms of C(X) induced by $(x,r) \mapsto (x, \lambda r)$, $\lambda > 0$ are called the DILATIONS. The dilation by λ multiplies the length of any curve by λ . In the N-body case this dilation corresponds to the standard Euclidean dilations of polygons.

EXAMPLES: The cone over a sphere of radius 1 is a Euclidean space. The cone over a circle of radius less than 1 is one of the standard cones constructed by rolling up a piece of paper.

The case of N=3 of the three-body problem is especially simple. Then $S=\mathbf{CIP^1}=\mathbf{S^2}(\frac{1}{2})$ is the two-sphere of radius $\frac{1}{2}$. C(S) is homeomorphic to \mathbb{R}^3 , but the metric on it is singular at the origin, which represents the cone point. This case was known to Lemaitre [?], [?], and Deprit-Delie [?], [?]. It is implicit in the qualitative picture of Moeckel [?] for three-body dynamics, and in a paper of Saari [?]. It is also the central ingredient in an number of recent papers of Hsiang [?], [?], and in the author's paper [?] in this journal.

We will be searching for **collision-free** orbits. The binary collisions, in which two of the N bodies co-incide, define distinguished subvarieties of codimension 2 in C. There are $\binom{N}{2}$ such subvarieties, labelled Σ_{ij} for the masses i, j which collide. We will denote their union by Σ . It consists of all possible collision configurations. A collision free reduced periodic curve is then a closed curve in

$$C^* = C \setminus \Sigma$$
.

GOAL: Given a free homotopy class in C^* , find a reduced periodic collision-free solution to Newton's equations for the motion of N bodies which realizes this solution.

The set of free homotopy classes of any path-connected space X is in one-toone correspondence with the set of conjugacy classes of its fundamental group,

 $\pi_1(X)$. It is fairly well-known that the fundamental group of C^* is the projective colored braid group, by which we mean the quotient of the colored braid group by its center. (The colored braid group on N strands is the fundamental group of the N-body configuration space Q minus collisions. It is the normal subgroup of the usual braid group which corresponds to each braid returning to its starting point. See Birman [?] for a detailed description of braid groups.) Hence, our goal is to find a reduced periodic solution realizing any given conjugacy class in the projective colored braid group.

The situation is again particularly simple when N=3. The collision set Σ corresponds to three rays through the origin in \mathbb{R}^3 . So C^* is homotopic to the two-sphere minus three points. The resulting fundamental group is the free group on two letters.

Our method of attack is the direct method of the calculus of variations, applied to curves on C. We will assume throughout that the potential energy $V:Q\to I\!\!R$ of the N-body problem is invariant under rigid motions. In this case it defines a function on C, which we will denote by the same symbol V. Any curve $\gamma(t)\in Q$ defines a curve $c(t)\in C$ of shapes. If γ satisfies Newton's equations, then c satisfies a second-order differential equation on C, which we call the reduced Newton's equations. These equations are parameterized by the (constant) value of the curve's angular momentum J. The equations are markedly simpler when J=0, for in this case they are Newton's equations on C: $\nabla_{\dot{c}}\dot{c}=-\nabla V$ where ∇ is the Levi-Civita connection on C. See, for

example, Reinsch-Littlejohn [?] or Marsden-Ratiu [?]. These are the Euler-Lagrange equations on C for the functional

$$c \to A(c) = \int \left[\frac{1}{2} \|\dot{c}(t)\|^2 - V(c(t))\right] dt.$$

of paths on C. We will call this the "reduced action functional". (If $J \neq 0$ then one must add an effective potential term, and a magnetic force to these equations. The strength of the magnetic force is proportional to J and the corresponding magnetic two-form is not exact. It follows that when $J \neq 0$ there is no well-defined global action functional on C.)

REFINED GOAL: Given a free homotopy class in C^* , find a reduced, periodic, zero-angular momentum, collision-free solution in this homotopy class which minimizes the reduced action among all loops in C^* in this class.

Theorem 1 below asserts that we can achieve this goal, provided we restrict the homotopy classes to a large (essentially dense) subset of the set of all classes, and provided we restrict the potential to be a "strong-force" potential. Theorem 1, combined with Theorem 2 which provides a precise description of the allowable classes, forms our main results.

These two restrictions are imposed in Theorem 1 order to contend with the two difficulties which arise in applying the direct method of the calculus of variations. These difficulties correspond to the two types of non-compactness of C^* . The first non-compactness, or "infinity" (or end) is the usual spatial infinity. This infinity is approached by a curve c(t) in C whenever there is some pair ij of masses whose Euclidean distance $r_{ij}(t)$ tends to infinity. The other type of

"infinity" in C^* is the collision locus. It is approached whenever $r_{ij}(t) \to 0$ for some pair ij. We overcome the first difficulty by making an assumption on the free homotopy class of c. This assumption corresponds precisely to Gordon's ([?]) notion of being "tied" to a singularity. We overcome the second difficulty by a kind of a cheat, also found in Gordon, and very popular ever since, which is the "strong-force" assumption on the potential V. We will describe these two assumptions in more detail momentarily.

COMPARISONS WITH THE LITERATURE: Two of the central ideas of our paper, that of being "tied", and the strong-force assumption, are due to Gordon. So in a sense our work is an appendix to Gordon. In this regard, our main contribution is Theorem 2 which gives explicit braid-theoretic criteria which insure that a free homotopy class in C^* is "tied" in Gordon's sense. Our criteria is purely homological, meaning that it is a condition of the image of the free homotopy class in the first homology group $H_1(C^*)$. We make a second addition to Gordon's work which is specific to the three-body problem. It occupies §4. There we re-interpret our Theorem 1 in terms of syzygies (eclipses), thus setting the stage for a possilbe symbolic dynamics for the three-body problem. A final contribution, is that we deal with reduced periodic orbits, whereas almost all of the existing variational literature on the N-body problem, (see for example [?], [?], [?] and references therein) including Gordon's, focuses on the strictly periodic reduced. One exception is the interesting paper of Sbano [?] which investigates the relation between collisions and action minimization for

the Newtonian 1/r potential.

