ELECTRONIC RESEARCH ANNOUNCEMENTS
OF THE AMERICAN MATHEMATICAL SOCIETY
Volume 1, Issue 1, 1995

ON NON-SEPARATING SIMPLE CLOSED CURVES IN A COMPACT SURFACE

FENG LUO

(Communicated by Walter Neumann)

ABSTRACT. We introduce a semi-algebraic structure on the set $\mathcal{S}$ of all isotopy classes of non-separating simple closed curves in any compact oriented surface and show that the structure is finitely generated. As a consequence, we produce a natural finite dimensional linear representation of the mapping class group of the surface. Applications to the Teichmüller space, Thurston's measured lamination space, the harmonic Beltrami differentials, and the first cohomology group of the surface are discussed.

§1. MAIN RESULTS

Given a compact oriented surface of positive genus Σ, let $\mathcal{S} = \mathcal{S}(\Sigma)$ be the set of all isotopy classes of non-separating simple closed unoriented curves in Σ. We introduce two relations, *orthogonal* and *disjoint* in $\mathcal{S}$ as follows. Two classes α and β in $\mathcal{S}$ are said to be *orthogonal,* denoted by $\alpha \perp \beta$, if they have representing simple closed curves a and b intersecting transversely at one point (in this case, we also say that a is *orthogonal* to b and denote it by $a \perp b$). Two classes α and β in $\mathcal{S}$ are said to be *disjoint* if they have representing simple closed curves a and b so that $a \cap b = \phi$. Our goal is to study $\mathcal{S}$ under these two relations.

Given two orthogonal simple closed curves p and q, define the product pq of p and q to be $D_p(q)$ where D_c is the positive Dehn twist about the simple closed curve c. Geometrically, pq and qp are obtained from $p \cup q$ by by breaking the intersection into two embedded arcs as in figure 1.

Clearly the isotopy classes of pq and qp depend only on the isotopy classes of p and q. We use $[p]$ to denote the isotopy class of a simple closed curve p and define the product $[p][q]$ to be $[pq]$ when $[p] \perp [q]$. Our first result states that $\mathcal{S}$ is finitely generated in the product. Actually, a stronger form of the finiteness result holds. It is on the stronger form that we will focus. To this end, we introduce the following definition.

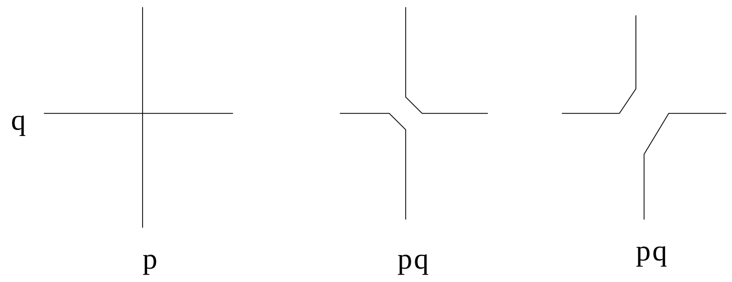
Definition. Given a subset $\mathcal{X}$ of $\mathcal{S}$, the *derived set* $\mathcal{X}'$ of $\mathcal{X}$ is $\mathcal{X} \cup \{\alpha\beta | \alpha, \beta$, and $\beta\alpha$ are in $\mathcal{X}$ $\}$. Inductively define $\mathcal{X}^n$ to be the derived set of $\mathcal{X}^{n-1}$ for $n > 1$. We define $\cup_{n=1}^{\infty} \mathcal{X}^n$ to be the set *generated by* $\mathcal{X}$, and denoted it by $\mathcal{X}^\infty$.

Our first theorem is the following.

Received by the editors April 22, 1995, and, in revised form, March 22, 1995.
1991 *Mathematics Subject Classification.* 57.
Key words and phrases. Simple closed curve, surface.

Theorem 1. *If Σ is a compact orientable surface of positive genus, then there is a finite subset $\mathcal{F}$ of isotopy classes of non-separating simple closed curves so that $\mathcal{F}^\infty = \mathcal{S}(\Sigma)$.*

The next theorem is an analogue of Dehn-Nielsen's theorem that the mapping class group of a closed surface is the outer automorphism group of the fundamental group of the surface. To be more precise, given any orientation preserving self-homeomorphism ϕ of Σ, ϕ induces a bijective map of $\mathcal{S}$ by the formula $\phi_*([a]) = [\phi(a)]$. Clearly, ϕ_* preserves disjointness, orthogonality and the product.



