The Information Metric on Rational Maps

Michael K. Murray

CONTENTS

Mathematics Subject Classification: 58E20, 58E15, 53C20

The information metric is a construction in statistics which can be used to define a (possibly degenerate) metric on various moduli spaces such as those of instantons and harmonic maps. This metric is shown to be nondegenerate for the space of harmonic maps of the two-sphere onto itself of any degree.

1. INTRODUCTION

The information metric arises in statistics as a metric on a manifold of probability distributions [Rao 1945]. Its construction is very simple and can be applied, in principle, to any manifold that parametrises a set of probability distributions or measures. However, in this general setting, the information metric may be degenerate.

Such manifolds can be taken, for example, as the minimum sets of variational problems. Here each point of the minimum set has associated to it an energy density function, which after appropriate scaling can be regarded as a probability density function.

The example I would ultimately like to understand is the information metric on the moduli space of instantons on a four-manifold. The reader is referred to [Donaldson and Kronheimer 1990] for details and definitions. This space is of particular importance in the study of four-manifolds, where, among other things, it leads to the celebrated theorem of Donaldson on the non-smoothability of certain four-manifolds, which leads, in turn, to the existence of exotic $\mathbb{R}^4$'s. Except for instantons of charge one on the four-sphere, essentially nothing is known about the structure of the information metric on this manifold—even whether it is nondegenerate. However, for charge-one instantons on the four-sphere the moduli space is the five-ball and

1058-6458/96 $0.50 per page

symmetry considerations [Hitchin 1988] show that the information metric is the hyperbolic metric.

In an attempt to develop some intuition for the problem just described, I have studied its analogue for rational maps from the two-sphere to itself. These are harmonic maps of the two-sphere to itself of minimum energy and positive degree. The information metric turns out to be degenerate in directions corresponding to compositions of the rational map with isometries of the target sphere. This defines an action of $\mathrm{SO}(3)$ on the space of rational maps and the appropriate space to consider is the quotient of the space of rational maps by this action. These isometries play a role analogous to gauge transformations for instantons.

For example, when the map has degree one, the space of rational maps is isomorphic to $\mathrm{PSL}(2,\mathbb{C})$ and the space of interest is $\mathrm{PSL}(2,\mathbb{C})/\mathrm{SO}(3) = \mathrm{SL}(2,\mathbb{C})/\mathrm{SU}(2)$, which is a well-known model for hyperbolic three-space. The information metric in this case is indeed the hyperbolic metric, because of symmetry considerations completely analogous to the instantons of charge one on the four-sphere—namely, the metric is determined, up to scale, by the fact that it is invariant under the group of conformal transformations of the two-sphere or four-sphere as the case may be (see the end of Section 4).

Section 2 defines the information metric in its traditional context, and Section 3 details the extension to manifolds of minima. Sections 4 and 5 specialize to the case of rational maps. Section 6 gives a computer proof that for maps of degree two the information metric is nondegenerate, Section 7 an analytic proof that works for maps of any degree.

The latter proof of course makes the former redundant, but I have included it as an interesting example of the power of symbolic computation in mathematical experimentation, and because it provided the essential motivation for the search for the analytic proof for maps of any degree. There was no other source of intuition, other than the simple, highly symmetric cases like charge-one maps and instantons, that suggested that the information metric would be nondegenerate.

Of course, the analytic proof is, in its turn, motivation to investigate further the instanton case. This is work in progress and I return to it in the Conclusion. For those familiar with instantons, let me remark here that the information metric is most definitely not the same as the L^2 metric on the moduli space. For instance, in the case of instantons on the four-sphere of charge one, the information metric has negative curvature and infinite volume, whereas the L^2 metric has positive curvature and finite volume [Groisser and Parker 1987].

In the conclusion I also discuss some possible generalisations to other spaces of harmonic and holomorphic maps.