TIED CLASSES: We now describe Gordon's notions of "tied". Let Y be a non-compact complete Riemannian metric space and $\Sigma \subset Y$ a non-compact subvariety which we wish to avoid. Let α be a free homotopy class of loops in $Y^* = Y \setminus \Sigma$. Let us say that a sequence of loops $c_n \subset X$ tends to infinity if we can choose points $P_n \in c_n$ such that the $P_n \to \infty$. (When we say " $P_n \to \infty$ " we mean, that if we fix any base point $0 \in Y$ then $d(0, P_n) \to \infty$.)

Definition 1 We say that " α is tied to Σ " if whenever $c_n \in \alpha$ is a sequence of representative loops tending to infinity, then the lengths, $\ell(c_n)$, tend to ∞ .

EXAMPLE: In the planar three-body problem Y=C(S), with $S=S^2(\frac{1}{2})$. Σ is the union of three rays through the cone point 0. Any loop which encircles just one of the rays can be pushed out to infinity along that axis without changing its length. Hence the class represented by such a loop is NOT tied to Σ . On the other hand, the class of a "figure 8", a loop which winds around two of the three rays, one clockwise and the other counterclockwise, forms a tied class. For its length is always greater than $2Rsin(\phi)$ where R is its maximum distance from the cone point, and ϕ is the angular distance on S between the two points which represent the two rays. Note that $0 < \phi \le \pi/4$ because of the 1/2 in $S = S^2(1/2)$.) Also note that 2ϕ is the infimum of the lengths of the loops representing the given class in $S^* = S$ \ the three points. Compare with the estimate in the last paragraph of the proof of lemma 3 of §3.

REMARK: It is amusing to note that exactly this example is the penultimate

example of Gordon's article. However, Gordon mistakenly asserted that it is not related to any physical problem.

This concept of being tied is useful in studying natural mechanical systems on Y. We suppose that the kinetic energy is the one defined by the metric on Y, and that the potential energy is a nonpositive function V which has singularities on Σ : $V(y_n) \to -\infty$ as $y_n \to \Sigma$. Newton's equations on Y are $\nabla_{\dot{y}}\dot{y} = -\nabla V$. The action functional for studying Newton's equation on Y is $A(y(\cdot)) = \int \frac{1}{2} ||\dot{y}||^2 - V(y) dt$ as usual. We fix a **tied** class, and try to minimize A over all loops $c: [0,T] \to Y^*$, c(0) = c(T), which represent this class. We claim that any minimizing sequence $c_n \in \alpha$ cannot tend to ∞ . For if it did, then its length $\ell_n = \ell(c_n)$ would go to infinity by the definition of "tied". And by Cauchy-Schwartz $(\int ||\dot{y}||^2)T \geq \ell_n^2$. Since $-V \geq 0$ it follows that $A(c_n) \to \infty$, contradicting the assumption that c_n was a minimizing sequence.

STRONG FORCES: We now describe the "strong force" assumption on V, introduced by Gordon to exclude collisions.

Definition 2 A potential $V: Y \to (-\infty, 0]$ satisfies the strong force law if there is some positive constant c such that $V(y) < -c/\rho^2$ whenever $\rho = d(y, \Sigma)$ is sufficiently small, i.e. whenever the configuration y is sufficiently close to collision.

EXAMPLES: Standard N body potentials are the sum of two body potential V_{ij} : $V = \Sigma V_{ij}(r_{ij})$ where r_{ij} is the Euclidean distance between the ith and jth body, and the sum is over all distinct pairs ij. In this case the strong force

law is equivalent to $V_{ij}(r) < -c/r^2$ for r small enough. This is because on the shape space C, we have $d(y, \Sigma_{ij}) = k_{ij}r_{ij}$ for some positive constant k_{ij} depending only on the particle masses. The Maneff potential [?] is perhaps the most popular of the strong force potentials. (See also Diacu [?] and references therein.) The standard Newtonian potential is **not** a strong force law!

Finite action curves in a strong force law must avoid collision. The argument is simple. Let $\gamma[a,b]$ be an arc of a curve in Y with $\gamma(b) \in \Sigma$, and γ lying within the range in which the strong force assumption holds. Then $A(\gamma) = \frac{1}{2} [\int_{\gamma} ds^2 + \int_{\gamma} \frac{2c}{\rho^2} dt]$. But $ds \geq d\rho$ and so $ds^2 + \frac{2c}{\rho^2} \geq 2\sqrt{2c}d\rho/\rho$. The integral of the latter quantity diverges logarithmically as $\rho \to 0$ so that $A(\gamma) = +\infty$.

Remark: Gordon's original strong force assumption is that $-V \ge |\nabla U|^2$ holds in some neighbborhood of Σ , and for some function U which tends to ∞ at Σ . Variants of this assumption are rampant in the recent literature on the calculus of variations applied to the N body problem. These typically are found as the last item in a list of half-a-dozen or so assumptions on the potential. Although these assumptions are more general, we find them to be more opaque, with no natural examples to justify this greater generality.

We can now state our first theorem

Theorem 1 Suppose that the N-body potential satisfies the strong force condition: $-V \ge c/r^2$ whenever the distance r between any two of the masses is sufficiently close. Suppose that the free homotopy class α on C^* is tied to the collision locus. Then for any positive period T there is a reduced periodic colli-

sion free solution to the N-body problem which represents the homotopy class α , and which has period T. It is obtained by minimizing the reduced action over the class α .

Given what we have said regarding our two assumptions, the proof is a standard exercise in the calculus of variations. It is really embedded within Gordon's paper, but for completeness we sketch it here.

