Mobius Energy of Knots and Unknots
Author(s): Michael H. Freedman, Zheng-Xu He and Zhenghan Wang
Source: *Annals of Mathematics*, Second Series, Vol. 139, No. 1 (Jan., 1994), pp. 1-50
Published by: Mathematics Department, Princeton University
Stable URL: https://www.jstor.org/stable/2946626

Annals of Mathematics, **139** (1994), 1–50

Möbius energy of knots and unknots

By MICHAEL H. FREEDMAN, ZHENG-XU HE AND ZHENGHAN WANG*

Contents

Introduction

The behavior of a charged loop of string is a traditional subject for tea-time speculation. There are many tantalizing questions about extremal configurations and dynamics. Recently O'Hara [O'H1-3] began to give this discussion a foundation by describing several "potential energies" for a $\mathcal{C}^2$ loop in 3-space and reminding topologists of the method physicists use—regularization—to make their defining integrals converge. Among these methods, potential energies and their close relatives, the most interesting seems to be the energy E defined in the following paragraph.

Let $\gamma = \gamma(u)$ be a rectifiable curve in $\mathbb{R}^3$, where u belongs to an interval of $\mathbb{R}$ or the circle $\mathbb{S}^1$. For any pair of points $\gamma(u)$, $\gamma(v)$, denote by $D(\gamma(u), \gamma(v))$ the distance between them on the curve; i.e., the minimum of the lengths of subarcs of γ with one endpoint at $\gamma(u)$ and the other at $\gamma(v)$. We define the

*Research by Michael Freedman is partially supported by N.S.F. Grant DMS 89-01412 and the Minnesota Geometry Center, which is supported by N.S.F. Grant DMS 89-20161. Research by Zheng-Xu He is partially supported by N.S.F. Grant DMS 90-06954. Research by Zhenghan Wang is partially supported by N.S.F. Grant DMS 89-01412.

Part of this work was announced at the NATO Advanced Research Workshop in "Topological Fluid Dynamics" at the Institute of Theoretical Physics, University of California, Santa Barbara, November, 1991 (see [FrH2]), and another part appeared in the Bulletin of the A.M.S.

energy of the curve γ *relative to the point* $\gamma(u)$ to be the following integral:

$$\mathrm{E}(\gamma, \gamma(u)) = \int \left\{ \frac{1}{|\gamma(v) - \gamma(u)|^2} - \frac{1}{D(\gamma(v), \gamma(u))^2} \right\} |\dot\gamma(v)| \, \mathrm{d}v. \tag{0.1}$$

Note that the function $u \mapsto \mathrm{E}(\gamma, \gamma(u)) \in [0, \infty]$ is measurable. The following integral

$$\mathrm{E}(\gamma) = \int \mathrm{E}(\gamma, \gamma(u)) |\dot\gamma(u)| \, \mathrm{d}u \tag{0.2}$$

will be called the *energy* of the curve γ.

By equations (0.1) and (0.2) we obtain

$$\mathrm{E}(\gamma) = \iint \left\{ \frac{1}{|\gamma(v) - \gamma(u)|^2} - \frac{1}{D(\gamma(v), \gamma(u))^2} \right\} |\dot\gamma(u)| \, |\dot\gamma(v)| \, \mathrm{d}u \, \mathrm{d}v. \tag{0.3}$$

The following lemma is immediate.

LEMMA 0.1. (i) $\mathrm{E}(\gamma, \gamma(u))$ *and* $\mathrm{E}(\gamma)$ *do not depend on the parametrization or orientation of the curve.*

(ii) *Let* $\mathcal{T}\colon \mathbb{R}^3 \to \mathbb{R}^3$ *be an affine similarity with a linear expansion equal to* s. *Then* $s\mathrm{E}(T \circ \gamma, T \circ \gamma(u)) = \mathrm{E}(\gamma, \gamma(u))$ *and* $\mathrm{E}(T \circ \gamma) = \mathrm{E}(\gamma)$. □

A fundamental property of $\mathrm{E}(\gamma)$ is a form of the Möbius invariance.

THEOREM 2.1. *Let* γ *be a simple closed curve in* $\mathbb{R}^3$ *and let* T *be a Möbius transformation of* $\mathbb{R}^3 \cup \{\infty\}$. *The following statements hold:*

(i) *If* $T \circ \gamma \subseteq \mathbb{R}^3$, *then* $\mathrm{E}(T \circ \gamma) = \mathrm{E}(\gamma)$.

(ii) *If* $T \circ \gamma$ *passes through* ∞, *then* $\mathrm{E}((T \circ \gamma) \cap \mathbb{R}^3) = \mathrm{E}(\gamma) - 4$.

We see that E is a regularization of $1/r^2$-potential energy[1]; and in the case of closed curves it differs only in normalization from the $1/r^2$-potential energy of O'Hara, $\mathrm{E}_{\mathrm{O'Hara}} = 1/2\mathrm{E} - 2$ (see [O'H1]). Note that there are competing candidates for the exponent equal to -2 in the definition of E. For example, the newtonian potential in $\mathbb{R}^3$ has an exponent equal to -1. When the exponent is strictly larger than -3, finite values are obtained for smooth simple loops. Exponents smaller than or equal to -2 yield energies that blow up as a simple loop γ begins to acquire a double point, thus creating an infinite energy barrier to a change of topology. The exponent -2 is the largest exponent where a divergence is obtained if two distinct strands of γ cross. To appreciate this it is sufficient to consider the contribution to (unregularized)

[1]The usual newtonian potential in $\mathbb{R}^3$ is $1/r$.

energy from a unit-speed arc α of an x-axis arc and a unit-speed arc β of a y-axis near the origin:

$$\iint_{\text{disk}} \frac{1}{(\sqrt{x^2+y^2})^2}\,\mathrm{d}x\,\mathrm{d}y = \iint_{\theta,\rho} \frac{1}{\rho}\,\mathrm{d}\rho\,\mathrm{d}\theta = \infty.$$

Such a barrier would not exist for the newtonian potential in $\mathbb{R}^3$. Similarity and Möbius invariance are, of course, special to the exponent -2.

In fact, if the energy of a curve γ is finite and if u is the arc-length parameter, then $u \mapsto \gamma(u)$ must be a (topologically) tame bi-Lipschitz embedding (via Lemma 1.2 and Theorem 4.1). On the other hand, if γ is a simple closed curve in $\mathbb{R}^3$ whose curvature is uniformly bounded, then the energy of γ is bounded (via Proposition 1.5).

Thus it is natural to address questions in the theory of knots and links[2] in terms of the energy of embedded curves in their isotopy classes. In this direction O'Hara proved that given simultaneous upper bounds on several geometric quantities, namely energy, length, and the L^2 norm of the curvature, only finitely many knot types can occur. We drop the hypotheses on length and the L^2 norm of the curvature to prove the following theorem:

THEOREM 3.3. *Let γ be a simple closed curve in $\mathbb{R}^3$ and let $c([\gamma])$ denote the topological crossing number of the knot type $[\gamma]$ of γ. Then*

$$2\pi c([\gamma]) + 4 \le \mathrm{E}(\gamma).$$

Since the number $K(n)$ of distinct knots of at most n crossings satisfies

$$2^n \le K(n) \le 2 \cdot 24^n$$

(see [S], [T], [W]), the number of knot types with representatives below a given energy threshold can be bounded by an exponential. Precisely we have the following corollary:

COROLLARY 3.5. *The number of (isomorphism classes of) knots that can be represented by curves of $\mathrm{E} \le \mathrm{M}$ is bounded by $2 \cdot (24^{-4/2\pi}) \cdot (24^{1/2\pi})^{\mathrm{M}} \simeq (0.264)(1.658)^{\mathrm{M}}$.*

In our normalization all (round) circles have $\mathrm{E}(\text{circle}) = 4$, and this is the smallest possible value for the energy of closed curves in $\mathbb{R}^3$ (via Corollary 2.2). On the other hand, if a closed curve γ satisfies $\mathrm{E}(\gamma) < 6\pi + 4$, then γ is unknotted (via Corollary 3.4). At this point it is interesting to compare the energy $\mathrm{E}(\gamma)$ with the total curvature $TK(\gamma) = \int |(\gamma'(u)/|\gamma'(u)|)'|\,\mathrm{d}u$. Clearly

[2]In this article, a knot or link means a topologically tame knot or link.

$TK(\text{circle}) = 2\pi$ and, according to Milnor [Mi], $TK(\gamma) \leq 4\pi$ implies that γ is unknotted. However the functional TK is less coercive than E, since unlike Corollary 3.4 there are infinitely many 2-bridge knots, all having representatives with $TK = 4\pi + \epsilon$ for any given $\epsilon > 0$.

Given a knot K, one may seek a loop γ_K of knot type K with minimal energy. In this article the existence of extremal functions will be established for irreducible knots.

THEOREM 4.3. *Let K be an irreducible knot. There exists a simple loop $\gamma_K: \mathbb{S}^1 \to \mathbb{R}^3$ with knot type K such that $\mathrm{E}(\gamma_K) \leq \mathrm{E}(\gamma)$ for any other simple loop $\gamma: \mathbb{S}^1 \to \mathbb{R}^3$ of the same knot type.*

A simple loop γ in $\mathbb{R}^3$ of finite energy is called a *locally extremal loop* if $\mathrm{E}(\gamma) \leq \mathrm{E}(\gamma^*)$ for any simple loop γ^*, which is ambiently isotopic to γ and which is contained in some neighborhood of γ. We conjecture that any locally extremal loop is smooth ($\mathcal{C}^\infty$). By some elementary geometrical argument we will prove the following $\mathcal{C}^{1,1}$-regularity theorem:

THEOREM 5.4. *Let γ be a locally extremal loop in $\mathbb{R}^3$. Then, in arc-length parametrization, $\gamma(s)$ is a $\mathcal{C}^{1,1}$ function.*

The organization of the article is already clear from the table of contents. We start in Section 1 with some elementary properties. In Section 2 we prove the Möbius invariance of the energy. In Section 3 we discuss the crossing number of knots and the average crossing number of curves in 3-space. We show that the average crossing number of a closed curve is bounded by its energy up to some multiple. Using this property, we prove in Section 4 that curves of finite energy are (topologically) tame. We also prove the existence of the extremal curves, which minimizes the energy in the family of loops representing any given irreducible knot. In Section 5 the extremal curves are shown to be in the class $\mathcal{C}^{1,1}$. We derive some variational formulas for the gradient of the energy in Section 6, while Section 7 contains some remarks on the energy of links. Finally in Section 8 we consider the extremal problem when the exponent in the definition of energy is no longer equal to -2.

1. Elementary properties

We will use X to denote either an interval of $\mathbb{R}$ or $\mathbb{S}^1$. Let $\gamma: X \to \mathbb{R}^3$ be a rectifiable curve. This means that γ admits a locally integrable first-order derivative $\dot{\gamma} = \gamma'$. In case of closed curves we will identify $\mathbb{S}^1$ with $\mathbb{R}/\ell\mathbb{Z}$, where $\ell > 0$; and we may regard γ as a periodic function defined on $\mathbb{R}$ with period ℓ.

Moreover, if the arc-length parametrization is used, then ℓ equals the length of the curve and

$$\text{(1.1)} \qquad \mathrm{E}(\gamma) = \int_{-\ell/2}^{\ell/2} \mathrm{d}x \int_{x-\ell/2}^{x+\ell/2} \left\{ \frac{1}{|\gamma(y)-\gamma(x)|^2} - \frac{1}{|y-x|^2} \right\} \mathrm{d}y.$$

Example 1.1. The energy of a circle is 4. To show this let us assume that the radius of the circle is 1. Let $\gamma_0 \colon \mathbb{R}/(2\pi\mathbb{Z}) \to \mathbb{R}^3$ be an arc-length parametrization of the circle. If $|y-x| \leq \pi$, then

$$|\gamma_0(y) - \gamma_0(x)| = 2\sin(|y-x|/2).$$

Thus by equation (1.1) we have

$$\begin{aligned} \mathrm{E}(\gamma_0) &= \int_{-\pi}^{\pi} \mathrm{d}x \int_{-x-\pi}^{x+\pi} \left\{ \frac{1}{\left[2\sin\frac{|y-x|}{2}\right]^2} - \frac{1}{(y-x)^2} \right\} \mathrm{d}y \\ &= 4\pi \int_0^{\pi} \left[\frac{1}{4\sin^2(y/2)} - \frac{1}{y^2} \right] \mathrm{d}y \\ &= 2\pi \int_0^{\pi/2} \left[\frac{1}{\sin^2 u} - \frac{1}{u^2} \right] \mathrm{d}u \\ &= 2\pi \left(-\cot u + \frac{1}{u} \right) \Big|_0^{\pi/2}. \end{aligned}$$

Since $\cot u = 1/u - (1/3)u + \cdots$, we have $\cot u - 1/u \to 0$ as $u \to 0$. It follows that

$$\text{(1.2)} \qquad \mathrm{E}(\gamma_0) = 4.$$

A map $f\colon X \to Y$ between metric spaces is called L-Lipschitz if the distance $(f(u), f(v)) \leq \mathrm{L}$ distance (u, v), for all $u, v \in X$. It is called L-bi-Lipschitz if f is L-Lipschitz and its inverse $f^{-1}\colon f(X) \to X$ exists and is also L-Lipschitz.

LEMMA 1.2. *Let $\gamma(u)$ be a rectifiable curve in $\mathbb{R}^3$ parametrized by the arc length. If $\mathrm{E}(\gamma)$ is finite, then the mapping $\gamma\colon X \to \mathbb{R}^3$ is* L*-bi-Lipschitz with the bi-Lipschitz constant $\mathrm{L} = \mathrm{L}(\gamma)$ depending only on $\mathrm{E}(\gamma)$. Furthermore $\mathrm{L}(\gamma)$ converges to* 1 *when $\mathrm{E}(\gamma)$ tends to* 0.

Proof. The first part of the lemma is due to O'Hara, but for the convenience of the readers we provide a somewhat different approach that proves both statements of the lemma.

Since u is the arc-length parameter, the mapping $u \mapsto \gamma(u)$ is Lipschitz with $\mathrm{L} = 1$. We need to show that there is a constant L depending only on

$E(\gamma)$ such that for any u_1 and u_2, with $u_2 - u_1 = D(\gamma(u_1), \gamma(u_2))$, we have

$$u_2 - u_1 \le L|\gamma(u_2) - \gamma(u_1)|. \tag{1.3}$$

Since the energy is invariant under affine similarities (via part (ii) of Lemma 0.1), we may assume that $D(\gamma(u_1), \gamma(u_2)) = u_2 - u_1 = 4$ and, further, that $u_1 = -2$ and $u_2 = 2$. Then inequality (1.3) reduces to

$$4 \le L|\gamma(2) - \gamma(-2)|. \tag{1.4}$$

Let $t \in (0, 2)$. Then

$$\begin{aligned} E(\gamma) &\ge \int_{-2}^{2} dx \int_{-2}^{2} \left\{ \frac{1}{|\gamma(y) - \gamma(x)|^2} - \frac{1}{|y - x|^2} \right\} dy \\ &\ge \int_{-2}^{-2+t} dx \int_{2-t}^{2} \left\{ \frac{1}{|\gamma(y) - \gamma(x)|^2} - \frac{1}{|y - x|^2} \right\} dy \\ &\ge \int_{-2}^{-2+t} dx \int_{2-t}^{2} \left\{ \frac{1}{(|\gamma(2) - \gamma(-2)| + 2 - y + x + 2)^2} \right. \\ &\qquad \left. - \frac{1}{|2 - t - (-2 + t)|^2} \right\} dy. \end{aligned} \tag{1.5}$$

By letting $t = 1$, we deduce that

$$\begin{aligned} E(\gamma) &\ge \log\left[\frac{(|\gamma(2) - \gamma(-2)| + 1)^2}{|\gamma(2) - \gamma(-2)|(|\gamma(2) - \gamma(-2)| + 2)} \right] - \frac{1}{4} \\ &\ge \log\left[\frac{1}{2|\gamma(2) - \gamma(-2)|} \right] - \frac{1}{4}. \end{aligned}$$

Thus

$$|\gamma(2) - \gamma(-2)| \ge \frac{1}{2e^{1/4}e^{E(\gamma)}}.$$

So inequality (1.4) holds for some $L \le 2e^{1/4}e^{E(\gamma)}$.

Next, by letting $t = (4 - |\gamma(2) - \gamma(-2)|)/8$ in inequality (1.5), we obtain

$$\begin{aligned} E(\gamma) &\ge \int_{-2}^{-2+t} dx \int_{2-t}^{2} \left\{ \frac{1}{(|\gamma(2) - \gamma(-2)| + 2 - y + x + 2)^2} \right. \\ &\qquad \left. - \frac{1}{|2 - t - (-2 + t)|^2} \right\} dy \\ &\ge \int_{-2}^{-2+t} dx \int_{2-t}^{2} \left\{ \frac{1}{(|\gamma(2) - \gamma(-2)| + 2t)^2} - \frac{1}{|4 - 2t|^2} \right\} dy \\ &= \frac{2(4 + |\gamma(2) - \gamma(-2)|)(4 - |\gamma(2) - \gamma(-2)|)^3}{(4 + 3|\gamma(2) - \gamma(-2)|)^2(12 + |\gamma(2) - \gamma(-2)|)^2} \\ &\ge \left(\frac{4 - |\gamma(2) - \gamma(-2)|}{16} \right)^3. \end{aligned}$$

This implies that the bi-Lipschitz constant L goes to 1 as $E(\gamma)$ tends to 0. □

For a rectifiable curve γ, the energy integrand in equation (0.3) is almost everywhere defined. If $E(\gamma)$ is finite, then this integrand is an L^1 function of u and v. So for every $\epsilon' > 0$ there is a δ' such that, for any subarc $\gamma_{\delta'}$ of γ of length δ',

$$E(\gamma_{\delta'}) < \epsilon'. \tag{1.6}$$

Using Lemma 1.2, we obtain the following corollary:

COROLLARY 1.3. *Given γ with $E(\gamma)$ finite and given any $\epsilon > 0$, there exists some $\delta > 0$ so that any subarc γ_δ of length δ is a $(1+\epsilon)$-bi-Lipschitz embedding under the arc-length parametrization.* □

Note that δ of Corollary 1.3 may depend on the curve γ.

Although a bound on the energy implies the bi-Lipschitz property for a curve, it does not imply that the curve is continuously differentiable. The reader may try to find an arc that has an infinitely turning tangent (i.e., many spirals) approaching some interior point, while having arbitrarily small energy. Hint: Fit together infinitely many segments s_i of circles with ϵ/i degrees of arc and rapidly decreasing radius r_i. Since $E(s_i) = O(\epsilon^2/i^2)$ and the off-diagonal contributions to energy are very small when r_i decreases quickly, the total energy can be made $O(\epsilon)$, while the total turning equal to $\epsilon\sum_{i=1}^{\infty} 1/i$ is infinite.

The following lemma is obvious, but useful:

LEMMA 1.4. *Let γ be a rectifiable curve and let γ_k, $k = 1, 2, \ldots$, be subarcs of γ with disjoint interiors. Then*

$$\sum_k E(\gamma_k) \leq E(\gamma). \tag{1.7}$$ □

We will end this section by showing that, for a simple closed curve γ, $E(\gamma)$ is finite if $\ddot{\gamma}$ is only $L^{2+\delta}$-integrable for some $\delta > 0$. In particular, if γ is $\mathcal{C}^{1,1}$, then $E(\gamma) < \infty$. The proof is quite lengthy, but the argument is relevant to Section 6.

PROPOSITION 1.5. *Let γ be a simple closed curve in $\mathbb{R}^3$ whose second derivative is $L^{2+\delta}$-integrable for some $\delta > 0$. Then $E(\gamma)$ is finite.*

Proof. Let u be an arc-length parameter for γ. By assumption, $\ddot{\gamma}(u)$ is $L^{2+\delta}$-integrable. Then $|\dot{\gamma}(u)| = 1$ and

$$\begin{aligned} \gamma(v) - \gamma(u) &= (v-u)\int_0^1 \dot{\gamma}(u + t(v-u))\,dt \\ &= (v-u)\dot{\gamma}(u) + (v-u)^2 \int_0^1 (1-t)\ddot{\gamma}(u+t(v-u))\,dt. \end{aligned} \tag{1.8}$$

It then follows that

$$
\begin{aligned}
&|\gamma(v)-\gamma(u)|^2 \\
&\quad = (v-u)^2\left[1+2(v-u)\int_0^1 (1-t)\big\langle\dot\gamma(u),\ddot\gamma(u+t(v-u))\big\rangle\,\mathrm{d}t\right. \\
&\qquad\left. +(v-u)^2\left|\int_0^1 (1-t)\ddot\gamma(u+t(v-u))\,\mathrm{d}t\right|^2\right],
\end{aligned}
\tag{1.9}
$$

where $\langle\ ,\ \rangle$ denotes the inner product in $\mathbb{R}^3$.