2. THE INFORMATION METRIC IN STATISTICS

The information metric arises in statistics in the following manner. Interest lies in a submanifold P of the space of all probability measures of some probability space Ω. To avoid the technicalities involved in describing what that might mean, it is usual to assume that P is parametrised by coordinates θ so it has the form:

$$P = \{p(\ ,\theta)\mu \mid \theta \in U \subset \mathbb{R}^n\},$$

where μ is a measure on Ω, U is an open subset of $\mathbb{R}^n$ and $p(\ ,\theta)$ is a probability density function on Ω, for each $\theta \in U$. It will become clear below that various conditions need to be imposed on the differentiability of p as a function of θ and on the integrability of p and its partial derivatives. These conditions will be ignored in this motivational discussion but addressed later for the case of rational maps.

The information metric is defined in the θ coordinates by

$$g_{ij} = \int_\Omega \frac{1}{p}\frac{\partial p}{\partial \theta^i}\frac{\partial p}{\partial \theta^j}\mu.$$

Various more invariant formulations are possible. We can suppress the parameters θ and replace the coordinate vector fields by two tangent vectors X

and Y in the tangent space to P at $p\mu$: the metric applied to these two vectors is then

$$g_p(X,Y) = \int_\Omega \frac{1}{p} X(p)Y(p)\mu,$$

where $X(p)$ is the function on Ω whose value at $x \in \Omega$ is the result of applying X to the function on P of density at x (i.e., to $p\mu \mapsto p(x)$). Finally we can suppress the dependence on the tangent vectors by letting $dp\,dp$ be the tensor that when applied to the two tangent vectors X and Y yields $X(p)Y(p)$: then

$$g_p = \int_\Omega \frac{1}{p}\, dp\, dp\, \mu.$$

Although it is not important for the applications below, it is worth remarking that the information metric is defined in statistics in a slightly different manner. There the so-called "log-likelihood" is defined by $l(\theta) = \log p(\theta)$, and the metric is then $g_{ij} = E_p((\partial l/\partial\theta^i)(\partial l/\partial\theta^j))$, where $E_p(f)$ is the expectation of the random variable f with respect to the density p, that is, $E_p(f) = \int_\Omega f p\mu$. For more details see [Amari 1985; Murray and Rice 1993].

The information metric has two, related, properties that will be important below. First, if the measure μ is changed to another equivalent measure $(\exp f)\mu$, the metric does not change. To see this, let $\mu' = (\exp f)\mu$ and suppose $p'\mu' = p\mu$; then $p = (\exp f)p'$ and

$$\frac{1}{p'}\, dp'\, dp'\mu' = \frac{1}{p}\, dp\, dp\, \mu, \tag{2.1}$$

so that the metric is the same for either choice of measure. This means that the information metric really lives on the space of measures. A more direct definition is possible [Murray and Rice 1993].

Secondly, the information metric enjoys an important invariance property. A transformation of Ω induces a transformation of the space of measures on Ω. Assuming this transformation fixes P and acts on it as a diffeomorphism, it is easy to check that this action is isometric. This is the property that proves that on the space of instantons of charge one or of rational maps of degree one the information metric coincides with the hyperbolic metric.

An interesting example of the information metric for both geometers and statisticians is the space of normal densities. Here Ω is the real line and the mean and variance identify the space of all normal densities with the upper half-plane in $\mathbb{R}^2$. A straightforward calculation shows that the information metric on this upper half-plane is the Poincaré metric [Rao 1945; Amari 1985].

3. THE INFORMATION METRIC ON MODULI SPACES

Consider a general variational problem. Assume that Ω is a Riemannian manifold and let φ be some type of object on Ω whose energy can be minimised. For instance, φ may be a connection on a bundle over Ω, or a map to another Riemannian manifold. Denote by $E(\varphi)$ the energy of φ, and assume that it is the integral of an energy density $e(\varphi)$:

$$E(\varphi) = \int_\Omega e(\varphi)\,\mathrm{vol},$$

where vol is the Riemannian volume form on Ω. Assume that the minima of E form a manifold P. Then the construction of the information metric can be applied, because all points of P have the same total mass (integral). If X and Y are tangent vectors at $\varphi \in P$ and $X(e)$ and $Y(e)$ are the rates of change of e in the corresponding directions, then