SKETCH OF PROOF: Let $c_n \in \alpha$ be any sequence of loops $c_n : [0,T] \to C^*$, $c_n(0) = c_n(T)$, such that $A(c_n)$ tends to the infimum of A over α . This infimum is a non-negative number, since we have assumed that $-V \geq 0$ and so the action is always positive. Since the class α is tied, and the actions A(c) are bounded, the sequence c_n must lie in some bounded subset $r \leq k$ of C. Closed bounded subsets of C are compact. By the Arzela-Ascoli theorem, the sequence has a C^0 -convergent subsequence: $c_n \to c_0$. By the lower-semicontinuity of A, c_0 has no collisions, for otherwise $A(c_n)$ would tend to ∞ according to the strong force assumption. Since $c_n \to c_0$ in the C^0 topology, the curve c_0 also represents α . By standard arguments in the calculus of variations, c_0 satisfies the Euler-Lagrange equations away from Σ . Since it never touches Σ it satisfies the equations everywhere.

2 Tied and Tangled Classes

Theorem 1 would have little content if we did not have some description of the tied classes. In particular we will want to know that this set of classes is non-empty! Our second theorem gives a computable sufficient condition for a class in C^* to be tied. As a corollary we can show that "almost all" classes are tied.

W recall that there is a canonical map $F: \pi^1 = \pi^1(C^*) \to H_1(C^*)$, the Hurewicz map, which identifies $H_1(C^*)$ with the Abelianization $\pi^1/[\pi^1, \pi^1]$ of π^1 . Since F(gh) = F(hg), F maps each conjugacy class, or free homotopy class, $\alpha \subset \pi^1(C^*)$ to a single element, denoted $[\alpha]$ of $H_1(C^*)$.

We will need an explicit description of $H_1(C^*)$. This is achieved by using the canonical fibration $G \to Q^* \to C^*$ where $Q^* = Q \setminus \{collisions\}$ and $Q = (\mathbb{R}^2)^n$. The fundamental group of Q^* is the colored braid group, CB_N , on N braids. Its center Z is the infinite cyclic subgroup generated by the single element σ which corresponds to rigidly rotating all N masses once around. (Again, see Birman, [?].) It follows from this and the homotopy exact sequence that $\pi_1(C^*) = CB_N/Z$. (We are indebted to A. Knutson for this observation.)

The homology, $H_1(Q^*)$ of CB_N is very well-understood. (See Arnol'd [?].) It has a canonical basis which is generated by the "tight binary stars". By the ij binary we mean the loop in Q^* obtained by having mass i and mass j move around each other in a circle, while all other masses are fixed, and far away. The corresponding homology classes, a_{ij} , form a \mathbb{Z} basis for $H_1(Q^*)$. The image of the generator σ of the center \mathbb{Z} under the Hurewicz map is the sum $[\sigma] = \Sigma a_{ij}$

of all of these basis elements. It follows from all this and elementary algebraic topology that $H_1(C^*)$ is the group generated by the a_{ij} with the single relation $\Sigma a_{ij} = 0$.

A basis for $H_1(C^*)$ is formed by singling out any pair i_0j_0 of masses, and deleting it from the list of a_{ij} . Now let α be a free homotopy class in C^* and $[\alpha] \in H_1(C^*)$ its image in homology. Expand $[\alpha]$ in terms of the basis:

$$[\alpha] = \sum_{ij \neq i_0 j_0} c_{ij} a_{ij}, \tag{1}$$

thus obtaining a collection of integers, c_{ij} . We associate to this expansion a graph $\Gamma_{i_0j_0}$ on N vertices. The vertices are labelled by the integers $1, \ldots, N$ which label the N masses. There is an edge connecting i to j if and only if $c_{ij} \neq 0$.

Theorem 2 If each of the $\binom{N}{2}$ graphs $\Gamma_{i_0j_0}$ associated to α is connected, then α is a tied class in the sense of Gordon.

Theorem 2 is proved in the final section of the paper.

Definition 3 The classes $[\alpha]$ which satisfy the hypothesis of Theorem 2 will be called HOMOLOGICALLY TANGLED. The complementary set of classes in $H_1(C^*)$ will be called HOMOLOGICALLY SEPERABLE.

Lemma 1 (ubiquity of tangled classes) Suppose that, for some choice $\{a_{ij}: ij \neq i_0 j_0\}$ of canonical basis, no more than N-3 of the $\binom{N}{2}-1$ coefficients c_{ij} in the expansion (??) of $[\alpha]$ vanish, and that no more than N-3 of them are equal. Then $[\alpha]$ is a homologically tangled class.

EXAMPLE 1: The class with coefficients $\{c_{ij}\}=\{1,2,3,\ldots\binom{N}{2}-1\}$ is intertwined.

EXAMPLE 2: In the case of the three body problem, there are three a_{ij} s corresponding to the three points on the two-sphere. A class is tied if when it is expanded as $[\alpha] = c_{12}a_{12} + c_{23}a_{23}$ then $c_{12} \neq 0$, $c_{23} \neq 0$, and $c_{12} \neq c_{23}$. One can check that these are **necessary and sufficient** conditions for a class to be tied. In other words, the untied classes for N = 3 are precisely those classes represented by the powers of any one of the three tight binaries a_{ij} .

DENSITY OF THE TANGLED CLASSES: Fix a canonical basis. The number of classes for which all of the $|c_{ij}|$ are less than or equal to a number R grows like R^M where $M = \binom{N}{2} - 1$ is the rank of $H_1(C^*)$. The number of these same classes which **do not** satisfy the hypotheses of the above lemma grows like $R^{M-(N+3)}$. Consequently, the density of "tied" homology classes is 1. Here "density" is defined, as in number theory, to be the asymptotic ratio of the number of **tied** classes with $|c_{ij}| < R$ to the set of **all** classes with $|c_{ij}| < R$, as $R \to \infty$.

