J. DIFFERENTIAL GEOMETRY
35 (1992) 689-710

ISOSPECTRALITY IN THE FIO CATEGORY

STEVEN ZELDITCH

0. Introduction

Compact riemannian manifolds (M_1, g_1), resp. (M_2, g_2), are called isospectral if there exists a unitary operator $U: L^2(M_1) \to L^2(M_2)$ which intertwines their Laplacians: $U\Delta^{(1)}U^* = \Delta^{(2)}$. At this time, quite a variety of (nonisometric) isospectral pairs have been constructed. On the other hand, all of these pairs are quite special: to the author's knowledge, each known pair has a common riemannian cover, and frequently a common quotient. These observations raise the questions:

(Q1)—Are isospectral manifolds locally isometric? Do they have a common riemannian cover?[1]

(Q2)—Is a generic metric spectrally determined (i.e., not nontrivially isospectral to another)? Is a metric with simple length spectrum spectrally determined?

There exist few positive results on these problems at present. Our purpose in this paper is to show that they can be solved (affirmatively) if we restrict the isospectral problem to the FIO (Fourier Integral Operator) category. At least, we will show this for (M, g) of dimension $d = 2$ and curvature $K < 0$. These dimension and curvature restrictions represent the current state of knowledge on the isometry problem for conjugate geodesic flows ([3], [4], [17]; see below); they should become relaxed as this knowledge develops further.

Isospectral Laplacians Δ_1 and Δ_2 will be called isospectral in the FIO category (or, Fourier-isospectral for short) if there exists a unitary FIO U intertwining them as above. More precisely, U will be assumed to lie in the Hörmander space $I^0(M_1 \times M_2, C)$ for some closed, embedded canonical relation $C \hookrightarrow \dot{T}^*M_1 \times \dot{T}^*M_2^-$, such that $C \circ C^t$ is a clean composition (see §1). To prevent confusion, we emphasize that C is not assumed to be the graph of a symplectic diffeomorphism (even locally). Indeed, our first step (§2–3) will be to characterize the canonical relations

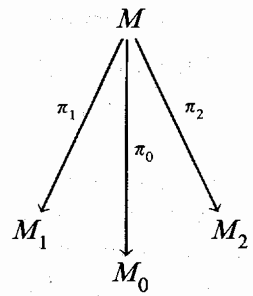
Received April 17, 1990 and, in revised form, March 7, 1991.

[1] A counterexample has recently been found by C. Gordon.

underlying unitary FIO's and in particular those FIO's which intertwine a pair of Laplacians.

Our original motivation for studying Fourier-isospectrality came from the observation (with A. Uribe) that Sunada's isospectral Laplacians can be intertwined by unitary FIO's. To give some idea of the kinds of canonical relations that come up in isospectral theory, let us recall that his isospectral pairs (M_1, M_2) fit into a diagram

(0.1)



of finite normal covers. Let H_i be the covering group for π_i, and G the covering group for π_0. Sunada observes that if $L^2(G/H_1)$ and $L^2(G/H_2)$ are unitarily equivalent G-modules, then, for any metric g_0 on M_0, $\pi_1^*(g_0)$ will be isospectral to $\pi_2^*(g_0)$. We add the following observation: from a unitary intertwining kernel $A(g)$ between these modules, one can construct such a kernel between the Laplacians (§5). The resulting operator is essentially just the weighted sum $\sum_{g\in G} A(g)\pi_{2*}T_g\pi_1^*$ of Radon transforms between M_1 and M_2 (T_g is the translation associated to g). The corresponding canonical relation is thus the union (for $g \in G$) of the conormal bundles $N^*(\text{graph}(\pi_2 \circ g \circ \pi_1^{-1}))$ to graphs of the indicated correspondences.

Sunada's examples form in a certain sense the main class of known isospectral pairs: for, the pairs come in families of positive functional dimension equal to the dimension of the M_i. Moreover, the metric need not be locally homogeneous, or have any local isometrics. By comparison, the other known examples still use rather special metrics (e.g., flat [16], spherical [12], hyperbolic [19], or (partially) locally solvable [8], [5]).

The next most robust examples are those of DeTurck-Gordon (especially [5]) and Gordon-Wilson. In particular, DeTurck-Gordon construct isospectral pairs (in fact, continuous families) of quotients $\widetilde{M}/\Gamma$, where $\widetilde{M}$ carries an action by a nilpotent Lie group G such that G/Γ and $\widetilde{M}/\Gamma$ are compact. The metric on $\widetilde{M}$ only needs to be invariant under a certain subgroup ΓH (see [5, Proposition 4.1]). Hence, their examples also come in families of positive functional dimension, although not of the dimension of $\widetilde{M}$.

DeTurck-Gordon explicitly construct intertwining operators between their Laplacians [5, Theorem 2.1]. Recently, F. Marhuenda made a microlocal analysis of (at least some of) these intertwining operators [15]. They turn out to be a singular FIO's associated to cleanly intersecting canonical relations in the sense of Guillemin, Melrose, and Uhlmann. In particular, they have well-defined and composable principal symbols.

Thus, many of the robust (i.e., highly deformable) known isospectral pairs live in the FIO category—broadly enough interpreted. We do not presently know which of the other examples are Fourier-isospectral (although the transplantation examples of Buser and Berard almost certainly are). However, the remaining examples appear to be isolated among isospectral pairs, and hence may be considered sporadic. So, at least according to our present knowledge, Fourier-isospectrality provides a kind of boundary between generic and sporadic isospectralities. It would be very desirable to have an a priori understanding of this (i.e., not confined to studying examples). As we will see, the fundamental isospectral problems can, to a large degree, be reduced to this question of how generically an isospectrality is Fourier.