$$g_\varphi(X,Y) = \int_\Omega \frac{1}{e} X(e)Y(e)\mu.$$

If this integral exists, it defines a possibly degenerate metric on P, that is a positive, symmetric two-tensor on P. For this metric to be nondegenerate it is necessary and sufficient that $X(e)$ never be identically zero on Ω unless X is zero. This is equivalent to asking that the energy density, considered as a map

$$e : P \to C^\infty(\Omega, \mathbb{R}),$$

be an immersion (have an injective derivative). In the cases considered by statisticians this condition is built in as an assumption, namely that the so-called "scores" $\partial l/\partial\theta^i = (\partial p/\partial\theta^i)/p$ be linearly independent as functions on Ω.

4. RATIONAL MAPS

The reader is referred to [Eells and Lemaire 1983], for example, for the basic results stated here. We denote by Rat_k the space of rational maps of S^2 of degree k, that is, maps that are given, under identification of S^2 with the extended complex plane by stereographic projection, by

$$f(z) = \frac{p(z)}{q(z)},$$

where p and q are polynomials of degree at most k without common factors, at least one being of degree k. These maps are conformal except at branch points, that is, the derivative $df(x)$, considered as a linear map from the tangent space at $x \in S^2$ to the tangent space at $f(x)$, is a scalar multiple of an orthogonal transformation with respect to the (spherical) inner product on both tangent spaces. We denote by $|df(x)|$ this scalar factor. It is well-known that rational maps are minima of the *Dirichlet energy* functional

$$E(f) = \int_{S^2} |df|^2 \,\mathrm{vol}_{S^2},$$

where vol_{S^2} is the spherical volume (or area) form. We denote the corresponding energy density $|df|^2$ by $\varepsilon(f)$.

It is convenient to work in the complex plane instead of in S^2. In $\mathbb{C}$, the Riemannian metric induced from S^2 by stereographic projection has element of length $2/(1+|z|^2)$ times the Euclidean element of length. Therefore

$$\varepsilon(f)(z) = \frac{|f'(z)|^2}{(1+|f(z)|^2)^2}(1+|z|^2)^2,$$

where f' is the usual complex derivative of f regarded as a meromorphic function. Also,

$$\mathrm{vol}_{S^2} = \frac{4}{(1+|z|^2)^2}\,\mathrm{vol}_{\mathbb{C}},$$

where $\mathrm{vol}_{\mathbb{C}} = \frac{1}{2} i\, dz \wedge d\bar{z}$ is the usual volume element in $\mathbb{C}$. Thus $E(f)$ is the integral of the energy density

$$e(f) = \frac{4|f'|^2}{(1+|f|^2)^2}$$

over $\mathbb{C}$ with respect to $\mathrm{vol}_{\mathbb{C}}$. Note that $e(f)$ is finite at every point in $\mathbb{C}$, because $\varepsilon(f)$ is.

The information metric is the symmetric tensor defined by

$$g = \int_{S^2} \frac{1}{\varepsilon}\, d\varepsilon\, d\varepsilon \;\mathrm{vol}_{S^2} \tag{4.1}$$

if this integral exists. The integrand is finite except at the points where $df = 0$—the branch points of f—which are finite in number. It follows that the convergence of (4.1) only has to be checked around each of these points; moreover we can assume that $\infty \in S^2$ is not a branch point, so using (2.1) we can write

$$g = \int_{\mathbb{C}} \frac{1}{e} de\, de \;\mathrm{vol}_{\mathbb{C}},$$

and we just have to check that this integral converges near the points where $e = 0$. Using the representation $f = p/q$, where p and q are polynomials, we have

$$e(f) = \left| 4\frac{qp' - pq'}{|p|^2 + |q|^2} \right|^2 = |\alpha|^2, \tag{4.2}$$

where $\alpha = 4(qp' - pq')/(|p|^2 + |q|^2)$ is finite everywhere in $\mathbb{C}$. Then

$$de = 2\,\mathrm{Re}(\alpha\, d\bar{\alpha}) = 2\,\mathrm{Re}(\bar{\alpha}\, d\alpha),$$

so that

$$\frac{de\, de}{e} = \frac{4\,\mathrm{Re}(\alpha\, d\bar{\alpha})\,\mathrm{Re}(\bar{\alpha}\, d\alpha)}{\alpha\bar{\alpha}}.$$

By the Cauchy–Schwartz inequality this is bounded above by $4\, d\alpha\, d\bar{\alpha}$, which gives the required result.