Since $\ddot\gamma(u)\in L^{2+\delta}$, we have by Hölder inequality

$$
\begin{aligned}
\int_0^1 |\ddot\gamma(u+t(v-u))|\,\mathrm{d}t &= \frac{1}{(v-u)}\int_u^v |\ddot\gamma(w)|\,\mathrm{d}w \\
&\le \frac{1}{|v-u|}\|\ddot\gamma\|_{2+\delta}|v-u|^{(1+\delta)/(2+\delta)} \\
&\le \mathrm{C}\frac{|v-u|^\alpha}{|v-u|^{1/2}},
\end{aligned}
\tag{1.10}
$$

where $\|\ddot\gamma\|_{2+\delta}$ is the $L^{2+\delta}$-norm of $|\ddot\gamma(u)|$, $\alpha=(1+\delta)/(2+\delta)-1/2=\delta/(2(2+\delta)>0$, and C denotes some constant independent of u and v.

Combining formulas (1.9) and (1.10), we get

$$
\begin{aligned}
&\frac{1}{|\gamma(v)-\gamma(u)|^2} \\
&\quad = \frac{1}{(v-u)^2}\left[1-2(v-u)\int_0^1 (1-t)\big\langle\dot\gamma(u),\ddot\gamma(u+t(v-u))\big\rangle\,\mathrm{d}t\right. \\
&\qquad\left. +\mathcal{O}\big((v-u)^{1+2\alpha}\big)\right] \\
&\quad = \frac{1}{(v-u)^2}-\frac{2}{(v-u)}\int_0^1 (1-t)\big\langle\dot\gamma(u),\ddot\gamma(u+t(v-u))\big\rangle\,\mathrm{d}t \\
&\qquad +\mathcal{O}\left(\frac{1}{|v-u|^{1-2\alpha}}\right),
\end{aligned}
\tag{1.11}
$$

where $\mathcal{O}(x)$ denotes some quantity with $|\mathcal{O}(x)|\le \mathrm{K}|x|$ for some constant $\mathrm{K}>0$ independent of u and v.

Therefore, for any $\epsilon>0$,

$$
\begin{aligned}
&\iint\limits_{|v-u|\ge\epsilon}\left\{\frac{1}{|\gamma(v)-\gamma(u)|^2}-\frac{1}{D(\gamma(u),\gamma(v))^2}\right\}\mathrm{d}v\,\mathrm{d}u \\
&\quad = \iint\limits_{|v-u|\ge\epsilon}\left\{\frac{1}{|\gamma(v)-\gamma(u)|^2}-\frac{1}{(v-u)^2}\right\}\mathrm{d}u\,\mathrm{d}v
\end{aligned}
\tag{1.12}
$$

$$(1.12)\qquad = -\iint\limits_{|v-u|\geq\epsilon} \left[\frac{2}{v-u}\int_0^1 (1-t)\big\langle\dot\gamma(u),\ddot\gamma(u+t(v-u))\big\rangle\,\mathrm{d}t + \mathcal{O}\Big(\frac{1}{|v-u|^{1-2\alpha}}\Big)\right]\mathrm{d}u\,\mathrm{d}v$$
$$= -\iint\limits_{|w|\geq\epsilon} \frac{2}{w}\left(\int_0^1 (1-t)\big\langle\dot\gamma(u),\ddot\gamma(u+tw)\big\rangle\,\mathrm{d}t\right)\mathrm{d}u\,\mathrm{d}w + \mathrm{C}_\epsilon,$$

where C_ϵ is uniformly bounded as $\epsilon\to 0^+$.

It remains to show that the integral on the right-hand side of equation (1.12) is uniformly bounded. Denote this integral by I_ϵ. Since $|\dot\gamma(u)| = 1$, we have $\langle\dot\gamma(u),\ddot\gamma(u)\rangle = 0$, a.e. Then

$$I_\epsilon = \iiint\limits_{\substack{|w|\geq\epsilon\\ 0\leq t\leq 1}} \frac{2(1-t)}{w}\big\langle\dot\gamma(u),\ddot\gamma(u+tw)\big\rangle\,\mathrm{d}t\,\mathrm{d}u\,\mathrm{d}w$$
$$= \iiint\limits_{\substack{|w|\geq\epsilon\\ 0\leq t\leq 1}} \frac{2(1-t)}{w}\big\langle\dot\gamma(u)-\dot\gamma(u+tw),\ddot\gamma(u+tw)\big\rangle\,\mathrm{d}t\,\mathrm{d}u\,\mathrm{d}w$$
$$= \iint\limits_{\substack{|w|\geq\epsilon\\ 0\leq t\leq 1}} \frac{2(1-t)}{w}\left[\int_u \big\langle\dot\gamma(u)-\dot\gamma(u+tw),\ddot\gamma(u+tw)\big\rangle\,\mathrm{d}u\right]\mathrm{d}t\,\mathrm{d}w.$$

By Hölder inequality

$$|\dot\gamma(u)-\dot\gamma(u+tw)| \leq \int_u^{u+tw} |\ddot\gamma(v)|\,\mathrm{d}v \leq \mathcal{O}\big(|w|^{\frac12+\alpha}\big);$$

hence

$$\left|\int_u \big\langle\dot\gamma(u)-\dot\gamma(u+tw),\ddot\gamma(u+tw)\big\rangle\,\mathrm{d}u\right| \leq \mathcal{O}\big(|w|^{\frac12+\alpha}\big)\int_u |\ddot\gamma(u+tw)|\,\mathrm{d}u$$
$$= \mathcal{O}\big(|w|^{\frac12+\alpha}\big)\|\ddot\gamma\|_1$$
$$= \mathcal{O}\big(|w|^{\frac12+\alpha}\big).$$

It then follows that

$$|I_\epsilon| \leq \int_{|w|\geq\epsilon} \frac{2}{|w|}\mathcal{O}\big(|w|^{\frac12+\alpha}\big)\,\mathrm{d}w$$
$$\leq \int_w \mathcal{O}\big(|w|^{\alpha-\frac12}\big)\,\mathrm{d}w,$$

which is finite. □

2. Möbius invariance of energy

The energy $\mathrm{E}(\gamma)$ is independent of parametrization and is unchanged if γ is changed by a similarity of $\mathbb{R}^3$. In this section we prove a fundamental property of $\mathrm{E}(\gamma)$: invariance under Möbius transformation. In order to state the following theorem more conveniently, we will make the convention that the energy of any nonrectifiable curve is infinite.

THEOREM 2.1. *Let γ be a simple closed curve in $\mathbb{R}^3$ and let T be a Möbius transformation of $\mathbb{R}^3 \cup \{\infty\}$. The following statements hold:*
(i) *If $T \circ \gamma \subseteq \mathbb{R}^3$, then* $\mathrm{E}(T \circ \gamma) = \mathrm{E}(\gamma)$.
(ii) *If $T \circ \gamma$ passes through ∞, then* $\mathrm{E}((T \circ \gamma) \cap \mathbb{R}^3) = \mathrm{E}(\gamma) - 4$.

Recall that the Möbius transformations of the 3-sphere $\mathbb{S}^3 = \mathbb{R}^3 \cup \{\infty\}$ are the 10-dimensional group of angle-preserving diffeomorphisms generated by inversion in 2-spheres. If T is the inversion of $\mathbb{R}^3 \cup \{\infty\}$ in the 2-sphere $\{x \in \mathbb{R}^3 \colon |x-a| = r\}$, where $a \in \mathbb{R}^3$ and $r > 0$, then T is defined by $T(x) = a + (r/|x-a|)^2 (x-a)$.

As an immediate application, we find that the circles have least energy. Actually it was the attempt to prove the extremal properties of the circles that lead to the discovery of Theorem 2.1.

COROLLARY 2.2. *For any loop $\gamma \colon \mathbb{S}^1 \to \mathbb{R}^3$, the energy satisfies $\mathrm{E}(\gamma) \geq 4$, where equality holds if and only if γ is a circle.*

Proof. Assume that γ is simple; otherwise $\mathrm{E}(\gamma) = +\infty$. Let T be a Möbius transformation that maps some point of γ to ∞. Then $\mathrm{E}(\gamma) = 4 + \mathrm{E}((T \circ \gamma) \cap \mathbb{R}^3)$. Clearly $\mathrm{E}((T \circ \gamma) \cap \mathbb{R}^3) \geq 0$ and equality holds if and only if $(T \circ \gamma) \cap \mathbb{R}^3$ is a straight line. □

To prove Theorem 2.1 we need the following lemma:

LEMMA 2.3. *Let $\gamma(u) \colon \mathbb{R}/\ell\mathbb{Z} \to \mathbb{R}^3$ be a closed curve such that the positive real function $|\dot\gamma(u)|$ is Lipschitz in u. Then, for any $\epsilon > 0$,*

$$(2.1) \qquad -\iint\limits_{|u-v|\geq\epsilon} \frac{1}{D(\gamma(u),\gamma(v))^2} |\dot\gamma(u)|\,|\dot\gamma(v)|\,\mathrm{d}u\,\mathrm{d}v = 4 - \frac{2\ell}{\epsilon} + \mathcal{O}(\epsilon).$$

Note that if $\gamma(u)$ is parametrized by arc length, then the error term in equation (2.1) vanishes for small ϵ.

Proof. Let L be the length of γ. We assume that ϵ is small enough. Then the left-hand side of (2.1) is equal to

$$(2.2)\qquad \begin{aligned} &-\int_{u\in\mathbb{R}/\ell\mathbb{Z}}\left[\int_{|v-u|\geq\epsilon}\frac{|\dot\gamma(v)|}{D(\gamma(u),\gamma(v))^2}\,\mathrm{d}v\right]|\dot\gamma(u)|\,\mathrm{d}u\\ &\qquad=\int_{u\in\mathbb{R}/\ell\mathbb{Z}}\left[\frac{4}{L}-\frac{1}{\epsilon_+}-\frac{1}{\epsilon_-}\right]|\dot\gamma(u)|\,\mathrm{d}u,\end{aligned}$$

where

$$\epsilon_+=\epsilon_+(u)=D\big(\gamma(u+\epsilon),\gamma(u)\big)=\int_u^{u+\epsilon}|\dot\gamma(t)|\,\mathrm{d}t$$

and

$$\epsilon_-=\epsilon_-(u)=D\big(\gamma(u),\gamma(u-\epsilon)\big)=\int_{u-\epsilon}^{u}|\dot\gamma(t)|\,\mathrm{d}t.$$

Since $|\dot\gamma(u)|$ is Lipschitz, $(\mathrm{d}/\,\mathrm{d}u)|\dot\gamma(u)|\in L^\infty(\mathbb{R}/\ell\mathbb{Z})$. By an identity of calculus we have

$$\epsilon_+=\epsilon|\dot\gamma(u)|+\epsilon^2\int_0^1(1-t)\left(\frac{\mathrm{d}}{\mathrm{d}u}|\dot\gamma(u+\epsilon t)|\right)\,\mathrm{d}t$$

and

$$\epsilon_-=\epsilon|\dot\gamma(u)|-\epsilon^2\int_0^1(1-t)\left(\frac{\mathrm{d}}{\mathrm{d}u}|\dot\gamma(u-\epsilon t)|\right)\,\mathrm{d}t.$$

It follows that

$$(2.3)\qquad \begin{aligned}\frac{1}{\epsilon_+}&=\frac{1}{|\dot\gamma(u)|\epsilon}\frac{1}{\left[1+\frac{\epsilon}{|\dot\gamma(u)|}\int_0^1(1-t)\left(\frac{\mathrm{d}}{\mathrm{d}u}|\dot\gamma(u+\epsilon t)|\right)\mathrm{d}t\right]}\\ &=\frac{1}{|\dot\gamma(u)|\epsilon}\left[1-\frac{\epsilon}{|\dot\gamma(u)|}\int_0^1(1-t)\left(\frac{\mathrm{d}}{\mathrm{d}u}|\dot\gamma(u+\epsilon t)|\right)\,\mathrm{d}t+\mathcal{O}(\epsilon^2)\right]\\ &=\frac{1}{|\dot\gamma(u)|\epsilon}-\frac{1}{|\dot\gamma(u)|^2}\int_0^1(1-t)\left(\frac{\mathrm{d}}{\mathrm{d}u}|\dot\gamma(u)+\epsilon t)|\right)\,\mathrm{d}t+\mathcal{O}(\epsilon).\end{aligned}$$

Similarly

$$(2.4)\qquad \frac{1}{\epsilon_-}=\frac{1}{|\dot\gamma(u)|\epsilon}+\frac{1}{|\dot\gamma(u)|^2}\int_0^1(1-t)\left(\frac{\mathrm{d}}{\mathrm{d}u}|\dot\gamma(u-\epsilon t)|\right)\,\mathrm{d}t+\mathcal{O}(\epsilon).$$

Then by equation (2.2)

$$(2.5)\qquad \begin{aligned}&-\iint\limits_{|u-v|\geq\epsilon}\frac{1}{D(\gamma(u),\gamma(v))^2}|\dot\gamma(v)|\,|\dot\gamma(u)|\,\mathrm{d}v\,\mathrm{d}u\\ &\qquad=4-\int_{u\in\mathbb{R}/\ell\mathbb{Z}}\left(\frac{1}{\epsilon_+}+\frac{1}{\epsilon_-}\right)|\dot\gamma(u)|\,\mathrm{d}u\end{aligned}$$

$$= 4 - \int_{u\in\mathbb{R}/\ell\mathbb{Z}} \frac{2}{\epsilon}\,\mathrm{d}u + \int_{u\in\mathbb{R}/\ell\mathbb{Z}} \int_0^1 (1-t)\frac{1}{|\dot\gamma(u)|} \left[\frac{\mathrm{d}}{\mathrm{d}u}|\dot\gamma(u+\epsilon t)| - \frac{\mathrm{d}}{\mathrm{d}u}|\dot\gamma(u-\epsilon t)|\right] \mathrm{d}t\,\mathrm{d}u + \mathcal{O}(\epsilon). \tag{2.5}$$

Note that

$$\int_{u\in\mathbb{R}/\ell\mathbb{Z}} \frac{1}{|\dot\gamma(u+\epsilon t)|}\frac{\mathrm{d}}{\mathrm{d}u}|\dot\gamma(u+\epsilon t)|\,\mathrm{d}u = \int_{u\in\mathbb{R}/\ell\mathbb{Z}} \mathrm{d}\big(\log|\dot\gamma(u+\epsilon t)|\big) = 0.$$

Thus equation (2.5) is equal to

$$4 - \frac{2\ell}{\epsilon} + \int_0^1 (1-t) \left[\int_{u\in\mathbb{R}/\ell\mathbb{Z}} \left(\frac{1}{|\dot\gamma(u)|} - \frac{1}{|\dot\gamma(u+\epsilon t)|}\right)\frac{\mathrm{d}}{\mathrm{d}u}|\dot\gamma(u+\epsilon t)|\,\mathrm{d}u - \int_{u\in\mathbb{R}/\ell\mathbb{Z}} \left(\frac{1}{|\dot\gamma(u)|} - \frac{1}{|\dot\gamma(u-\epsilon t)|}\right)\frac{\mathrm{d}}{\mathrm{d}u}|\dot\gamma(u-\epsilon t)|\,\mathrm{d}u\right] \mathrm{d}t.$$

Since

$$\frac{1}{|\dot\gamma(u)|} - \frac{1}{|\dot\gamma(u\pm\epsilon t)|} = \frac{|\dot\gamma(u\pm\epsilon t)| - |\dot\gamma(u)|}{|\dot\gamma(u)|\,|\dot\gamma(u\pm\epsilon t)|} = \mathcal{O}(\epsilon),$$

equation (2.5) equals

$$4 - \frac{2\ell}{\epsilon} + \mathcal{O}(\epsilon) + \mathcal{O}(\epsilon),$$

which proves the lemma. □

For any closed curve $\gamma\colon \mathbb{R}/\ell\mathbb{Z} \to \mathbb{R}^3$ and for each $\epsilon > 0$ define

$$\mathrm{E}_\epsilon(\gamma) = \iint\limits_{|u-v|\geq\epsilon} \frac{1}{|\gamma(u)-\gamma(v)|^2}|\dot\gamma(u)|\,|\dot\gamma(v)|\,\mathrm{d}u\,\mathrm{d}v + 4 - \frac{2\ell}{\epsilon}. \tag{2.6}$$

Then by Lemma 2.3 we have

$$\mathrm{E}_\epsilon(\gamma) = \iint\limits_{|u-v|\geq\epsilon} \left[\frac{1}{|\gamma(u)-\gamma(v)|^2} - \frac{1}{D(\gamma(u),\gamma(v))^2}\right] |\dot\gamma(u)|\,|\dot\gamma(v)|\,\mathrm{d}u\,\mathrm{d}v + \mathcal{O}(\epsilon).$$

We deduce the following:

COROLLARY 2.4. *If $|\dot\gamma(u)|$ is Lipschitz in u, then*

$$E(\gamma) = \lim_{\epsilon\to 0^+} E_\epsilon(\gamma). \tag{2.7}$$

□

Proof of Theorem 2.1. Without loss of generality, we may assume that T is an inversion in a 2-sphere. Let us start with the case where $T(\gamma) \subset \mathbb{R}^3$. Let $u \in \mathbb{R}/\ell\mathbb{Z}$ be the arc-length parameter of γ. Then $|\dot\gamma(u)| = 1$ and $|(T\circ\gamma)'(u)| = |\mathrm{T}'(\gamma(u))|$ are both Lipschitz functions in u, where $\mathrm{T}'(\cdot)$ denotes the differential of T. Hence $\mathrm{E}(\gamma) = \lim_{\epsilon\to 0^+} \mathrm{E}_\epsilon(\gamma)$ and $\mathrm{E}(T\circ\gamma) = \lim_{\epsilon\to 0^+} \mathrm{E}_\epsilon(T\circ\gamma)$.

By a calculation using the law of cosines, we have

$$\frac{|\mathrm{T}'(\gamma(u))|\,|\mathrm{T}'(\gamma(v))|}{|T(\gamma(u)) - T(\gamma(v))|^2} = \frac{1}{|\gamma(u)-\gamma(v)|^2}. \tag{2.8}$$

Integrating this over the region $|u-v| \geq \epsilon$, we obtain $\mathrm{E}_\epsilon(T\circ\gamma) = \mathrm{E}_\epsilon(\gamma)$ (see (2.6)). Hence $\mathrm{E}(T\circ\gamma) = \mathrm{E}(\gamma)$ by Corollary 2.7.

Next let us assume that $T(\gamma)$ passes through ∞. Let $\gamma_1 = T(\gamma)\cap\mathbb{R}^3$. If γ_1 is not rectifiable, then neither is γ, and then $\mathrm{E}(\gamma) = \mathrm{E}(\gamma_1) = \infty$. Thus we may assume that γ_1 is rectifiable. Let $u \in \mathbb{R}$ be the arc-length parameter for γ_1.

Clearly $\gamma - T^{-1}(\infty) = T^{-1}(\gamma_1)$ is a rectifiable curve. It has finite length if and only if γ is rectifiable. If γ is not rectifiable, then its energy $\mathrm{E}(\gamma)$ is infinite, and in this case we need to show that the energy of γ_1 is also infinite. We will delay the proof of this until later. For now let us consider the case when γ is rectifiable.

Let $N \in (100,\infty)$ and let $\epsilon \in [0,.001)$. Define

$$\mathrm{E}_\epsilon^N(\gamma_1) = \iint\limits_{\substack{u\in[-N,N]\\ |u-v|\geq\epsilon}} \left[\frac{1}{|\gamma_1(u)-\gamma_1(v)|^2} - \frac{1}{D(\gamma_1(u),\gamma_1(v))^2}\right] \cdot |\dot\gamma_1(u)|\,|\dot\gamma_1(v)|\,\mathrm{d}u\,\mathrm{d}v. \tag{2.9}$$

Since $D(\gamma_1(u),\gamma_1(v)) = |u-v|$ and $|\dot\gamma_1(u)| = 1$, we have for positive ϵ

$$\begin{aligned} \mathrm{E}_\epsilon^N(\gamma_1) &= \iint\limits_{\substack{u\in[-N,N]\\ |u-v|\geq\epsilon}} \frac{|\dot\gamma_1(u)|\,|\dot\gamma_1(v)|}{|\gamma_1(u)-\gamma_1(v)|^2}\,\mathrm{d}u\,\mathrm{d}v \\ &\quad - \int_{-N}^{N}\left[\int_{|v-u|\geq\epsilon}\frac{1}{(v-u)^2}\,\mathrm{d}v\right]\mathrm{d}u \\ &= \iint\limits_{\substack{u\in[-N,N]\\ |u-v|\geq\epsilon}} \frac{|\dot\gamma_1(u)|\,|\dot\gamma_1(v)|}{|\gamma_1(u)-\gamma_1(v)|^2}\,\mathrm{d}u\,\mathrm{d}v - \frac{4N}{\epsilon}. \end{aligned} \tag{2.10}$$

Now let γ be parametrized in such a way that, for any $u \in [-N-1, N+1] \subset \mathbb{R}/\ell\mathbb{Z}$ (choose $\ell > 2N+2$), we have $\gamma(u) = T^{-1}\circ\gamma_1(u)$. Similarly define $\mathrm{E}_\epsilon^N(\gamma)$ by equation (2.9), where γ_1 is replaced by γ. Let L be the length of γ.