Let us now turn to the main results of the this paper. In answer to (Q1), we have:

(4.1) **Theorem.** *Let (M_1, g_1) and (M_2, g_2) be a pair of Fourier-isospectral surfaces. If (M_1, g_1) is nonpositively curved, then the (M_i, g_i) posses a common, finite riemannian cover.*

In answer to (Q2), we have:

(4.2) **Theorem.** *Let (M_1, g_1) be a negatively curved surface with simple length spectrum. If (M_1, g_1) is Fourier isospectral to (M_2, g_2), then it is isometric to (M_2, g_2).*

(Here, $\mathrm{Lsp}(M, g)$ is the length spectrum: the set of lengths of closed geodesics. Simplicity means at most one geodesic has a given length.)

The proofs of these theorems contain two main ingredients. The first is a symbolic analysis of Fourier-isospectrality (§§1–3). In the case of surfaces, our result is:

Lemma (see Corollary 3.7(b) and Proposition 3.8). *Let (M_1, g_1) and (M_2, g_2) be Fourier-isospectral compact surfaces. Then:*

(i) *there exists a common finite cover $p_i: M \to M_i$;*

(ii) *there is a common cyclic cover $q_i: Q \to S_i^* M$ ($S_i^* M$ being the unit cotangent bundle for $p_i^*(g_i)$), such that q_i only unwinds the circles $S_i^* M_m$;*

(iii) *there is a diffeomorphism* $\Phi: Q \to Q$ *so that the correspondence* $q_1 \circ \Phi \circ q_1^{-1}$ *conjugates the geodesic flows* $\tilde{G}_i^t$ *of* $p_i^*(g_i)$.

Thus, Fourier-isospectral compact surfaces nearly have smoothly (even symplectically) conjugate geodesic flows: the flows are conjugate up to certain finite cyclic covers.

The second main point is to determine when such near conjugacy implies local isometry. The crucial ingredients here are the recent results of Croke [3], Croke-Fathi-Feldman [4], and Otal [17] on the marked length spectrum of a nonpositively curved surface. This is the function L_g: $\hat{\pi}_1(M_1) \to \mathbf{R}^+$ on free homotopy classes of loops, which assigns to $\hat{\gamma} \in \hat{\pi}_1(M)$ the common length $L(\gamma)$ of the closed geodesics γ for g in $\hat{\gamma}$ (γ is unique if $K < 0$). We will use:

Theorem A [4]. *Let* M *be a closed surface and* g_1, g_2 *metrics on* M, *with* g_1 *of nonpositive curvature and* g_2 *without conjugate points. If* g_1 *and* g_2 *have the same marked length spectrum, then they are isometric.*

As we will see, the Lemma implies that if (M_1, g_1) is negatively curved (say) and (M_2, g_2) is Fourier-isospectral to (M_1, g_1), then (M_2, g_2) has no conjugate points and has the same marked length spectrum as (M_1, g_1). Hence Theorem A will imply Theorem 4.1.

In sum, our point in this paper is that many of the principal questions in isospectral theory (such as (Q1) and (Q2) can be reduced, at least for broad enough classes of metrics, to the solvability of the isospectral equation

$$(0.2) \qquad \begin{cases} (\Delta_x^{(1)} - \Delta_y^{(2)})U(x, y) = 0, \\ U^*U - I = 0 \end{cases}$$

by Lagrangian distribution $U \in \mathscr{D}'(M_1 \times M_2)$. Actually, since our results depend only on a symbolic analysis, it would suffice to solve (0.2) to leading order.

The main problems suggested by this work seem to be the following: First, for what class of metrics does symbolic isospectrality imply local isometry? (Note that Zoll spheres are always symbolically isospectral [20].) Second, how generically are isospectral pairs generically Fourier-isospectral? Third, does isospectrality generically imply the microlocal solvability of (0.1) along products of $\alpha \times \beta$ of closed geodesics of $M_1 \times M_2$ with $L(\alpha) = L(\beta)$? (In other words, can one find a solutions $U_{\alpha\times\beta}$ so the left sides in (0.2) have wavefront set disjoint from $\alpha \times \beta$, resp. $\alpha \times \alpha$, $\beta \times \beta$.) This question is closely related to Weinstein's conjecture that the spectrum determines the Birkhoff-Moser canonical forms for the Poincaré maps associated to each closed geodesic γ (see [7]).

1. Fourier-isospectrality and symbolic isospectrality

Recall that the respective Laplacians $\Delta^{(1)}$ and $\Delta^{(2)}$ of compact riemannian manifolds (M_1, g_1) and (M_2, g_2) are *Fourier-isospectral* if there exists an FIO $U: L^2(M_1) \to L^2(M_2)$ so that

$$\text{(1.1)} \qquad \begin{aligned} &\text{(i)} \quad U\Delta^{(1)}U^* = \Delta^{(2)}, \\ &\text{(ii)} \quad UU^* \quad = U^*U = \mathrm{Id}. \end{aligned}$$

By FIO, we mean that the (Schwartz) kernel $U(x, y)$ lies in a space $I^0(M_1 \times M_2, C)$ for some closed, embedded, canonical relation $C \hookrightarrow \dot{T}^*M_1 \times \dot{T}^*M_2^-$. C is understood to be homogeneous (i.e., to be invariant under the free $\mathbb{R}^+$-action on $\dot{T}^*M_1 \times \dot{T}^*M_2$). We will also assume that $C \circ C^t$ and $C \circ C^t$ are clean compositions [11, III, 21.2.14]. Here, $C^t = \{(y, \eta, x, \xi): (x, \xi, y, \eta) \in C\}$ is the transposed canonical relation of C.