So the information metric is a well-defined symmetric two-tensor on the space of rational maps. It is not hard to see that it is, in fact, degenerate! Changing the map f by rotating the target sphere leaves the energy density unchanged, so if X is a tangent vector to Rat_k in the direction of such a rotation we have $de(X) = 0$ and hence $g(X, X) = 0$. We can remove this degeneracy by quotienting the space of rational maps by this action of the rotation group. We set

$$\mathcal{M}_k = \mathrm{Rat}_k \,/\, \mathrm{SO}(3)$$

and consider the energy density as defined on elements of $\mathcal{M}_k$, obtaining a map

$$e : \mathcal{M}_k \to C^\infty(\mathbb{C}).$$

The information metric is defined in the same way as before.

If we compose $f : S^2 \to S^2$ with a fractional linear transformation X of the domain, the result is again a new rational map $f \circ X$. This defines an action of $\mathrm{SL}(2, \mathbb{C})$ on Rat_k that clearly descends to $\mathcal{M}_k$. It is straightforward to check that $e(f \circ X) = e(f) \circ X|X'|^2$. It follows from the second property of the information metric discussed in Section 2 that this metric is invariant under the action of $\mathrm{SL}(2, \mathbb{C})$.

Rational maps of degree one are just fractional linear transformations. After dividing by the action of $\mathrm{SO}(3)$ this gives

$$\mathcal{M}_1 = \mathrm{SL}(2, \mathbb{C}) / \,\mathrm{SO}(3)$$

and the information metric is $\mathrm{SL}(2, \mathbb{C})$ invariant. The only possibility is that $\mathcal{M}_1$, equipped with the information metric, is isometric to hyperbolic three-space.

5. PARAMETRISATION OF RATIONAL MAPS OF DEGREE k

So far we have not defined a manifold structure on either Rat_k or $\mathcal{M}_k$. Indeed, as one is the quotient of the other, there might be cause for concern unless the group action is free. We define this structure now. Recall that $f \in \mathrm{Rat}_k$ can be written as a quotient of polynomials $p_0 + \cdots + p_k z^k$ and $q = q_0 + \cdots + q_k z^k$ without common factors, at least one of p_k and q_k being nonzero. The criterion for p and q to have a common zero is algebraic in the coefficients $p_0, \ldots, p_k, q_0, \ldots, q_k$ (see [Milnor 1993, p. 69], for example). Interpreting this $(2k+1)$-tuple of coefficients as the homogeneous coordinates of a point in complex projective space $\mathbb{CP}^{2k+1}$, we see that Rat_k can be identified with the complement of an algebraic subset of $\mathbb{CP}^{2k+1}$, which is open and so has a natural structure as a $(2k + 1)$-dimensional complex manifold or a $(4k + 2)$-dimensional real manifold. For $k > 0$ any rational map is onto, so $\mathrm{SO}(3)$ acts freely on Rat_k and $\mathcal{M}_k$ is a manifold of real dimension $4k - 1$.

