MSC 2010 : 70F10, 37N05, 70G45

THE HYPERBOLIC PLANE, THREE-BODY PROBLEMS, AND MNËV'S UNIVERSALITY THEOREM.

RICHARD MONTGOMERY

ABSTRACT. We show how to construct the hyperbolic plane with its geodesic flow as the reduction of a three-problem whose potential is proportional to I/Δ^2 where I is the moment of inertia of this triangle whose vertices are the locations of the three bodies and Δ is its area. The reduction method follows [11]. Reduction by scaling is only possible because the potential is homogeneous of degree -2. In trying to extend the assertion of hyperbolicity to the analogous family of planar N-body problems with three-body interaction potentials we run into Mnëv's astounding universality theorem which implies that the extended assertion is doomed to fail.

Part 1

1. A Reverse Engineering Question. Summary of First part of paper.

In an earlier work [11] I showed how a planar three-body problem with a $1/r^2$ pair potential leads, via symmetry reduction, to a complete Riemannian metric on the shape sphere minus three points, and that the Gaussian curvature of this metric is negative everywhere except at two points. This work suggests a reverse engineering question. *Can I design a three-body potential whose analogous symmetry reduction yields the standard hyperbolic plane?* Here I answer this question affirmatively by constructing such a potential. See eq (4) and theorem 1 below.

I construct this potential by combining the Jacobi-Maupertuis metric perspective on mechanics with two facts. Fact one: the hyperbolic plane is realized by the "Jemisphere model", which is the usual spherical metric multipled by the conformal factor $1/z^2$ where z is height above the equator. Fact two: the height coordinate z on the shape sphere ([12]) is proportional to the signed area Δ of triangles.

2. Set-up and Main Result.

Newton's equations corresponding to a three-body potential $V : \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ are

$$m_a \ddot{q}_a = -\nabla_a V, a = 1, 2, 3, \tag{1}$$

The positive constants m_a are the values of the masses. The $q_a \in \mathbb{R}^2 = \mathbb{C}$ are the instantaneous locations of the point masses in the plane. And $\nabla_a = \frac{\partial}{\partial q_a}$ indicates the gradient with respect to q_a. Henceforth we will assume that the potential V is

- (A1) invariant under rigid motions, and
- (A2) homogeneous of degree -2

(A1) implies that Newton's equations (1) are invariant under rigid motions and admit the usual conserved quantities of physics: energy $H(q, v) = K(v) + V(q)$

Date: September 15, 2016.

(where K is the kinetic energy), linear momentum P, and angular momentum J See appendix A for analytic expressions of these invariants.

(A2) yields the scale symmetry $q(t) \mapsto \lambda q(\lambda^{-1}t)$ taking solutions of Newton's equations to solutions. It also yields the virial (or Lagrange-Jacobi) identity

$$\ddot{I} = 4H$$

where

$$I = \Sigma m_a m_b r_{ab}^2 / \Sigma m_a ; r_{ab} = |q_a - q_b| \tag{2}$$

is the moment of inertia function. Thus, this virial identity says that $\dot{I}$ is a "sub-invariant": $\dot{I}$ is conserved on the energy level set $\{H = 0\}$.

(A1) and (A2) imply that the group G of rigid motions and scalings acts by symmetries for our Newton's equations (1): group elements take solutions to solutions. The resulting G-action on phase space also preserves the common zero level set of all our invariants $H, P, J, \dot{I}$. Thus the quotient space of the common zero level set $\{H = J = P = \dot{I} = 0\}$ by G is a space on which Newton's equations yields a well-defined induced dynamics. We call the dynamics on this quotient space the "reduced dynamics". This reduced dynamics is best understood using the 'shape sphere'.

Write $\mathbb{C}_{coll} \subset \mathbb{C}^3$ for the triple collision variety $\{q = (q_1, q_2, q_3) \in \mathbb{C}^3 : q_1 = q_2 = q_3\}$. Then the shape sphere S^2 is the quotient space $\mathbb{C}^3 \setminus \mathbb{C}_{coll}$ by G. Its points represent oriented similarity classes of triangles. (See ([12]) or Appendix B for details.) We call the corresponding quotient map the "shape projection":

$$\Pi : \mathbb{C}^3 \setminus \mathbb{C}_{coll} \xrightarrow{G} S^2 = \text{ shape sphere} \tag{3}$$

Theorem 2 below asserts that when we project solutions lying in our invariant submanifold to the shape sphere we get geodesics on that sphere relative to a metric conformal to the standard round metric with a conformal factor $\lambda^2 = -IV$.

The shape sphere S^2 sits inside $\mathbb{R}^3$ as $w_1^2 + w_2^2 + w_3^2 = 1$ where w_1, w_2, w_3 are standard linear coordinates. Thus the standard round metric is the restriction of $dw_1^2 + dw_2^2 + dw_3^2$ to the sphere. Multiplying this metric by the conformal factor $1/(w_3)^2$ yields a metric on either hemisphere of the sphere which is isometric to the standard hyperbolic plane. This model of the hyperbolic plane is referred to as the 'Jemisphere" model. (See [2].) Our main result asserts that we will achieve the Jemisphere model by reduction if we take our three-body potential to be

$$V(q) = -\frac{\gamma \text{ (moment of inertia)}}{(\text{area})^2} = -\frac{\gamma I(q)}{\Delta(q)^2}. \tag{4}$$

where $\Delta(q) = \frac{1}{2}(q_2 - q_1) \wedge (q_3 - q_1)$ is the signed area of the triangle with vertices $q = (q_1, q_2, q_3)$. (The constant $\gamma > 0$ is a physical constant needed to make the units of the potential that of energy, so that γ has units of $(\text{length})^4/(\text{time})^2$.)

Theorem 1. *If we take (4) as potential then any solution to Newton's equations (1) for which the conserved quantities H, J, P and $\dot{I}$ are all zero projects, by the shape space projection, onto a geodesic for the Jemisphere model of hyperbolic geometry. Conversely, every such geodesic is the projection of such a solution to these Newton's equations.*

3. Reduction of the Jacobi-Maupertuis Metric

I will prove Theorem 1 by "Riemannian reduction" - forming the Riemannian submersion of a Riemannian manifold onto its quotient space by a subgroup of its isometry group. To start the method I will need the Jacobi-Maupertuis [JM] principle, a reformulation of Newton's equations as geodesic equations for a certain energy-dependent Riemannian metric. Theorem 1 could be proved, instead, using classic symplectic reduction, with a twist coming from the the "sub-invariance' of $\dot{I}$. The heart of either method is the same.