PROOF OF LEMMA 2: Let λ stand for any of the indices $\{ij\}$ which indexes our generators a_{ij} of $H_1(C^*)$. Thus, the single relation among the generators is written Σa_{λ} , and a typical basis is $\{a_{\lambda} : \lambda \neq \lambda_0\}$. When we change bases from $\{a_{\lambda} : \lambda \neq \lambda_0\}$ to $\{a_{\lambda} : \lambda \neq \lambda_1\}$ then the coefficients c_{λ} in the expansion of $[\alpha]$ change from $c_{\lambda}, \lambda \neq \lambda_0$ to $\bar{c}_{\lambda}, \lambda \neq \lambda_1$ where

$$\bar{c}_{\lambda} = c_{\lambda} - c_{\lambda_0}, for \lambda \neq \lambda_0$$

$$\bar{c}_{\lambda_0} = -c_{\lambda_1}.$$

Set $I_{\Delta} = \{\lambda \neq \lambda_0 : c_{\lambda} = c_{\lambda_1}\}$, $I_0 = \{\lambda \neq \lambda_0 : c_{\lambda} = 0\}$, and $\bar{I}_0 = \{\lambda \neq \lambda_1 : \bar{c}_{\lambda} = 0\}$. By a simple combinatorial argument, if $|I_{\Delta}| \leq |I_0|$ then $|\bar{I}_0| \leq |I_0|$. Our assumption is that $|I_0|, |I_{\Delta}| \leq N - 3$. It follows that in any basis at most N - 3 of the coefficients c_{ij} vanish.

So any of the corresponding graphs $\Gamma = \Gamma_{i_1j_1}$ are obtained from the complete graph on N vertices by deleting at most N-2 edges. The extra 1 in N-2=N-3+1 occurs because $\Gamma_{i_1j_1}$ has no edge i_1j_1 . The lemma now follows immediately from the following elementary graph theory lemma, whose proof we leave to the reader.

Lemma 2 If a certain graph is obtained by removing k edges from the complete graph on N vertices, and if that graph is disconnected, the we must have removed $k \geq N-1$ edges.

3 Minimizing over untied classes

What happens when we try to minimize over an **untied** class α ? Let us suppose that the potential is of the STANDARD TYPE, meaning that $V = \sum V_{ij}(r_{ij})$ is a sum over two-body potentials V_{ij} , with each $V_{ij} < 0$, and satisfying $V_{ij} \to 0$ as $r_{ij} \to 0$. The Newtonian potential $V_{ij} = -m_i m_j / r_{ij}$ satisfies this requirement, as does any inverse power law, and the Maneff potential.

Proposition 1 If the potential is of the above standard type, and if the class α is untied then it is impossible to realize the infimum of the action over loops in

this class.

Proving this proposition requires a better understanding of precisely what the untied classes are. We begin by working on Q^* , where the elements of $\pi_1(Q^*)$ correspond to pure braids. We first think of a colored braid in terms of the usual model, where it is a tangle of N colored strings, each joining a colored bead of its color on a 'bottom plane' to a similarly colored bead lying directly above this bead on a 'top' plane. Try to grab a subset of k < N of these strings and pull them away from all the rest. If this is possible, then we will say that the braid is "not fully tangled", or that it can be "partially untangled". Otherwise we will say that it is "fully tangled". These properties are invariant under conjugation, so make sense on conjugacy classes of braids. An element or conjugacy class in $\pi_1(C^*)$ is said to "fully tangled" if **every one** of its inverse images in $\pi_1(Q^*)$ is fully tangled.

We want an N-body description of these notions. Recall the relation between the 'tangle-of-strings' and N-body motion model of the braid group. Let $q(t) = (q_1(t), \ldots, q_N(t)), 0 \le t \le T$ be a collision-free periodic motion of the N bodies, with $q_j(t) = (x_j(t), y_j(t)) \in \mathbb{R}^2$. Replace the motion of the jth body by its **graph** $(x_j(t), y_j(t), t)$ in space-time, and think of t as the height variable. In this way we have described the jth string in the braid represented by the loop q(t). (Since the q_j do not collide, the corresponding strings never touch.) The process of untangling the braid into two sub-braids corresponds to finding a moving curve, asymptotic to a fixed line, which splits the plane into two halves,

such that some of the bodies always lie in one half, and the others lie in another half. After a homotopy, we may suppose that this moving curve is the fixed line ℓ .

Definition 4 A loop $q(t) = (q_1(t), \ldots, q_N(t))$ in Q^* can be GEOMETRICALLY SEPARATED INTO TWO STRANDS if we can find a fixed line $\ell \subset \mathbb{R}^2$ and a partition of the the mass labels $\{1, 2, 3, ...N\}$ into two nonempty subsets $I, J, I \cup J = \{1, 2, ..., N\}$ such that the bodies $q_i(t), i \in I$ lie strictly on one side of the line, and the $q_j(t), j \in J$ lies strictly on the other side, for all time t. A free homotopy class in Q^* is said to be geometrically separated into two strands if it has a representative curve which is geometrically separated into two strands. A free homotopy class in C^* is geometrically separated into two strands if one of its lifts to Q^* is geometrically separated into two strands.

Lemma 3 The following properties are equivalent for a conjugacy class α in either $\pi_1(C^*)$ or $\pi_1(Q^*)$

- (1) α is represented by a fully tangled braid
- ullet (2) α cannot be geometrically separated
- (3) α is tied to the collisions

PROOF OF THE LEMMA: The equivalence between (1) and (2) follows from the preceding discussion. Details are left to the reader.

We prove the equivalence between (2) and (3) for classes in Q^* . Let α be a geometrically separable class. Let $q(t) = (q_1(t), \dots, q_N(t))$ be a loop in

 Q^* which represents α and which is separated by the line ℓ into two clusters $\{q_i:i\in I\}$ and $\{q_j:j\in J\}$, as in definition 4. Let n be the normal to the line, with n pointing into the half plane containing the I masses. Translate all of the I masses s units in the n direction and all of the J masses s units in the -n direction. This defines a homotopy corresponding to pulling the two groups of strings apart from each other. It does not change the LENGTH of the path q(t) but it has the property that $r_{ij} \to \infty$ as $s \to \infty$, whenever $i \in I$ and $j \in J$. Thus the class we started with is untied in Gordon's sense. The reverse implication can be proved by reversing our arguments.