Because $\mathrm{SO}(3)$ acts transitively on the sphere with stabiliser $\mathrm{SO}(2) = U(1)$, the quotient $\mathcal{M}_k = \mathrm{Rat}_k \,/\, \mathrm{SO}(3)$ factors through Rat_k^b, the space of *based rational maps* of degree k—those taking, say, ∞ to 0. Then $\mathcal{M}_k = \mathrm{Rat}_k^b \,/ U(1)$, where $U(1)$ is the stabiliser of 0. In the current situation, the condition that ∞ maps to 0 is equivalent to $p_k = 0$, so Rat_k^b can be identified with $\mathrm{Rat}_k \cap \{p_k = 0\}$. By the same argument as in the preceding paragraph, Rat_k^b is the (open) complement of an algebraic subset of $\mathbb{CP}^{2k}$, or indeed of $\mathbb{C}^{2k}$ because we can normalize the coefficient q_k to be 1. The action of $U(1)$ on Rat_k^b of which $\mathcal{M}_k$ is the quotient is given by

$$\exp(i\theta)(p, q) = (\exp(i\theta)p, q).$$

This action is free.

Rather than show that e is an immersion as a map from $\mathcal{M}_k$, it is simplest to consider it as a map from Rat_k^b and show that the only tangent vectors in the kernel of the tangent map to e are those pointing in the direction of the $U(1)$ fibres. Note that at a point (p, q) the tangent vectors in the $U(1)$ fibre direction are those of the form $v = (i\mu p, 0)$ where μ is a real number.

Two proofs of the fact that e is an immersion shall be given. The first, in Section 6, is a computer experiment with the case $k = 2$. This was what motivated the search for the analytic proof that is given in the Section 7.

6. THE COMPUTER PROOF FOR k = 2

Let

$$f(z)=\frac{p(z)}{q(z)}=\frac{az+b}{z^2+cz+d}$$

be our rational map. The condition that p and q do not have a common zero is easily computed by substituting the known zero of p into q: we get

$$b^2+da^2-abc\neq 0. \tag{6.1}$$

So Rat_k^b can be identified with the open subset of $\mathbb{C}^4$ where $b^2+da^2-abc\neq 0$. The space $\mathcal{M}_2$ is the quotient of this open set under the action of $U(1)$ defined by

$$\exp(i\theta)(a,b,c,d)=(\exp(i\theta)a,\ \exp(i\theta)b,\ c,d).$$

To find the kernel of the tangent map to the energy we look at the exterior derivative de of the energy density (4.2). This differential form is a quotient of polynomials in z and $\bar z$ with denominator $(p\bar p+q\bar q)^3$. To find out when it vanishes it suffices to examine the numerator

$$n=(p\bar p+q\bar q)^3\,de,$$

a polynomial in z and $\bar z$ whose coefficients are linear combinations of the differential forms da, db, dc, dd and their complex conjugates. Applying this to a tangent vector v gives a polynomial $n(v)$ in z and $\bar z$, and we must show that if its coefficients are all zero we necessarily have $v=(ita,itb,0,0)$ for t real.

This calculation is impractical if approached the wrong way, because of the number of terms of $n(v)$: when expanded using the symbolic manipulation program Maple [Char et al. 1991], this polynomial is some 300 lines long. One way to proceed is this: pick out the leading coefficient of $n(v)$—a quite simple expression—and set it to zero. This imposes a relation between $da(v)$ and $d\bar a(v)$. Then substitute this into $n(v)$, simplify the result and pick out the next leading coefficient. This imposes a new relation. Continue on in this way, in the hope that the desired condition on v is forced before the relations make the polynomial vanish. This indeed occurs, and hence e is an immersion and the information metric is nondegenerate on $\mathcal{M}_2$.

Although the calculation could done with pen and paper, symbolic manipulation software such as Maple makes the process much less painful and decreases the probability of error. Because some readers may be as unfamiliar with use of such systems as I was when this work was started, I will detail what happens. Let $v=(v_a,v_b,v_c,v_d)$; then there are two cases:

Case 1: $a\neq 0$. The leading coefficient of $n(v)$ is $v_a\bar a+\bar v_a a$, and this vanishes if and only if $v_a=ita$ for some real number t.