Then by a similar argument to that in Lemma 2.3, we have

$$
\begin{aligned}
\mathrm{E}_\epsilon^N(\gamma) = & \iint\limits_{\substack{u\in[-N,N]\\|u-v|\geq\epsilon}} \frac{|\dot\gamma(u)|\,|\dot\gamma(v)|}{|\gamma(u)-\gamma(v)|^2}\,\mathrm{d}u\,\mathrm{d}v \\
& - \int_{-N}^{N}\left[\int_{|v-u|\geq\epsilon} \frac{|\dot\gamma(v)|}{D(\gamma(u),\gamma(v))^2}\,\mathrm{d}v\right]|\dot\gamma(u)|\,\mathrm{d}u \\
= & \iint\limits_{\substack{u\in[-N,N]\\|u-v|\geq\epsilon}} \frac{|\dot\gamma(u)|\,|\dot\gamma(v)|}{|\gamma(u)-\gamma(v)|^2}\,\mathrm{d}u\,\mathrm{d}v \\
& - \int_{-N}^{N}\left(\frac{1}{\epsilon_+}+\frac{1}{\epsilon_-}\right)|\dot\gamma(u)|\,\mathrm{d}u + \frac{4}{L}\int_{-N}^{N}|\dot\gamma(u)|\,\mathrm{d}u \\
= & \iint\limits_{\substack{u\in[-N,N]\\|u-v|\geq\epsilon}} \frac{|\dot\gamma(u)|\,|\dot\gamma(v)|}{|\gamma(u)-\gamma(v)|^2}\,\mathrm{d}u\,\mathrm{d}v \\
& - \frac{4N}{\epsilon} + \mathcal{O}(\epsilon) + \frac{4}{L}\int_{-N}^{N}|\dot\gamma(u)|\,\mathrm{d}u,
\end{aligned}
\tag{2.11}
$$

where ϵ_+ and ϵ_- are defined as in equation (2.2) and $\mathcal{O}(\epsilon)$ may depend on N.

On the other hand, equation (2.8) implies that

$$
\iint\limits_{\substack{u\in[-N,N]\\|u-v|\geq\epsilon}} \frac{|\dot\gamma_1(u)|\,|\dot\gamma_1(v)|}{|\gamma_1(u)-\gamma_1(v)|^2}\,\mathrm{d}u\,\mathrm{d}v = \iint\limits_{\substack{u\in[-N,N]\\|u-v|\geq\epsilon}} \frac{|\dot\gamma(u)|\,|\dot\gamma(v)|}{|\gamma(u)-\gamma(v)|^2}\,\mathrm{d}u\,\mathrm{d}v.
$$

Combining the last three formulas, we have

$$
\mathrm{E}_\epsilon^N(\gamma_1) - \mathrm{E}_\epsilon^N(\gamma) = -\frac{4}{L}\int_{-N}^{N}|\dot\gamma(u)|\,\mathrm{d}u + \mathcal{O}(\epsilon).
$$

By letting $\epsilon \to 0$, we obtain

$$
\mathrm{E}_0^N(\gamma_1) - \mathrm{E}_0^N(\gamma) = -\frac{4}{L}\,\mathrm{length}\left(T^{-1}\circ\gamma_1([-N,N])\right)
\tag{2.12}
$$

(recall that $\gamma(u) = T^{-1}\gamma_1(u)$ for $u \in [-N-1, N+1]$). By the definition of energy, $\lim_{N\to+\infty}\mathrm{E}_0^N(\gamma_1) = \mathrm{E}(\gamma_1)$ and $\lim_{N\to+\infty}\mathrm{E}_0^N(\gamma) = \mathrm{E}(\gamma)$. It follows that $\mathrm{E}(\gamma_1)$ is finite if and only if $\mathrm{E}(\gamma)$ is finite. In case they are finite, part (ii) of Theorem 2.1 follows from equation (2.12) if we let $N \to +\infty$.

It remains to show that the energy of γ_1 is infinite if γ is not rectifiable. Assuming this, we see that $T^{-1}(\gamma_1) = \gamma - T^{-1}(\infty)$ has infinite length. This means that either $T^{-1}\circ\gamma_1((-\infty,-1))$ or $T^{-1}\circ\gamma_1((1,\infty))$ has infinite length.

Without loss of generality let us assume the former. Then

$$\begin{aligned}
\mathrm{E}(\gamma_1) &\geq \iint\limits_{\substack{u\in(-\infty,-1]\\ v\in[0,1]}} \left[\frac{1}{|\gamma_1(u)-\gamma_1(v)|^2} - \frac{1}{|u-v|^2}\right] \mathrm{d}u\,\mathrm{d}v \\
&= \iint\limits_{\substack{u\in(-\infty,-1]\\ v\in[0,1]}} \frac{\mathrm{d}u\,\mathrm{d}v}{|\gamma_1(u)-\gamma_1(v)|^2} - \log 2 \\
&= \iint\limits_{\substack{u\in(-\infty,-1]\\ v\in[0,1]}} \frac{|(T^{-1}\circ\gamma_1)'(u)|\,|(T^{-1}\circ\gamma_1)'(v)|}{|T^{-1}\circ\gamma_1(u) - T^{-1}\circ\gamma_1(v)|^2}\,\mathrm{d}u\,\mathrm{d}v - \log 2 \\
&\geq \delta \int_{u\in(-\infty,-1]} |(T^{-1}\circ\gamma_1)'(u)|\,\mathrm{d}u - \log 2 \\
&= \infty,
\end{aligned}$$

where $\delta > 0$ is some constant independent of u. The proof of Theorem 2.1 is thus complete. □

3. Energy bounds the average crossing number

In this section we will show that the average crossing number of a closed curve γ is bounded by the energy up to some multiple. As a direct consequence, given any positive constant M, there are only finitely many ambient isotopy classes of embeddings of $\mathbb{S}^1$ that can be represented by a curve of $\mathrm{E} \leq \mathrm{M}$.

Recall that (see [FrH1], pp. 196-197) the *average crossing number* $c(\gamma)$ of a rectifiable curve $\gamma\colon X \to \mathbb{R}^3$ (over itself) is

$$c(\gamma) = c(\gamma,\gamma) = \frac{1}{4\pi}\iint\limits_{X\times X} \frac{|(\dot\gamma(x),\dot\gamma(y),\gamma(y)-\gamma(x))|}{|\gamma(y)-\gamma(x)|^3}\,\mathrm{d}x\,\mathrm{d}y. \tag{3.1}$$

In the numerator of the integrand we see the absolute value of the scalar triple product of three vectors. If the curve γ is simple and has bounded curvature, the integrand is actually bounded, since near the diagonal of $X \times X$ the numerator undergoes a double degeneracy (see inequality (3.9) below).

For any planar curve $\eta\colon X \to \mathbb{R}^2$, the number of self-crossings of η is just 1/2 the cardinality (= a natural number or ∞) of the subset $\{(x,y) \in X \times X; x \neq y, \eta(x) = \eta(y)\}$. Let $\gamma\colon X \to \mathbb{R}^3$ be a rectifiable curve. For any θ in the sphere $\mathbb{S}^2$ of unit vectors in $\mathbb{R}^3$, let $P_\theta\colon \mathbb{R}^3 \to \mathbb{R}^2$ be the orthogonal projection in the direction θ ($P_\theta(\theta) = 0$). Let $n(\gamma;\theta)$ be the number of self-crossings of the planar curve $P_\theta \circ \gamma$. We have the following characterization of the average crossing number:

LEMMA 3.1. *Let $\gamma\colon X \to \mathbb{R}^3$ be a simple rectifiable curve. Then the average crossing number of γ is just the number of self-crossings of its θ-projections averaged over $\theta \in \mathbb{S}^2$:*

$$c(\gamma) = \frac{1}{4\pi} \iint\limits_{\theta\in\mathbb{S}^2} n(\gamma;\theta)\,\mathrm{d}S, \tag{3.2}$$

where $\mathrm{d}S$ *denotes the area form on* $\mathbb{S}^2$.

Proof. Consider the map $F\colon X \times X -$ diagonal $\to \mathbb{S}^2$ defined by

$$F(x,y) = \frac{\gamma(x)-\gamma(y)}{|\gamma(x)-\gamma(y)|}.$$

The Jacobian of this map satisfies,

$$|\det(\mathrm{d}F)| = \frac{|(\dot\gamma(x),\dot\gamma(y),\gamma(x)-\gamma(y))|}{|\gamma(x)-\gamma(y)|^3}, \quad \text{a.e. } (x,y)\in X\times X - \text{diagonal.}$$

(Compare [Ar] and [FrH1].) Hence $c(\gamma)$ is the area of the unsigned images of F (counting multiplicities) divided by 4π (see equation (3.1)). On the other hand, for any θ, the number of self-crossings of the planar curve $P_\theta \circ \gamma$, $n(\gamma;\theta)$, is equal to the cardinality of $F^{-1}(\theta)$. This implies that the area of the unsigned images of f is equal to the integral of $n(\gamma;\theta)$ (cf. [Fe]). Hence we have equation (3.2). □

For a knot K in $\mathbb{R}^3$, the topological notion, *crossing number* of K, is defined to be the minimum of $n(\gamma;\theta)$, where γ is any simple closed curve in the isotopy class of K and θ is any unit vector in 3-space. If γ is a tame embedding of $\mathbb{S}^1$, then the corresponding knot will be denoted by $[\gamma]$. Moreover, if γ is rectifiable, then by Lemma 3.1 the crossing number of the knot $[\gamma]$ is bounded by the average crossing number of γ.

It is convenient to broaden our picture from closed curves to proper rectifiable embeddings of $\mathbb{R}$ in $\mathbb{R}^3$. Such embeddings will also be called *proper rectifiable lines* in $\mathbb{R}^3$.

THEOREM 3.2. *For any proper rectifiable line $\gamma\colon \mathbb{R} \to \mathbb{R}^3$, the average crossing number satisfies*

$$c(\gamma) \le \frac{1}{2\pi}\mathrm{E}(\gamma). \tag{3.3}$$

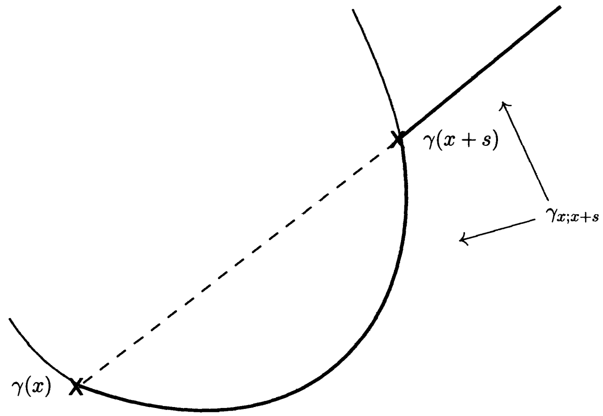

FIGURE 3.1

Proof. Let $\gamma(x)$ be parametrized by arc length. For any fixed $x \in \mathbb{R}$ define an associated auxiliary function $\mathrm{G}_x\colon \mathbb{R} \to \mathbb{R}$ as follows:

$$\begin{aligned}\mathrm{G}_x(s) = &\int_x^{x+s} \left\{ \frac{1}{|\gamma(y)-\gamma(x)|^2} - \frac{1}{(y-x)^2} \right\} \mathrm{d}y \\ &+ \int_{x+s}^{\pm\infty} \left\{ \frac{1}{[|\gamma(x+s)-\gamma(x)| + |y-x-s|]^2} - \frac{1}{(y-x)^2} \right\} \mathrm{d}y \\ = &\int_x^{x+s} \left\{ \frac{1}{|\gamma(y)-\gamma(x)|^2} - \frac{1}{(y-x)^2} \right\} \mathrm{d}y \pm \frac{1}{|\gamma(x+s)-\gamma(x)|} - \frac{1}{s}.\end{aligned}$$

The upper limit of integration for y is $+\infty$ if $s > 0$ and $-\infty$ if $s < 0$. Similarly the $\pm$ sign is $+$ if $s > 0$ and $-$ if $s < 0$. For $s > 0$, $\mathrm{G}_x(s)$ is the energy relative to $\gamma(x)$ of the curve $\gamma_{x;x+s}$, which we obtain from $\gamma\,|_{[x,x+s]}$ by joining a ray in the direction of the vector $\gamma(x+s) - \gamma(x)$ (see equation (0.1) and Figure 3.1). If $s < 0$, then $\mathrm{G}_x(s)$ is the negative of the energy relative to $\gamma(x)$ of a similar curve in the opposite direction. Clearly $\mathrm{G}_x(\pm 0) = 0$ and G_x is continuous as a function in s.

The key to the proof is to imagine the dynamical process, where by a forward (and then a backward), the ray from $\gamma(x)$, which is tangent to γ, is wrapped around γ, giving the family of curved "rays" $\gamma_{x;x+s}$ along the way. We then find an estimate for the integrand of equation (3.1) in terms of quantities such as $\mathrm{dG}_x(s)/\mathrm{d}s$ (see inequality (3.10)). This integrates to give an estimate of $c(\gamma)$ in terms of the regularized energy.

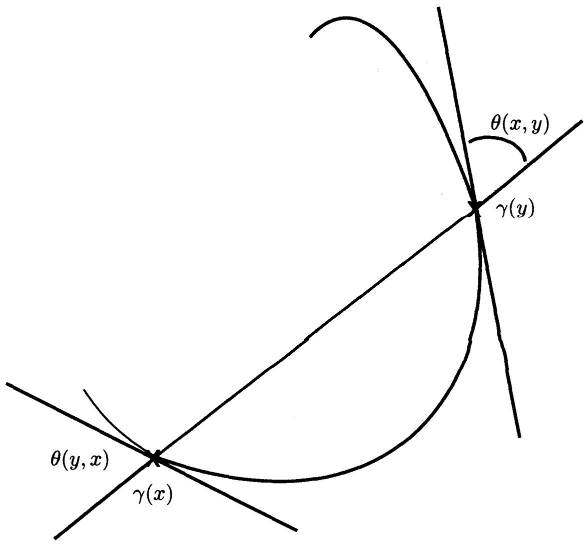

FIGURE 3.2

For a.e. $x \in \mathbb{R}$, G_x is an absolutely continuous function. We have for a.e. x, $s \in \mathbb{R}$,

$$\frac{\mathrm{dG}_x(s)}{\mathrm{d}s} = \frac{1}{|\gamma(x+s)-\gamma(x)|^2}\left[1 - \frac{\partial}{\partial(\pm s)}|\gamma(x+s)-\gamma(x)|\right]. \tag{3.4}$$

For any $x \in \mathbb{R}$ and a.e. $y \in \mathbb{R}$ let us define $\theta(x,y) \in [0,\pi]$ by

$$\theta(x,y) = \begin{cases} \text{angle between } \dot\gamma(y) & \text{and } \frac{\gamma(y)-\gamma(x)}{|\gamma(y)-\gamma(x)|}, \quad \text{if } x \le y, \\ \text{angle between } -\dot\gamma(y) & \text{and } \frac{\gamma(y)-\gamma(x)}{|\gamma(y)-\gamma(x)|}, \quad \text{if } x \ge y, \end{cases} \tag{3.5}$$

(see Figure 3.2).

Then for a.e. $s > 0$,

$$\begin{aligned} \frac{\partial}{\partial s}|\gamma(x+s)-\gamma(x)| &= \frac{(\dot\gamma(x+s), \gamma(x+s)-\gamma(x))}{|\gamma(x+s)-\gamma(x)|} \\ &= \cos\theta(x, x+s), \end{aligned} \tag{3.6}$$

and hence,

$$\begin{aligned} 1 - \frac{\partial}{\partial s}|\gamma(x+s)-\gamma(x)| &= 1 - \cos\theta(x,x+s) \\ &= 2\sin^2\frac{\theta(x,x+s)}{2}. \end{aligned} \tag{3.7}$$

A similar argument shows that the above equality also holds for a.e. $s < 0$. By equations (3.4) and (3.7) we have

$$\frac{\mathrm{dG}_x(s)}{\mathrm{d}s} \geq 2\sin^2\frac{\theta(x,x+s)}{2}\cdot\frac{1}{|\gamma(x+s)-\gamma(x)|^2}, \quad \text{for a.e. } s\in\mathbb{R}. \tag{3.8}$$

On the other hand, for a.e. $x, y \in \mathbb{R}$

$$\begin{aligned}\left|\left(\dot\gamma(x),\dot\gamma(y),\frac{\gamma(y)-\gamma(x)}{|\gamma(y)-\gamma(x)|}\right)\right| &\leq \sin\theta(x,y)\sin\theta(y,x)\\ &\leq 4\sin\frac{\theta(x,y)}{2}\sin\frac{\theta(y,x)}{2}\\ &\leq 2\sin^2\frac{\theta(x,y)}{2}+2\sin^2\frac{\theta(y,x)}{2}\end{aligned} \tag{3.9}$$

(see (3.5)). Combining inequalities (3.8) and (3.9), we obtain

$$\frac{|(\dot\gamma(x),\dot\gamma(y),\gamma(y)-\gamma(x))|}{|\gamma(y)-\gamma(x)|^3} \leq \left\{\frac{\partial \mathrm{G}_x(y-x)}{\partial y}+\frac{\partial \mathrm{G}_y(x-y)}{\partial x}\right\}. \tag{3.10}$$

Therefore

$$\begin{aligned}&\iint \frac{|(\dot\gamma(x),\dot\gamma(y),\gamma(y)-\gamma(x))|}{|\gamma(y)-\gamma(x)|^3}\,\mathrm{d}x\,\mathrm{d}y\\ &\leq \iint\left\{\frac{\partial \mathrm{G}_x(y-x)}{\partial y}+\frac{\partial \mathrm{G}_y(x-y)}{\partial x}\right\}\mathrm{d}x\,\mathrm{d}y\\ &=\int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{+\infty}\frac{\partial \mathrm{G}_x(y-x)}{\partial y}\,\mathrm{d}y\\ &\quad+\int_{-\infty}^{+\infty}\mathrm{d}y\int_{-\infty}^{+\infty}\frac{\partial \mathrm{G}_y(x-y)}{\partial x}\,\mathrm{d}x\\ &=2\int_{-\infty}^{+\infty}(\mathrm{G}_x(\infty)-\mathrm{G}_x(-\infty))\,\mathrm{d}x.\end{aligned} \tag{3.11}$$

Note that, for any $-\infty < t^- < 0 < t^+ < \infty$, we have

$$\begin{aligned}\mathrm{G}_x(t^+)-\mathrm{G}_x(t^-) &\leq \int_{-\infty}^{\infty}\left\{\frac{1}{|\gamma(x+s)-\gamma(x)|^2}-\frac{1}{|s|^2}\right\}\mathrm{d}s\\ &=\mathrm{E}(\gamma,\gamma(x)).\end{aligned}$$

It follows that $\mathrm{G}_x(\infty)-\mathrm{G}_x(-\infty)\leq \mathrm{E}(\gamma,\gamma(x))$. Then formula (3.11) implies that

$$\iint \frac{|(\dot\gamma(x),\dot\gamma(y),\gamma(y)-\gamma(x))|}{|\gamma(y)-\gamma(x)|^3}\,\mathrm{d}x\,\mathrm{d}y \leq 2\int_{-\infty}^{\infty}\mathrm{E}(\gamma,\gamma(x))\,\mathrm{d}x.$$

But the right-hand side in the above formula is equal to $2\mathrm{E}(\gamma)$ by equation (0.2); so inequality (3.3) follows. □

Note that any proper rectifiable line $\gamma_1 : \mathbb{R} \to \mathbb{R}^3$ gives rise to a simple closed curve $\hat{\gamma}_1$ in $\mathbb{S}^3 = \mathbb{R}^3 \cup \{\infty\}$ which passes through the point at infinity. In case $\hat{\gamma}_1$ is tame, its knot type, denoted by $[\hat{\gamma}_1]$ or $[\gamma_1]$, is well defined. It is elementary to show that

$$\text{crossing number}([\gamma_1]) \leq c(\gamma_1). \tag{3.12}$$

THEOREM 3.3. *Let γ be a simple closed curve in $\mathbb{R}^3$ and let $c([\gamma])$ denote the crossing number of the knot type $[\gamma]$ of γ. Then*

$$2\pi c([\gamma]) + 4 \leq \mathrm{E}(\gamma).$$

In Section 4 we will see that any simple closed curve of finite energy is tame and hence defines a (finite) knot type.