The JM principle applies to each fixed energy level. At energy level $H = 0$, the principle asserts that zero-energy solutions to Newton's equations (1) are, up to reparameterization, geodesics for the Jacobi-Maupertuis [JM] metric

$$ds^2_{JM} = U(q)|dq|^2_E, \text{ where } U = -V$$

on $\mathbb{C}^3 \setminus \{U = \infty\}$. Here

$$|dq|^2_E = K(dq) := m_1|dq_1|^2 + m_2|dq_2|^2 + m_3|dq_3|^2$$

denotes the Euclidean inner product on $\mathbb{C}^3$ associated to kinetic energy, and viewed as a Riemannian metric. (I will assume that $V < 0$ everywhere, so that $U > 0$, otherwise, simply restrict to the portion $\{U > 0\}$ of configuration space.)

Assumptions (A1) and (A2) imply that the JM metric is invariant under the group G generated by rigid motions and scalings. Consequently, this metric can be pushed down to the shape sphere. We denote this pushed-down metric by $\Pi_* ds^2_{JM}$, It is a metric on the shape sphere minus the singularity locus, the latter being the projection by Π of the locus $\{U = \infty\}$. (Note that because U is nonzero and homogeneous of negative degree, the locus $\{U = \infty\}$ contains the triple collision line $\mathbb{C}_{coll} \subset \mathbb{C}^3$ described earlier.) Thus, up to reparametrization, the projection to the shape sphere of the solutons to Newton's equations for which $H = J = P = \dot{I} = 0$ are geodesics for this pushed-down metric on the shape sphere.

Some words are in order regarding pushing down a metric. The fibers of Π are the orbits of G so that the differential $d\Pi_q$ at a configuration point $q \in \mathbb{C}^3$ (with $U(q) \neq \infty$) has as its kernel the tangent space $T_q(Gq)$ at q to the G orbit through q. See figure 1. Declaring this differential to be a linear isometry between the *orthogonal complement* to $T_q(Gq)$ and the tangent space of the base space downstairs defines the pushed-down metric, and turns Π into a *Riemannian submersion.* (Because G acts isometrically, the pushed down metric at $s \in S^2$ is independent of the point $q \in \Pi^{-1}(s)$ at which we perform this computation.) It follows that Π maps geodesics for ds^2_{JM} *orthogonal to G-orbits* onto geodesics for $\Pi_* ds^2_{JM}$. Mechanically speaking, a curve in $\mathbb{C}^3$ is orthogonal to the G-orbits through its points if and only if its angular momentum J, linear momentum P and 'scale momentum' $\dot{I}$ are zero at each point along the curve. See Appendix A below, or chapter eleven of [13] for details. Thus Π maps the solutions of (1) which have $H = J = P = \dot{I} = 0$ onto geodesics for $\Pi_* ds^2_{JM}$.

The standard round metric ds^2_{round} on S^2 arises from the above JM reduction procedure if the potential we start with is $V = -1/I$ where I is the moment of inertia as per eq (2). This fact follows immediately from eq (43) of [12]. (See also Appendix B here.) For a general potential V satisfying (A1) and (A2) above, the function $IU = -IV$ is invariant under G and as such can be viewed as a function on the shape sphere, which, for clarity we will write as $\Pi_*(IU)$. Thus, upon writing

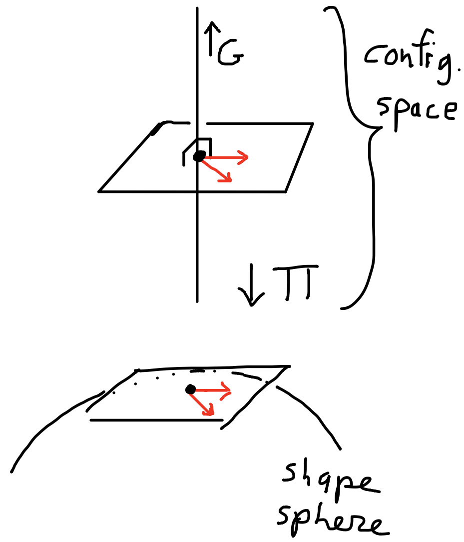

FIGURE 1. The shape space projection is a Riemannian submersion.

$ds^2_{JM} = (IU)(\frac{1}{I}ds^2_E)$ we see that $\Pi_* ds^2_{JM}$ is a metric on the shape sphere, conformal to the round metric, with conformal factor $\Pi_*(IU)$. We have proved:

Theorem 2. *Suppose the potential V satisfies (A1) and (A2) above. Then, up to reparameterization, the projections by Π of those solutions to Newton's equations (1) having conserved quantities all zero ($H = J = P = \dot{I} = 0$) are geodesics on the shape sphere S^2 for the metric conformal to the round metric with conformal factor $\Pi_*(IU)$, the push-down of the G-invariant function IU to the shape sphere. (Here $U = -V$.) The domain for this metric is $\{0 < \Pi_* IU < \infty\} \subset S^2$ Conversely, every such geodesic is the projection of such a solution to these Newton's equations.*

PROOF OF THEOREM 1. Theorem 1 is a corollary of theorem 2, the expression for the Jemisphere metric described above, and knowledge of the meaning of the height coordinate w_3 on the shape sphere. The Jemisphere metric ds^2 is $\frac{1}{w_3^2} ds^2_{round}$ where ds^2_{round} is the restriction of the ambient Euclidean metric $dw_1^2 + dw_2^2 + dw_3^2$ to the sphere. Theorem 2 asserts that we will get the Jemisphere metric by taking a potential V such that $\Pi_*(IV) = -\frac{1}{w_3^2}$, or $IU = \Pi^*(\frac{1}{Iw_3^2})$. This yields our formula for the potential , eq (4), provided

$$\Pi^* w_3 = c\Delta/I \tag{5}$$

for some constant c. This expression for $\Pi^* w_3$ is indeed valid. See eq (14) and the discussion around it in Appendix B.