We have departed here from the notation C^{-1} in [11, IV, 25.2] to emphasize that $C \circ C^t$ need not be the diagonal relation. We will depart from the customary notational conventions of FIO theory in a few other ways as well. For one, we will view $C = \dot{T}^*M_1 \times \dot{T}^*M_2$ as a relation (or correspondence) from $\dot{T}^*M_1$ to $\dot{T}^*M_2$ rather than the reverse. Hence, we will compose relations in the usual set-theoretic way: relations $C_1 \subseteq T^*X \times T^*Y$ and $C_2 \subseteq T^*Y \times T^*Z$ will compose as $C_2 \circ C_1 \subset T^*X \times T^*Z$. Further, we will not twist canonical relations as in [11, III, 21,2,9] or [11, IV, 25.2]. These and future departures are necessary in order to conform to conventions standard outside of FIO theory, and hopefully are transparent enough not to cause confusion.

The principal symbol of U will be denoted by σ_U. It is a section of $\Omega_C^{1/2} \otimes M_C$, where $\Omega_C^{1/2}$ is the bundle of $1/2$-densities on C and M_C is the Maslov bundle (a flat, trivializable hermitian line bundle over C). Our assumptions that U has order 0 means that σ_U is homogeneous of order $m/2$, where $m = \dim M_1$ $(= \dim M_2)$; i.e., $\sigma_U \in S^{m/2}(C, \Omega_C^{1/2} \otimes M_C)$.

The isospectral equations (1.1) imply a corresponding set of equations for the principal symbol data (C, σ_U). To state them, we first introduce some terminology and notation. G_i^t will denote the geodesic flow on $\dot{T}^*M_i$ $(= T^*M_i \backslash 0)$ generated by the norm function $|\xi|_i$ of the metric. The product flow $G_1^t \times G_2^{-t}$ is then the flow on $\dot{T}^*M_1 \times \dot{T}^*M_2^-$ of the Hamilton vector field H_f of the difference Hamiltonian $f(x_1, \xi_1, x_2, \xi_2) = |\xi_1|_1 - |\xi_2|_2$ (recall that $\dot{T}^*M^-$ is $\dot{T}^*M$ equipped with $-\omega$, ω being the canonical symplectic form). Further, $\circ$ denotes composition for canonical relations or symbols, $\overline{\sigma}^t_U$ denotes the adjoint symbol, $\Delta_{\dot{T}^*M}$ is the diagonal in $\dot{T}^*M \times \dot{T}^*M$, and μ is its canonical 1/2-density.

The pair (C, σ_U) determines what Guillemin-Uribe and Weinstein call a morphism in the symplectic category: this is the category whose objects are symplectic manifolds X, Y and whose morphisms are canonical relations $C \subset X \times Y^-$ equipped with 1/2-densities $\sigma \in C^\infty(\Omega_C^{1/2})$ ([10], [21]). Following their terminology, we will have:

(1.2) **Definition.** A morphism (C, σ) from $\dot{T}^*M_1$ to $\dot{T}^*M_2$ is *unitary* if $C \circ C^t$ and $C^t \circ C$ are clean compositions, and if

$$\text{(i)}\quad \Delta_{\dot{T}^*M_1} \subset C^t \circ C,$$

$$\text{(ii)}\quad \overline{\sigma}^t_U \circ \sigma_U = \begin{cases} \mu_1 & \text{on } \Delta_{\dot{T}^*M_1}, \\ 0 & \text{on } C^t \circ C\Delta_{\dot{T}^*M_1} \end{cases}$$

(similarly for $C \circ C^t$ and $\sigma_U \circ \overline{\sigma}^t_U$).

The simplest example of a unitary morphism is the graph Γ_χ of a symplectic diffeomorphism $\chi: \dot{T}^*M_1 \to \dot{T}^*M_2$, equipped with its natural graph 1/2-density (χ and all other maps are understood to be homogeneous). The intertwining operators in §0 provide other examples (see §5).

We will also have:

(1.3) **Definition.** A morphism (C, σ) from $\dot{T}^*M_1$ to $\dot{T}^*M_2$ is an *intertwining morphism* between the geodesic flows G_i^t n $\dot{T}^*M_i$ if:

(i) $(G_1^t \times G_2^{-t})(\operatorname{supp}\sigma_U) = \operatorname{supp}\sigma_U$.

(ii) $(G_1^t \times G_2^{-t})^*(\sigma_U) = \sigma_U$.

A special case is again the graph Γ_χ of a symplectic diffeomorphism χ, such that $\chi \circ G_1^t \circ \chi^{-1} = G_2^t$. There are other examples (see §3).

Finally, we will have:

(1.4) **Definition.** Compact riemannian manifolds (M_1, g_1) and (M_2, g_2) are *symbolically isospectral* if there exists a unitary morphism (C, σ) which intertwines their geodesic flows.

We then have the simple

(1.5) **Proposition.** *Fourier-isospectrality implies symbolic isospectrality.*

Proof. Let U be the unitary intertwining operator in (1.1). Modulo one technical problem, (1.1)(ii) immediately implies that (C, σ_U) is a unitary morphism. The technical problem is that $C^t \circ C$ and $C \circ C^t$ need not be embedded relations, so the usual definitions of $1/2$-densities on them and of the composition formula for $\overline{\sigma}^t_U \circ \sigma_U$ need to be modified (compare [11, IV, 25.2.3]). This complication occurs in the Sunada example, so is quite essential. We will deal with it in the appendix to this section (see (A1.8)).