Proof. If the knot type $[\gamma]$ is represented by a proper rectifiable line γ_1 in $\mathbb{R}^3$, then by inequality (3.12) and Theorem 3.2

$$c([\gamma]) \leq c(\gamma_1) \leq \frac{1}{2\pi}\mathrm{E}(\gamma_1).$$

According to Theorem 2.1, the energy will increase exactly by 4 if a Möbius transformation is used to move the proper rectifiable line γ_1 off infinity and into a closed curve γ. □

Since an essential knot must have 3 or more crossings, we obtain the following corollary:

COROLLARY 3.4. *Any rectifiable loop with energy less than $6\pi + 4 \simeq$ 22.84954 is unknotted.* □

Computer experiments of [Ah], as reported in [O'H3] and independently by Steve Bryson, yield an essential knot (a trefoil) with energy $\simeq 74$. Following these experiments, D. Kim, R. Kusner and G. Stengle [Ki] analytically solved for the minimum energy of linear (2,3)-torus knots on (circular) tori of revolution in $\mathbb{S}^3$. They obtained $\mathrm{E}_{\min} \approx 74.41204$.

According to Tutte [T], the number of "rooted" planar graphs with n edges is

$$\mathrm{T}_r(n) = \frac{2(2n)!3^n}{n!(n+2)!}. \tag{3.13}$$

Thus the smaller number $\mathrm{T}(n)$ of isomorphism classes of planar embeddings of graphs with n edges satisfies

$$\mathrm{T}(n) \leq 2 \cdot 12^n. \tag{3.14}$$

Given a knot diagram $\mathcal{D}$ with n-crossings, let $\mathcal{D}^+$ be the closed union of complementary regions for which an arc to infinity meets the diagram in an odd number of points. The union $\mathcal{D}^+$ is a thickening of a planar graph G with n edges. Furthermore this correspondence

$$\{\mathcal{D}, n \text{ crossings}\} \mapsto \{G, \text{ with } n \text{ edges}\}$$

is at most 2^n to 1. Thus the number of knot diagrams with exactly n crossings is bounded by $2^n(\mathrm{T}(n)) \leq 2(24^n)$. If a knot type is represented by a diagram with fewer than n crossings, then nugatory crossings may be added to make the number of crossings exactly n. Therefore the number $K(n)$ of knot diagrams with at most n crossings must satisfy

$$2^n \leq K(n) \leq 2(24^n). \tag{3.15}$$

Theorem 3.3 and inequality (3.15) may be combined to yield the following:

COROLLARY 3.5. *The number of (isomorphism classes of) knots that can be represented by curves of* $\mathrm{E} \leq \mathrm{M}$ *is bounded by* $2 \cdot (24^{-4/2\pi}) \cdot (24^{1/2\pi})^{\mathrm{M}} \simeq (0.264)(1.658)^{\mathrm{M}}$. □

The idea of using Tutte's results to obtain an upper bound on the number of knots occurs in a manuscript of Welsh [W]. Sumners, by studying two bridge knots, has shown that the number of distinct knots of n crossings grows at least as fast as 2^n (see [S]).

The following theorem gives a direct estimate for the average crossing number of a simple closed curve in terms of its energy. Such estimates will be useful in proving the tameness of curves of finite energy. The proof is a little more involved than the previous theorem.

THEOREM 3.6. *For any simple rectifiable curve* $\gamma\colon \mathbb{S}^1 \to \mathbb{R}^3$, *the average crossing number satisfies*

$$c(\gamma) \leq \frac{11}{12\pi}\mathrm{E}(\gamma) + \frac{1}{\pi}. \tag{3.16}$$

Proof. By Lemma 0.1 we may assume that the length ℓ of γ is 2. Suppose that $\gamma\colon \mathbb{R}/2\mathbb{Z} \to \mathbb{R}^3$ is parametrized by arc length. Then by equation (1.1)

$$\mathrm{E}(\gamma) = \int_{-1}^{1} dx \int_{x-1}^{x+1} \left\{ \frac{1}{|\gamma(y) - \gamma(x)|^2} - \frac{1}{(y-x)^2} \right\} dy.$$

Let $x \in \mathbb{R}$. Define an associated auxiliary function $\mathrm{G}_x\colon [-1,1] \to \mathbb{R}$ by

$$\begin{aligned} \mathrm{G}_x(s) = & \int_x^{x+s} \left\{ \frac{1}{|\gamma(y)-\gamma(x)|^2} - \frac{1}{(y-x)^2} \right\} dy \\ & + \int_{x+s}^{x\pm 1} \left\{ \frac{1}{[|\gamma(x+s)-\gamma(x)| + |y-x-s|]^2} - \frac{1}{(y-x)^2} \right\} dy. \end{aligned} \tag{3.17}$$

The upper limit of integration for y is $x+1$ if $s \in [0,1]$ and $x-1$ for $s \in [-1,0]$. For $s=0$, either definition gives $\mathrm{G}_x(0)=0$. For $s \in [0,1]$, $\mathrm{G}_x(s)$ is the energy relative to $\gamma(x)$ of the curve $\gamma_{x;x+s}$, which we obtain from $\gamma\,|_{[x,x+s]}$ by joining a straight segment of length $1-s$ in the direction of the vector $\gamma(x+s)-\gamma(x)$ (similar to Figure 3.1). If $s \in [-1,0]$, then $\mathrm{G}_x(s)$ is the negative of the energy relative to $\gamma(x)$ of a similar curve in the opposite direction.

For a.e. $x \in \mathbb{R}$ and a.e. $s \in [0,1)$ we have

$$\begin{aligned} \frac{\mathrm{dG}_x(s)}{\mathrm{d}s} &= \int_{x+s}^{x+1} \frac{\partial}{\partial s}\left\{\frac{1}{[|\gamma(x+s)-\gamma(x)|+y-x-s]^2}\right\} \mathrm{d}y \\ &= \int_{x+s}^{x+1} \frac{(-2)[\frac{\partial}{\partial s}|\gamma(x+s)-\gamma(x)|-1]}{[|\gamma(x+s)-\gamma(x)|+y-x-s]^3}\,\mathrm{d}y \\ &= \left[1-\frac{\partial}{\partial s}|\gamma(x+s)-\gamma(x)|\right] \\ &\quad \cdot \left[\frac{1}{|\gamma(x+s)-\gamma(x)|^2} - \frac{1}{[|\gamma(x+s)-\gamma(x)|+1-s]^2}\right]. \end{aligned} \tag{3.18}$$

For any $x \in \mathbb{R}$ and a.e. $y \in \mathbb{R}$ with $|x-y|<1$, let us define $\theta(x,y) \in [0,\pi]$ as in equation (3.5) (see Figure 3.2). Then, as above, for a.e. $s \in [0,1)$

$$\begin{aligned} 1-\frac{\partial}{\partial s}|\gamma(x+s)-\gamma(x)| &= 1-\cos\theta(x,x+s) \\ &= 2\sin^2\frac{\theta(x,x+s)}{2}. \end{aligned} \tag{3.19}$$

Now we will restrict $s \in [0,1/2]$. Then $1-s \geq 1/2 \geq s \geq |\gamma(x+s)-\gamma(x)|$. Then equations (3.18) and (3.19) yield

$$\frac{\mathrm{dG}_x(s)}{\mathrm{d}s} \geq 2\sin^2\frac{\theta(x,x+s)}{2}\cdot\frac{3}{4}\cdot\frac{1}{|\gamma(x+s)-\gamma(x)|^2}. \tag{3.20}$$

A similar argument shows that the above inequality also holds for a.e. $s \in [-1/2,0]$.

On the other hand, for a.e. $x,y \in \mathbb{R}$ with $|x-y|<1$

$$\left|\left(\dot\gamma(x),\dot\gamma(y),\frac{\gamma(y)-\gamma(x)}{|\gamma(y)-\gamma(x)|}\right)\right| \leq 2\sin^2\frac{\theta(x,y)}{2}+2\sin^2\frac{\theta(y,x)}{2} \tag{3.21}$$

(see inequality (3.9)). Combining inequalities (3.20) and (3.21), we obtain

$$\frac{|(\dot\gamma(x),\dot\gamma(y),\gamma(y)-\gamma(x))|}{|\gamma(y)-\gamma(x)|^3} \leq \frac{4}{3}\left\{\frac{\partial\mathrm{G}_x(y-x)}{\partial y}+\frac{\partial\mathrm{G}_y(x-y)}{\partial x}\right\}, \tag{3.22}$$

whenever $|y-x| \le 1/2$. Therefore

$$\begin{aligned}
&\iint\limits_{\substack{|y-x|\le 1/2\\ (x,y)\in\mathbb{R}/2\mathbb{Z}\times\mathbb{R}/2\mathbb{Z}}} \frac{|(\dot\gamma(x),\dot\gamma(y),\gamma(y)-\gamma(x))|}{|\gamma(y)-\gamma(x)|^3}\,\mathrm{d}x\,\mathrm{d}y \\
&\le \frac{4}{3} \iint\limits_{|y-x|\le 1/2} \left\{ \frac{\partial \mathrm{G}_x(y-x)}{\partial y} + \frac{\partial \mathrm{G}_y(x-y)}{\partial x} \right\} \mathrm{d}x\,\mathrm{d}y \\
&= \frac{4}{3}\int_{-1}^{1} \mathrm{d}x \int_{x-1/2}^{x+1/2} \frac{\partial \mathrm{G}_x(y-x)}{\partial y}\,\mathrm{d}y \\
&\quad + \frac{4}{3}\int_{-1}^{1} \mathrm{d}y \int_{y-1/2}^{y+1/2} \frac{\partial \mathrm{G}_y(x-y)}{\partial x}\,\mathrm{d}x \\
&= \frac{4}{3}\int_{-1}^{1} (\mathrm{G}_x(1/2)-\mathrm{G}_x(-1/2))\,\mathrm{d}x \\
&\quad + \frac{4}{3}\int_{-1}^{1} (\mathrm{G}_y(1/2)-\mathrm{G}_y(-1/2))\,\mathrm{d}y \\
&= \frac{8}{3}\int_{-1}^{1} (\mathrm{G}_x(1/2)-\mathrm{G}_x(-1/2))\,\mathrm{d}x.
\end{aligned} \tag{3.23}$$

By equation (3.18), G_x is nondecreasing on $[0,1]$. Similarly it is nondecreasing on $[-1,0]$. Thus $\mathrm{G}_x(1/2)-\mathrm{G}_x(-1/2) \le \mathrm{G}_x(1)-\mathrm{G}_x(-1)$; and by equations (3.23) (and (3.17)) we have

$$\begin{aligned}
&\iint\limits_{|y-x|\le 1/2} \frac{|(\dot\gamma(x),\dot\gamma(y),\gamma(y)-\gamma(x))|}{|\gamma(y)-\gamma(x)|^3}\,\mathrm{d}x\,\mathrm{d}y \\
&\le \frac{8}{3}\int_{-1}^{1} \{\mathrm{G}_x(1)-\mathrm{G}_x(-1)\}\,\mathrm{d}x = \frac{8}{3}\mathrm{E}(\gamma).
\end{aligned} \tag{3.24}$$

On the other hand,

$$\begin{aligned}
&\iint\limits_{1\ge|y-x|\ge 1/2} \frac{|(\dot\gamma(x),\dot\gamma(y),\gamma(y)-\gamma(x))|}{|\gamma(y)-\gamma(x)|^3}\,\mathrm{d}x\,\mathrm{d}y \\
&\le \iint\limits_{1\ge|y-x|\ge 1/2} \frac{\mathrm{d}x\,\mathrm{d}y}{|\gamma(y)-\gamma(x)|^2} \\
&= \iint\limits_{1\ge|y-x|\ge 1/2} \left\{ \frac{1}{|\gamma(y)-\gamma(x)|^2} - \frac{1}{(y-x)^2} \right\} \mathrm{d}y \\
&\quad + \int_{-1}^{1} \mathrm{d}x \int_{1\ge|y-x|\ge 1/2} \frac{\mathrm{d}y}{(y-x)^2}
\end{aligned} \tag{3.25}$$

$$\leq \iint_{\mathbb{R}/2\mathbb{Z}\times\mathbb{R}/2\mathbb{Z}} \left\{ \frac{1}{|\gamma(y)-\gamma(x)|^2} - \frac{1}{(y-x)^2} \right\} \mathrm{d}x\,\mathrm{d}y \tag{3.25}$$

$$+ \int_{-1}^{1} \left\{ \int_{x-1}^{x-\frac{1}{2}} \frac{\mathrm{d}y}{(y-x)^2} + \int_{x+\frac{1}{2}}^{x+1} \frac{\mathrm{d}y}{(y-x)^2} \right\} \mathrm{d}x$$

$$= \mathrm{E}(\gamma) + 4.$$

Combining (3.24) and (3.25), we obtain

$$\iint_{\mathbb{R}/2\mathbb{Z}\times\mathbb{R}/2\mathbb{Z}} \frac{|(\dot\gamma(x), \dot\gamma(y), \gamma(y)-\gamma(x))|}{|\gamma(y)-\gamma(x)|^3} \mathrm{d}x\,\mathrm{d}y \leq \frac{11}{3}\mathrm{E}(\gamma) + 4, \tag{3.26}$$

which implies inequality (3.16) by equation (3.1). □

It follows from Theorems 3.6 and 2.2 that there is a universal constant $\mathrm{C}_0 > 0$ such that for any simple closed curve γ

$$c(\gamma) \leq \mathrm{C}_0\mathrm{E}(\gamma). \tag{3.27}$$

The best constant for C_0 in inequality (3.27) is not known.

The energy of an arc bounds the average crossing number of a subarc in its interior. Considering the middle fifth of arcs, we obtain the following lemma:

LEMMA 3.7. *Let $\gamma\colon [-5\delta, 5\delta] \to \mathbb{R}^3$ be a unit-speed rectifiable arc and let γ^- be the restriction of γ to $[-\delta, \delta]$. Then*

$$c(\gamma^-) = \frac{1}{4\pi} \iint_{u,v\in[-\delta,\delta]} \frac{|(\dot\gamma(v), \dot\gamma(u), \gamma(v)-\gamma(u))|}{|\gamma(v)-\gamma(u)|^3} \mathrm{d}u\,\mathrm{d}v \leq \frac{2}{3\pi}\mathrm{E}(\gamma).$$

Proof. We may assume that $\delta = 1/4$. For any $x \in [-\delta, \delta] = [-1/4, 1/4]$ and $s \in [-1, 1]$ define $\mathrm{G}_x(s)$ as in the proof of Theorem 3.6 (equation (3.17)). Then the estimate (3.22) holds for a.e. x, $y \in [-1/4, 1/4]$. Carrying out the integration analogous to formulas (3.23) and (3.24) gives the result. □

4. Tameness and minimizers for knots

In Section 2 we found that the circles minimize the energy among all closed curves in $\mathbb{R}^3$. We now take up the issue of minimizing the energy within a knot type. Since rectifiable loops may be topologically "wild", it is interesting that the finiteness of the energy implies that tameness of the curve.

We recall that an embedding γ of X into a 3-manifold M^3 (in our case M is either $\mathbb{R}^3$ or $\mathbb{S}^3$) is *tame* if there is a topological ambient isotopy which deforms the curve into some smooth embedding of X into M^3; or equivalently there exists an extension of the embedding $X = X \times \{0\} \hookrightarrow M^3$ to an embedding $X \times \mathbb{R}^2 \hookrightarrow M^3$ (see [Mo]).

THEOREM 4.1. *Suppose that X is an open interval or $\mathbb{S}^1$. Any curve $\gamma\colon X \to \mathbb{R}^3$ with finite energy is tame.*

Proof. By a classical result of Bing [Bi] it is enough to show that γ is locally tame. That is, for any point on γ there is an open and tame subarc of γ that contains the point. By the remark before Corollary 1.3 there is some $\delta_0 > 0$ such that the energy of any subarc of γ with length $\leq 5\delta_0$ is bounded by some small positive number, say, $\epsilon = 0.02$. Let w be any point on γ. Let $\delta \leq \delta_0$ be a positive number such that γ contains a subarc $\eta_{w;5\delta}$ of length 5δ with the middle point at w. This is possible, since X is either an open interval or $\mathbb{S}^1$. Let $\eta_{w;\delta}$ be the subarc of $\eta_{w;5\delta}$ of length δ which also has a middle point at w. By Lemma 3.7 and our assumption on δ we deduce that the average crossing number of $\eta_{w;\delta}$ is at most $2\epsilon/(3\pi) \leq .01$. Using Lemma 3.1, we see that $n(\eta_{w;\delta};\theta)$ vanishes for some $\theta \in \mathbb{S}^2$. That is, the orthogonal projection P_θ takes the arc $\eta_{w;\delta}$ into a simple arc in the plane. A classical result in complex analysis says that any simple arc in a plane is tame; i.e., the embedding of the arc extends to an embedding of the product arc $\times\, \mathbb{R}$ into $\mathbb{R}^2$. Thickening the product structure in the plane with θ-parallel lines gives a taming of the arc $\eta_{w;\delta}$ in $\mathbb{R}^3$. Hence $\eta_{w;\delta}$ is tame. Since w is arbitrary, we conclude that γ is locally tame and, hence, tame. □

The following compactness property will be quite useful in showing the existence of extremal curves.

LEMMA 4.2. *Let $\gamma_i\colon [-N_i, N_i] \to \mathbb{R}^3$ be a sequence of rectifiable curves of uniformly bounded energy. Assume that the curves are all parametrized by arc length and that $N = \lim_{i\to\infty} N_i$ exists and is positive. If $\gamma_i(0)$ is a bounded sequence of points, then there is a subsequence γ_{i_k} of γ_i, which converges locally uniformly to a rectifiable simple curve $\gamma\colon [-N, N] \to \mathbb{R}^3$. Moreover*

$$\mathrm{E}(\gamma) \leq \liminf_{i_k\to\infty} \mathrm{E}(\gamma_{i_k}).$$

A similar result holds for sequences of closed curves.

Proof. We may assume that $\lim_{i_k\to\infty} \mathrm{E}(\gamma_{i_k})$ exists. By Lemma 1.2 all γ_i are uniformly bi-Lipschitz. On the other hand, $\gamma_i(0)$ are uniformly bounded, so we may use Ascoli's theorem to conclude that some subsequence of γ_i, say, γ_{i_k}, converges locally uniformly to a bi-Lipschitz mapping $\gamma\colon [-N, N] \to \mathbb{R}^3$.

Denote the energy integrand of equation (0.3) by $\mathrm{G}_\gamma(u, v)$. Since

$$D\bigl(\gamma(u), \gamma(v)\bigr) \leq \liminf_{i_k\to\infty} D\bigl(\gamma_{i_k}(u), \gamma_{i_k}(v)\bigr)$$

and

$$|\dot{\gamma}(u)| \le \liminf_{i_k \to \infty} |\dot{\gamma}_{i_k}(u)|,$$

for a.e. $u \neq v$ we have

$$\mathrm{G}_\gamma(u,v) \le \lim_{i_k\to\infty} \inf \mathrm{G}_{\gamma_{i_k}}(u,v).$$

Since $\mathrm{G}\gamma_{i_k}$ is a nonnegative function, by Fatou's lemma

$$\begin{aligned}\mathrm{E}(\gamma) &= \iint \mathrm{G}_\gamma(u,v)\,du\,dv \le \lim_{i_k\to\infty} \inf \iint \mathrm{G}_{\gamma_{i_k}}(u,v)\,du\,dv \\ &= \lim_{i_k\to\infty} \mathrm{E}(\gamma_{i_k}).\end{aligned} \tag{4.1}$$

This proves the lemma in the first case. It is easy to see that the same proof works for sequences of closed curves. □

As a functional defined in the space of all curves, E is nonnegative and (by Lemma 4.2 and its proof) lower semicontinuous in the topology of locally uniform convergence.