Finally, the argument on C^* follows directly from the fact that $\pi_1(C^*)$ is the colored braid group modulo its center, that the center is generated by rotations, and that rotations are unseen, and hence do not affect length, on C^* . Details are left to the reader.

QED

PROOF OF PROPOSITION 1: Given an untied class α , we follow the above 'untangling' procedure. That is, we separate the bodies into two subsets I and J and translate them out to infinity. Let A_I denote the action

$$A_I(q) = \int \frac{1}{2} \sum_{a \in I} m_a ||\dot{q}_a||^2 - \sum_{i_1, i_2 \in I} V_{i_1 i_2}(r_{i_1 i_2})$$

and similarly for A_J . Since $-V_{ij} > 0$ we have $A(q) > A_I(q) + A_J(q)$ for any **path** q(t) in Q. We will show that **if** the infimum were realized it would have to satisfy $A(q) = A_I(q) + A_J(q)$ thus obtaining a contradiction.

Since α is untangled, we can write a group element a which represents it as $a = a_I a_J$. These two elements commute. Here a_I is the sub-braid representing the tangling of the braids of the subset I. It can be realized as a loop q_I in $Q_I = \mathbb{R}^2_{i_1} \times \ldots \times \mathbb{R}^2_{i_l} \setminus \{collisions\}$ where the subscripts i_j parameterize I and also label the masses involved. Let A_{I*} denote the infimum of $A_I(q_I)$ over all loops representing the conjugacy class a_I . Let $\epsilon > 0$ be a given small number. We may choose the loop q_I so that $A_I(q_I) = A_{I*} + \epsilon$. Similarly, we can define A_{J*} and find a loop q_J . Now, after possibly translating q_I rigidly by some amount we can put these separate loops together into a single loop in Q^* which represents α and which has the property that the masses $q_i(t), i \in I$ are separated from the masses $q_j(t)$, $j \in J$ by a fixed line, as in definition 2. We now follow the previous procedure of translating the I-cluster away from the J cluster, along the normal direction. The resulting homotopy leaves the action unchanged, except for the cross terms , $-V_{ij}(r_{ij}), i \in I, j \in J$. As we let the two clusters tend to infinity these cross terms drop out and the total action converges to $A_I(q_I) + A_J(q_J) = A_{I*} + A_{J*} + 2\epsilon$. Since ϵ was arbitrary, we see that $A_* = A_{I*} + A_{J*}$, and consequently the infimum of A cannot be realized by any actual loop.

The idea for the problem in C^* is essentially the same, but rotations have to be kept track of. QED

EXAMPLE: Consider the case where α is represented by a single generator, say a_{12} , for the three-body problem. Recall this corresponds to a tight binary.

Gordon [?] has shown that any Kepler orbit realizes the infimum A_{12*} of the action for the two-body problem over the single generator of the homotopy group for that problem. (The configuration of the two-body problem minus collisions retracts onto the circle.) It follows that the infimum of $A(\gamma)$ over the $\gamma \in a_{12}$ tends to the Kepler action A_{12*} and is not realized by any actual loop. Rather it is approached by a sequence γ_i of actual orbits in which masses 1 and 2 move in a Kepler orbit, and 3 stays fixed. The distance d_i between the 12-center of mass system and 3 tends to infinity as $i \to \infty$.

4 The Case of Three Bodies

When N=3 we can give a more detailed description of the qualitative nature of our minimizers. This is possible because of the greater simplicity of the fundamental group and of the geometry of the shape space C^* in this case.

The N=3 shape space is the cone over the sphere S of radius 1/2. This points of this sphere represent similarity classes of oriented triangles. The equator $E \subset S$ represents the set of similarity classes of collinear triangles. E has three marked points, the three binary collisions. These divide E into three arcs. A point of E, or a point of the plane $C(E) \subset C = C(S)$ will be called a syzygy, eclipse, or collinear configuration.

We can introduce standard spherical coordinates (r, θ, ϕ) on C such that the metric is $dr^2 + \frac{1}{4}r^2(d\phi^2 + sin^2\phi d\theta^2)$. The cone point is defined by r = 0 and

represents triple collision. The sphere S is identified with the locus $\{r=1\}$. C(E) is defined by $\phi = \pi/2$. The height coordinate $h = rcos(\phi)$ is proportional to the **signed area** of the corresponding triangle. (The proportionality constant depends only on the ratios three masses.) Any improper isometry of the plane, such as $(x,y) \mapsto (x,-y)$, reverses the orientation of triangles and induces an isometric involution

$$\tau:C\to C$$

of C. In coordinates $\tau((r, \theta, h)) = (r, \theta, -h)$. The fixed point set of this isometric involution is C(E). Therefore C(E) is a totally geodesic submanifold. This involution τ also leaves any standard potential $V = \Sigma V_{ij}(r_{ij})$ invariant, since the r_{ij} are invariant under the full group of isometries of the plane, including the improper ones. Hence τ preserves the action of paths in C, and maps solutions of the reduced Newton's equations to solutions.

Any curve in C without triple collisions can be projected radially onto the sphere S of similarity classes. We call this projected curve its "spherical image".

Theorem 3 Suppose that a closed curve in $C^* = C \setminus \text{collisions}$ realizes our refined goal, namely it minimizes the reduced action over all loops in C^* lying in its homotopy class and having its period. Then the spherical image of this curve intersects the equator a finite number of times. The consecutive points of intersection must lie on different arcs of the equator.

SKETCH OF PROOF: The curve must satisfy a differential equation, so cannot

have an infinite number of intersection points with C(E) unless it lies on C(E). But this is impossible since it represents a nontrivial homotopy class.