4. Consequences for solutions.

There are two Jemispheres, $w_3 > 0$ and $w_3 < 0$. A solution cannot switch from one to the other, since the potential blows up as $w_3 \to 0$. The reflection $(w_1, w_2, w_3) \mapsto (w_1, w_2, -w_3)$ maps one Jemisphere to the other, isometrically. Choose one of these Jemispheres, corresponding to whether the initial signed area of the triangle is positive or negative.

Every geodesic in that Jemisphere is realized as the intersection of that corresponding hemisphere with a plane orthogonal to the equatorial plane $w_3 = 0$, which is to say, a plane of the form $Aw_1 + Bw_2 = const.$. In terms of the original Newton's equations, this means that we can describe the solutions for which all the invariants are zero as follows. Any such solution has $I = const.$. Call the particular constant of such a solution I_0. Project the solution to the shape space $\mathbb{R}^3$. It will lie on the sphere $|w| = I_0/2$. This projected solution forms a half circle made by intersecting the initial hemisphere of the sphere with one of these planes. (See figure 2.)

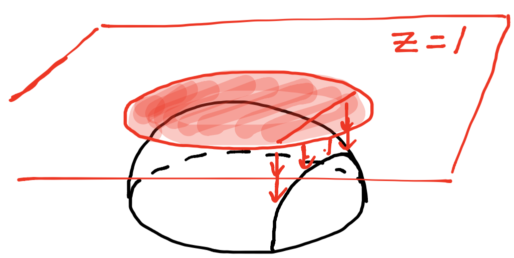

Figure 2. Projection along the vertical is an isometry between the Klein and Jemisphere models of the Hyperbolic Plane.

Solutions $q(t) \in \mathbb{C}^3$ to our Newton's equations for which $H = \dot{I} = 0$, and $I > 0$ initially will exist as long as the denominator $\Delta^2 = cw_3^2$ in the potential does not become zero. So extend such a solution to its maximum domain $(a, b) \subset \mathbb{R}$ of time definition and project it to shape space. It will begin and end on the equatorial plane $\Delta = 0$, which is to say that its limiting configurations, $\lim_{t \to a} q(t)$ and $\lim_{t \to b} q(t)$, will be collinear configurations. The angle between the lines containing these two limiting collinear configurations can be computed using the area formula, as per the periodicity proof for the figure eight solution given in [3].

Part 2

5. Interlude and Motivation.

Papers [11] and [6] suggested the approach taken here to building the "mechanical hyperbolic metric" of theorem 1. In [11] I applied the reductions of the present paper to the "strong-force" $1/r^2$ potential $-(1/r_{12}^2 + 1/r_{13}^2 + 1/r_{23}^2)$ in the equal mass three-body case to obtain a metric on the shape sphere minus its three binary collision points, i.e. on the topologist's pair-of-pants. The main theorem there is that this metric is complete with negative Gauss curvature everywhere except at two points, these points corresponding to equilateral triangles. As a corollary, the figure eight solution for that potential is unique up to isometry and scaling.

Connor Jackman began his thesis work trying to extend the $N = 3$ hyperbolicity just described to the case $N = 4$ of the equal mass 4-body problem. The shape sphere is now replaced by $\mathbb{CP}^2$ – the space of oriented similarity classes of planar quadrilaterals. The negative of the potential is now the sum of the six reciprocal distances squared. Jackman [6] discovered a surprise. The resulting 4-body JM metric, which he constructed in the way we have described here, has mixed curvature: some 2-planes have positive curvature, some negative. Indeed the equal mass 4 body problem contains invariant subproblems , the collinear and parallelogram subproblems, which correspond to totally geodesic two-dimensional surfaces within $\mathbb{CP}^2$. Jackman proved that the Gaussian curvatures of these surfaces are negative, but that the sectional curvatures of the two-planes orthogonal to the surfaces are positive.

The tricks of [11] and [6], added to the perspective of the present paper, apply to any negative isometry-invariant potential of homogeneity -2 on the planar N-body configuration space $\mathbb{C}^N$. The result is a metric on $\mathbb{CP}^{N-2} \setminus \Sigma$ where Σ is the set on which $V = -\infty$. That JM metric has the form $U ds^2_{FS}$ where $U = -V$ and ds^2_{FS} is the canonical Fubini-Study metric on complex projective space. If for example V is the negative of the sum of the reciprocal distances, then this singularity locus is a hyperplane arrangement, so we get some metric, conformal to FS, defined on the complement of this hyperplane arrangement.

6. More bodies and the appearance of Mnëv's Universality Theorem.

Reflecting on Jackman's jump from N = 3 to N=4 in [6], one wonders what will happen if we take the analogous jump for our "3-point" potential (4). To define the new N-body potential sum our 3-point potential over all triples of bodies. Thus, consider N point masses $q_1, q_2, \ldots, q_N$ in the plane. For each choice of 3 indices i, j, k out of $\{1, 2, \ldots, N\}$, let $I(i, j, k)$ be the moment of inertia of the triangle formed by vertices q_i, q_j, q_k and let $\Delta(i, j, k)$ be the signed area of this triangle.

Then our proposed N-body potential is

$$V_N = -\gamma \Sigma_{i,j,k} \frac{I(i,j,k)}{\Delta(i,j,k)^2} := -U_N, \tag{6}$$

the sum being over all three-element subsets $\{i,j,k\} \subset \{1,2,\ldots,N\}$.

V_N blows up on the real hypersurface

$$\Sigma_N := \{q \in (\mathbb{C})^N : U_N = \infty\} = \bigcup_{i,j,k} \{\Delta(i,j,k) = 0\} \tag{7}$$

Let us call a planar N-gon *generic* if no three vertices are collinear. Then Σ_N consists of the locus of all non-generic planar N-gons.