Now let K be a nontrivial knot. Let $\gamma_1, \gamma_2, \ldots$ be a sequence of loops in $\mathbb{R}^3$ of knot-type K with energy approaching the infimum. We may assume that each curve has unit length and that all curves lie in some fixed bounded region. Then we may use Lemma 4.2 to obtain a subsequence that converges uniformly to some closed curve γ_∞. However the knot type of the limit curve may not be the same as the curves in the sequence. There is the possibility that representatives of K may "pull tight" and disappear in the limit. To keep the limit γ_∞ in the same knot class we may suitably deform the curves by Möbius transformations. This would work only for irreducible knots. The Möbius transformations can only be used to prevent the degeneration of one component in case the knot is a connected sum of two. The "size" of the two summands cannot be controlled at the same time. For example, if the "pull-tight" phenomena occur near two points that are just a unit spherical distance away, no Möbius transformation of $\mathbb{S}^3$ can have a large derivative (in the spherical norm) near both points.

The main result of this section is the following existence theorem:

THEOREM 4.3. *Let K be an irreducible knot. There exists a simple loop $\gamma_K \colon \mathbb{S}^1 \to \mathbb{R}^3$ with knot-type K such that $\mathrm{E}(\gamma_K) \le \mathrm{E}(\gamma)$ for any other simple loop $\gamma \colon \mathbb{S}^1 \to \mathbb{R}^3$ of the same knot type.*

We lay the groundwork for the proof with the following definition and lemmas:

Definition. Let $B \subset \mathbb{S}^3$ be a closed topological ball; i.e., B is a subset homeomorphic to the unit ball $U = \{w \in \mathbb{R}^3; |w| \le 1\}$. We say that a loop $\gamma \subset$

$\mathbb{R}^3$ *has its knot type captured in the topological ball* B if $(\mathbb{S}^3 - \text{int}(B); \gamma - \text{int}(B))$ is homeomorphic to the unknotted pair (U, J), where $J = \{(u, 0, 0); 0 \leq u \leq 1\}$.

The following property follows from 3-manifold topology:

LEMMA 4.4. *If a nontrivially knotted loop* γ *has its knot type captured in each of the closed topological balls* B_1 *and* B_2, *then* $B_1 \cap B_2$ *is nonempty.*

Proof. By contradiction, assume that $B_1 \cap B_2 = \emptyset$. Then $(B_2; \gamma \cap B_2)$ is a subpair of $(\mathbb{S}^3 - \text{int}(B_1); \gamma - \text{int}(B_1))$. By the assumption on B_1 we deduce that $(B_2; \gamma \cap B_2)$ can be embedded into the unknotted pair (U, J). This means that $(B_2; \gamma \cap B_2)$ is also unknotted (i.e., homeomorphic to (U, J)). Since B_2 captures the knot type of γ, it follows that γ is unknotted, a contradiction. □

LEMMA 4.5. *For any nontrivially knotted tame loop* γ *in* $\mathbb{S}^3$ *there is some* $\epsilon > 0$ *such that no closed topological ball of spherical diameter* $\leq \delta$ *captures the knot type of the loop.*

Proof. Let Q_k, $1 \leq k \leq n$, be a finite number of tame open topological balls whose union contains γ such that, for each k, $\overline{Q_k}$ is some regular tubular neighborhood of a closed subarc of γ. These topological balls exist, since γ is tame and, hence, locally tame. Let $B_k = \mathbb{S}^3 - Q_k$. Then each B_k captures the knot type of γ. Now let B be any closed topological ball that captures the knot type of γ. By Lemma 4.4, B intersects each $B_k = \mathbb{S}^3 - Q_k$. Hence B cannot be contained in Q_k. As $B \cap \gamma \neq \emptyset$ and the open sets Q_k cover γ, we deduce that B cannot have an arbitrarily small spherical diameter. □

In the following, the term *diameter* will mean euclidean diameter.

LEMMA 4.6. *For any manifold* $M > 0$ *there is some* $\delta = \delta(M) > 0$ *such that the following statement holds: If* γ *is a nontrivially knotted loop with* $\mathrm{E}(\gamma) \leq M$, *then there is a Möbius transformation* T *of* $\mathbb{S}^3$, *which takes* γ *to a loop* $\gamma^* = T(\gamma) \subset \mathbb{R}^3$ *so that* $\text{length}(\gamma^*) = 1$ *and no closed topological ball of diameter* $\leq \delta$ *captures the knot type of* γ^*.

Proof. Let p be a point on γ where the tangent exists. Let T_1 be the inversion in a unit sphere centered at p. Then $T_1(p) = \infty$ and, hence, $\hat{\gamma}_1 = T_1 \circ \gamma$ is a closed curve in $\mathbb{S}^3$, which passes through ∞. Clearly the knot type of $\hat{\gamma}_1$ can be captured by a closed topological ball contained in $\mathbb{R}^3$. On the other hand, by Lemma 4.5 there is some $\epsilon_1 > 0$ such that no topological ball of spherical diameter $\leq \epsilon_1$ can capture the knot type of $\hat{\gamma}_1$. It follows that, for some $\epsilon > 0$, no topological ball in $\mathbb{R}^3$ of diameter $\leq \epsilon$ can capture the knot type of $\hat{\gamma}_1$. So by rescaling and translating the curve, we may assume that $\hat{\gamma}_1$ is captured by a closed topological ball B_0, with $B_0 \subset U$ the unit ball centered at the origin, but is not captured by any closed topological ball in $\mathbb{R}^3$ of diameter $\leq 1/2$.

Let $\gamma_1 = \hat{\gamma}_1 - \{\infty\}$. Theorem 2.1 implies that $\mathrm{E}(\gamma_1) = \mathrm{E}(\gamma) - 4 \leq M - 4$. Then for some δ_1 depending only on M, there is a (round) sphere in $\mathbb{R}^3$ disjoint from $\hat{\gamma}_1$ such that its radius is at least δ_1 and the distance from 0 to its center is at most 1. In fact, if no such δ_1 exists, then there would exist a sequence of rectifiable lines whose energy is bounded by $M - 4$, such that the limiting curve would fill up the ball of radius 1 centered at 0. That is certainly a contradiction to Lemma 4.2.

Let S be such a sphere. Let I_S be the inversion on S. We claim that there is some δ_2 depending only on M such that the knot type of the curve $I_S \circ \hat{\gamma}_1$ is not captured in any closed topological ball of diameter $\leq \delta_2$. In fact let B be a closed topological ball that captures the knot type of $I_S \circ \hat{\gamma}_1$. Then, as $(I_S)^{-1} = I_S$, the topological ball $I_S(B)$ captures the knot type of $\hat{\gamma}_1$. It follows that either $\infty \in I_S(B)$ or the diameter of $I_S(B)$ is at least $1/2$. On the other hand, by Lemma 4.4, $I_S(B) \cap B_0 \neq \emptyset$ and hence $I_S(B) \cap U \neq \emptyset$. As the radius of S is bounded from below by δ_1 and the center of S lies in U, conclude that the diameter of $B = I_S(I_S(B))$ is bounded from below by some δ_2.

Clearly $I_S \circ \hat{\gamma}_1$ is contained in the interior of the sphere S. So by Lemma 1.2 its length should be bounded by a constant depending on M. Rescaling this curve if necessary, we may obtain a curve γ^* of unit length that satisfies the conditions of the lemma. $\square$

Proof of Theorem 4.3. As above, let $\gamma_1, \gamma_2, \ldots$ be a sequence of loops in $\mathbb{R}^3$ of knot-type K with energy approaching to the infimum. Since $\lim_{k\to\infty} \mathrm{E}(\gamma_k)$ exists, there is some uniform bound M on $\mathrm{E}(\gamma_k)$. By Lemma 4.6 and Theorem 2.1 we may assume that all loops have unit length and that there is some $\delta > 0$ such that no closed topological ball in $\mathbb{R}^3$ of diameter $\leq \delta$ captures the knot type of γ_k. We may further assume that all curves are contained in some bounded region. Thus by Lemma 4.2, after subtracting a subsequence, we may assume that γ_k converges uniformly to a limit loop $\gamma_K \subset \mathbb{R}^3$ with

$$\mathrm{E}(\gamma_K) \leq \lim_{i\to\infty} \mathrm{E}(\gamma_i) = \text{energy infimum}. \tag{4.2}$$

We will show some (and hence all) γ_k are ambiently isotopic to γ_K. This will imply that the knot type is conserved in the limit and, by (4.2), γ_K is the energy minimizer.

By Theorem 4.1 we know that γ_K is tame. So there is some regular tubular neighborhood $\mathcal{M}$ (homeomorphic to $\mathbb{S}^1 \times \mathbb{R}^2$) of γ_K in $\mathbb{R}^3$. Corollary 1.3 allows the construction of an exceedingly thin, locally straight, closed, regular tubular neighborhood $\mathcal{N} \subset \mathcal{M}$ of γ_K in $\mathbb{R}^3$. The idea is to take an even number of points on γ_K separated by some small $\ell > 0$ in arc length and to fit together thin solid cylinders of length approximately ℓ whose axes are the straight line segments between adjacent points. Corollary 1.3 also shows that the ratio width/length of these cylinders can be taken to approach zero as $\ell \to 0$. The "fitting together" can be done in various ways. Rather than give a detailed construction which the interested reader could provide, we will be content to give the specifications $\mathcal{N}$ should satisfy.

We require $\mathcal{N}$ to be the union of adjacent closed cylinders C_n, $n = 1, 2, \ldots, N = 2j$, so that

(1) $N/2 = j \geq 700M + 2$.

(2) Each C_{even} is an isometric embedding of $D_r^2 \times [0, \ell]$ and each C_{odd} is some 1.01-bi-Lipschitz embedding of $D_r^2 \times [0, \ell]$. Here D_r^2 denotes the closed disk of radius r in the plane.

(3) The ratio r/ℓ is so small that within any cylinder C_n and for each k, the arcs of $\gamma_k \cap C_n$ have total length $\leq 1.1\ell$.

(4) $(700M)(1 + 1.1)\sqrt{\ell^2 + 4r^2} \leq \delta$.

Condition (3) can be achieved since each γ_k has energy bounded by M.

Let us show that $\mathcal{N}$ is a regular tubular neighborhood of γ_K. That is, γ_K is ambiently isotopic to the core curve of $\mathcal{N}$. The fundamental group of $\mathcal{M} - \gamma_K$ may be written as a free product with amalgamation over $\pi_1(\partial \mathcal{N}) \cong \mathbb{Z} \oplus \mathbb{Z}$:

$$\mathbb{Z} \oplus \mathbb{Z} \cong \pi_1(\mathcal{M} - \gamma_K) \cong \pi_1(\mathcal{N} - \gamma_K) \underset{\pi_1(\partial N)}{*} \pi_1(\mathcal{M} - \operatorname{int}(\mathcal{N})). \tag{4.3}$$

It follows that both summands on the right are isomorphic to $\mathbb{Z} \oplus \mathbb{Z}$. In particular $\pi_1(\mathcal{N} - \gamma_K) \cong \mathbb{Z} \oplus \mathbb{Z}$. Then it follows from 3-manifold topology that the pair $(\mathcal{N}, \gamma_K)$ is homeomorphic to the standard pair $(\mathbb{S}^1 \times \mathbb{R}^2, \mathbb{S}^1 \times \{0\})$. Hence $\mathcal{N}$ is a regular tubular neighborhood of γ_K.

By the uniform convergence of $\{\gamma_k\}$, after dropping a finite number of curves γ_k, we may assume that $\gamma_k \subseteq \mathcal{N}$. It remains to show that γ_k is ambiently isotopic to γ_K in $\mathcal{N}$. We fix the index k.

Call a cylinder C_n, n even, *good* if

$$\mathrm{E}\left(\gamma_k \cap \left(\bigcup_{m=n-3}^{n+3} C_m\right)\right) < 0.01, \tag{4.4}$$

and *bad* otherwise. By Lemma 1.4 there are at most $700M$ bad cylinders. On a good cylinder, Lemma 3.6 and conditions (2) and (3) above assure us that the average crossing number satisfies

$$c(\gamma_k \cap C_n) \leq 0.02/(3\pi) \leq 0.005.$$

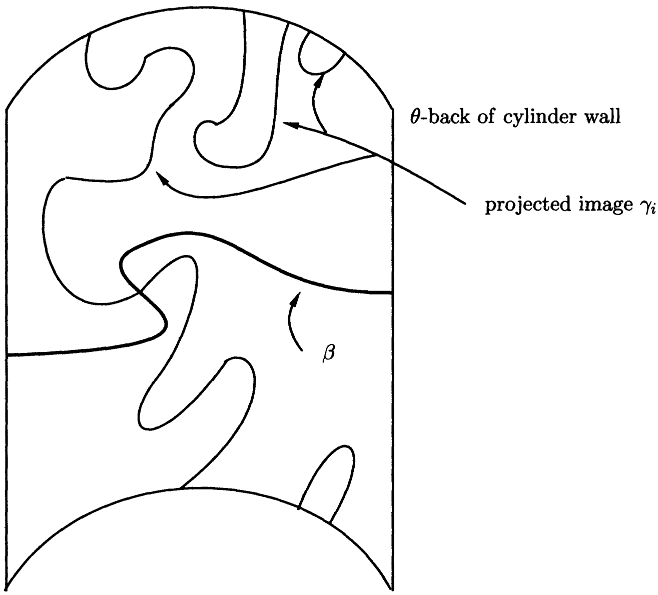

FIGURE 4.1

By Lemma 3.1 it follows that there is some orthogonal projection π_θ with θ almost perpendicular to the axis of $C_n = D_r^2 \times [\ell, \ell]$, which projects the subarc $\gamma_i \cap (D_r^2 \times [0.1\ell, 0.9\ell])$ injectively into the cylinder wall $\partial D_r^2 \times [-\ell, \ell]$ of C_n. In particular, the projection lines to the cylinder wall of, say, $\gamma_k \cap (D_r^2 \times [0.2\ell, 0.8\ell])$, do not meet any other part of γ_k (see condition (3)).

There is exactly one spanning arc $\gamma'_{k,n}$ of γ_k in the subcylinder $D_r^2 \times [0.2\ell,\ 0.8\ell]$ of C_n, where a "spanning arc" means an arc in the subcylinder which joins the two faces $D_r^2 \times \{0.2\ell, 0.8\ell\}$. By condition (3) all other arcs of $\gamma_i \cap (D_r^2 \times [0.2\ell,\ 0.8\ell])$ have length smaller than or equal to 0.1ℓ. Consequently there is an arc β that spans the θ-back wall of C_n and meets the projected image of γ_k transversally in exactly one point. Then $\overline{\beta} = \pi_\theta^{-1}(\beta)$ is an embedded disk separating $C_n \subseteq \mathcal{N}$ and meeting γ_k transversally in one point. (See Figure 4.1.)

Now consider the separating disks $\overline{\beta}_n$ for each good even index n. Since there can be at most $700M$ bad cylinders, by condition (1) there exist at least two separating disks. These disks divide $\widetilde{\mathcal{N}}$ into many small topological balls, which by condition (4) all have diameter $\leq \delta$. The curve γ_k meets each of the small topological balls in an arc, which is unknotted in the topological ball by the irreducibility of $[\gamma_k]$ and the fact that the knot type of γ_k is not captured in a closed topological ball of diameter $\leq \delta$. Thus γ_k is an unknotted core curve in $\mathcal{N}$. This implies that γ_K, also a core curve in $\mathcal{N}$, has the same knot type as γ_k. The proof of Theorem 4.3 is thus complete. $\square$

5. Regularity of extremal loops

Let $\gamma\colon \mathbb{S}^1 \to \mathbb{R}^3$ be a loop of finite energy. We call γ an *extremal loop* if $E(\gamma) \le E(\gamma^*)$ for any other loop γ^* which is ambiently isotopic to γ. If $E(\gamma) \le E(\gamma^*)$ holds for those loops γ^* which are ambiently isotopic to γ and are contained in some small neighborhood of γ, then γ is called a *locally extremal loop.* In this section we will show that any locally extremal loop γ is in $\mathcal{C}^{1,1}$. That is, in arc-length parametrization, the function $s \mapsto \gamma(s)$ is differentiable and the derivative $\gamma'(s)$ is Lipschitz in s.

Let us first introduce an elementary reflection argument. Let H be a (closed) halfspace bounded by a (geometric) plane, or a (closed, round) ball, in $\mathbb{R}^3$. We say that H touches a loop γ at w_1 and w_2, if $w_1, w_2 \in \gamma \cap \partial H$ and γ is disjoint from the interior of H. Suppose that H touches γ at w_1 and w_2; then the two touching points will divide the loop into two (closed) subarcs, say, γ_1 and γ_2. We may assume that the length of γ_1 is less than or equal to the length of γ_2. Let $\tilde{\gamma}_1$ be the image of γ_1 under the inversion in ∂H. Form a new loop $\gamma_2 \cup \tilde{\gamma}_1$ and denote it by γ_H. Clearly, $\gamma_H \subset \mathbb{R}^3$.

If H is a halfspace, then $\tilde{\gamma}_1$ is just the mirror image of γ_1 in the plane ∂H. As the subarcs γ_1 and γ_2 are on the same side of the plane ∂H, the energy integrand in equation (0.3) corresponding to γ_H is pointwise less than or equal to that corresponding to γ. Thus

$$E(\gamma_H) \le E(\gamma), \tag{5.1}$$

where equality holds if and only if either $\gamma_1 \subset \partial H$ or $\gamma_2 \subset \partial H$. Since any ball is Möbius equivalent to a halfspace, and since the energy is invariant under Möbius transformations (via part (1) of Theorem 2.1), the above property also holds in case H is a ball.

For a point b on γ let $\mathbf{W}_b$ be the set of points $w \in \mathbb{R}^3$ such that distance(w, b) = distance(w, γ). In other words, $b \in \mathbf{W}_b$ if and only if it is the center of a ball (of radius $|w - b|$) that touches the curve at b. If γ is differentiable at b, then $\mathbf{W}_b$ is contained in the normal plane to the curve through b. This property can be extended to the case where γ has finite energy, although the curve may not admit a normal plane at b.

LEMMA 5.1. *Let b be a point on a loop γ of finite energy. Then the set $\mathbf{W}_b$ is contained in some plane through b.*

Proof. Suppose that $\mathbf{W}_b$ is not contained in a line through b; otherwise our lemma would be trivially true. Let $q_1, q_2 \in \mathbf{W}_b$ such that q_1, q_2 and b are not colinear. Then the circle obtained as the intersection of spheres $\mathbb{S}^2_{|q_1-b|}(q_1) \cap \mathbb{S}^2_{|q_2-b|}(q_2)$ determines a "tangent direction" through b. Let P be the plane through b normal to this tangent direction. If some $q_3 \in \mathbf{W}_b$ fails to be in this plane, then the complement of the three open balls bounded by $\mathbb{S}^2_{|q_1-b|}(q_1)$, $\mathbb{S}^2_{|q_2-b|}(q_2)$ and $\mathbb{S}^2_{|q_3-b|}(q_3)$, respectively, has some neighborhood for b that is contained in a (round) cone of angle $< \pi$ with vertex at b. Since γ is a closed curve contained in the complement of these three balls, and since the curve passes through b, we deduce that γ contains two subarcs, both of which have an endpoint at b and are contained in the cone. Then it is elementary to show that the energy integrand in equation (0.3) integrated over pairs of points on the two subarcs is infinite (see also Corollary 1.3). Therefore the energy of γ is infinite, contradicting our assumption. Thus all points of $\mathbf{W}_b$ are contained in the plane P. □

A loop γ is in $C^{1,1}$ if and only if there is some constant $r_0 > 0$ such that, for any $b \in \gamma$, the set $\mathbf{W}_b$ contains a (planar) disk with radius r_0 and center at b. This is what we will show for a locally extremal loop. The following lemma is crucial.

LEMMA 5.2. *Let γ be a locally extremal loop. There is $\delta > 0$ so that any ball B of radius $r \leq \delta$ with* $\operatorname{int}(B) \cap \gamma_K = \emptyset$ *meets γ in a connected (possibly empty) set $B \cap \gamma_K$.*

Proof. Suppose that γ minimizes the energy in the ϵ_0-neighborhood of the curve, where $\epsilon_0 > 0$. That is, $\mathrm{E}(\gamma) \leq \mathrm{E}(\gamma^*)$ for any loop γ^* that is ambiently isotopic to γ and that is contained in the ϵ_0-neighborhood of γ.

As γ is a closed curve with finite energy, there is some $\delta > 0$, $2\delta < \epsilon_0$, such that the energy of any subarc of γ of length $\leq 5\delta$ is bounded by 0.0002 (see the remark before Corollary 1.3). Hence, by Lemma 3.7, the average crossing number of any subarc of γ of length $\leq \delta$ is at most $0.0004/(3\pi) < 0.0001$. On the other hand, using Corollary 1.3, we may choose δ so small that, given a pair of points w_1 and w_2 on the curve with $\operatorname{distance}(w_1, w_2) \leq 2\delta$, the shorter component of $\gamma - \{w_1, w_2\}$ is almost straight, i.e., its bi-Lipschitz constant is close to 1, say, $\mathrm{L} \leq 1 + 10^{-10}$.