Next, rewrite (1.1)(i) in the form $(\Delta^{(1)}_x - \Delta^{(2)}_y)U(x, y) = 0$. View $(\Delta^{(1)}_x - \Delta^{(2)}_y)U(x, y)$ as the composition of a ΨDO on $L^2(M_1 \times M_2)$ and an FIO from $\mathbb{C}$ to $L^2(M_1 \times M_2)$. As an FIO of order 2, its principal symbol is $f\sigma_U \in S^{m/2+2}(C, \Omega^{1/2}_C \otimes M_C)$. So $f\sigma_U = 0$ and, since C is Lagrangian, H_f must be tangent to C on $\operatorname{supp}(\sigma_U)$.

As an FIO of order 1, its principal symbol is $i^{-1}\mathscr{L}_{H_f}(\sigma_U)$ (see [11, IV, 25.2.4]; note that the subprincipal symbol of $\Delta^{(1)}_x - \Delta^{(1)}_x$ is zero). Hence $\mathscr{L}_{H_f}(\sigma_U) = 0$, proving 1.3(i)–(ii).

Remarks. (1) Observe that we have not assumed $C = \operatorname{supp}(\sigma_U)$. This temporarily leaves open the possibility that $\operatorname{supp}(\sigma_U)$ might be any closed invariant subset of C for $G^t_1 \times G^{-t}_2$ which can support a smooth function. Actually, we will show in §2 that the unitarity condition forces $\operatorname{supp}(\sigma_U)$ to be a closed Lagrangian manifold without boundary. At that point, it will be most sensible to require $C = \operatorname{supp}(\sigma_U)$.

(2) Suppose conversely that (C, σ) is a symbolic isospectrality between (M_1, g_1) and (M_2, g_2). Then $U\Delta_1 U^* - \Delta_2$ is an FIO of order 0 for any $U \in I^0(M_1 \times M_2, C)$ with $\sigma_U = \sigma$. In some cases, this conclusion can be significantly improved. For example, Weinstein has proved that if C is the graph of a symplectic diffeomorphism, then $|\lambda_n(M_1, g_1) - \lambda_{n+k}(M_2, g_2)| = O(1)$, as $n \to \infty$ for some integer k [20]. Here $k = \operatorname{ind}(U)$ is the index of U (completely mysterious at present). Weinstein's proof does not immediately generalize to C which are not graphs.

Appendix to §1

We need to discuss symbol composition when the various simplifying assumptions in [11, III, 25.2.3] and elsewhere are dropped. Hopefully, our discussion will also make §§2–3 accessible to those not already familiar with FIO theory.

Let $X_j = \dot{T}^* M_j$, and assume for simplicity that $\dim M_j = m$ $(j = 1, 2, 3)$. Also, let $C_j \hookrightarrow X_j \times X^-_{j+1}$ be a pair of closed, embedded, canonical relations $(j = 1, 2)$. The composition $C_2 \circ C_1 \subset X_1 \times X^-_3$ is just the usual set-theoretic composition of relations [11, III, 21.2.12]). It is said to be *clean* if the following fiber product is clean:

(A1.1)
$$\begin{array}{ccc} C_1 & \longleftarrow & F \\ \pi_1 \downarrow & & \downarrow \\ X_2 & \underset{\pi_2}{\longleftarrow} & C_2 \end{array}$$

where $F = \{(c_1, c_2) \in C_1 \times C_2 : \pi_1(c_1) = \pi_2(c_2)\}$, and $\pi_2 : X_j \times X_{j+1} \to X_2$ is the natural projection. Cleanliness of (A1.1) means that F is a disjoint union $\bigsqcup_j F_j$ of closed, embedded submanifolds of $C_1 \times C_2$ (of possibly varying dimensions d_j), and that the tangent diagram at each $f \in F$ is also a fiber product.

Now let $p: F \to C_2 \circ C_1$ be the natural projection; i.e., the restriction to F of the projection $\pi_1 \times \pi_3 : X_1 \times X_2^- \times X_2 \times X_3^- \to X_1 \times X_3^-$ onto the outer factors. If (A1.1) is clean, then p is a map of constant rank $2m$ (indeed, $dp_f(T_f F)$ is always Lagrangian). Hence p is a local fibration to its image (compare [11, III, 21.2.14]).

In general, p will fail to be a global fibration due to self-intersections in $C_2 \circ C_1$ (example: the Sunada intertwining relations). In order to compose symbols, we will require that these self-intersections be clean. More precisely, let $\{V_j\}$ be a finite (homogeneous) cover of F so that $p|_{V_j}$ is a fibration onto its image (note that $F/\mathbb{R}^+$ is compact). The images $B_j \stackrel{\text{def}}{=} p(V_j)$ are then open, embedded submanifolds of $X_1 \times X_3$, whose union is $C_2 \circ C_1$. We will refer to them as the "branches" of $C_2 \circ C_1$ (relative to the cover).

In general, let us call a map $\varphi: M \to N$ of constant rank between two manifolds a *clean local fibration* (CLF) if there is a cover of M for which the associated branches B_j intersect cleanly (i.e., $B_j \cap B_k$ is a submanifold of N and $T_b(B_j \cap B_k) = T_b(B_j) \cap T_b(B_k)$). We then say:

(A1.2) **Definition.** The composition $C_2 \circ C_1$ is *extra-clean* if (A1.1) is clean, and $p: F \to C_2 \circ C_1$ is a CLF.