Since V_N is homogeneous of degree -2 and invariant under our group G of rigid motions of the plane, we can play the same Jacobi-Maupertuis plus reduction game which we just played in proving Theorem 1 to arrive at a metric on $\mathbb{CP}^{N-2} \setminus \Sigma_N$, this metric being the push down of the metric $U_N ds^2_E$ from $\mathbb{C}^N \setminus \Sigma_N$

$$ds^2_{JM} = \Pi_*(U_N ds^2_E) \tag{8}$$

by the 'shape projection'

$$\Pi : \mathbb{C}^N \setminus \mathbb{C}_{tot} \to \mathbb{CP}^{N-2}$$

which takes a labelled planar N-gon $(q_1, q_2, \ldots, q_N) \in \mathbb{C}^N$ to its "shape" $\Pi(q) \in \mathbb{CP}^{N-2} = (\mathbb{C}^N \setminus \mathbb{C}_{tot})/G$. The line $\mathbb{C}_{tot} \subset \Sigma_N \subset \mathbb{C}^N$ which we must delete to insure that Π is well-defined is the total collision locus consisting of those q for which all the q_i are equal. By abuse of notation we use the same symbol Σ_N for the set of non-generic N-gons upstairs and for its projection $\Pi(\Sigma_N)$ downstairs in $C\mathbb{P}^{N-2}$:

$$\Sigma_N = \{\Pi(q) \in \mathbb{CP}^{N-2} : q \text{ a non-generic N-gon}\} \tag{9}$$

Upon deleting Σ_N, the remaining space becomes disconnected. [1] So for each N, we have a collection of metrics on the components $C_1, C_2, \ldots$ of our manifolds.

Let us be bold then and ask

Q1. Is this induced metric (8) hyperbolic on each component of $\mathbb{CP}^{N-2} \setminus \Sigma_N$?

Before answering I establish a few properties of these metrics.

Property 1. Each metric is complete. To prove this, suppose we approach, within one component, a point p_∞ on the smooth locus of Σ_N. Such a point is a zero of one $\Delta = \Delta(i,j,k)$ and no other and is also a regular point of this Δ (with Δ viewed either upstairs or downstairs). Write $f = \Delta(i,j,k)/\sqrt{I}$ where $I = |q|^2$ is the total moment of inertia, or squared length up on $\mathbb{C}^N$. We use I because the Fubini-Study metric ds^2_{FS} is the projection of $\frac{1}{I} ds^2_E$ by Π. Then p_∞ is a regular value of f and $U_N = (\frac{1}{f^2} + O(1))$ near p_∞ so that our metric has the form $ds^2 = (\frac{1}{f^2} + O(1)) ds^2_{std}$ with ds^2_{std} being either ds^2_E or ds^2_{FS} depending on whether we are working upstairs or downstairs. One now verifies without difficulty that $ds \geq k|df|/f + O(1)$ (for some $k > 0$) as we approach p_∞. (This inequality holds as an inequality between functions on the tangent bundle.) This implies that if $c(s)$, $0 \leq s < s_0$ is any curve lying in our component and approaching p_∞ as $s \to s_0$ then the length of $c[0,s]$ grows like $|\log(f(c(s))|$ and thus becomes infinite. At singular points of Σ_N this length blow-up estimate only gets better: the lengths

[1] The number of components grows super-exponentially with N, like N^N. See the last entry in table 5.6.2 of [4].

of the curves blow up like $|\log(f_1(c(s))| + |\log(f_(c(s))| + \ldots$ with one f_i for each branch $\Delta(k, l, m) = 0$ passing through p_∞.

Property 2. Each metric is asymptotically negatively curved. Indeed this curvature condition holds for any metric of the form $(\frac{1}{f^2} + O(1))ds^2_{std}$ as we approach a regular zero p_∞ of f. The sectional curvatures as we approach p_∞ are of the form $-\|\nabla f\|^2 + O(f)$. See [8] for this computation.

Property 2 gives us some optimism regarding a 'yes' answer to Q1. It seems conceivable that the negative sectional curvatures coming from the terms of our potential (eq (6)) might conspire to yield a sectional curvature everywhere -1 off of Σ_N. But an astounding theorem of Mnëv answers Q1 with a resounding **no!**. Indeed we have

Theorem 3. *[Mnëv; [9]] Let X be any connected semialgebraic set. Then there is an integer N and a component C of the space $\mathbb{CP}^{N-2} \setminus \Sigma_N$ of similarity classes of generic N-gons (see eqs. (7) and (9) above) such that $C \times \mathbb{R}^s$ is homeomorphic to $X \times \mathbb{R}^e$ for some integers s, e. In particular X and C are homotopic.*

See figure 3 for an attempt to illustrate Mnëv's theorem.

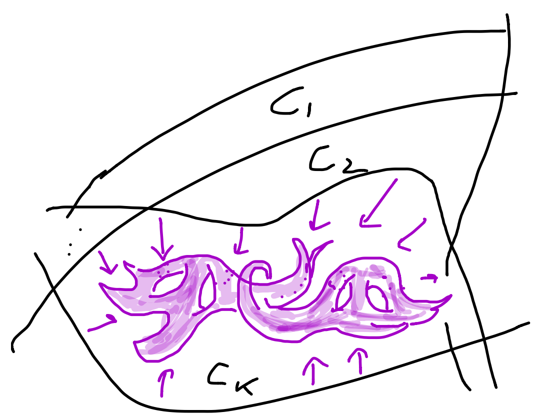

FIGURE 3. The components of the shape space of generic N-gons can retract onto arbitrarily complicated compact manifolds or singularities.

To see how Mnëv's theorem yield the answer 'no' to question Q1, recall by Nash and Weierstrass that any compact manifold M can be realized as an algebraic subvariety of $\mathbb{R}^N$ for some N. We can thicken the variety a bit by changing the equalities $g = 0$ to inequalities $-\epsilon < g < \epsilon$ and in this way arrive at an open semi-algebraic set which retracts onto a smooth variety representing any given smooth compact manifold type M. In particular all possible algebraic topological invariants are represented by our components C. Now take for X in Mnëv's theorem a space whose homotopy type prevents it from admitting a complete hyperbolic structure or indeed any complete metric of non-negative sectional curvature. For example, by Cartan-Hadamard, if C were complete with non-negative sectional curvatures then its universal cover must be simply connected, and hence C must be aspherical. So if we take X to be a space with the homotopy type of a sphere, or indeed any compact simply connected manifold we would exclude the corresponding C which represents X from admitting a complete metric of non-negative sectional curvature and in particular a hyperbolic structure. Or we could choose X so that its fundamental group fails to satisfy one of the many restrictive properties which the fundamental group of a complete hyperbolic manifold must satisfy such as the Haageruup property or the property that all of its finitely generated subgroups are residually finite. (See for example [5] for a more exhaustive list of such properties.) Thus we have proved:

Theorem 4. *For sufficiently large N there exist components C of the space $\mathbb{CP}^{N-2} \setminus \Sigma_N$ of generic planar N-gons (see eqs. (7) and (9) which admit no complete metric whose sectional curvatures are everywhere non-positive.*

After receiving the resounding slap of a 'no' to our Q1 from Mnëv, let us relax and ask a more modest question.