If the curve had an subarc whose endpoints lay on the same arc of E, then we could obtain a new curve by applying the isometric involution τ to this arc, while leaving the rest of the curve the same. The result would be a closed curve in the same homotopy class, with the same action. But this new curve dips down to touch the plane C(E), and then "bounces" back up without crossing C(E). This contradicts its minimality. (It also contradicts the fact that the new curve is also a minimizer and so must satisfy the reduced Newton's equations.) QED

We can re-interpret theorems 1 and 3 in terms of a symbolic dynamics in which each free homotopy class is encoded by its sequence of syzygies, or equatorial crossings. This encoding is essentially the same as that of Morse [?].

Label the three arcs of the equator E between the three collision points with the letters A, B, and C. Recall that E divides the sphere into two hemispheres, according to whether the triangles are positively or negatively oriented with respect to a fixed orientation of the plane (i.e. according to the sign of h). Let A_+ stand for any oriented arc s(t) lying wholly in the upper hemisphere and approaching the ecliptic arc $A \subset E$. Said more carefully, A_+ stands for any arc s(a,b] for which $s(b) \in A$ and s(a,b) lies in the upper hemisphere. Similarly, A_- denotes an arc approaching A from below. We use analogous symbols B_+ , B_- , C_+ , C_- for B and C and in this way obtain an alphabet with six symbols. We can use these symbols to generate unique representatives of all the free homotopy classes. For example, A_+B_- represents the generator a_{12} encircling the 12 collision. And the word $B_+C_-B_+A_-$ corresponds to a figure 8.

Now A_+A_- represents a contractible arc, and so does A_+B_+ . These suggest that we impose certain rules of grammar.

Rule 1: We do not allow words for which two of the same letters XX appear in a row (whatever their \pm subscripts).

RULE 2: Whenever the pair $X_{\epsilon}Y_{\delta}$ occurs in a word the subscripts $\epsilon, \delta = \pm$ have opposite parity: $\delta = -\epsilon$.

Note that theorem 3 is a variational counterpart of these rules.

Since the words are meant to represent the free homotopy class of a loop, they must have an even number of elements. And they should be viewed as cyclic words, meaning that the same class is represented by any cyclic permutation of this word. For example:

$$A_{+}B_{-}A_{+}B_{-}C_{+}B_{-} = B_{-}A_{+}B_{-}A_{+}B_{-}C_{+}.$$

Both rules of grammar must hold for each cyclic permutation of the word. This

is equivalent to insisting on

Rule 3: The beginning and ending letters of the word are different.

In this manner we can encode each free homotopy class as a cyclic word in our six letter alphabet, subject to the above rules of grammar. We leave it as an exercise to the reader to show that this representation is unique.

Recall from above that the "untied" classes are simply those classes which are represented by a powers of a single generator. If we want to insure that a word does not represent a power of a generator, we merely need to add the condition that all three letters A, B and C appear in our word. We can now restate theorem 1 and 3 as follows:

Corollary 1 (Symbolic Dynamics) Let a finite word be given in our alphabet, subject to the above Rules 1, 2 and 3 of grammar, and such that all three letters A, B and C occur in this word. Let a period of time be given. Suppose that the potential satisfies the strong force hypothesis of Theorem 1. Then there is a reduced periodic three-body solution following the sequence of syzygies specified by this word. This solution has zero angular momentum and minimizes the action among all collision-free reduced periodic loops following this sequence of syzygies in this time.

5 Proof of theorem 2: tied and tangled classes

The purpose of this section is to prove theorem 2, and to better understand the tied and the tangled classes.

We begin with some general facts regarding tied classes on cones. Suppose C(X) is the cone over a compact Riemannian manifold X. Let $\Sigma = C(D)$ play the role of the collision locus, where $D \subset X$ is a closed subset. Radial dilation defines a retraction of $C(X)^* := C(X) \setminus C(D)$ onto $X^* = X \setminus D$. In particular, free homotopy classes in $C(X)^*$ are in one-to-one correspondence with free homotopy classes in X^* . We will say that such a class in X^* is a vanishing class if it has a representative loop c for which there is a homotopy $c_s \subset X^*$, $0 \le s < b$ for which the lengths of the c_s tend to zero as $s \to b$. A vanishing class is one which can be represented by a loop which can be shrunken to an arbitrarily small size without hitting the singular locus D. If D is a finite union of smooth submanifolds then this is equivalent to saying that the infimum of the lengths $\ell(c)$ of the loops c in X^* which represent α is zero.

EXAMPLE: Suppose that X is the two-sphere and $D = \{x_1, \ldots, x_M\}$ is a finite collection of points in X. The homotopy group of X^* is generated by M letters, with each letter a_i being represented by counterclockwise loop c_i which encircles x_i and no other point. (There is a single relation $a_1a_2 \ldots a_M = 1$, which means the group is freely generated by any choice of M-1 of these M letters.) A class is a vanishing class if and only if it can be represented as a power $(a_i)^k$ of a single generator.

Lemma 4 A free homotopy class in $C(X)^*$ is tied to Σ if and only if its corresponding class in X^* is not a vanishing class.

PROOF: Suppose α is a vanishing class, represented by a "vanishing homotopy", c_R . We may suppose that $0 \leq R < \infty$ and that $\ell(c_R) = 1/R$. X is isometrically embedded in $C(X) = X \times \mathbb{R}/\sim$ by setting the $r \in \mathbb{R}$ factor equal to 1, so that c can be identified with the curve (c(t), 1). The dilated curves $Rc_R(t) := (c_R(t), R)$ have length 1 and lie a distance R from the origin. As $R \to \infty$ the curves c_R are pushed off to infinity while staying in the same homotopy class. Hence α is not a tied class.

Conversely, suppose that α is not a tied class. Then the infimum of the lengths of all curves $c \subset X^*$ representing α is some positive number δ . Let $\gamma \subset C(X)^*$ denote an arbitrary representative of α and let R denote the maximum distance of a point on γ from the cone point 0. We claim that

$$\ell(\gamma) \ge 2Rm$$

where

$$m = \sin(\delta/2)$$
, if $\delta \le \pi$, $m = 1$ if $\delta \ge \pi/2$.