The orientation is the right handed orientation in the plane
Figure 1

Theorem 2. *If h is a bijective map of the set of all isotopy classes of non-separating simple closed curves in a closed orientable surface so that h preserves disjointness, orthogonality and the product, then h is induced by an orientation preserving self-homeomorphism of the surface.*

§2. MOTIVATIONS

Theorems 1 and 2 are well known for the torus T^2. As usual, choose an oriented meridian m and an oriented longitude l for T^2. Each element in $\mathcal{S}$ is uniquely represented by $\pm(p[m] + q[l])$ in the first homology group where p and q are two relatively prime integers. By assigning the rational number p/q to the class, we identify $\mathcal{S}$ with $\mathbf{Q} \cup \{\infty\} \subset \mathbf{R} \cup \{\infty\}$ which is considered to be the natural boundary of the hyperbolic upper half-plane. Two classes p/q and r/s are orthogonal if and only if $ps - qr = \pm 1$. Furthermore, if $\beta \perp \gamma$, then $\beta\gamma$ and $\gamma\beta$ are symmetric with respect to the hyperbolic reflection about the geodesic ending at β and γ. Thus by the well known modular picture, one sees that $\mathcal{S}$ is generated by 0, 1, and ∞ as in figure 2.

To see theorem 2 for the torus, let us take a bijective map h of $\mathcal{S} = \mathbf{Q} \cup \{\infty\}$ preserving the orthogonality. Thus $h(0)$, $h(1)$ and $h(\infty)$ are three pairwise orthogonal rational numbers. Therefore, there is an element $\psi \in \mathrm{GL}(2, \mathbf{Z})$ acting on $\mathbf{Q} \cup \{\infty\}$ as Möbius transformations so that $\psi(0) = h(0)$, $\psi(1) = h(1)$ and $\psi(\infty) = h(\infty)$.

Since GL(2, $\mathbf{Z}$) is the mapping class group of the torus, we may assume at that h leaves the three curves 0, 1 and ∞ fixed. Thus by the modular pictures above, the bijection h leaves each rational number fixed.