Q2. For $N = 4$ does the induced metric ds^2_{JM} on the (14) components of $\mathbb{CP}^2 \setminus \Sigma_4$ have strictly negative sectional curvatures?

Finally one might ask:

Q3. Do potentials of the form of eq (4) or (6) arise in any physical or chemical problems of interest?

In the nuclear physics and physical chemistry literature one finds various references to "three-body forces" – forces arising from potentials that, like ours, cannot be written as a sum of two-body potentials. I did not find any of these of our form (eq 4) or the more general form $f(I, \Delta)$.

7. Confessional

I teach a geometry class almost every year whose audience consists in good measure of future high school mathematics teachers. I feel obligated to teach the rudiments of hyperbolic geometry. But when I'm done, almost every student leaves that class with no understanding of hyperbolic geometry beyond a vague mental cartoon of Escherish shapes which get really really small as they approach some line. The present paper began as an attempt to do better by providing some natural inroad to hyperbolic geometry. I have failed these students. Nevertheless, I hope my readers find this excursion interesting.

Appendix A. Invariants.

The kinetic energy K, being a positive definite real quadratic form on velocities, defines an inner product $\langle \cdot, \cdot \rangle$ on the N- body configuration space $\mathbb{C}^N$:

$$K(v) := \frac{1}{2}\langle v, v\rangle = \frac{1}{2}\Sigma_{a=1}^N m_a |v_a|^2. \tag{10}$$

Specifically:

$$\langle q, v\rangle = m_1 q_1 \cdot v_1 + m_2 q_2 \cdot v_2 + m_3 q_3 \cdot v_3 + \ldots + m_N q_N \cdot v_N.$$

where the dot product $q_a \cdot v_i a = Re(q_i \bar{v}_i)$ is the usual dot product in $\mathbb{R}^2 \cong \mathbb{C}$. We call this inner product the "mass-inner product" or sometimes the kinetic energy metric. The total energy H is given by

$$H = K(v) + V(q) \tag{11}$$

The total angular momentum J, and linear momentum P are given by

$$P = \Sigma m_a v_a, J = \Sigma m_a q_a \wedge v_a.$$

and are conserved by the dynamics: i.e. constant along solutions to eq (1).

The virial, or Lagrange-Jacobi identity, asserts that

$$\ddot{I} = 4H$$

and is a consequence of the homogeneity assumption (A2) on V. It follows that if $H = 0$ and $\dot{I} = 0$ then $I = const$. Thus $\dot{I}$ is a 'sub-invariant': it is invariant on the subvariety $\{H = 0\}$ of phase space.

The tangent space $T_q(Gq)$ to the G-orbit at q is sum of $\mathbb{C}_{coll} := \{q : q_1 = q_2 = \ldots = q_N\}$ and $\mathbb{C}q$, the $\mathbb{C}$-span of q. From this fact one computes without difficulty that

$$v \perp T_q(Gq) \iff J(q, v) = P(v) = \dot{I}(q, v) = 0.$$

This equivalence is the essential connection between the Riemannian submersion Π and our Newtonian dynamics as described in the third paragraph of section 3. For more on this relation between geometry and mechanics see part 2 of the book [13], especially chapter 14.

Appendix B. Projections to the shape space and sphere

Here we review the shape sphere and the shape space. This information can be found in a condensed version in [3] and in a leisurely fashion in [12].

The shape sphere is the space of oriented similarity classes of planar triangles, while the shape *space* is the space of oriented *congruence* classes of planar triangles. Topologically, the shape sphere is S^2 and shape space is $\mathbb{R}^3$. The shape sphere is both a submanifold of, and a (sub-) quotient of, the shape space.

Our group G of translations, rotations, and scalings can be built up from a sequence of normal subgroups

$$\mathbb{R}^2 \subset Iso_+(\mathbb{R}^2) \subset G, \tag{12}$$

whose corresponding quotient groups are

$$Iso_+(\mathbb{R}^2)/\mathbb{R}^2 = S^1; G/Iso_+(\mathbb{R}^2) = \mathbb{R}^+.$$

Here $\mathbb{R}^2$ denotes the group of translations of the plane, $Iso_+(\mathbb{R}^2)$ is the group of rigid motions of the plane, S^1 is the group of rotations of the plane and $\mathbb{R}^+$ is the group of scalings. The shape space is the quotient space $\mathbb{C}^3/Iso_+(\mathbb{R}^2)$, while the shape sphere is the sub-quotient $(\mathbb{C}^3 \setminus \mathbb{C}_{coll})/G$ of $\mathbb{C}^3$ by G. These spaces and the relations between them are best understood by implementing the projections, in order, for each subgroup in our list of normal subgroups above:

$$\mathbb{C}^3 \underset{\mathbb{R}^2}{\longrightarrow} \mathbb{C}^2 \underset{S^1}{\longrightarrow} \mathbb{R}^3 - \underset{\mathbb{R}^+}{\longrightarrow} S^2, \tag{13}$$

where the final dotted arrow is used to indicate that the domain of this map is not all of $\mathbb{R}^3$, but rather $\mathbb{R}^3 \setminus \{0\}$. The shape projection of eq (3), central to the statement of theorem 1, is the composition of these three projection maps.