With this assumption, the tangent planes to the branches of $C_2 \circ C_1$ never coincide. Hence the manifold $\Lambda_{C_2 \circ C_1}$ of such tangent planes is an embedded submanifold of the Lagrangian Grassmannian $\Lambda(X_1 \times X_3^-)$. Here, for any symplectic manifold S, $\Lambda(S)$ is the bundle over S whose fiber at $s \in S$ is the Grassmannian $\Lambda(T_s S)$ of Lagrangian planes of $T_s S$.

The natural projection from $\Lambda(X_1 \times X_3^-)$ to $X_1 \times X_3^-$ restricts to $\Lambda_{C_2 \circ C_1}$ to determine an immersion $i_{C_2 \circ C_1}\colon \Lambda_{C_2 \circ C_1} \to X_1 \times X_3^-$. It is the parametrization of $C_2 \circ C_1$ by its tangent planes.

Let Λ_{C_1}, resp. Λ_{C_2}, similarly denote the manifold of tangent planes to C_1, resp. C_2. The corresponding maps i_{C_j} are now diffeomorphisms. Hence, we may view the fiber product F above as a submanifold of $\Lambda_{C_1} \times \Lambda_{C_2}$. We may also factor the projection p as $i_{C_2 \circ C_1} \circ \psi$, where:

(A1.3) **Definition.** $\psi\colon F \to \Lambda_{C_2 \circ C_1}$ is the map $\psi(\lambda_1, \lambda_2) = \lambda_2 \circ \lambda_1$ (i.e., the composition of these subspaces of $T_{(x_1, \xi_1, x_2, \xi_2)}(X_1 \times X_2^-)$, resp. $T_{(x_2, \xi_2, x_3, \xi_3)}(X_2 \times X_3^-)$ [11, III, 21.2.12]).

If $C_2 \circ C_1$ is extra clean, then each $\psi|_{F_j}$ is a fibration to its image.

We now define a composition law for 1/2-densities: it is a natural bilinear map

$$\text{(A1.4)} \qquad \circ\colon \Omega^{1/2}_{C_2} \otimes \Omega^{1/2}_{C_1} \to \Omega^{1/2}_{\Lambda_{C_2 \circ C_1}}.$$

First, identify $\Omega^{1/2}_{C_j}$ with $\Omega^{1/2}_{\Lambda_{C_j}}$. A 1/2-density σ_j on C_j is thus a family $\{\sigma_j(\lambda)\}$ of 1/2-densities, with $\sigma_j(\lambda) \in |\lambda|^{1/2}$ ($|W|^s$ denotes the space of s-densities on a vector space W). The exterior tensor product $(\sigma_2 \boxtimes \sigma_1)_{(\lambda_1 \times \lambda_2)}$ is then an element of $|\lambda_1 \times \lambda_2|^{1/2}$. In a natural way, it determines a gadget $(\sigma_2 \boxtimes \sigma_1)_{(\lambda_1 \times \lambda_2)} \in |V_{\lambda_1 \times \lambda_2}| \otimes |\lambda_2 \circ \lambda_1|^{1/2}$, where $\lambda_1 \times \lambda_2 \in F$, and where $V_{\lambda_1 \times \lambda_2}$ is the vertical subspace of $T_{\lambda_1 \times \lambda_2} F$ (tangent to the fibers of ψ). Since it plays an important role in §§2–3, we give a brief and rather plebian description of it (see [6, §5] or [11, III, §25] for more details).

Let $S_j = T_{(x_j, \xi_j)} X_j$ $(j = 1, 2, 3)$. Then $\lambda_1 \times \lambda_2 \subset S_1 \times S_2^- \times S_2 \times S_3^-$ and $T_{\lambda_1 \times \lambda_2} F$ is the subspace of vectors (u, v, v, w). The fiber F_λ over $\lambda \in \Lambda_{C_2 \circ C_1}$ is the set $\{\lambda_1 \times \lambda_2 \colon \lambda_2 \circ \lambda_1 = \lambda\}$, and its tangent space $V_{\lambda_1 \times \lambda_2}$ is the space of $(0, v, v, 0)$'s. Under $(0, v, v, 0) \mapsto v$, it may be identified with a subspace $V \subset S_2$.

Let $\tau\colon \lambda_1 \times \lambda_2 \to S_2$ be the map $\tau(u, v_1, v_2, w) = v_2 - v_1$. Also let $\alpha\colon T_{\lambda_1 \times \lambda_2} F \to \lambda_2 \circ \lambda_1$ be $\alpha(u, v_1, v_1, w) = (u, w)$. Using that λ_j is Lagrangian in $S_j \times S_{j+1}^-$ one easily shows that $V = (\operatorname{im} \tau)^\perp$ [6, §5]. So the symplectic form ω_2 of S_2 defines a nonsingular pairing between V and $S_2 / \operatorname{im} \tau$.

We now define $(\sigma_2 \boxtimes \sigma_1)_{(\lambda_1 \times \lambda_2)}$ for a basis (ν, γ) of $V_{\lambda_1 \times \lambda_2} \times \lambda_2 \circ \lambda_1$. Here ν is a basis $\{(0, v_i, v_i, 0) \colon i = 1, \cdots, e\}$ for $V_{\lambda_1 \times \lambda_2}$, corresponding to a basis $\nu = \{v_i\}$ of V, and $\gamma = \{(u_i, w_i), i = 1, \cdots, 2m\}$ is a basis of $\lambda_2 \circ \lambda_1$.