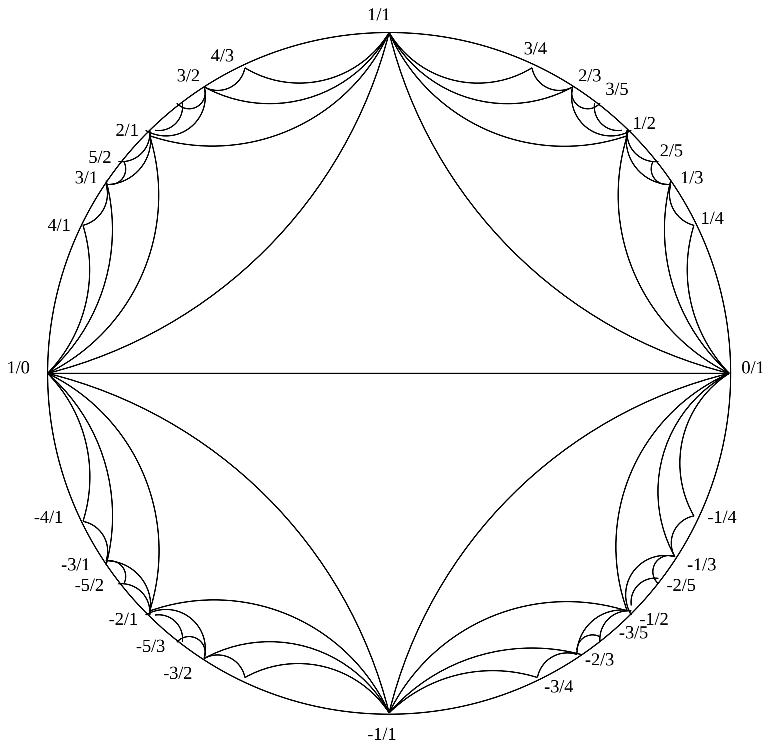


Figure 2

The other motivation of the theorem 1 comes from Thurston's compactification of the Teichmüller space and the trace formula $\mathrm{tr}(XY) + \mathrm{tr}(X^{-1}Y) = \mathrm{tr}(X)\mathrm{tr}(Y)$ for X and Y in SL(2, $\mathbf{C}$) (see [FLP]). If two elements A and B in $\pi_1(\Sigma)$ have representatives in the free homotopy class $[a]$, $[b]$ in $\mathcal{S}$ so that $[a] \perp [b]$, then the products AB and $A^{-1}B$ are represented by the classes $[ab]$ and $[ba]$. As a consequence of the trace formula, the hyperbolic lengths of the classes A, B, AB and $A^{-1}B$ with the property above satisfy the following nonlinear relation: $\cosh(l_{AB}/2)+ \cosh(l_{A^{-1}B}/2) = 2\cosh(l_A/2)\cosh(l_B/2)$, i.e., it satisfies the following,

$$f(\alpha\beta) + f(\beta\alpha) = 2f(\alpha)f(\beta), \quad \alpha \perp \beta. \tag{1}$$

Now if we consider a degenerate family of hyperbolic metrics which tends projectively to a simple closed curve c (or more generally a measured lamination) in the Thurston's compactification of the Teichmüller space, then the relation (1) degenerates to $I([a],[c]) + I([b],[c]) = \max(I([ab],[c]), I([ba],[c]))$, i.e., it satisfies the following,

$$f(\alpha) + f(\beta) = \max(f(\alpha\beta), f(\beta\alpha)), \quad \alpha \perp \beta. \tag{2}$$

where I($\cdot, \cdot$) is the geometric intersection number. Since equation (2) is piecewise linear ($\max(x,y) = 1/2(x + y + |x - y|)$), this may be one reason that the action of the mapping class group on the measured lamination space is piecewise linear.

Since both the Teichmüller space and the measured lamination space are finite dimensional, this prompts us to ask about the existence of finite generators for $\mathcal{S}$.

§3. SOME CONSEQUENCES

As a consequence of theorem 1, we produce a finite dimensional linear representation of the mapping class group $\mathcal{G} = \mathcal{G}(\Sigma)$ of the surface Σ. Recall that the mapping class group $\mathcal{G}(\Sigma)$ is defined to be Homeo(Σ)/isotopies and $\mathcal{G}$ acts naturally on $\mathcal{S}$ by permuting the isotopy classes. The finite dimensional linear representations of $\mathcal{G}$ are constructed as follows. Take V to be the vector space $\mathbf{C}^{\mathcal{S}}$ of all complex valued functions on $\mathcal{S}$, i.e., V $= \{f : \mathcal{S} \to \mathbf{C}\}$. The mapping class group $\mathcal{G}$ acts naturally linearly on V by permuting $\mathcal{S}$ coordinates. Let a, b, c and d be four given complex numbers with $cd \neq 0$. Define a linear subspace $V_{a,b,c,d}$ of V by the equations

$$af(\alpha) + bf(\beta) = cf(\alpha\beta) + df(\beta\alpha), \quad \alpha \perp \beta. \tag{3}$$

By theorem 1, $V_{a,b,c,d}$ is finite dimensional and is clearly invariant under the linear action of the mapping class group.

Corollary. . *If the genus g of the surface is bigger than one, then the only space $V_{a,b,c,d}$ which contains a non-constant function is $V_{2a,2a,a,a} = V_{2,2,1,1}$. Furthermore, the dimension of $V_{2,2,1,1}$ is at least $(2g + n)(2g + n - 1)/2$ where g is the genus and n is the numbers of boundary components of the surface.*

The basic idea of the proof is the following. If a simple closed curve s is orthogonal to another simple closed curve t, then there is a universal relation that $[s(ts)] = [t]$ (and also $[(st)s] = [t]$). This relation $[s(ts)] = [t]$ implies the well known braid relation $D_s D_t D_s = D_t D_s D_t$ mod(isotopy) in the mapping class group. By iterating the relation $[s(ts)] = [t]$ several times for three pairwise orthogonal simple closed curves which are not in a torus with a hole, we obtain that $2a = 2b = c = d$. To show that $V_{2,2,1,1}$ contains non-constant solutions, we take the square of algebraic intersection number $f(\alpha) = i^2(\alpha, \gamma)$ for a fixed class γ in $\mathcal{S}$. One sees easily that f satisfies the skein relation $f(\alpha\beta) + f(\beta\alpha) = 2f(\alpha) + 2f(\beta)$ whenever $\alpha \perp \beta$. Since for all choice of γ in $\mathcal{S}$, these functions span a linear subspace of dimension at least $(2g + n)(2g + n - 1)/2$ in $V_{2,2,1,1}$ (the actual dimensions of the subspace is $(2g + n)(2g + n - 1)/2$ if $n \neq 0$ and is $g(2g + 1)$ if $n = 0$), the estimate on the dimension follows.