It follows from this that if $R \to \infty$ then $\ell(\gamma) \to \infty$, and so α is indeed tied to Σ . To prove the inequality, consider the "radial" projection $c \in X$ of γ . (Thus $\gamma(t) = (c(t), r(t))$ and c is identified with (c, 1).) The curve c represents α in X^* , and the original curve γ lies in the cone $C(c) \subset C(X)$ over the curve c. This is a two-dimensional cone, so its geometry is Euclidean and we can use

trigonometry. Within this cone, γ is a curve whose farthest point is a distance R from the origin, and which subtends an angle of measure $\phi = \ell(c)$, the length of c. By basic trigonometry, we have $\ell(\gamma) \geq 2R\sin(\phi/2)$ if $\phi < \pi/2$, and $\ell(\gamma) \geq 2R$ otherwise. Since $\ell(c) \geq \delta$ we have the desired inequality and the proof is complete. QED

To approach the proof of Theorem 2, we ask the reader to first take a moment to reflect on the definitions of "homologically tangled" and "homologically seperated" back in §2. To proceed, we will need a more geometric characterization of these concepts. Let $Q_0 \subset Q$ denote the subspace consisting of those configurations whose center of mass is at the origin. Thus $q = (q_1, \ldots, q_N) \in Q_0$ if and only if $\sum m_a q_a = 0 \in \mathbb{R}^2$. Here the $m_a, a = 1, \ldots, N$ are the masses of the bodies, and the $q_a \in \mathbb{R}^2$ represent the positions of the bodies. It is clear that $Q/G = Q_0/SO(2)$, where SO(2) denotes the group of rotations about the center of mass. And so $C^* \cong Q_0^*/G$ where

$$Q_0^* = \{ q \in Q_0 : q_i \neq q_j \text{ for all pairs } i, j \}.$$

Set

$$\Delta_{ij} = \{ q \in Q_0 : q_i = q_j \} \subset Q_0 \tag{2}$$

so that $Q_0^* = Q_0 \setminus \Delta_{ij}$. (Note the varieties Σ_{ij} which we delete from C to form C^* are Δ_{ij}/G .)

Lemma 5 Let E denote the edge set of the graph $\Gamma = \Gamma_{i_0j_0}(\alpha)$. The graph is

disconnected if and only if

$$\cap_{ij \in E} \Delta_{ij} = \{0\} \subset Q_0. \tag{3}$$

PROOF OF LEMMA:. Suppose the graph is connected. Then there is a circuit of edges passing through all N vertices, and this circuit is a subset of the edges set E. We will show that the intersection of the Δ_{ij} over only the edges in this circuit is $\{0\}$. For if a point q is in this intersection, then $q_1 = q_2 = \dots q_N$. But $\Sigma m_i q_i = 0$ and $\Sigma m_i q_i = M q_1$ where $M = \Sigma m_i$ so that all the $q_i = 0$. The point q of intersection must be the N-tuple collision point q.

Conversely, if the graph is disconnected, then the vertices fall into at least two components, so we can partition them into two nonempty subsets $I, J \subset \{1, 2, ..., N\}$, $I \cup J = \{1, 2, ..., N\}$ with no vertex in I connected to one in J by an edge. Set $M_I = \sum_{i \in I} m_i$, $M_J = \sum_{j \in J} m_i$. Pick $r, s \in \mathbb{R}^2$, $r, s \neq 0$ such that $M_I r + M_J s = 0$ Then the element $q = (q_1, ..., q_N)$ with $q_i = r, i \in I$, $q_j = s, j \in J$ is a nonzero vector lying in the intersection.

Lemma 6 If the free homotopy class α represents a vanishing class in S^* then its image $[\alpha]$ in the first homology of C^* is homologically separable into two strands.

PROOF OF THEOREM 2: Theorem 2 is the contrapositive of the implication just asserted, according to lemma 4.

PROOF OF THE LEMMA:

Let $\Sigma_{ij} \subset S = \mathbf{CIP^{N-2}}$ denote the binary collision loci in S. These can be

obtained by taking the images of the collision loci $\Delta_{ij} \subset Q_0 \subset Q$ in C under the quotient map $Q_0 \to Q_0/(SO(2)) = C$ and then intersecting these images with $S \subset C$. They are hyperplanes in the complex projective space S. Lemma 5 translates to assert that the class α is homologically separated if and only if for some i_0j_0 we have

$$\bigcap_{ij\in E} \Sigma_{ij} \neq \emptyset \tag{4}$$

Suppose that α is a vanishing class, so that the infimum of the lengths of the curves representing it is zero. Then there is a sequence of loops γ_n in S^* which represent the class α and whose diameters tend to zero. By compactness of S we can find a subsequence of paths which tend to a fixed point $P \in S$. P cannot lie in all of the Σ_{ij} since the intersection of all of them is empty. Hence there is at least one, say $\Sigma_{i_0j_0}$ which it does not lie in. Consider the corresponding basis $\{a_{ij}: ij \neq i_0j_0\}$ for $H_1(C^*)$ and the associated graph $\Gamma = \Gamma_{i_0j_0}(\alpha)$, with its edge set E We will show that $P \in \cap_{ij \in E} \Sigma_{ij}$. This will prove that the intersection is non-empty and the class $[\alpha]$ is homologically separated, according to the inequality (4).