Assign the variable v_a the value ita and the variable $\bar v_a$ the value $-it\bar a$. Simplify the resulting expression and extract the coefficient of $z^4\bar z^3$, which is

$$2a(\bar v_b+it\bar b-\bar v_c\bar a).$$

As $a\neq 0$ this expression vanishes only if $v_b=itb+v_ca$ and $\bar v_b=-it\bar b+\bar v_c\bar a$. Assign these values and simplify; the coefficient of $z^4\bar z^2$ is

$$3(\bar v_c\bar c a\bar a-\bar v_d a\bar a-\bar v_c a\bar b)=3a\bar a(\bar v_c\bar c-\bar v_d-\bar v_c\bar b/\bar a).$$

Again using $a\neq 0$, deduce that $v_d=v_c(c-b/a)$ and $\bar v_d=\bar v_c(\bar c-\bar b/\bar a)$. Substitute this into what remains of $n(v)$ and obtain the coefficient of $z^4\bar z$:

$$4\bar v_c\frac{a}{\bar a}(-\bar a\bar b\bar c+\bar b^2+\bar d\bar a^2).$$

Now use (6.1) to derive $\bar v_c=0$ and $v_c=0$, then work back to get $v_d,\bar v_d=0$ and finally

$$v_b=it\bar b,\qquad \bar v_b=-itb.$$

Hence $v=(ita,itb,0,0)$ as required.

Case 2: $a=0$. Here (6.1) takes the form $b\neq 0$. Proceeding as in the first case, compute the leading coefficient of $n(v)$ as $2v_a\bar b$, so $v_a=0$. Assign the value zero to v_a and simplify; the next leading coefficient is

$$4\bar v_b b+4v_b\bar b,$$

so $v_b = itb$ for some real t and $\bar{v}_b = -it\bar{b}$. Assign these values and simplify; the leading coefficient becomes $-6\bar{v}_c b\bar{b}$, so $\bar{v}_c = 0$ and $v_c = 0$. The next leading coefficient has the form $-8\bar{v}_d b\bar{b}$, so $v_d = 0$. Hence $v = (0, itb, 0, 0) = (ita, itb, 0, 0)$, as required.

7. ANALYTIC PROOF FOR ARBITRARY k

So far all the holomorphic maps we have been considering have have been defined on all of S^2 and when originally submitted this paper contained an analytic proof that the energy density of such maps is an immersion. It was suggested by a referee that this result should, in fact, be local, that is, true for maps defined on any open subset of S^2. This turns out to be the case. Perhaps referees are an even better source of inspiration than computers!

We give an analytic proof of the local result, followed on the next page by a more geometric one, suggested by Curt McMullen.

Consider a family of holomorphic maps f_t from a connected open subset V of S^2 into S^2, parametrised by t in a ball of radius ε in $\mathbb{R}^n$. We shall assume that the dependence on t is smooth and, of course, that the derivative of the energy $e(f_t)$ is zero at $t = 0$. Identify S^2 with the extended complex plane and fix a point z_0 in V. Since $\mathrm{SU}(2) \to S^2$ is a locally trivial principal bundle we can, by shrinking ε if necessary, lift the function $t \mapsto f_t(z_0)$ to a function $t \mapsto g_t(z_0)$ into $\mathrm{SU}(2)$. The map $(g_t(z_0))^{-1} f_t$ now maps z_0 to 0 and has the same energy as f_t, so we may as well assume at the outset that $f_t(z_0) = 0$.

Let $f = f_0$ and let $\dot{f}$ denote the derivative of f_t with respect to t at $t = 0$. We will show that $\dot{f} = i\mu f$, where μ is a real constant. Note that because f and $\dot{f}$ are holomorphic and V is connected this is equivalent to showing that $\mathrm{Re}(\dot{f}\bar{f}) = 0$ in some neighbourhood of z_0. Indeed, if this is true, $\dot{f}/(if)$ is real and meromorphic in a neighbourhood of z_0, and therefore equals a constant μ, so that $\dot{f} - i\mu f = 0$ in this neighbourhood and hence on all of V. To pick the neighbourhood note that, by continuity, after a possible decrease in ε, we can find an open disk U about z_0 that contains no pole of f_t for all t with $|t| < \varepsilon$. From now on we work in U.