First, lift γ to $\overline{\gamma} \overset{\text{def}}{=} \{(u_i, 0, 0, w_i)\} \subset T_{\lambda_1 \times \lambda_2} F$, so that $\{\nu, \overline{\gamma}\}$ is a basis for $T_{\lambda_1 \times \lambda_2} F$. Choose a partial basis $\beta = \{(u_k, v_{2_k}, v_{1_k}, w_k), k = 1, \cdots, 2m - e\}$ of $\lambda_1 \times \lambda_2$ so that $\boldsymbol{\beta} \overset{\text{def}}{=} \{v_{2_k} - v_{1_k}\}$ is a basis for $\operatorname{im} \tau$. Then $(\nu, \overline{\gamma}, \beta)$ is a basis for $\lambda_1 \times \lambda_2$. $(\sigma_2 \boxtimes \sigma_1)_{(\lambda_1 \times \lambda_2)}(\nu, \overline{\gamma}, \beta)$ is then well defined but depends on β. To cancel this dependence, we let $\gamma^* = \{v_j^* : j = 1, \cdots, e\}$ be elements of S_2 so that $\omega_2(v_i, v_j^*) = \delta_{ij}$. γ^* is uniquely determined modulo $\operatorname{im} \tau$. Then set:

(A1.5) **Definition.**

$$(\sigma_2 \times \sigma_1)_{(\lambda_1 \times \lambda_2)}(\nu, \gamma) = \frac{\sigma_2 \boxtimes \sigma_1(\nu, \overline{\gamma}, \beta)}{|\omega_2^m|^{1/2}(\nu^*, \boldsymbol{\beta})}.$$

The right side is independent of β, and defines a mixed density in $|V_{\lambda_1 \times \lambda_2}| \otimes |\lambda_2 \circ \lambda_1|^{1/2}$.

Finally, $(\sigma_2 \circ \sigma_1)_\lambda \in |\lambda|^{1/2}$ is given by:

(A1.6) **Definition.** $(\sigma_2 \circ \sigma_1)_\lambda = \int_{F_\lambda} (\sigma_2 \times \sigma_1)_{\lambda_1 \times \lambda_2}$ $(\lambda \in \Lambda_{C_2 \circ C_1})$.

Extending this composition law is a natural bilinear map (Symbol composition):

$$\text{(A1.7)} \qquad \circ : (\Omega_{C_2}^{1/2} \otimes M_{C_2}) \times (\Omega_{C_1}^{1/2} \otimes M_{C_1}) \to \Omega_{\Lambda_{C_2 \circ C_1}}^{1/2} \otimes M_{\Lambda_{C_2 \circ C_1}},$$

where M is the Maslov line bundle. $M_{\Lambda_{C_2 \circ C_1}}$ is defined precisely as in the embedded case [9, IV], as is the identity $i^*(M_{C_2} \boxtimes M_{C_1}) \cong \psi^*(M_{\Lambda_{C_2 \circ C_1}})$, where $i : F \hookrightarrow \Lambda_1 \times \Lambda_2$ is the inclusion (cf. [6, 5.3]). The resulting formula for $\circ$ is just as in (A1.6) except that σ_j is replaced by $\sigma_j \otimes r_j$ (r_j being a Maslov factor), and $\sigma_2 \times \sigma_1$ is replaced by $(\sigma_2 \times \sigma_1) \otimes i^*(r_2 \boxtimes r_1)$. In the future, σ_j will denote a (1/2-density) $\otimes$ (Maslov factor), and the formula in Definition A1.6 will be used for principal symbol composition. (Principal symbols are homogeneous sections of these bundles.)

Now suppose $A_j \in I^0(M_j \times M_{j+1}, C_j)$ is a Lagrangian kernel ($j = 1, 2$), and suppose $C_2 \circ C_1$ is an extra clean composition. The composition kernel $A_2 \circ A_1(x, y)$ can thus be written as a (locally) finite sum of oscillatory integrals $I_j = \int \alpha_j e^{i\phi_j}$, where the phase functions ϕ_j parametrize the branches B_j of $C_2 \circ C_1$. The principal symbol of I_j is then a section of $\Omega_{B_j}^{1/2} \otimes M_{B_j}$. These local symbols piece together to form a global section $\sigma_{A_2 \circ A_1}$ of the bundle $\Omega^{1/2} \otimes M$ along the immersion $i_{C_2 \circ C_1}$. Hence, $\sigma_{A_2 \circ A_1}$

can be identified with a section of $\Omega^{1/2}_{\Lambda_{C_2 \circ C_2}} \otimes M_{C_2 \circ C_1}$. The usual composition formula, $\sigma_{A_2 \circ A_1} = \sigma_{A_2} \circ \sigma_{A_1}$, then holds in the sense of Definition A1.6 and (A1.7); indeed, it can be localized to open sets where $i_{C_2 \circ C_1}$ is an embedding, and hence can be reduced to the embedded case [11, IV, 25.2.3].

Finally, we complete the proof of Proposition 1.5:

(A1.8) **Addendum to Proposition 1.5.** The symbols $\sigma_{U^* U}$ and $\sigma_{U^*} \circ \sigma_U$ are now well defined as sections of $\Omega^{1/2}_{B_j} \otimes M_{B_j}$ over $\Lambda_{C_2 \circ C_1}$, and $\sigma_{U^* U} = \overline{\sigma}^t_U \circ \sigma_U$. We further transport μ_1 and $\Delta_{\dot{T}^* M_1}$ to the manifold $\Lambda_1 \subset \Lambda(x_1 \times X_1^-)$ of tangent planes to $\Delta_{\dot{T}^* M_1}$. The unitarity condition (1.1)(ii) then implies that, as symbols along submanifolds of $\Lambda(X_1 \times X_1^-)$, $\sigma_{U^* U} = \mu_1$. It follows that Λ_1 is a connected component of $\Lambda_{C_2 \circ C_1}$, making 1.2(i) more precise. It also follows that $\overline{\sigma}^t_U \circ \sigma_U = \mu_1$ on Λ_1, and $\overline{\sigma}^t_U \circ \sigma_U = 0$ on $\Lambda_{C_2 \circ C_1} \backslash \Lambda_1$, making 1.2(ii) more precise. Similarly for UU^*.