Theorem 1 may give rise to a characterization of the length spectrum of a hyperbolic metric in a closed surface. Given a closed orientable surface Σ, let Teich(Σ) be the Teichmüler space of Σ. For each equivalence class $[d]$ of hyperbolic metric in Teich(Σ), we produce a function $f_{[d]}$ from $\mathcal{S}$ to $\{\mathbf{R}|t > 1\}$ by setting $f_{[d]}(\alpha) = \cosh(l_d(\alpha)/2)$ where $l_d(\alpha)$ is the length of the geodesic in the class α in the metric d. By the remark above, we have

$$f_{[d]}(\alpha\beta) + f_{[d]}(\beta\alpha) = 2f_{[d]}(\alpha)f_{[d]}(\beta), \quad \alpha \perp \beta. \tag{4}$$

The set $\mathcal{T}$ of all functions from $\mathcal{S}$ to $\{\mathbf{R}|t > 1\}$ satisfying (1) is finite dimensional by theorem 1. It is well known that the map from Teich(Σ) to $\mathcal{T}$ sending $[d]$ to $f_{[d]}$ is injective and continuous. It is natural to ask whether the map is an onto map.

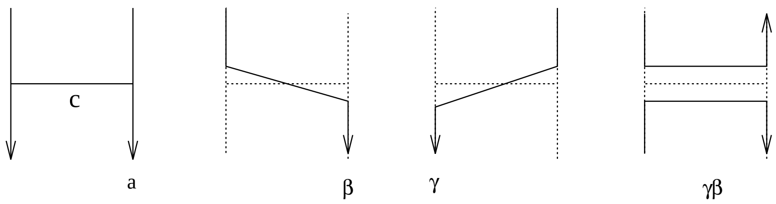
One possible approach to the above problem is to find the dimension of $\mathcal{T}$. Since harmonic Beltrami differentials are deformations of the hyperbolic metrics in the Teichmüller space, given a function f satisfying the relation (1) above, it is tempting to call a function g satisfying

$$g(\alpha\beta) + g(\beta\alpha) = 2f(\alpha)g(\beta) + 2f(\beta)g(\alpha), \quad \alpha \perp \beta. \tag{5}$$

a *"harmonic Beltrami differential"* in the *"conformal structure"* given by f. Denote $T_f(\mathcal{T})$ the linear space (the tangent space of $\mathcal{T}$ at f) of all functions g satisfying (5). One would expect to have a Riemann-Roch theorem which calculates the dimension of $T_f(\mathcal{T})$.

§4. SKETCH OF THE PROOFS

The main step in the proof of theorem 1 is an induction on a semi-norm $||\Sigma|| = 3g + n$ where Σ is a compact orientable surface of genus g with n boundary components. We show that there are finitely many subsurfaces Σ_i in Σ of smaller norms so that non-separating simple closed curves in Σ_i generate $\mathcal{S}(\Sigma)$. The main technical difficult is due to the fact that separating simple closed curves in Σ_i may become non-separating in Σ.



The orientation is the right handed orientation in the plane
Figure 3

We uses several induction steps to achieve the above result. In each step of the induction process, we construct a 1-complex G in Σ and induct on a semi-norm $||\alpha||_G$ for $\alpha \in \mathcal{S}$ where $||\alpha||_G = inf\{|G \cap a||a \in \alpha\}$, i.e., the geometric intersection number. Given a class $\alpha \in \mathcal{S}$, we would like to find two classes β and γ in $\mathcal{S}$ so that $\beta \perp \gamma$, $\alpha = \beta\gamma$ and the semi-norms of β, γ and $\gamma\beta$ are smaller than $||\alpha||_G$. To this end, let us take $a \in \alpha$ so that $||\alpha||_G = |G \cap a|$ and find an arc c in Σ so that $c \cap a = \partial c$ and c approaches both end points of ∂c from different sides of a. Indeed, assuming c has been constructed, we obtain three new non-separating simple closed curves representing β, γ and $\gamma\beta$ as in the following figure 3. Furthermore, $\alpha = \beta\gamma$. We call this an H-reduction on the curve a.