Let B be a ball of radius $r \leq \delta$ that touches γ at points w_1 and w_2. As above, let γ_1 and γ_2 be the closed subarcs of γ with endpoints at w_1 and w_2, with $\operatorname{length}(\gamma_1) \leq \operatorname{length}(\gamma_2)$; and let $\tilde{\gamma}_1$ be the image of γ_1 under the inversion in the sphere ∂B. We have $\gamma_B = \gamma_2 \cup \tilde{\gamma}_1$. Since $2r \leq 2\delta < \epsilon_0$, and since $\tilde{\gamma}_1$ is contained in the ball B touching γ, γ_B is contained in the ϵ_0-neighborhood of γ. We still need to show that γ_H is isotopic to γ. By inequality (5.1) and the extremal property of the energy of γ, this would imply that $\mathrm{E}(\gamma_H) = \mathrm{E}(\gamma)$; and, therefore, either γ_1 or γ_2 is contained in ∂B. Thus $\gamma \cap \partial B$ is connected.

By our assumption on δ, γ_1 is L-bi-Lipschitz with $\mathrm{L} \leq 1 + 10^{-10}$. Then the spherical angle between any two points in $\gamma_1 \cap \partial B$ (seen from the center of B) must be very small, ≤ 0.01. In particular, it follows that the spherical angle between w_1 and w_2 on ∂B is at most 0.01, thus distance$(w_1, w_2) \leq 0.01r$. As γ_1 is L-bi-Lipschitz, we have length$(\gamma_1) \leq 0.01r\mathrm{L} < 0.02r$.

Without loss of generality, we may assume that B is centered at 0 and that $w_1 = (0, 0, r)$. Let β be the subarc of γ of length r with its middle point at w_1. Clearly γ_1 is contained in the middle 25^{th} of β and, therefore, by the L-bi-Lipschitz property the points on $\gamma - \beta$ are at least $r/3$ away from the points on γ_1 or $\tilde{\gamma}_1$. Let $\beta_B = (\beta - \gamma_1) \cup \tilde{\gamma}_1$. Since $r \leq \delta$, the average crossing number $c(\beta)$ is at most 0.0001. Using estimate (5.1) and Lemma 3.7 again, we may also obtain $c(\beta_B) \leq 0.0001$. It follows by Lemma 3.1 that for some unit vector θ, which is 0.1-close to the vector $(0, 0, 1)$, the orthogonal projection in the direction of θ maps both β and β_B injectively into the plane perpendicular to θ. Since β and β_B are contained in the ball of radius $r/2$ centered at $w_1 = (0, 0, 1)$, each projection line through a point of β or β_B passes through a unique point on the upper half of the sphere ∂B. Using these projection lines, we deduce that there is some simple arc γ_1^* (or $\tilde{\gamma}_1^*$) on the upper half of the sphere ∂B, with endpoints w_1 and w_2, such that β (or β_B) is ambiently isotopic to $(\beta - \gamma_1) \cup \gamma_1^*$ (or $(\beta - \tilde{\gamma}_1) \cup \tilde{\gamma}_1^*$, respectively). Moreover these isotopies can be chosen so that they leave the points in $\gamma - \beta$ fixed and, therefore, give isotopies of γ with $\gamma_2 \cup \gamma_1^*$, and γ_B with $\gamma_2 \cup \tilde{\gamma}_1^*$, respectively. Now it is elementary to see that $\gamma_2 \cup \gamma_1^*$ and $\gamma_2 \cup \tilde{\gamma}_1^*$ are ambiently isotopic. It follows that γ is ambiently isotopic to γ_B, thus completing the proof of the lemma. □

LEMMA 5.3. *Under the assumptions of Lemma* 5.2, *for each b in γ, the set $\mathbf{W}_b$ contains a planar disk of radius δ with center b.*

Proof. Let $\mathcal{N}_\delta$ be the open δ-neighborhood of γ. Let the set $G \colon \mathcal{N}_\delta \to \gamma$ be a multivalued function defined as follows: For any point $q \in \mathcal{N}$, Lemma 5.2 tells us that the points of γ closest to q are a closed arc. We define $G(q)$ to be the set of all points on this arc. Clearly, if $p \in \gamma$ and $q \in \mathcal{N}_\delta$, then $p \in G(q)$ if and only if $q \in \mathbf{W}_p$. Thus $G^{-1}(b)$ is the intersection of $\mathbf{W}_b$ with the open δ-ball centered at b.

By Lemma 5.1, $\mathbf{W}_b$ and hence $G^{-1}(b)$ are contained in a plane P_b. Let D_δ be the open disk in P_b of radius δ and center b. We will show that $D_\delta = G^{-1}(b)$ and, hence, that $\mathbf{W}_b$ contains the closure of D_r (as $\mathbf{W}_b$ is closed).

By contradiction let us assume that $D_\delta \neq G^{-1}(b)$. Then there is some point $w \in D_\delta - G^{-1}(b)$. Let $\epsilon = \text{distance}(w, \partial D_\delta) = \delta - |w - b|$. Let β be the subarc of γ of length ϵ and with its middle point at b. Let w_1 and w_2 be the endpoints of β. Let $i = 1$ or 2. Construct an arc α_i as follows: If w_i is not in P_b, let α_i be the line segment joining w and w_i; and if $w_i \in P_b$, then choose α_i to be an arc joining w and w_i obtained from the line segment from w to w_i by slightly perturbing its interior off P_b. Clearly their union $\alpha = \alpha_1 \cup \alpha_2$ can

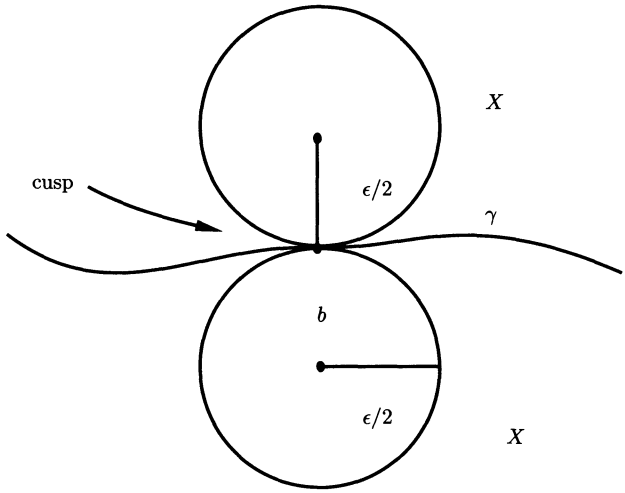

FIGURE 5.1

be kept inside $\mathcal{N}_\delta$. Since $\alpha \cap P_b$ contains at most the three points w, w_1 and w_2 and these three points do not belong to $G^{-1}(b)$, we have $\alpha \cap G^{-1}(b) = \emptyset$. Hence $G(\alpha) \subset \gamma - \{b\}$.

On the other hand, since the image by G of any point is connected (beginning a closed arc or a point), and since the "graph" $\bigcup_q \{q\} \times G(q) \subset \mathcal{N}_\delta \times \gamma$ (where $q \in \mathcal{N}_\delta$) is connected and closed, it is elementary to show that G maps the connected arc α onto a connected subset of γ. The diameter of this image $G(\alpha)$ is at most $\epsilon + 2\delta \leq 3\delta$ and $w_1, w_2 \in G(\alpha)$. Then we deduce that the length of $\gamma - \beta$ is at most 3δ and subsequently the total length of γ is $\leq 3\delta + \epsilon \leq 4\delta$, a contradiction because δ was chosen to be small. □

We are now ready to prove the regularity theorem.

THEOREM 5.4. *Let γ be a locally extremal loop in $\mathbb{R}^3$. Then in arc-length parametrization, $\gamma(s)$ is a $C^{1,1}$ function.*

Proof. Let δ be the constant in Lemma 5.2. Let r be a fixed number with $\delta/2 < r < \delta$. At any point $b \in \gamma$ let C_b be the boundary circle of the planar disk of radius r centered at b, which is contained in $\mathbf{W}_b$ (see Lemma 5.3). Let X_b be the union of all closed balls of radius r centered at points in C_b. Then clearly $\gamma \cap X_b = \{b\}$. Locally γ is constrained to pass through an axially symmetrical quadratic cusp. (See Figure 5.1.)

By Lemma 1.2 the norm of the difference quotients $(\gamma(s_0)-\gamma(s_i))/(s_0-s_i)$ converges to 1 if $s_i \searrow s_0$. If $\gamma(s_0) = b$, then the passage of γ through the cusp at b implies that the spherical angle of the difference quotients converges to some unit vector perpendicular to $\mathbf{W}_b$. Thus $\gamma(s)$ has a derivative $\gamma'(s)$ for any s.

The r-disks $D_{\gamma(t)}$ are mutually disjoint. It follows that $r \cdot$ angle $(\gamma'(s), \gamma'(s+\Delta s)) \leq$ distance$(\gamma(s), \gamma(s+\Delta s)) \leq \Delta s$. Since $\|\gamma'\| = 1$, it follows that $(\gamma'(s+\Delta s) - \gamma'(s))/\Delta s \leq \frac{1}{r}$ for Δs small enough. Hence $v(s)$ is a Lipschitz function with Lipschitz constant $1/r \leq 2/\delta$. □

6. Remarks on the gradient flow

Let $\gamma\colon \mathbb{R}/\ell\mathbb{Z} \to \mathbb{R}^3$ be a simple closed $C^{1,1}$ curve. Then for any $C^{1,1}$-function $h\colon \mathbb{R}/\ell\mathbb{Z} \to \mathbb{R}^3$ consider the curve family $\gamma_t\colon \mathbb{R}/\ell\mathbb{Z} \to \mathbb{R}^3$ defined by

$$\gamma_t(u) = \gamma(u) + th(u), \tag{6.1}$$

where $t \in \mathbb{R}$. For t close to 0, each γ_t is a simple closed $C^{1,1}$-curve and, hence, $\mathrm{E}(\gamma_t)$ is finite (see Proposition 1.5). If $t \mapsto \mathrm{E}(\gamma_t)$ is differentiable at $t = 0$, then the derivative is called the *gradient of* E *at* γ in the direction of h and is denoted by

$$\nabla_h \mathrm{E}(\gamma) = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}(\gamma_t)\Big|_{t=0}. \tag{6.2}$$

LEMMA 6.1. *Let $\gamma\colon \mathbb{R}/\ell\mathbb{Z} \to \mathbb{R}^3$ be a simple closed $C^{1,1}$ curve and let $h\colon \mathbb{R}/\ell\mathbb{Z} \to \mathbb{R}^3$ be any $C^{1,1}$ function. Then $\nabla_h \mathrm{E}(\gamma)$ exists and is given by*

$$\begin{aligned}\nabla_h \mathrm{E}(\gamma) = 2 \iint\limits_{v\in\mathbb{R}/\ell\mathbb{Z}\, u\in\mathbb{R}/\ell\mathbb{Z}} &\left[\frac{\langle \dot\gamma(u), \dot h(u)\rangle}{|\dot\gamma(u)|^2} - \frac{\langle \gamma(v)-\gamma(u), h(v)-h(u)\rangle}{|\gamma(v)-\gamma(u)|^2}\right] \\ &\cdot \frac{|\dot\gamma(v)|\,|\dot\gamma(u)|}{|\gamma(v)-\gamma(u)|^2}\,\mathrm{d}u\,\mathrm{d}v.\end{aligned} \tag{6.3}$$

Remark. The integral in (6.3) is usually not absolutely convergent near the diagonal. Such an integral is defined in general by

$$\iint f(u,v)\,\mathrm{d}u\,\mathrm{d}v = \lim_{\epsilon\to 0^+} \iint\limits_{|u-v|\geq\epsilon} f(u,v)\,\mathrm{d}u\,\mathrm{d}v, \tag{6.4}$$

whereas the limit exists. It is left to the reader to check the convergence of the integral in equation (6.3) in this sense.

Proof. By Lemma 2.3 we have

$$\mathrm{E}_\epsilon(\gamma_t) = \iint\limits_{|u-v|\geq\epsilon} \frac{|\dot\gamma_t(u)|\,|\dot\gamma_t(v)|}{|\gamma_t(u)-\gamma_t(v)|^2}\,\mathrm{d}u\,\mathrm{d}v + 4 - \frac{2\ell}{\epsilon}$$

and

$$\lim_{\epsilon\to 0^+} \mathrm{E}_\epsilon(\gamma_t) = \mathrm{E}(\gamma_t).$$

Then

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}_\epsilon(\gamma_t)\\
&= \iint\limits_{|u-v|\geq\epsilon} \frac{-2\langle\gamma_t(v)-\gamma_t(u), h(v)-h(u)\rangle}{|\gamma_t(v)-\gamma_t(u)|^4}|\dot\gamma_t(u)|\,|\dot\gamma_t(v)|\,\mathrm{d}u\,\mathrm{d}v\\
&\quad+\iint\limits_{|u-v|\geq\epsilon}\Bigg[\frac{|\dot\gamma_t(v)|}{|\gamma_t(v)-\gamma_t(u)|^2}\frac{\langle\dot\gamma_t(u),\dot h(u)\rangle}{|\dot\gamma_t(u)|}\\
&\qquad+\frac{|\dot\gamma_t(u)|}{|\gamma_t(v)-\gamma_t(u)|^2}\frac{\langle\dot\gamma_t(v),\dot h(v)\rangle}{|\dot\gamma_t(v)|}\Bigg]\,\mathrm{d}u\,\mathrm{d}v\\
&= 2\iint\limits_{|u-v|\geq\epsilon}\Bigg[\frac{\langle\dot\gamma_t(u),\dot h(u)\rangle}{|\dot\gamma_t(u)|^2}-\frac{\langle\gamma_t(v)-\gamma_t(u),h(v)-h(u)\rangle}{|\gamma_t(v)-\gamma_t(u)|^2}\Bigg]\\
&\qquad\cdot\frac{|\dot\gamma_t(u)|\,|\dot\gamma_t(v)|}{|\gamma_t(u)-\gamma_t(v)|^2}\,\mathrm{d}u\,\mathrm{d}v.
\end{aligned}\tag{6.5}$$

Since γ_t and h are $\mathcal{C}^{1,1}$ functions, it follows that

$$\gamma_t(v)-\gamma_t(u) = \dot\gamma_t(u)(v-u)+(v-u)^2\int_0^1(1-\lambda)\ddot\gamma_t(u+\lambda(v-u))\,\mathrm{d}\lambda,$$

$$h(v)-h(u) = \dot h(u)(v-u)+(v-u)^2\int_0^1(1-\lambda)\ddot h(u+\lambda(v-u))\,\mathrm{d}\lambda.$$

Then

$$\begin{aligned}
&\big\langle\gamma_t(v)-\gamma_t(u),h(v)-h(u)\big\rangle\\
&\quad=\big\langle\dot\gamma_t(u),\dot h(u)\big\rangle(v-u)^2\\
&\qquad+(v-u)^3\int_0^1(1-\lambda)\big\langle\ddot\gamma_t(u+\gamma(v-u)),\dot h(u)\big\rangle\,\mathrm{d}\lambda\\
&\qquad+(v-u)^3\int_0^1(1-\lambda)\big\langle\dot\gamma_t(u),\ddot h(u+\lambda(v-u))\big\rangle\,\mathrm{d}\lambda\\
&\qquad+\mathcal{O}((v-u)^4).
\end{aligned}$$

A similar expression holds for $|\gamma_t(v)-\gamma_t(u)|^2 = \langle\gamma_t(v)-\gamma_t(u),\gamma_t(v)-\gamma_t(u)\rangle$. It follows that

$$\begin{aligned}
&\frac{\big\langle\gamma_t(v)-\gamma_t(u),h(v)-h(u)\big\rangle}{|\gamma_t(v)-\gamma_t(u)|^2}\\
&\quad=\Bigg\{\big\langle\dot\gamma_t(u),\dot h(u)\big\rangle+(v-u)\int_0^1(1-\lambda)\Big[\big\langle\ddot\gamma_t(u+\lambda(v-u)),\dot h(u)\big\rangle
\end{aligned}$$

$$+ \left\langle \dot\gamma_t(u), \ddot h(u+\lambda(v-u)) \right\rangle \Big] \, d\lambda + \mathcal{O}((v-u)^2) \Big\} \cdot \frac{1}{|\dot\gamma_t(u)|^2}$$

$$\cdot \Big\{ 1 - \frac{2(v-u)}{|\dot\gamma_t(u)|^2} \int_0^1 (1-\lambda) \left\langle \ddot\gamma_t(u+\lambda(v-u)), \dot\gamma_t(u) \right\rangle d\lambda$$

$$+ \mathcal{O}((v-u)^2) \Big\}$$

$$= \frac{\langle \dot\gamma_t(u), \dot h(u) \rangle}{|\dot\gamma_t(u)|^2} + \frac{(v-u)}{|\dot\gamma_t(u)|^2}$$

$$\cdot \int_0^1 (1-\lambda) \Big[\left\langle \ddot\gamma_t(u+\lambda(v-u)), \dot h(u) \right\rangle + \left\langle \dot\gamma_t(u), \ddot h(u+\lambda(v-u)) \right\rangle$$

$$- 2 \langle \dot\gamma_t(u), \dot h(u) \rangle \left\langle \ddot\gamma_t(u+\lambda(v-u)), \frac{\dot\gamma_t(u)}{|\dot\gamma_t(u)|^2} \right\rangle \Big] \, d\lambda$$

$$+ \mathcal{O}((v-u)^2).$$

Hence

$$\frac{\langle \dot\gamma_t(u), \dot h(u) \rangle}{|\dot\gamma_t(u)|^2} - \frac{\langle \gamma_t(v) - \gamma_t(u), h(v) - h(u) \rangle}{|\gamma_t(v) - \gamma_t(u)|^2}$$

$$= (v-u) \int_0^1 (1-\lambda)$$

$$\cdot \Big[- \left\langle \ddot\gamma_t(u+\lambda(v-u)), \frac{\dot h(u)}{|\dot\gamma_t(u)|^2} \right\rangle - \left\langle \dot\gamma_t(u), \frac{\ddot h(u+\lambda(v-u))}{|\dot\gamma_t(u)|^2} \right\rangle \tag{6.6}$$

$$+ 2 \langle \dot\gamma_t(u), \dot h(u) \rangle \left\langle \ddot\gamma_t(u+\lambda(v-u)), \frac{\dot\gamma_t(u)}{|\dot\gamma_t(u)|^4} \right\rangle \Big]$$

$$\cdot \, d\lambda + \mathcal{O}((v-u)^2).$$

Similarly

$$\frac{|\dot\gamma_t(u)| \, |\dot\gamma_t(v)|}{|\gamma_t(u) - \gamma_t(v)|^2} = \frac{1}{(u-v)^2} + \mathcal{O}\left(\frac{1}{|u-v|} \right). \tag{6.7}$$

Then by equations (6.6) and (6.7)

$$\left[\frac{\langle \dot\gamma_t(u), \dot h(u) \rangle}{|\dot\gamma_t(u)|^2} - \frac{\langle \gamma_t(v) - \gamma_t(u), h(v) - h(u) \rangle}{|\gamma_t(v) - \gamma_t(u)|^2} \right] \cdot \frac{|\dot\gamma_t(u)| \, |\dot\gamma_t(v)|}{|\gamma_t(v) - \gamma_t(u)|^2}$$

$$= \frac{1}{(v-u)} \int_0^1 (1-\lambda)$$

$$\cdot \Big[- \left\langle \ddot\gamma_t(u+\lambda(v-u)), \frac{h(u)}{|\dot\gamma_t(u)|^2} \right\rangle - \left\langle \ddot h(u+\lambda(v-u)), \frac{\dot\gamma_t(u)}{|\dot\gamma_t(u)|^2} \right\rangle \tag{6.8}$$

$$+ 2\langle\dot\gamma_t(u), \dot h(u)\rangle \left\langle \ddot\gamma_t(u+\lambda(v-u)), \frac{\dot\gamma_t(u)}{|\dot\gamma_t(u)|^4}\right\rangle\Bigg] \mathrm{d}\lambda + \mathcal{O}(1). \tag{6.8}$$