2. Unitary morphisms

A canonical relation $C \subset \dot{T}^* M_1 \times \dot{T}^* M_2$ will be called *unitarizable* if there exists a symbol $\sigma \in (\Omega^{1/2}_C \otimes M_C)$ so that $\operatorname{supp}(\sigma) = C$ and so that (C, σ) is a unitary morphism. What kinds of C are unitarizable? Canonical graphs Γ_χ clearly are, but the Sunada intertwining relations (see §§0 and 5) give nongraph examples. They are, however, local canonical graphs, and one might suspect that (at least for embedded C) they have to be.

(2.1) **Proposition.** *Let C be an embedded (but not necessarily connected) unitarizable canonical relation in $\dot{T}^* M_1 \times \dot{T}^* M_2$. Then the natural projections*

$$\begin{array}{ccccc} & & C & & \\ & \swarrow^{\pi_1} & & \searrow^{\pi_2} & \\ \dot{T}^* M_1 & & & & \dot{T}^* M_2 \end{array}$$

are finite, $\mathbb{R}^+$-homogeneous covers.

Proof. Let F be the fiber product in (A1.1) with $C_1 = C$ and $C_2 = C^t$. Thus, $F = \{(x_1, \xi_1, y, \eta : y, \eta, x_2, \xi_2) : (x_i, \xi_i, y, \eta) \in C\}$. As in the Appendix to §1, F is the disjoint union $\bigsqcup_j F_j$ of embedded submanifolds in $C \times C^t$, and the maps $\psi|_{F_j}: F_j \to \Lambda_{C^t \circ C}$ are fibrations (Definition A1.3).

Let $F_\Delta = (i_{C^t \circ C} \circ \pi)^{-1}(\Delta_{\dot{T}^* M_1})$, where $i_{C^t \circ C}$ is, we recall, the immersion $\Lambda_{C^t \circ C} \to X_1 \times X_1^-$ taking a tangent plane to its point of tangency. Thus, $F_\Delta = \{(x, \xi, y, \eta : y, \eta, x, \xi) : (x, \xi, y, \eta) \in C\}$, and it is obvious that π_1 is a finite cover if and only if $i_{C^t \circ C} \circ \psi : F_\Delta \to \Delta_{\dot{T}^* M_1}$ is one.

Since C is unitarizable, there is a unitary symbol σ on C with $\operatorname{supp}(\sigma) = C$. By Addendum A1.8, the diagonal Λ_1 is then a connected component of $\Lambda_{C^t \circ C}$. Let $F_\Delta^0 = \psi^{-1}(\Lambda_1)$, and let $\psi_\Delta^0 = \psi|_{F_\Delta^0}$. Then $\psi_\Delta^0 : F_\Delta^0 \to \Lambda_1$ is a fibration. The theorem clearly reduces to the

(2.2) *Claims.*

(i) ψ_Δ^0 is a finite cover.

(ii) $F_\Delta^0 = F_\Delta$.

Proof. (2.2)(i) A point $f \in F_\Delta$ is of the form $f = (z, z^t) \in C \times C^t$, where $z = (x, \xi, y, \eta)$ and $z^t = (y, \eta, x, \xi)$. Identifying $C \times C^t$ with $\Lambda_C \times \Lambda_{C^t}$ as in (A1), such an f corresponds to a product $\lambda_0 \times \lambda_0^t$, where $\lambda_0 = T_z C$ is a Lagrangian plane in $S_1 \times S_2^-$ ($S_1 = T_{(x,\xi)}(\dot{T}^* M_1)$, $S_2 = T_{(y,\eta)}(\dot{T}^* M_2)$). F_Δ^0 then consists of the $\lambda_0 \times \lambda_0^t$ in F_Δ satisfying $\lambda_0^t \circ \lambda_0 = \lambda_\Delta$, where $\lambda_\Delta \subset S_1 \times S_2^-$ is the diagonal plane.

As with any Lagrangian subspace $\lambda_0 \subset S_1 \times S_2$, there are symplectic orthogonal decompositions $S_1 = S_{11} \oplus S_{12}$ and $S_2 = S_{21} \oplus S_{22}$ so that $\lambda_0 = \lambda_{01} \oplus G_0 \oplus \lambda_{02}$, with λ_{0j} Lagrangian in S_{jj}, and with G_0 the graph of a symplectic linear map $S_{12} \to S_{21}$ [11, IV, 25.3.6]. For $\lambda_0 \times \lambda_0^t \in F_\Delta^0$, the only possibility is that $\lambda_0 = G_0$ and $\lambda_{01} = \lambda_{02} = \{0\}$. Consequently, the vertical space $V_{\lambda_0 \times \lambda_0^t}$ for ψ_Δ^0 is $\{0\}$: indeed, it is the diagonal in $\lambda_{02} \times \lambda_{02}$ (cf. Definition A1.3). Thus, ψ_Δ^0 is a proper local diffeomorphism, proving (i).