Now if the arc c satisfies $\partial c = c \cap G$, then clearly $||\beta||_G$, $||\gamma||_G$ and $||\gamma\beta||_G$ are all less than $||\alpha||_G$.

Let us illustrate this by considering the case of a torus, i.e., the modular picture is generated by hyperbolic reflections on the three sides of the ideal triangle with vertices $0 = [l]$, $1 = [ml]$ and $\infty = [m]$. Given a class $\alpha = \pm(p[m]+q[l])$ where p and q are relatively prime integers, define a norm $||\alpha|| = |p|+|q| = \inf\{|a\cap m|+|a\cap l||a \in \alpha\}$. Let $a \in \alpha$ be a representative so that $||\alpha|| = |a \cap m| + |a \cap l|$ where $|a \cap l| = |p|$ and $|a \cap m| = |q|$. Then all intersection points in $a \cap m$ (and in $a \cap l$ respectively) have the same intersection signs. Thus, if one of the numbers $|a \cap m|$ or $|a \cap l|$ is bigger than 1, say $|a \cap l| \geq 2$, than there are two adjacent intersection points x and y in the curve l so that there is an arc c in l joining x and y with $c \cap a = \partial c$ and $c \cap m = \phi$. Then, the H-reduction on the curve a at c produces two classes β and γ so that $\beta \perp \gamma$, $\alpha = \beta\gamma$ and the norms of β, γ and $\gamma\beta$ are smaller than $||\alpha||$. Finally, if both p and q are at most 1, then α is one of the four classes $[m]$, $[l]$, $[ml]$ and $[lm]$. Thus the result follows.

There are two steps in the proof of theorem 1. In the first step, we show that there are finitely many non-separating simple closed curves $\{c_1, ..., c_k\}$ in $\mathcal{S}$ so that $\mathcal{X} = \{\alpha \in \mathcal{S}|\alpha$ is disjoint from one of $[c_i]\}$ satisfies $(\mathcal{X})^\infty = \mathcal{S}$. This step is relatively easy to achieve. The next step is the major step in which we replace these curves c_i by subsurfaces each of them has only one boundary component. Knowing this, we apply the induction on the semi-norm $||\Sigma||$ of the subsurface and end the proof.

To prove theorem 2, we show that the Lickorish-Humphries basis $F_0 = \{[a_1], ..., [a_{2g+1}]\}$ (as a subset of $\mathcal{S}$ see figure 4) for the mapping class group of a closed surface of genus g satisfies the following property.

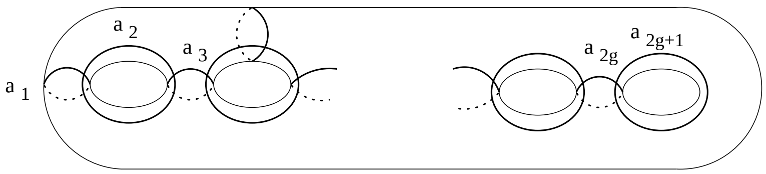


Figure 4 A Lickorish-Humphries basis

Let F_n be the set $\{\alpha|\alpha = \beta\gamma$, where β and γ are in F_{n-1} $\} \cup F_{n-1}$. Then $\mathcal{S} = \cup_{n=1}^{\infty} F_n$. Now, any two Lickorish-Humphries bases are related by a self homeomorphism of the surface and a bijective map of $\mathcal{S}$ preserving disjointness and orthogonality sends a Lickorish-Humphries basis to a Lickorish-Humphries basis. Thus the result follows. This is an analogue of the modular group case where we use two curves 0 and ∞ as basis. The basic technique of the proof is the same as that used in theorem 1. We do not have to consider the third curve $\gamma\beta$ here but we do need to produce the minimal set of generators in the new sense.