Let ω_{ij} , $ij \neq i_0 j_0$, be the basis for the 1st cohomology group which is dual to the basis we have just chosen for the 1st homology. The classes ω_{ij} can be represented as differential one-forms, in which case the integers c_{ij} are given by

$$c_{ij} = \int_{\gamma_n} \omega_{ij}.$$

Let w_1, \ldots, w_{N-2} be complex affine coordinates for $S \setminus \Sigma_{i_0 j_0}$ so that Σ_{ij} is

defined by the equation $w_1 = 0$. Then $\omega_{ij} = \frac{1}{2\pi i} \frac{dw_1}{w_1}$. We will show that if $c_{ij} \neq 0$ (for some $\{ij\} \neq \{i_0j_0\}$) then $P \in \Sigma_{ij}$. Our proof is by contraposition. Suppose $P \notin \Sigma_{ij}$. Let B be a small ball centered at P disjoint from Σ_{ij} . With respect to the coordinates w_i above we have, for n sufficiently large, $\gamma_n(t) = (w_1(t), w_2(t), \dots, w_{N-2}(t))$ with $w_1(t)$ lying in a small disc **disjoint from the origin**. It follows from the above representations for c_{ij} and ω_{ij} and the Cauchy integral formula that $c_{ij} = 0$.

6 Acknowledgements

I would like to thank Cesar Castilho, Chris Golé, Franchesco Fasso, and Alan Weinstein for useful conversations, and Allen Knutson for lessons on the projective braid group. I would especially like to thank Wu-Yi Hsiang for suggesting the bulk of the problems here, and for inspiration. This work was supported by NSF grant DMS-9400515 and a University of Santa Cruz Faculty Research Grant. I would also like to thank Jean-Claude Hausmann and the University of Geneva for hospitality during the writing, and CIMAT in Guanajuato, Mexico, during the rewriting.

7 Bibliography

References

- A. Ambrosetti, V. Zelati, Periodic Solutions of Singular Lagrangian Systems, Birkhäuser, Boston, 1993.
- [2] V.I. Arnol'd, The Cohomology of the Colored Braid Group, Math Zametki5, no. 2, 227-231; English translation: Math. Notes. 5, 138-140 (1969).
- [3] A. Bahri and P. Rabinowitz, Periodic solutions of Hamiltonian systems of 3-body type, Annales Inst. Henri Poincaré, Analyse non linéaire, v. 8, no. 6,561-649, (1991).
- [4] A. Deprit and A. Delie, Régularisation du problème des trois corps, Arch. for Rat. Mech., 12, 325-353, (1962).
- [5] A. Deprit, et A. Delie, Courbure de la métrique riemannienne du probème plan des trois corps, Bull. Cl. Sc. Acad. Belg., no. 5, 48, 652-658, (1962).
- [6] J. Birman, Braids, Links and Mapping Class Groups, Princeton Univ. Press, Annals of Math Studies v. 88, (1975).
- [7] F. Diacu, The planar isosceles problem for Maneff's gravitational law J. Math. Phys., v. 34, no. 12, 5671-5690, (1993).
- [8] W. Gordon, A Minimizing Property of Keplerian Orbits, Am. J. Math, 99, no. 5, 961-071, (1977).

- [9] W-Y Hsiang, Geometric Study of the Three-Body Problem I, report PAM-620, Center for Pure and Applied Math, University California, Berkeley, 1994.
- [10] W-Y Hsiang and E. Straume, , Kinematic Geometric of triangles with given mass distribution, report PAM-636, Center for Pure and Applied Math, University California, Berkeley, 1995.
- [11] W-Y Hsiang, Kinematic Geometric of mass triangles and the three-body problem in quantum mechanics, report PAM-682, Center for Pure and Applied Math, University California, Berkeley, 1996.
- [12] T. Iwai, A geometric setting for classical molecular dynamics, Ann. Inst. Henri Poincare, vol. 47, no. 2, pp 199-219. (1987).
- [13] T. Iwai, A geometric setting for internal motions of the quantum three-body problem, J. Math. Phys., v. 28, 1315-1226, (1987).
- [14] G. Lemaitre, Coordonnées symétriques dans le problème des trois corps, Bull. Cl. Sc. Acad. Belg. no. 5, 38, 582-592, 1218-1234, (1952).
- [15] G. Lemaitre Regularization of the Three Body Problem, in Vistas in Astronomy, 1, 207-215, (1955).
- [16] T. Levi-Civita, Sur La Réqularisation du Problème des Trois Corps, Acta Math., v. 42, 99-144; esp. p. 105, eq.(12); (1921).

- [17] R.G. Littlejohn, RG and M. Reinsch, Gauge fields in the separation of rotations and internal motions in the n-body problem, Reviews of Modern Physics, 69 no. 1, 213-275, (1997).
- [18] P. Majer and S. Terracini, Periodic solutions to some problems of n-body type, Arch. Rat. Mech., 124, no. 4, 381-404, (1993).
- [19] G. Maneff, La gravitation et l'énergie au zéro, Comptes Rendus, v. 190, 1374-1377, (1930).
- [20] J. Marsden and T. Ratiu, Introduction to Mechanics with Symmetry, Springer-Verlag, 1994.
- [21] R. Moeckel, Some qualitative features of the three-body problem in Hamiltonian Dynamical Systems, proceedings, AMS Contemporary Math, 81, ed. K. Meyer and C. Saari, 1-22, American Math. Society, (1988).
- [22] R. Montgomery, The geometric phase of the three-body problem, Nonlinearity, 9, no. 5, 1341-1360, (1996).
- [23] M. Morse, A Fundamental Class of Geodesics on Any Closed Surface of Genus Greater Than One, Trans. AMS, v. 26 (1924).P, History of Modern Physics and Astronomy v. 13, (1993).
- [24] H. Poincaré, Les Methodes Nouvelles de la Mécanique Céleste v. 1, ch. 3, Paris, Gauthier-Villars et fils, (1892).
 see also

- introduction to the English translation by D. Goroff, p. I 42. AIP, History of Modern Physics and Astronomy v. 13, (1993).
- [25] D. Saari, Symmetry in n-particle systems, in Hamiltonian Dynamical Systems, proceedings, AMS Contemporary Math, Vol. 81, ed. K. Meyer and C. Saari, 23-42, esp. 26-27, American Math. Society, (1988).
- [26] L. Sbano, On the Lack of Coercivity of the reduced Action-functional for zero total angular momentum in the planar three body problem, SISSA/ISAS preprint 185/94/FM, (1994).