$\mathbb{C}^3 \to \mathbb{C}^2$. The translation group $\mathbb{R}^2$ acts by $(q_1, q_2, q_3) \mapsto (q_1 + q, q_2 + q, q_3 + q)$ and as such is the action of $\mathbb{C}_{coll} \subset \mathbb{C}^3$ on $\mathbb{C}^3$ by vector addition. Thus the quotient space is simply the two dimensional quotient vector space $\mathbb{C}^3/\mathbb{C}_{coll}$. We can identify this quotient with the orthogonal complement to $\mathbb{C}_{coll}$. This orthogonal complement is the space of center-of-mass zero configurations: $\mathbb{C}_{coll}^{\perp} = V_{cm} = \{q \in \mathbb{C}^3 : m_1 q_1 + m_2 q_2 + m_3 q_3 = 0\}$, provided we use the mass inner product of appendix A to compute orthogonality. The origin of $\mathbb{V}_{cm}$ then corresponds to the triple collision line $\mathbb{C}_{coll}$. By choosing an orthonormal frame for $\mathbb{V}_{cm}$ we choose an explicit Hermitian linear isomorphism between $\mathbb{V}_{cm}$ and $\mathbb{C}^2$ with its standard Hermitian structure. In [3], section 5, this linear isomorphism was called " $\mathcal{J}$" after "Jacobi coordinates".

$\mathbb{C}^2 \to \mathbb{R}^3$. The rotation group S^1 now acts on $\mathbb{V}_{cm} = \mathbb{C}^2$ in the standard way coming from complex multipliction: $(Z_1, Z_2) \mapsto (e^{i\theta} Z_1, e^{i\theta} Z_2)$. That the quotient space $\mathbb{C}^2/S^1$ is the shape space $\mathbb{R}^3$ can be seen by using the invariants $|Z_1|^2, |Z_2|^2$ and $Z_1 \bar{Z}_2$. The quotient map is essentially the famous Hopf map: $(Z_1, Z_2) \mapsto (\frac{1}{2}(|Z_1|^2 - |Z_2|^2, Re(Z_1 \bar{Z}_2), Im(Z_1 \bar{Z}_2)) = (w_1, w_2, w_3)$ These image coordinates (w_1, w_2, w_3) are linear coordinates for $\mathbb{R}^3$. The coordinate w_3 has the direct geometric meaning as described above:

$$w_3 = \mu \Delta; \mu^2 = \frac{m_1 m_2 m_3}{m_1 + m_2 + m_3}$$

where Δ is the signed area of the oriented triangle represented by the point (w_1, w_2, w_3) of shape space. In [3], section 5 this quotient map was called the Hopf map and denoted $\mathcal{K}$.

It is useful to look at the shape space from the point of view of invariant theory. Albouy observed that the space of all S^1 invariant real polynomials on a complex two-dimensional vector space (our $\mathbb{V}_{cm}$) are generated by the *quadratic* homogeneous polynomials which in turn form a real four-dimensional vector space. This vector space has for basis w_1, w_2, w_3 and I. Said in terms of the original three-body configuration space $\mathbb{C}^3$, these four functions are a basis for the space of real homogeneous quadratic $Iso_+(\mathbb{R}^2)$-invariant polynomials There is a single relation among these invariants:

$$w_1^2 + w_2^3 + w_3^2 = \frac{I^2}{4}.$$

$\mathbb{R}^3 - \to S^2$. This final map is the quotient of shape space by the scaling group $\mathbb{R}^+$. A scaling element $\lambda \in \mathbb{R}^+$ acts on $\mathbb{C}^3$ by $q \mapsto \lambda q$ with corresponding action on shape space being $w \mapsto \lambda^2 w$, where we have written $w = (w_1, w_2, w_3)$. Consequently, we can identify the shape sphere with the sphere $\{|w|^2 = 1\} \subset \mathbb{R}^3$. The corresponding quotient map is radial projection which is the final broken arrow map $w \mapsto w/|w|$ of eq (13).

Lifting functions. In proving theorem 1 from theorem 2 we had use eq (5) for computing $\Pi^* w_3$ - the lift of w_3 to $\mathbb{C}^3$ as a G-invariant function. It would be more correct here to write $\Pi^* i^* w_3$ where $i : S^2 \to \mathbb{R}^3$ is the inclusion. The function w_3 itself is a linear coordinate on $\mathbb{R}^3$ and as such represents a quadratic $Iso_+(\mathbb{R}^2)$-invariant, namely $\mu\Delta$. Being quadratic, w_3 (or more properly $pr^* w_3$ where pr is the projection $\mathbb{C}^3 \to \mathbb{R}^3$) is not scale invariant and so not of the form $\Pi^* f$ for some function f on the shape sphere. But if we extend $i^* w_3$ off of our shape sphere $\{|w| = I/2\}$ so as to be homogeneous of degree 0, then we will have a function which is G-invariant. This homogeneous degree 0 extension of w_3 is $w_3/|w| = 2\mu\Delta/I$, which is now manifestly G-invariant. We have shown that

$$\Pi^* i^* w_3 = c\Delta/I \tag{14}$$

More bodies. If instead of three bodies we have N bodies in the plane, then the corresponding configuration space is $\mathbb{C}^N$ and its points are to be thought of as vertex-labelled N gons. The same sequence of groups (12) acts. The analogue of the sequence of projections (13) becomes

$$\mathbb{C}^N \to \mathbb{C}^{N-1} \to Cone(\mathbb{CP}^{N-2}) - \to \mathbb{CP}^{N-2}$$

We still have a mass metric associated to kinetic energy. Using this metric we identify $\mathbb{C}^{N-1}$ with center-of-mass zero configurations in $\mathbb{C}^N$. The metric's squared norm, restricted to $\mathbb{C}^{N-1}$, is the moment of inertia I. If we push down the induced metric on the sphere $\{I = 1\}$ to $\mathbb{CP}^{N-2}$ we get the standard Fubini-Study metric. The pushed-down metric on the shape space, $Cone(\mathbb{CP}^{N-2})$ is the cone over the Fubini-Study metric.

Appendix C. On Mnëv's Universality Theorem.

The version of Mnëv's theorem stated here (Theorem 3) is not the theorem found in Mnëv or in most subsequent expositions on Mnëv. In this appendix we state Mnëv's theorem in a form close to its original version and show how this version implies ours. Beyond performing this translation, this appendix is a discourse on his theorem, but not a proof. For proofs we recommend [16] or [9].

The soul of Mnëv's theorem and its proof rests on the classical projective geometric constructions of the basic arithmetic operations of $+, -, \times, \div$, often attributed to von Staudt. We refer the interested reader to the figures around p. 218 of [15], or the end of [7], or figures 15 and 16 towards the end of [16].