Therefore

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{E}_\epsilon(\gamma_t)\\
&= 2 \iint_{|v-u|\geq\epsilon} \frac{1}{(v-u)} \int_0^1 (1-\lambda)\\
&\quad\cdot\Bigg[-\left\langle \ddot\gamma_t(u+\lambda(v-u)), \frac{h(u)}{|\dot\gamma_t(u)|^2}\right\rangle - \left\langle \ddot h(u+\lambda(v-u)), \frac{\dot\gamma_t(u)}{|\dot\gamma_t(u)|^2}\right\rangle\\
&\quad + 2\langle\dot\gamma_t(u), \dot h(u)\rangle \left\langle \ddot\gamma_t(u+\lambda(v-u)), \frac{\dot\gamma_t(u)}{|\dot\gamma_t(u)|^4}\right\rangle\Bigg] \mathrm{d}\lambda\, \mathrm{d}v\, \mathrm{d}u + \mathcal{O}(1)\\
&= 2 \int_{u\epsilon\mathbb{R}/\ell\mathbb{Z}} \int_{\substack{w\epsilon[-\ell/2,\ell/2]\\ |w|\geq\epsilon}} \int_{\lambda=0}^{\lambda=1} \frac{(1-\lambda)}{w}\\
&\quad\cdot\Bigg[-\left\langle \ddot\gamma_t(u+\lambda w), \frac{h(u)}{|\dot\gamma_t(u)|^2}\right\rangle - \left\langle \ddot h(u+\lambda w), \frac{\dot\gamma_t(u)}{|\dot\gamma_t(u)|^2}\right\rangle\\
&\quad + 2\langle\dot\gamma_t(u), \dot h(u)\rangle \left\langle \ddot\gamma_t(u+\lambda w), \frac{\dot\gamma_t(u)}{|\dot\gamma_t(u)|^4}\right\rangle\Bigg] \mathrm{d}\lambda\, \mathrm{d}w\, \mathrm{d}u + \mathcal{O}(1)
\end{aligned} \tag{6.9}$$

$$\begin{aligned}
&= 2 \iiint_{x\epsilon\mathbb{R}/\ell\mathbb{Z}, w\epsilon[-\ell/2,\ell/2], |w|\geq\epsilon, \lambda\epsilon[0,1]} \frac{(1-\lambda)}{w}\\
&\quad\cdot\Bigg[-\left\langle \ddot\gamma_t(x), \frac{\dot h(x-\lambda w)}{|\dot\gamma_t(x-\lambda w)|^2}\right\rangle - \left\langle \ddot h(x), \frac{\dot\gamma_t(x-\lambda w)}{|\dot\gamma_t(x-\lambda w)|^2}\right\rangle\\
&\quad + 2\langle\dot\gamma_t(x-\lambda w), \dot h(x-\lambda w)\rangle \left\langle \ddot\gamma_t(x), \frac{\dot\gamma_t(x-\lambda w)}{|\dot\gamma_t(x-\lambda w)|^4}\right\rangle\Bigg]\\
&\quad\cdot \mathrm{d}\lambda\, \mathrm{d}w\, \mathrm{d}x + \mathcal{O}(1).
\end{aligned}$$

Since $\dot h$ and $\dot\gamma_t$ are Lipschitz, we have

$$\begin{aligned}
\frac{\dot h(x-\lambda w)}{|\dot\gamma_t(x-\lambda w)|^2} &= \frac{\dot h(x)}{|\dot\gamma_t(x)|^2} + \mathcal{O}(w),\\
\frac{\dot\gamma_t(x-\lambda w)}{|\dot\gamma_t(x-\lambda w)|^2} &= \frac{\dot\gamma_t(x)}{|\dot\gamma_t(x)|^2} + \mathcal{O}(w),\\
\frac{\dot\gamma_t(x-\lambda w)}{|\dot\gamma_t(x-\lambda w)|^4} &= \frac{\dot\gamma_t(x)}{|\dot\gamma_t(x)|^4} + \mathcal{O}(w),
\end{aligned}$$

and

$$\langle\dot\gamma_t(x-\lambda w), \dot h(x-\lambda w)\rangle = \langle\dot\gamma_t(x), \dot h(x)\rangle + \mathcal{O}(w).$$

Then by equation (6.9)

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}_\epsilon(\gamma_t)\\
&\quad = 2 \iiint\limits_{w\epsilon[-\ell/2,\ell/2],|w|\geq\epsilon} \frac{(1-\lambda)}{w}\\
&\qquad \cdot\left[-\left\langle \ddot{\gamma}_t(x), \frac{\dot{h}(x)}{|\dot{\gamma}_t(x)|^2}\right\rangle - \left\langle \ddot{h}(x), \frac{\dot{\gamma}_t(x)}{|\dot{\gamma}_t(x)|^2}\right\rangle\right.\\
&\qquad \left. + 2\langle\dot{\gamma}_t(x), \dot{h}(x)\rangle \left\langle \ddot{\gamma}_t(x), \frac{\dot{\gamma}_t(x)}{|\dot{\gamma}_t(x)|^4}\right\rangle\right] \mathrm{d}\lambda\, \mathrm{d}w\, \mathrm{d}x + \mathcal{O}(1) + \mathcal{O}(1)\\
&\quad = -2 \iiint\limits_{w\epsilon[-\ell/2,\ell/2],|w|\geq\epsilon} (1-\lambda)\frac{1}{w}\frac{\mathrm{d}}{dx}\frac{\langle\dot{\gamma}_t(x), \dot{h}(x)\rangle}{|\dot{\gamma}_t(x)|^2}\, \mathrm{d}\lambda\, \mathrm{d}w\, \mathrm{d}x + \mathcal{O}(1).
\end{aligned}$$

The triple integral on the right-hand side is 0. It follows that $\mathrm{d}/(\mathrm{d}t)\mathrm{E}_\epsilon(\gamma_t)$ is uniformly bounded and converges uniformly to the integral in equation (6.3) as $\epsilon \to 0^+$. □

The next lemma follows from the proof of Lemma 6.1.

LEMMA 6.2. *Let γ be a simple loop of class $\mathcal{C}^{1,1}$. As a functional in h, $\nabla_h\mathrm{E}(\gamma)$ is linear and bounded in $\mathcal{C}^{1,1}(\mathbb{S}^1;\mathbb{R}^3)$.* □

As a consequence of Lemma 6.2, we obtain the following:

COROLLARY 6.3. *Let γ_k be a sequence of simple loops that converge to γ in $\mathcal{C}^{1,1}$. Then $\mathrm{E}(\gamma) = \lim_{k\to\infty}\mathrm{E}(\gamma_k)$.* □

For a simple closed $\mathcal{C}^{1,1}$ curve $\gamma(u)$ let $\mathbb{P}_{\gamma(u)}\colon \mathbb{R}^3 \to \mathbb{R}^3$ be the orthogonal projection of $\mathbb{R}^3$ onto the normal vector plane to γ at $\gamma(u)$. Then

$$\mathbb{P}_{\gamma(u)}(w) = w - \frac{\langle w, \dot{\gamma}(u)\rangle\dot{\gamma}(u)}{|\dot{\gamma}(u)|^2}. \tag{6.10}$$

Let

$$\mathbf{N}_{\gamma(u)} = \frac{1}{|\dot{\gamma}(u)|}\frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{\dot{\gamma}(u)}{|\dot{\gamma}(u)|}\right). \tag{6.11}$$

Note that $\mathbf{N}_{\gamma(u)}$ is a.e. well defined and parallel to the principal normal if defined.

Assume that $\gamma\colon \mathbb{R}/\ell\mathbb{Z} \to \mathbb{R}^3$ is a simple closed $\mathcal{C}^{3,\alpha}$ curve for some $\alpha > 0$. Define $\mathrm{G}_\gamma\colon \mathbb{R}/\ell\mathbb{Z} \to \mathbb{R}^3$ by

$$\mathrm{G}_\gamma(u) = 2\int_{\mathbb{R}/\ell\mathbb{Z}}\left[\frac{2\mathbb{P}_{\gamma(u)}(\gamma(v)-\gamma(u))}{|\gamma(v)-\gamma(u)|^2} - \mathbf{N}_{\gamma(u)}\right]\frac{|\dot{\gamma}(v)|}{|\gamma(v)-\gamma(u)|^2}\,\mathrm{d}v, \tag{6.12}$$

where the right-hand-side integral is well defined, in the sense of equation (6.6), and $|\mathrm{G}_\gamma(u)|$ is uniformly bounded. Note that $\mathrm{G}_\gamma(u)$ is the L^2 gradient of E at γ.

LEMMA 6.4. *Let $\gamma\colon \mathbb{R}/\ell\mathbb{Z} \to \mathbb{R}^3$ be a simple closed $\mathcal{C}^{3,\alpha}$ curve and let $h\colon \mathbb{R}/\ell\mathbb{Z} \to \mathbb{R}^3$ be a $\mathcal{C}^{2,\alpha}$ function such that $\langle h(u), \dot\gamma(u)\rangle = 0$ for $u \in \mathbb{R}/\ell\mathbb{Z}$. Then*

$$\nabla_h \mathrm{E}(\gamma) = \langle \mathrm{G}_\gamma, h\rangle = \int_{\mathbb{R}/\ell\mathbb{Z}} \langle \mathrm{G}_\gamma(u), h(u)\rangle |\dot\gamma(u)|\,\mathrm{d}u. \tag{6.13}$$

Proof. By equation (6.5) we know that

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}_\epsilon(\gamma_t)|_{t=0}\\
&= 2\iint\limits_{|u-v|\geq\epsilon} \frac{\langle\dot\gamma(u), \dot h(u)\rangle}{|\dot\gamma(u)|^2}\frac{|\dot\gamma(u)|\,|\dot\gamma(v)|}{|\gamma(u)-\gamma(v)|^2}\\
&\quad - 2\iint\limits_{|u-v|\geq\epsilon} \frac{\langle\gamma(v)-\gamma(u), h(v)-h(u)\rangle}{|\gamma(v)-\gamma(u)|^4}|\dot\gamma(v)|\,|\dot\gamma(u)|\,\mathrm{d}u\,\mathrm{d}v\\
&= 2\int_v \left[\int_{|u-v|\geq\epsilon}\left\langle \frac{\dot\gamma(u)}{|\dot\gamma(u)|\,|\gamma(u)-\gamma(v)|^2}, \dot h(u)\right\rangle \mathrm{d}u\right]|\dot\gamma(v)|\,\mathrm{d}v\\
&\quad + 4\iint\limits_{|u-v|\geq\epsilon}\frac{\langle\gamma(v)-\gamma(u), h(u)\rangle}{|\gamma(v)-\gamma(u)|^4}|\dot\gamma(u)|\,|\dot\gamma(v)|\,\mathrm{d}u\,\mathrm{d}v.
\end{aligned}$$

Using integration by parts for the first integral on the right-hand side above, we obtain

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}_\epsilon(\gamma_t)\,|_{t=0}\\
&= -2\int_v |\dot\gamma(v)|\,\mathrm{d}v\int_{|u-v|\geq\epsilon}\left\langle\frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{\dot\gamma(u)}{|\dot\gamma(u)|}\right)\frac{1}{|\gamma(v)-\gamma(u)|^2}, h(u)\right\rangle\mathrm{d}u\\
&\quad + \int_u\left[\int_{|v-u|\geq\epsilon}\left\langle\frac{4(\gamma(v)-\gamma(u))}{|\gamma(v)-\gamma(u)|^4}, h(u)\right\rangle|\dot\gamma(v)|\,\mathrm{d}v\right]|\dot\gamma(u)|\,\mathrm{d}u\\
&= \int_u\int_{|v-u|\geq\epsilon}\left\langle\frac{-2\mathbf{N}_{\gamma(u)}}{|\gamma(v)-\gamma(u)|^2}h(u)\right\rangle|\dot\gamma(v)|\,\mathrm{d}v|\dot\gamma(u)|\,\mathrm{d}u\\
&\quad + \int_u\left\langle\int_{|v-u|\geq\epsilon}\frac{4\mathbb{P}_{\gamma(u)}(\gamma(v)-\gamma(u))}{|\gamma(v)-\gamma(u)|^4}|\dot\gamma(v)|\,\mathrm{d}v, h(u)\right\rangle|\dot\gamma(u)|\,\mathrm{d}u\\
&= 2\int_u\left\langle\int_{|v-u|\geq\epsilon}\left[\frac{2\mathbb{P}_{\gamma(u)}(\gamma(v)-\gamma(u))}{|\gamma(v)-\gamma(u)|^2} - \mathbf{N}_{\gamma(u)}\right]\right.\\
&\qquad\left.\frac{|\dot\gamma(v)|}{|\gamma(v)-\gamma(u)|^2}\,\mathrm{d}v, h(u)\right\rangle|\dot\gamma(u)|\,\mathrm{d}u.
\end{aligned}$$

It is elementary to show that the right-hand term converges to $\langle \mathrm{G}_\gamma, h\rangle$ as $\epsilon \to 0^+$. □

Remark. Using a more refined argument, one may show that Lemma 6.2 holds where $\ddot\gamma$ and $\ddot h$ are $L^{2+\delta}$-integrable for any $\delta > 0$.

One of the equivalences of the Smale conjecture (see [Ha2]) is the fact that the space of all smooth, unknotted, simple, closed loops is homotopy equivalent to the space of round loops ($\simeq RP^2$). This remarkable fact suggests that there may be some physical procedure, continuous in initial conditions, that will evolve a "tangled" but unknotted simple loop through embeddings to a round circle. Because $\mathrm{E}(\gamma)$ blows up if a self-crossing is approached, a function space flow

$$\frac{\partial\gamma}{\partial t} = -\mathrm{grad}(\mathrm{E}(\gamma)) \tag{6.14}$$

is a plausible candidate.

To give the right-hand side of equation (6.14) precise meaning it is necessary to have an inner product defined on the space of variations u to γ. Oded Schramm observed that if γ is not already a round circle, there exists a Möbius-invariant inner product $\langle\,,\,\rangle$ on the tangent bundle of $\mathbb{R}^3$ restricted to $\gamma\tau\mathbb{R}^3\,|_\gamma$. At $p \in \gamma$ the unit sphere in $\tau_p\mathbb{R}^3$ is the sphere $S_r \subset \mathbb{R}^3$, whose center is at p and with radius r, determined by the condition

$$\lim_{\substack{x\to p_+\\ y\to p_-}} D\left(\operatorname*{inv}_{S_r} x, \operatorname*{inv}_{S_r} y\right) - d\left(\operatorname*{inv}_{S_r} x, \operatorname*{inv}_{S_r} y\right) = 1. \tag{6.15}$$

Where x and y approach p along γ from opposite sides, inv_{S_r} is the inversion in S_r, D is the arc length along $\mathrm{inv}_{S_p}(\gamma)$ and d represents the euclidian distance. Now Schramm's inner product on variations becomes:

$$\langle\langle u, v\rangle\rangle = \int \langle u(t), v(t)\rangle \langle\, \mathrm{d}t,\, \mathrm{d}t\rangle_{\mathrm{dual}}^{-1/2}\, \mathrm{d}t. \tag{6.16}$$

Now the formal gradient is defined by

$$\langle\langle \mathrm{grad}\,\mathrm{E}, v\rangle\rangle = \mathrm{dE}(v). \tag{6.17}$$

Unfortunately we do not know any of the basic theory (existence, regularity, uniqueness, and convergence) for such equations. Also we do not know if the space of unknotted simple loops contains any E-critical points besides round circles. According to Hatcher there is no *topological* necessity for such points. The "tangled" unknot indicated in Figure 6.1 is an interesting unknotted initial condition for (6.14).

Figure 6.1 is an example of an (unknotted) proper knot diagram with 32 crossings. It cannot be connected to a round diagram by a family of knot

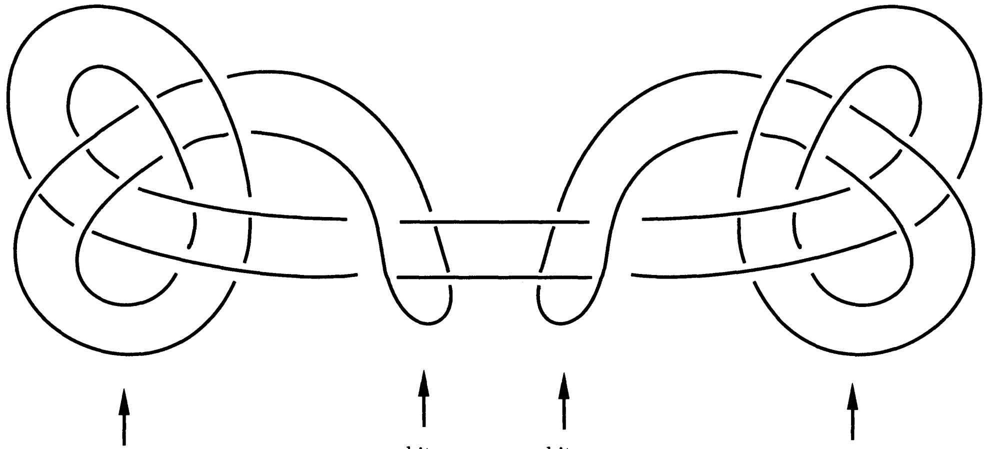

FIGURE 6.1

diagrams without the number of crossings increasing to at least 33. (*Proof*: Consider the three basic Reidemeister moves I, II, and III (cf. [Ka]). Since all triangles and bi-gons in the diagram alternate, no type II or III moves are available. There are no mono-gons, hence no type I moves. The initial move must be I^{-1} or II^{-1}, either of which increases the crossing number.) From the picture it looks like a maximum of 36 crossings must occur in the family. We wonder if "gradient flow" will shrink the "knob" relative to the "bite" and undo the tangle.

Because we have a conformally invariant gradient, we may normalize by keeping one point of γ at ∞. This leads to the consideration of the spaces $X_K = \{\pm$ vertically asymptotic simple lines with a finite knot type $K\}$. From the Appendix (equivalence 6) of [Ha2] and [Ha1] one finds the homotopy type of X_k to be $K(\pi, 1)$, where π is $\mathrm{Out}_\partial(\pi_1(\mathbb{S}^3 \backslash k))$ = the group of outer automorphisms of $\pi_1(\mathbb{S}^3 \backslash k)$, which agree with some inner automorphism on the peripheral subgroup. This group is trivial for the trivial knot type, but has elements of infinite order whenever $(\mathbb{S}^3 \backslash \dot{\mathcal{N}}(K))$ contains an essential torus (e.g., satellite knots) or an annulus, which is not a union of fibers in a Seifert fibered structure on $(\mathbb{S}^3, K)$. Also elements of finite order occur for knots with finite-order symmetries, which do not embed in circle actions. Thus for "knotted z-axes" the topology of the function space suggests that, for some knot types K, there should be critical points for E of all indexes = 0, 1, 2,... At present, no critical points of positive index are known to exist.

Because E is Möbius invariant, any γ admitting a symmetry group $\mathrm{G} \subset \mathrm{Möb}(\mathbb{S}^3)$ would retain this symmetry under a gradient flow. In [FrL] a purely topological investigation shows that there is no obstruction to symmetry preserving a homotopy from unknotted symmetric curves to round circles. It is still open whether a similar result holds for parameter families of symmetric unknots.

7. Energy of links

The energy can also be defined for links. Let $\gamma_1, \gamma_2, \ldots, \gamma_k : \mathbb{S}^1 \to \mathbb{R}^3$ be a link consisting of disjoint embeddings. The *total energy* of the link $(\gamma_1, \gamma_2, \ldots, \gamma_k)$ is defined to be

$$TE(\gamma_1, \gamma_2, \ldots, \gamma_k) = \sum_{i=1}^{k} \mathrm{E}\,(\gamma_i, \gamma_i) + \frac{1}{2} \sum_{i,j=1, i\neq j}^{k} \mathrm{E}(\gamma_i, \gamma_j), \tag{7.1}$$

where

$$\mathrm{E}(\gamma_i, \gamma_i) = \mathrm{E}(\gamma_i), \tag{7.2}$$

and for $i \neq j$,

$$\mathrm{E}(\gamma_i, \gamma_j) = \iint_{\mathbb{S}^1 \times \mathbb{S}^1} \frac{|\dot\gamma_i(u)| \cdot |\dot\gamma_j(v)|}{|\gamma_i(u) - \gamma_j(v)|^2}\, du\, dv. \tag{7.3}$$

Let $c(\gamma_i, \gamma_j)$ be the average crossing number of γ_i and γ_j (see [FrH1], pp. 196-197) so that

$$c(\gamma_i, \gamma_j) = \frac{1}{4\pi} \iint_{\mathbb{S}^1 \times \mathbb{S}^1} \frac{|(\dot\gamma_i(u), \dot\gamma_j(v) - \gamma_j(v) - \gamma_i(u))|}{|\gamma_j(v) - \gamma_i(u)|^3}\, du\, dv. \tag{7.4}$$

Then we have immediately

$$c(\gamma_i, \gamma_j) \leq \frac{1}{4\pi} \mathrm{E}(\gamma_i, \gamma_j) \quad \text{if } i \neq j. \tag{7.5}$$

By Theorem 3.6

$$c(\gamma_i, \gamma_i) \leq \frac{11}{12\pi} \mathrm{E}(\gamma_i, \gamma_i) + \frac{1}{\pi}. \tag{7.6}$$

As a consequence, we have the following:

COROLLARY 7.1. *Let $M > 0$. There are finitely many link types that can be represented by a set of disjoint embeddings of $\mathbb{S}^1$ into $\mathbb{R}^3$ with $TE \leq M$.*

Proof. Let $\gamma_1, \gamma_2, \ldots, \gamma_k : \mathbb{S}^1 \to \mathbb{R}^3$ be disjoint embeddings with $TE(\gamma_1, \gamma_2, \ldots, \gamma_k) \leq M$. As each loop contributes at least 4 to the energy, we deduce that the number of components k is bounded by $M/4$. On the other hand, inequalities (7.5) and (7.6) imply that

$$\sum_{i=1}^{k} c(\gamma_i, \gamma_i) + \frac{1}{2} \sum_{i,j=1, i\neq j}^{k} c(\gamma_i, \gamma_j) \leq \frac{11}{12\pi} M + \frac{k}{\pi}.$$

Thus the total crossing number of the link represented by $(\gamma_1, \gamma_2, \ldots, \gamma_k)$ is bounded by $11/(12\pi)M + k/\pi$, hence the corollary. □

As before, we are interested in the extremal configurations of links. An elementary computation for two planar concentric circles shows the infimum for the total energy of links is 8 while no link attains it. If only essential links are considered, the least energy configuration must be a Hopf link (see Corollary 7.4). If both (or even one) component is a round circle, the least energy configuration is located by Corollary 7.3. It is an open problem how to prove that this configuration has least energy in its isotopy class.