The energy of the map f_t can be written as

$$e(f) = \partial\bar{\partial} \log(1 + |f_t|^2),$$

where we use the shorthands $\partial = \partial/\partial z$ and $\bar{\partial} = \partial/\partial \bar{z}$. Then

$$\left. \frac{d}{dt} \log(1 + |f_t|^2) \right|_{t=0}$$

is in the kernel of $\partial\bar{\partial}$ on U. It follows that

$$\left. \frac{d}{dt} \log(1 + |f_t|^2) \right|_{t=0} = \varphi + \bar{\varphi}$$

for some analytic function φ. Computing the derivative with respect to t at $t = 0$ yields

$$2\,\mathrm{Re}(\bar{f}\dot{f}) = (1 + f\bar{f})(\varphi + \bar{\varphi}). \tag{7.1}$$

If we evaluate this at z_0 we find that

$$\varphi(z_0) + \bar{\varphi}(z_0) = 0.$$

If we differentiate (7.1) with respect to z any number of times and evaluate at z_0 we find that φ is a constant, so $\varphi + \bar{\varphi} = 0$. Hence we have $\mathrm{Re}(\bar{f}\dot{f}) = 0$ as required.

A similar approach can also be used to show that the energy density is injective. If two holomorphic functions f_1 and f_2 have the same energy density, then

$$\partial\bar{\partial} \log\left(\frac{1 + |f_1|^2}{1 + |f_2|^2}\right) = 0$$

and hence

$$\log\left(\frac{1 + |f_1|^2}{1 + |f_2|^2}\right) = \varphi + \bar{\varphi}$$

for holomorphic φ. Letting $\psi = \exp(\varphi)$ gives

$$1 + |f_1|^2 = (1 + |f_2|^2)\psi\bar{\psi}.$$

Calculating the derivatives of ψ at z_0 and using the fact that $f_1(z_0) = f_2(z_0)$ shows that ψ is a constant complex number of modulus 1. Hence $|f_1|^2 = |f_2|^2$ and it follows that $f_1 = \omega f_2$ for ω some constant of modulus 1, as required.

McMullen's geometric argument is the following. To see that the energy density is injective, note that a holomorphic map f of the type we are interested in is determined by its values on an open set, so we can always restrict to an open set U on which f is injective. If we now have two maps with the same energy density, composing one with the inverse of the other we get a conformal map between open subsets of S^2 that preserves the area form, and hence is a local isometry. It is well known that such a map must be a rotation.

To see that the energy density is injective, note that f induces a linear map from, on the one hand, the tangent space at f to the space of holomorphic maps $U \to S^2$, to, on the other hand, the space of holomorphic tangent vector fields on $f(U)$. Saying that the derivative of the energy map is zero along some tangent vector $\dot{f}$ is saying that the corresponding vector field preserves the metric of S^2. Such a vector field corresponds to the action of an element of the Lie algebra of the rotation group.

CONCLUSION

We have seen that the information metric on the space of rational maps modulo isometries is nondegenerate. As was discussed in the Introduction, the original reason for considering this problem was to provide a model for the information metric on the space of instantons. This work is still in progress. One partial result is that for the moduli spaces considered in Donaldson's theorem, that have a boundary looking like the four-manifold in question, it can be shown that the information metric in a neighbourhood of this boundary is nondegenerate.

In addition to the applications indicated of the information metric to instantons, one can consider generalisations to the space of harmonic maps between two Riemannian manifolds M and N, modulo the action of isometries of the space N. For definitions in what follows see for instance [Eells and Lemaire 1983]. The energy density of a map $f : M \to N$ is given by $e(f)(x) = |d_x f|^2$, where $d_x f$ is the tangent map from $T_x M$ to $T_{f(x)} N$ and the norm is taken using the metrics on the two tangent spaces. The energy of f is $E(f) = \int_M e(f) \operatorname{vol}_M$. If f_t is a family of maps from M to N, its tangent at $t = 0$ is a section of the pullback bundle $f^{-1}TN$ where $f = f_0$. If X is such a section, it is easy to show that the derivative of the energy density is

$$d_f e(X) = \langle df, f^{-1}\nabla X \rangle,$$

where $f^{-1}\nabla$ is the pullback connection and the inner product is that induced on $T^*M \otimes f^{-1}TN$ by the respective metrics. The proof in Section 4 that shows that the integral defining the information metric converges can now be generalised to this case. Of course in general one needs to consider carefully if there is a manifold of harmonic maps; I am indicating here only that there is a well-defined two-tensor on the space of all maps. Nevertheless this result is useful in two cases.