Consider the configuration space $(\mathbb{RP}^2)^{N+2}$ of $N+2$ points in the real projective plane $\mathbb{RP}^2$. Call such a configuration $(p_0, p_1, \ldots, p_{N+1})$ "generic" if it is generic in our sense: no three points p_i, p_j, p_k are collinear. The set of generic points is a Zariski open and dense subset $(\mathbb{RP}^2)^{N+2}$ on which the group $\mathbb{P}Gl(3,\mathbb{R})$ of projective transformations of the plane acts freely, provided $N \geq 2$, so that there are indeed 4 points. Call the corresponding quotient space the projective shape space. Using the sign of our areas $\Delta(i,j,k)$, Mnëv decomposed projective shape space into equivalence classes he called "open oriented combinatorial types". His equivalence relation is weaker than the equivalence relation of lying in the same topological component so that a single oriented combinatorial class might contain many topological components. An "open primary semialgebraic variety" is the solution space of polynomial inequalities with coeffients in the rational numbers. Two such sets are called "stably equivalent" if, after forming the product of each by Euclidean spaces of possibly different dimensions, the resulting spaces are birationally isomorphic by way of rational maps whose polynomial numerator and denominators have rational coefficients.

Theorem 5 (Mnëv). *Given any open primary semi-algebraic set X there is an N such that X is stably equivalent to some open oriented combinatorial class within the projective shape space for $N+2$ points.*

The main philosophical difference between his theorem and ours (theorem 3) is that his combinatorial classes, like general semi-algebraic sets, can contain many components.

Translating this Mnëv theorem to our Mnëv theorem. A basic theorem in projective geometry asserts that the group of projective transformations of the projective plane acts freely *and transitively* on the space of generic 4-tuples. This theorem allows Mnëv to realize projective shape space by the gauge-fixing artifice of insisting that the first 4 points p_1, p_2, p_3, p_4 of each $N+2$-tuple are four particular generic points, namely the points making up the standard projective frame, $p_0 = [0,0,1], p_x = [1,0,0], p_y = [0,1,0]$ and $p_E = [1,1,1]$, where $[x,y,z]$ are standard homogeneous coordinates for the projective plane. The line $p_x p_y$ corresponds to the usual line at infinity ($z = 0$) so we denote it as ℓ_∞. Then $\mathbb{RP}^2 \setminus \ell_\infty \cong \mathbb{R}^2 \cong \mathbb{C}$ corresponds to the standard affine plane ($z \neq 0$) with affine coordinates $X = x/z, Y = y/z$. The points p_0 and p_E of the projective frame lie in the affine plane, and by genericity, so do all the remaining vertices $p_5, p_6, \ldots, p_{N+2}$ of our gauge-fixed projective $N+2$-gon. The act of forgetting p_x and p_y defines a diffeomorphism of projective shape space onto a real Zariski open subset of an affine subspace of our N-body configuration space, $(\mathbb{R}^2)^N$; in other words, we send $(p_0, p_x, p_y, p_E, p_5, \ldots, p_{N+2})$ to $(p_0, p_E, p_4, p_5, \ldots, p_{N+2}) \in A_{N-2} \subset (\mathbb{R}^2)^N$. where $A_{N-2} = (\mathbb{R}^2)^{N-2} = \mathbb{C}^{N-2}$ is the affine space of N-tuples whose first two points equal p_0 and p_E.

This isomorphic copy of projective shape space is *almost* the shape space $\mathbb{CP}^{N-2} \setminus \Sigma_N$, of theorem 3. Recall that our shape space is the quotient space $(\mathbb{C}^N \setminus \Sigma_N)/G$ where G is the group of rotations, translations and scalings acting on the affine plane. Since G acts freely and transitively on *pairs* of distinct points of the affine plane, so we can again play the gauge fixing game by using the G-action to impose the condition that the first two points of our affine N-tuple are the points $(0,0) = 0$ and $(1,1) = 1+i$. This corresponds to the projective fixing of p_0 and p_E. Imposing this condition "uses up" all of G, thus identifying $\mathbb{CP}^{N-2} \setminus \Sigma_N$ with a Zariski open dense subset of the same affine $\mathbb{C}^{N-2} \subset \mathbb{C}^N$. **Mnëv's projective shape space is a Zariski open subset of ours**. Our condition of genericity is that no three points from our affine N-tuple (including p_0 and p_E now) are collinear which we expressed as $\Delta(i,j,k) \neq 0$ where $\Delta(i,j,k)$ is the (signed) area of the triangle. (Choose an orientation of the affine plane.) To these conditions Mnëv would add $X_i - X_j \neq 0$ and $Y_i - Y_j \neq 0$ which are the conditions that p_i, p_j, p_x and p_i, p_j, p_y are not collinear. Mnëv declared that two of his tuples were in the same "open oriented combinatorial type" if the signs of the functions $\Delta(i,j,k), X_i - X_j, Y_i - Y_j$ agree on the two tuples, for all indices i,j,k. Dropping the conditions on $X_i - X_j, Y_i - Y_j$ yields our shape space, and by ignoring these functions will not change the results of Mnëv's theorem. We will not go through the verification of this last assertion. In this way, we get all primary semi-algebraic subsets within our shape space, and taking components of these, get the result of theorem 3. **End of translation between theorems.**

Remark. The theorem as Mnëv originally stated it allows him to insist that some subset of the $\Delta(i,j,k) = 0$, which corresponds to forced collinearity conditions. In this way he is able to get polynomial equalities as well as inequalities, so the theorem asserts that in this way all semi-algebraic sets, open and closed are realized. For simplicity we have omitted going through this version as it makes the statements a bit more complicated and we do not need it.

Matroids and minors.