Let $\gamma\colon \mathbb{S}^1 \hookrightarrow \mathbb{R}^3$ be an embedding. Let $p(x, y, z)$ be a point in $\mathbb{R}^3 \backslash \mathbb{S}^1$. Then the energy of p with respect to γ is defined by

$$\mathrm{E}(\gamma, p) = \int_{\mathbb{S}^1} \frac{|\dot\gamma(u)|}{|p - \gamma(u)|^2} \, \mathrm{d}u.$$

Setting $\mathrm{d}s^2_{\mathrm{E}(\gamma)} = \mathrm{E}(\gamma, p)^2(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2)$, then $\mathrm{d}s^2_{\mathrm{E}(\gamma)}$ is a new metric on $\mathbb{R}^3\backslash\gamma$ which is conformal to the euclidean metric. In fact $\mathrm{d}s^2_{\mathrm{E}(\gamma)}$ completes to a smooth ($\mathcal{C}^\infty$) metric on $\mathbb{S}^3\backslash\gamma$. It is obvious that the energy of a curve in $\mathbb{S}^3\backslash\gamma$ with γ (i.e., the energy cross term) is exactly the length of the curve with respect to the new metric $\mathrm{d}s^2_{\mathrm{E}(\gamma)}$.

Now consider the case where γ is a round circle. Since the energy is Möbius invariant, we may assume that γ is $\{(x, y, 0)|x^2 + y^2 = 1\}$ in the $x - y$ plane. Then we have the next theorem.

THEOREM 7.2. *The sphere $\mathbb{S}^3\backslash\gamma$ with metric $\mathrm{d}s^2_{\mathrm{E}(\gamma)}$ is isometric to hyperbolic space crossing a circle, $\mathbb{H}^2 \times \mathbb{S}^1$, where $\mathbb{S}^1$ is the unit circle.*

Proof. By the Möbius invariance of energy, we may make the computation for $\gamma' = z$-axis. For any $p(x, y, z)$ not on the z-axis we have

$$\mathrm{E}(\gamma, p) = \int_{-\infty}^{+\infty} \frac{1}{x^2 + y^2 + (t - z)^2} \, \mathrm{d}t = \frac{\pi}{\sqrt{x^2 + y^2}}.$$

Hence

$$\mathrm{d}s^2_{\mathrm{E}(\gamma)} = \frac{\pi^2}{x^2 + y^2} \left(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\right).$$

Changing x, y to polar coordinates $\{\begin{smallmatrix} x=r\cos\theta \\ y=r\sin\theta \end{smallmatrix}$ yields

$$\begin{aligned} ds^2_{\mathrm{E}(\gamma)} &= \frac{\pi^2}{r^2}\left(dr^2 + r^2\, d\theta^2 + dz^2\right) \\ &= \pi^2\left(\frac{dr^2 + dz^2}{r^2} + d\theta^2\right) \end{aligned}$$

The first two terms describe the hyperbolic metric in the coordinates of the halfspace model. Transforming the z-axis back to γ in $\mathbb{R}^3$, we obtain the theorem. □

It is easy to check that the energy cross term of the z-axis, with any round circle lying on a plane $\{z = \text{const}\}$ and the origin on the z-axis, is $2\pi^2$. Since a round circle has energy 4, the total energy of this Hopf link is $8 + 2\pi^2$.

COROLLARY 7.3. *The least energy configuration for a Hopf link with both (or even one) component a round circle is obtained by any Möbius transformation of* $\{(z\text{-axis}), \gamma\}$ *into* $\mathbb{R}^3$. *All Hopf links obtained in this way have total energy* $8 + 2\pi^2$.

Remark. The obvious representative for the Hopf link

$$\{(x, y, 0) \,\|\, x^2 + y^2 = 1\} \cup \{(0, y, z) \,|\, (y-1)^2 + z^2 = 1\}$$

is not an extremal configuration.

Proof of Corollary 7.3. By Theorem 7.2, $\mathbb{S}^3 \backslash \gamma$ with the new metric is $\mathbb{H}^2 \times \mathbb{S}^1$. The energy cross term of a curve in $\mathbb{S}^3 \backslash \mathbb{S}^1$ with γ is its length in $ds^2_{\mathrm{E}(\gamma)}$. Fixing γ shows that the least energy configuration will be γ with the shortest closed geodesic that links γ, provided such a closed geodesic is round. But the factor circles of $\mathbb{H}^2 \times S$ are all the shortest, simply linking closed geodesics in $\mathbb{H}^2 \times S$ and round ones in $\mathbb{S}^3$. □

COROLLARY 7.4. *The absolute minima for total energy among essential links are topologically a Hopf link.*

Proof. By Corollary 7.3 an essential link L of the smallest possible total energy satisfies $\mathrm{E} \leq 8 + 2\pi^2$. By inequality (7.5) it is easy to check that it must have only two components. Hence

$$8 + 4\pi c(\gamma_1, \gamma_2) \leq T\mathrm{E}(\gamma_1, \gamma_2) \leq 8 + 2\pi^2.$$

Consequently $c(\gamma_1, \gamma_2) < \pi/2$. Therefore L is a Hopf link. □

8. $1/r^\alpha$-potential energy

The energy E can be thought of as an $1/r^2$-potential energy. In this section we make a few remarks on $1/r^\alpha$-potential energy in general.

Let $\gamma = \gamma(u)$ be a rectifiable curve in $\mathbb{R}^3$. Using the same notation as before, for any real number α define the $1/r^\alpha$-potential energy $\mathrm{E}^\alpha(\gamma)$ by the following integral:

$$\mathrm{E}^\alpha(\gamma) = \iint_{u,v} \left\{ \frac{1}{|\gamma(u)-\gamma(v)|^\alpha} - \frac{1}{D(\gamma(u),\gamma(v))^\alpha} \right\} |\dot\gamma(u)|\,|\dot\gamma(v)|\,du\,dv. \tag{8.1}$$

For simplicity we will assume that $\alpha > 0$ in the following. There are corresponding results for $\alpha \le 0$.

The following properties are immediate:

PROPOSITION 8.1. (1) $\mathrm{E}^\alpha(\gamma)$ *is independent of the parametrization or orientation of the curve.*

(2) $\mathrm{E}^\alpha(\gamma)$ *is invariant under euclidean isometry.*

(3) *If* $\mathbb{R}^3$ *is rescaled by a linear factor* s, *then* $\mathrm{E}^\alpha(s\gamma) = s^{2-\alpha}\mathrm{E}^\alpha(\gamma)$. □

PROPOSITION 8.2. *Let* γ *be a simple closed* C^3*-curve in* $\mathbb{R}^3$. *Then* $\mathrm{E}^\alpha(\gamma)$ *is finite for* $\alpha < 3$.

Proof. Let u be an arc-length parameter for γ. Using $|\dot\gamma(u)| = 1$ and

$$\gamma(v) - \gamma(u) = \dot\gamma(u)(v-u) + \frac{\ddot\gamma(u)}{2}(v-u)^2 + \mathcal{O}((v-u)^3),$$

we have

$$|\gamma(v)-\gamma(u)|^2 = (v-u)^2 + \mathcal{O}((v-u)^4). \tag{8.2}$$

Thus for any $\epsilon > 0$

$$\begin{aligned} &\iint_{|v-u|\ge\epsilon} \left\{ \frac{1}{|\gamma(v)-\gamma(u)|^\alpha} - \frac{1}{D(\gamma(v),\gamma(u))^\alpha} \right\} du\,dv \\ &= \iint_{|v-u|\ge\epsilon} \left\{ \frac{1}{|\gamma(v)-\gamma(u)|^\alpha} - \frac{1}{|v-u|^\alpha} \right\} du\,dv \\ &= \iint_{|v-u|\ge\epsilon} \mathcal{O}\left((v-u)^{2-\alpha}\right) du\,dv \\ &= \iint_{u,|w|\ge\epsilon} \mathcal{O}\left(w^{2-\alpha}\right) dw\,du. \end{aligned} \tag{8.3}$$

Thus, for $\alpha < 3$, $\mathrm{E}^\alpha(\gamma)$ is finite. □

On the other hand, it is easy to check that, for the unit circle and $\alpha \ge 3$, integral (8.1) is divergent.

We show that, for each $0 < \alpha < 3$, there exists a rectifiable, simple, closed curve γ_0 realizing the infimum of $1/r^\alpha$-potential energy among curves of length $= 1$.

PROPOSITION 8.3. *For* $0 < \alpha < 3$ *let* $\mathcal{C}$ *be the rectifiable, simple, closed curves of length* 1. *Then there exists a rectifiable, simple, closed curve* γ_0 *in* $\mathcal{C}$ *with* $\mathrm{E}^\alpha(\gamma_0) \leq \mathrm{E}^\alpha(\gamma)$ *for other* γ *in this class* $\mathcal{C}$.

Proof. Let $\{\gamma_i\colon \mathbb{S}^1 = \mathbb{R}/\mathbb{Z} \to \mathbb{R}^3, |\dot{\gamma}(u)| = 1\}$, $i = 1, 2, \ldots$, be a sequence of curves in $\mathcal{C}$ and $\mathrm{E}^\alpha(\gamma_i)$ approaching the infimum of $\mathrm{E}^\alpha(\gamma)$ over all curves with arc-length parametrization in $\mathcal{C}$. By translation we may assume that each curve passes through the origin of $\mathbb{R}^3$. As in Lemma 1.2 the functions γ_i are all 1-Lipschitz, so by Ascoli's theorem there is a convergent subsequence $\gamma_1, \gamma_2, \ldots$ (use the same indices) converging in $\mathcal{C}^0$ to a Lipschitz embedding γ_0.

Denote the energy integrand of equation (8.1) by $\mathrm{G}^\alpha_{\gamma_0}\colon \{u, v | u \neq v\} \to \mathbb{R}^+$. Clearly $\mathrm{G}^\alpha_{\gamma_0}$ is the pointwise limit of $\mathrm{G}^\alpha_{\gamma_i}$. Since $\mathrm{G}^\alpha_{\gamma_i}$ is a positive function, by Fatou's lemma

$$
\begin{aligned}
\mathrm{E}^\alpha(\gamma_0) &= \iint\limits_{u,v} \mathrm{G}^\alpha_{\gamma_0}\, du\, dv \leq \lim_{i\to+\infty} \iint\limits_{u,v} \mathrm{G}^\alpha_{\gamma_i}\, du\, dv \\
&= \lim_{i\to+\infty} \mathrm{E}(\gamma_i).
\end{aligned}
\tag{8.4}
$$

□

Next we attempt to determine the shape of the minimizers.

THEOREM 8.4. *For* $0 < \alpha < 3$ *any minimizer* γ_0 *for* $\mathrm{E}^\alpha(\gamma)$ *is a planar, convex, simple, closed curve.*

The main ingredient in the proof is the following reflection lemma. The proof is the same as in Section 5. Given a closed curve γ in $\mathbb{R}^3$, a plane P is a *plane of support* for γ if $\gamma \cap P \neq \phi$, $\gamma \subset H_P$, the positive closed halfspace determined by P.

LEMMA 8.5. *Let* γ *be a closed curve in* $\mathbb{R}^3$ *with* P *a plane of support such that* $\gamma \cap P$ *is disconnected. Let p and q belong to distinct connected components of* $\gamma \cap P$; *divide* γ *into two arcs,* α *and* β, $\alpha \cup \beta = \gamma$ *and* $\alpha \cap \beta = p \cup q$.

Let $\gamma' = \alpha \cup \bar{\beta}$, *where* $\bar{\beta}$ *is the reflection of* β *in* P. *Then* $\mathrm{E}^\alpha(\gamma') < \mathrm{E}^\alpha(\gamma)$.

Note that γ' *has the same length as* γ. □

Proof of Theorem 8.4. Consider the relation $\mathcal{R} \subset \mathbb{S}^2 \times \mathbb{S}^1$ given by $(\theta, s) \in \mathcal{R}$ if and only if there exists a plane of support P with $\theta \perp P$, $s \subset P \cap \gamma_0$ and $\gamma_0 \subset H_P$, the positive closed halfspace determined by P. Clearly $\mathcal{R}$ is a closed set.

For $\theta \in \mathbb{S}^2$, $\mathcal{R}^{-1}(\theta)$ is either the entire circle, is disconnected or has the Čech cohomology of a point (i.e., an interval or point). For $s \in \mathbb{S}^1$, $\mathcal{R}(s) \subset \mathbb{S}^2$ is geodesically convex and also has the Čech cohomology of a point. □

LEMMA 8.6. *Either γ_0 is planar or there exists a plane of support P with $\gamma_0 \cap P$ disconnected.*

Proof. Suppose neither conditions hold. Let S^- be the closed subset of $\mathbb{S}^1$ for which γ_0 has a plane of support. Both surjections $\mathcal{R} \to S^-$ and $\mathcal{R} \to \mathbb{S}^2$ satisfy the hypothesis of the Vietoris theorem (see [Br], p. 202); i.e., all point inverses have the Čech homology of a point and therefore induce isomorphisms on Čech cohomology. Since $\check{H}^2(S^-, \mathbb{Z}) \simeq 0 \not\simeq \mathbb{Z} \simeq \check{H}^2(\mathbb{S}^2, \mathbb{Z})$, we have a contradiction. □

LEMMA 8.7. *The curve γ_0 is planar.*

Proof. If not, by the previous lemma there is a plane of support P with $\gamma_0 \cap P$ disconnected. By Lemma 8.5 we may reduce the energy of γ_0, contradicting minimality. □

LEMMA 8.8. *The curve γ_0 is a convex plane curve.*

Proof. If γ_0 is not convex, let L be a line "of support" meeting γ_0 in a disconnected set. As in Lemma 8.5, write $\gamma_0 = \alpha \cup \beta$ with $\alpha \cap \beta$ a pair of points lying in different components of $\gamma_0 \cap L$. Again reflecting β in L would yield a lower energy curve $\bar{\gamma} = \alpha \cup \bar{\beta}$, contradicting minimality. □

Conjecture. For $0 < \alpha < 3$ any length $= 1$ minimizer of E^α energy is a round circle of radius $1/(2\pi)$.

In Lemma 1.2, we showed that, for $\alpha = 2$, finite energy implies bi-Lipschitz. This is also true for $2 < \alpha < 3$, and the proof is left as an exercise.

From the Introduction we know that, for $2 \le \alpha < 3$, $\mathrm{E}^\alpha(\gamma)$ blows up when a simple closed loop γ begins to acquire a double point. We reexamine this blowup using the following model: Given $0 < \epsilon \ll 1$, let

$$\gamma_1 : \ \{(x, 0, \epsilon) \mid |x| \le 1\},$$
$$\gamma_2 : \ \{(0, y, 0) \mid |y| \le 1\}.$$

Then, by an elementary computation, the crossing energy

$$\mathrm{E}^\alpha_\epsilon(\gamma_1, \gamma_2) = \iint\limits_{\gamma_1\gamma_2} \frac{1}{|\gamma_1(x) - \gamma_2(y)|^\alpha} \, \mathrm{d}x \, \mathrm{d}y$$

is $\mathrm{A}\epsilon^{2-\alpha} + \mathcal{O}(\epsilon^2) + \mathrm{B}$ when $\alpha \ne 2$, and $\mathrm{A}\log(1+\epsilon^{-2}) + \mathcal{O}(\epsilon^2) + \mathrm{B}$ when $\alpha = 2$ for some constants A and B.

It is clear that, when $2 \le \alpha < 3$,

$$E_\epsilon^\alpha(\gamma_1, \gamma_2) \to +\infty \quad \text{as } \epsilon \to 0^+,$$

and $E_\epsilon^\alpha(\gamma_1, \gamma_2)$ is still finite as $\epsilon \to 0^+$ if $0 < \alpha < 2$. But

$$\frac{d}{d\epsilon} E_\epsilon^\alpha(\gamma_1, \gamma_2) \to +\infty \quad \text{as } \epsilon \to 0^+$$

if $1 < \alpha \le 2$. Although, when $1 < \alpha < 2$, $E^\alpha(\gamma)$ does not blow up when γ crosses itself, the derivative does blow up. This suggests that if there exists a gradient flow for $1 < \alpha < 2$, then, because of this infinite-derivative barrier, self-crossing would never occur. For the newtonian potential in $\mathbb{R}^3$, $\alpha = 1$, we see no reason that a gradient flow should preserve embeddedness.

University of California, San Diego, La Jolla, California

Princeton University, Princeton, N. J. and University of California, San Diego, La Jolla, California

University of California, San Diego, La Jolla, California and University of Michigan, Ann Arbor, Michigan

References

[Ah] K. Ahara, Energy of a knot, screened at Topology Conf., Univ. of Hawaii, August, 1990, K.H. Dovermann, organizer.

[Ar] V.I. Arnold, The asymptotic Hopf invariant and its applications, Sel. Math. Sov. **5** (1986), 327-345 (English trans.).

[Bi] R.H. Bing, Locally tame sets are tame, Ann. of Math. **59** (1954), 145-158.

[Br] G.E. Bredon, *Sheaf Theory*, McGraw–Hill, New York, 1967.

[Fe] H. Federer, *Geometric Measure Theory*, Springer–Verlag, New York, 1969.

[FrH1] M.H. Freedman and Z.-X. He, Divergence-free fields: energy and asymptotical crossing number, Ann. of Math. **134** (1991), 189-229.

[FrH2] ———, Research announcement on the 'energy' of knots, in *Topological Aspects of the Dynamics of Fluids and Plasmas*, K. Moffatt et al., eds., Kluwer Academic Publ., Dordrecht, 1992, pp. 219-222.

[FrL] M.H. Freedman and F. Luo, Equivariant isotopy of unknots to round circles, Preprint.

[Ha1] A. Hatcher, Homeomorphisms of sufficiently large P^2-irreducible 3-manifolds, Topology **15** (1976), 343-347.

[Ha2] A. Hatcher, A proof of the Smale conjecture, Diff($\mathbb{S}^3$) $\simeq$ O(4), Ann. of Math. **117** (1983), 553-607.

[Ka] L. Kauffman, *On Knots*, Princeton Univ. Press, Princeton, 1987.

[Ki] D. Kim and R. Kusner, Torus knots extremizing the Möbius energy, Experimental Math. **2** (1993), 1-9.

[Mi] J.W. Milnor, On the total curvature of knots, Ann. of Math. **52** (1950), 248-257; Zbl. 37, 389; M.R. 12, 373.

[Mo] E.E. Moise, Affine structures in 3-manifolds, V. The triangulation theorem and Hauptvermutung, Ann. of Math. **56** (1952), 95-114.

[O'H1] J. O'Hara, Energy of a knot, Topology **30** (1991), 241-247.

[O'H2] ———, Family of energy functionals of knots, Topology Appl. **48** (1992), 147-161.

[O'H3] J. O'Hara, Energy functionals of knots, in *Topology Conf. Hawaii*, K.H. Dovermann, ed., World Scientific, 1992, pp. 201-214.

[S] D. Sumners, The growth of the number of prime knots, Math. Proc. Cambridge Phil. Soc. **102** (1987), 303-315.

[T] W.T. Tutte, A census of planar maps, Canadian J. of Math. **15** (1963), 249-271.

[W] D.J.A. Welsh, On the number of knots, Preprint.

(Received January 6, 1992)
(Revised June, 1992)