The first is when M is a Riemann surface and $N = S^2$. It was pointed out to me by a referee that the local result in Section 7 can be applied to show that the information metric is non-degenerate on the space of all holomorphic maps from M to S^2 modulo isometries of S^2.

The second is when $M = S^2$ and $N = \mathbb{CP}_n$. Now the space of interest is rational maps modulo $PU(n+1)$ acting as isometries of $\mathbb{CP}_n$. If we consider such a map locally we get a holomorphic map $f : U \to \mathbb{C}^{n+1}$ and we can proceed as in Section 7. The energy density is now

$$e(f) = \partial\bar{\partial} \log\langle (f, f\rangle,$$

where $\langle\ ,\ \rangle$ is the hermitian inner product on $\mathbb{C}^{n+1}$. We can assume that $f(z_0) = (1, 0, \dots, 0)$ and the same sort of argument shows that $\operatorname{Re}\langle f, \dot{f}\rangle = 0$ for all z. We need to show now that this is only possible if $\dot{f} = Xf$ for some constant skew-hermitian matrix X. This needs some linear algebra and the condition that f is full (that is, does not have its image in a proper subspace).

ACKNOWLEDGEMENTS

The financial support of the Australian Research Council is gratefully acknowledged. I would also like to thank Nick Buchdahl, Alan Carey and Mike Eastwood for many useful conversations.

This paper underwent several changes during the refereeing process. My thanks to the referee who suggested that my original result for maps on S^2 should, in fact, be a "local" result; to Curt McMullen for the argument at the end of Section 7; and to Silvio Levy, the editor, for his assistance.

REFERENCES

[Amari 1985] S. Amari, *Differential Geometrical Methods in Statistics*, Lecture Notes in Stat. **28**, Springer, Heidelberg, 1985.

[Char et al. 1991] Bruce W. Char et al., *Maple V Language Reference Manual*, Springer, New York, 1991.

[Donaldson and Kronheimer 1990] S. K. Donaldson and P. Kronheimer, *The Geometry of Four-Manifolds*, Mathematical Monographs, Oxford University Press, Oxford, 1990.

[Eells and Lemaire 1983] J. Eells and L. Lemaire, *Selected topics in harmonic maps*, Regional Conf. Series in Math. **50**, American Mathematical Society, Providence, RI, 1983.

[Groisser and Parker 1987] D. Groisser and T. H. Parker, "The Riemannian geometry of the Yang–Mills moduli space", *Commun. Math. Phys.* **112** (1987), 663–689.

[Hitchin 1988] N. J. Hitchin, "The geometry and topology of moduli spaces", pp. 1–48 in *Global Geometry and Mathematical Physics*, Lecture Notes in Mathematics **1451**, Springer, Heidelberg, 1988.

[Milnor 1993] J. Milnor, "Geometry and dynamics of quadratic rational maps", *Experimental Mathematics* **2** (1993), 37–82.

[Murray and Rice 1993] M. K. Murray and J. W. Rice, *Statistics and Differential Geometry*, Monographs on Stat. and Appl. Prob. **48**, Chapman and Hall, London, 1993.

[Rao 1945] C. R. Rao, "Information and the accuracy attainable in the estimation of statistical parameters", *Bull. Calcutta Math. Soc.* **37** (1945), 81–91.

Michael K. Murray, Department of Pure Mathematics, The University of Adelaide, Adelaide, SA 5005, Australia (mmurray@maths.adelaide.edu.au)

Received June 8, 1993; accepted in revised form December 8