Observe that the functions just described, $\Delta(i,j,k)$, $X_i - X_j, Y_i - Y_j$ are the 3 by 3 minors of the matrix

$$\begin{bmatrix} 1 & 0 & 0 & 1 & X_5 & X_6 & \dots & X_{N+2} \\ 0 & 1 & 0 & 1 & Y_5 & Y_6 & \dots & Y_{N+2} \\ 0 & 0 & 1 & 1 & 1 & 1 & \dots & 1 \end{bmatrix} \tag{15}$$

while if we forget the functions corresponding to the points p_x, p_y out at infinity and instead want to simply use the signed areas $\Delta(i,j,k)$ these are the 3 by 3 minors of the submatrix

$$\begin{bmatrix} 0 & 1 & X_5 & X_6 & \dots & X_{N+2} \\ 0 & 1 & Y_5 & Y_6 & \dots & Y_{N+2} \\ 1 & 1 & 1 & 1 & \dots & 1 \end{bmatrix} \tag{16}$$

The assignment $\{i,j,k\} \mapsto sign(\Delta(i,j,k))$ used to describe the oriented combinatorial type of Mnëv is an example of an "oriented matroid". A 'uniform oriented matroid of rank 3" on the index set $\{1,2,\dots,N\}$ is, according to one of a number of equivalent definitions, an assignment $\{i,j,k\} \mapsto M(i,j,k) \in \{+1,-1\}$ of a $\pm$ sign to each three element subset i,j,k of the index set. (This map is subject to certain constraints copying the Plücker relations.) See [1]. An affine realization of such a gadget M is a real matrix of the form of eq (16) such that the sign of the minors whose columns, indexed by this same index set agrees with the matroid's assignments. And a projective realization of M is a real amatirx of the form of eq (15) satisfying the same properties. The corresponding projective realization space consisting of all matrix realizations is then equal to the set of points having a given oriented combinatorial type in Mnëv's original theorem. The affine realization space corresponds to a combinatorial component containing some number of components as described in our version of Mnëv's theorem.

If we relax the assumption that the signs of the determinants of the minors are ± 1 by allowing some minors to be zero, then we are insisting the corresponding labelled triples become collinear. The corresponding combinatorial object is now a function $M :$ (r element subsets of $\{1,2,\dots,N\}) \to \{+1,-1,0\}$. (Again certain conditions are imposed on M which are sign versions of the Plücker relations.) Such an object is, by definition, a rank r oriented matroid, so that the adjective "uniform" applied to an oriented matroid excludes M from taking on the value 0, while the adjective "rank" describes the size r of the subsets of the index set to which the map applies. The original full (as opposed to 'open') version of the Mnëv theorem asserts that all algebraic and semi-algebraic real sets are stably isomorphic to the realization space for some oriented rank 3 matrix.

Acknowledgements

I am grateful to Persi Diaconis and Toshiaki Fujiwara for inspirational conversations during the formation of this piece, to Serge Tabachnikov and Richard Schwartz for conversations around the questions section, to Rafe Mazzeo and Jie Qing for emails around asymptotic hyperbolicity near the 'real divisor', to Gil Bor for a careful reading and useful questions, to Alain Albouy for intellectual orientation over the decades, and I am especially grateful to Misha Kapovich for pointing out and explaining Mnëv's work and its import. I thank NSF grant DMS-1305844 for essential support.

References

[1] A. Björner, M. Las Vergnas, B. Sturmfels, N. White, and G. Ziegler, **Oriented Matroids**, Encyc. of Math. and its Applications, v. 46, edited by G.C. Rota, Cambridge U. Press, 1993

[2] J. W. Cannon, W. J. Floyd, R. Kenyon and W. R Parry *Hyperbolic Geometry* in **Flavors of Geometry** MSRI Publications Volume 31, 1997

[3] A. Chenciner and R. Montgomery, *A remarkable periodic solution of the three-body problem in the case of equal masses*, with Alain Chenciner, Annals of Mathematics, **152**, 881-901 2000.

[4] J Goodman, *Pseudoline Arrangements*, ch. 5 of **Handbook of Discrete and Computational Geometry**, J Goodman and J O'Rourke, ed.,CRC Press, 1997, Boca Raton, Florida.

[5] M. Kapovich, **Hyperbolic Manifolds and Discrete Groups**, Progress in Mathematics, vol. 183, Birkhäuser, Boston, 2001.

[6] C. Jackman and R. Montgomery, *No Hyperbolic Pants for the Planar Four-Body Problem*, Pacific Journal of Mathematics 280-2 (2016), 401–410.

[7] S. H. Lee and R. Vakil, *Mnëv -Sturmfels Universality for Schemes* arxiv: 1202.3934v2

[8] R. Mazzeo, *Hodge Cohomology of Negatively Curved Manifolds*, Thesis, MIT, Cambridge MA, 1986.

[9] N. E. Mnëv, *Universality Theorems on the Classification Problem of Configuration Varieties and Convex Polytopes Varieties* in: Topology and geometry – Rohlin Seminar, 527-543, Lecture Notes in Math., 1346, Springer, Berlin, (1988).

[10] R. Montgomery, *Infinitely Many Syzygies Archives for Rational Mechanics*, v. 164 (2002), no. 4, 311–340, 2002.

[11] R. Montgomery, *Hyperbolic Pants fit a three-body problem* Ergodic Theory and Dynamical Systems, vol.25, 921-947, (2005).

[12] R. Montgomery, *The Three-Body Problem and the Shape Sphere*, Amer. Math. Monthly, v 122, no. 4, pp 299-321 , 2015; arXiv:1402.0841.

[13] R. Montgomery, **A tour of subriemannian geometries , their geodesics, and applications**, Mathematical Surveys and Monographs, vol. 91, American Math. Society, Providence, Rhode Island, 2002.

[14] J. Richter-Gebert and G Ziegler, *Oriented Matroids*, ; in the CRC Handbook on Discrete and Computational Geometry (J.E. Goodman, J. O?Rourke, eds.) pp. 111-132, CRC Press, Boca Raton, New York (1997).

[15] J. Richter-Gebert, *The Universality Theorems for Oriented Matroids and Polytopes* in Contempary Mathematics series, vol. 223, B. Chazelle, J. Goodman and R. Pollack, Editors, volume title Advances in Discrete and Computational Geometry, American Math. Society, Providence, Rhode Island, 1996.

[16] P. Shor, *Stretchability of Pseudolines in NP-Hard*, DIMACS Series in Discrete Math. and Theoretical Computer Science, vol. 4, 531-554,1991.

Mathematics Department, University of California, Santa Cruz, Santa Cruz CA 95064

E-mail address: rmont@ucsc.edu