§5. SOME OBSERVATIONS AND QUESTIONS

There is an interesting similarity between the relations (1), (2) and (3) above and the functional equations for the well known elementary functions from the set of integers $\mathbf{Z}$ to $\mathbf{R}$. Namely, the hyperbolic trigonometric functions $\cosh(\lambda t)$ satisfy

$f(x+y)+f(x-y)=2f(x)f(y)$ with $f(x)>1$, the absolute value functions $|\lambda t|$ satisfy $f(x)+f(y)=\max(f(x+y), f(x-y))$, and the square functions λt^2 satisfy $f(x+y)+f(x-y)=2f(x)+2f(y)$. Here λ is a parameter. It is easy to show that these are all non-constant solutions to the functional equations above. Note that $\cos(\lambda t)$ also satisfies $f(x+y)+f(x-y)=2f(x)f(y)$.

This suggests one to ask several questions.

Question 1. What is the dimension of the vector space $V_{2,2,1,1}$?

Question 2. Given a closed orientable surface Σ and a function $f: \mathcal{S} \to \{t \in \mathbf{R} | t > 1\}$ satisfies the relation $f(\alpha\beta)+f(\beta\alpha)=2f(\alpha)f(\beta)$ whenever $\alpha \perp \beta$, is it true that $f = f_{[d]}$ for some hyperbolic metric d in the surface?

It can be shown easily that there are non-constant solutions to the relation $f(\alpha\beta)+f(\beta\alpha)=2f(\alpha)f(\beta)$ which take some values equal to one.

In view of the linear skein relations for Jones type knot invariants, one may call a function in $V_{2,2,1,1}$ a two-dimensional "Jones invariant". More generally, one may define a 2-dimensional "Jones invariants" as follows. Take a finite collection of classes α_1,..., α_k in $\mathcal{S}$ (or more generally the set of all isotopy classes of simple closed curves in the surface) and finite collection of non-zero numbers d_1, ..., d_k. A 2-dimensional "Jones invariant" is a function $f: \mathcal{S} \to \mathbf{C}$ satisfying a linear skein relation:

$$\sum_{i=1}^{k} d_i f(\phi_*(\alpha_i)) = 0 \tag{6}$$

for all ϕ in the mapping class group of the surface. As an example, take three pairwise orthogonal simple closed curves a, b and c so that $a \cap b \cap c \neq \phi$ and $a \cup b \cup c$ is not in a torus with a hole. Let $\alpha = [a]$, $\beta = [b]$ and $\gamma = [c]$. Clearly each function f in $V_{2,2,1,1}$ satisfies the relation (7) below.

$$f(\alpha\beta)+f(\beta\alpha)+f(\beta\gamma)+f(\gamma\beta)+f(\gamma\alpha)+f(\alpha\gamma)=4f(\alpha)+4f(\beta)+4f(\gamma) \tag{7}$$

Is it true that the set of all solutions to the relation (7) forms a finite dimensional vector space?

Finally, suppose the surface Σ is closed and $\mathcal{G}$ is the mapping class group of the surface. If each function f in $V_{2,2,1,1}$ comes from the square of an algebraic intersection number, then the canonical Miller-Morita-Mumford class in $H^2(\mathcal{G})$ comes from the finite dimensional linear representation $V_{2,2,1,1}$. One may ask whether each Miller-Morita-Mumford class (or more generally each class in the cohomology group of $\mathcal{G}$) in $H^{2k}(\mathcal{G})$ comes from a finite dimensional linear representation of the group $\mathcal{G}$.

Acknowledgement. I would like to thank X.-S. Lin for many helpful discussions and Peter Landweber for careful reading of the manuscript. This work is supported in part by the NSF.

References

[De] Dehn, M., *Papers on group theory and topology, J. Stillwell, ed.*, Springer-Verlag, New York, 1987.

[FLP] Fathi, A., Laudenbach, F., Poènaru, V., *Travaux de Thurston sur les surfaces*, Astérisque (1979), 66-67.

[Li] Lickorish, W., *A representation of oriented combinatorial 3-manifolds*, Ann. Math. **72** (1962), 531-540.

[Hu] Humphries, *Generators for the mapping class group*, Lecture Notes in Math. **722** ((1979)), Springer, Berlin, 44-47.

[Lu1] Luo, F., *On non-separating simple closed curves in a compact surface, preprint.*

[Lu2] Luo, F., *On the mapping class groups of compact surfaces, in preparation.*

Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903

E-mail address: fluo@math.rutgers.edu