(2.2)(ii) Suppose to the contrary that $F_\Delta^0 \neq F_\Delta$, and let $\lambda_0 \times \lambda_0^t \in F_\Delta \backslash F_\Delta^0$. The unitary assumption on σ then implies that $\overline{\sigma}^t \circ \sigma$ must vanish on $\lambda_0 \times \lambda_0^t$. By Definition A1.6,

$$0 = \int_{F_{\lambda_0^t \circ \lambda_0}} (\overline{\sigma}^t \times \sigma)_{(\lambda \times \lambda^t)}(\cdot, \gamma) \tag{2.3}$$

for any basis γ of $\lambda_0 \times \lambda_0^t$. This leads to a contradiction, for any density of the form $(\overline{\sigma}^t \times \sigma)_{(\lambda \times \lambda^t)}(\cdot, \gamma)$ must be positive. Indeed, in view of Definition

A1.5, it suffices to show that $\overline{\sigma}^t \times \sigma$ is a positive density on any product $\lambda \times \lambda^t$. This possibility may be checked on any basis of $\lambda \times \lambda^t$; and of course we choose one of the form $\{(b, 0), (0, b^t)\}$, with b a basis of λ and b^t the corresponding one of λ^t (under the interchange map $s: S_1 \times S_2 \to S_2 \times S_1$). It is immediate from the definition of $\overline{\sigma}^t$ that $\overline{\sigma}^t(b^t) = \overline{\sigma}(b)$ (cf. [11, IV, 25.1.15 and 25.2.2]; σ^t is written $s^*\sigma^*$ there). Hence $\overline{\sigma}^t \times \sigma((b, 0), (0, b^t)) = |\sigma(b)|^2 > 0$, completing the proof of (2.3).

(2.4) **Corollary.** *Let $C \subset \dot{T}^*M_1 \times \dot{T}^*M_2$ be an embedded canonical relation, and let σ be a unitary symbol on C. Then* Supp(σ) *is a union of components of C.*

Proof. Supp(σ) is a finite, homogeneous cover of $\dot{T} * M_1$ and hence is a closed, boundaryless submanifold of C of full dimension. q.e.d.

This corollary explains Remark 1 of §1. Henceforth, a unitary morphism will be a pair (C, σ) as in Definition 1.2 with $C = \operatorname{Supp}\sigma$.

3. Unitary intertwining morphisms

A unitary morphism $C \subset \dot{T}^*M_1 \times \dot{T}^*M_2^-$ may be viewed as the graph, $C = \Gamma_\chi$, of a finitely multi-valued homogeneous symplectic correspondence $\chi: \dot{T}^*M_1 \to \dot{T}^*M_2$ ($\chi = \pi_2 \circ \pi_1^{-1}$ in the notation of Proposition 2.1). The invariance condition 1.3(i) on an intertwining morphism immediately translates into

$$\chi \circ G_1^t = G_2^t \circ \chi. \tag{3.1}$$

Thus, a UIM (unitary intertwining morphism) defines, up to some finite ambiguity, a symplectic conjugacy between the flows. We now resolve this ambiguity by passing to covers.

First, we give a more precise description of the covers $\pi_i: C \to \dot{T}^*M_i$ arising in Proposition 2.1).

(3.2) **Proposition.** *Let $\pi: C \to \dot{T}^*M$ be a finite, homogeneous cover. Then there exists a finite cover $p: \widetilde{M} \to M$ and a homogeneous cyclic cover $q: C \to \dot{T}^*M$ of $\dot{\mathbb{R}}^n$-bundles over $\widetilde{M}$ so that π factors as $C \xrightarrow{q} \dot{T}^*\widetilde{M} \xrightarrow{\tilde{p}} \dot{T}^*M$, where $\tilde{p}$ is the homogeneous cover induced by p, and q is a diffeomorphism if* $\dim M \geq 3$.

Proof. Let $\mathscr{V}$ be the foliation of $\dot{T}^*M$ by the (vertical) cotangent spaces $\dot{T}^*M$. The inverse image $\pi^{-1}\mathscr{V}$ is then a foliation of C by homogeneous manifolds. For each $L_{m,j} \in \pi^{-1}(\dot{T}_m^*M)$, $\pi: L_{m,j} \to \dot{T}_m^*M$ must be a homogeneous cover; so it is a cyclic of some degree d if $\dim M = 2$ or a diffeomorphism if $\dim M \geq 3$.

Let $\widetilde{M}$ be the leaf space $C/\pi^{-1}\mathscr{V}$. Since π is a homogeneous cover, $\widetilde{M}$ is a compact manifold, and the natural projection $\tilde{q}\colon C \to \widetilde{M}$ is an $\dot{\mathbb{R}}^n$-bundle. We may define $p\colon \widetilde{M} \to M$ so that the following diagram commutes:

$$\begin{array}{ccc} C & \xrightarrow{\pi} & \dot{T}^*M \\ {\scriptstyle \tilde{q}}\downarrow & & \downarrow \\ \widetilde{M} & \xrightarrow{p} & M \end{array} \tag{3.3}$$

It is easy to see that p is a cover; so it induces a homogeneous cover $\tilde{p}\colon \dot{T}^*\widetilde{M} \to \dot{T}^*M$.

Finally, we define a map $q\colon C \to \dot{T}^*\widetilde{M}$ so that the following diagram commutes:

$$\begin{array}{ccc} & C & \\ {\scriptstyle q}\swarrow & & \searrow{\scriptstyle \pi} \\ \dot{T}^*\widetilde{M} & \xrightarrow{\tilde{p}} & \dot{T}^*M \\ \downarrow & & \downarrow \\ \widetilde{M} & \xrightarrow{p} & M \end{array} \tag{3.4}$$

Precisely, for each $c \in C$, $\tilde{p}^{-1}(\pi(c))$ is a finite set $\{(\tilde{x}_j, \tilde{\xi}_j)\}$ of covectors in $\dot{T}^*\widetilde{M}$ with $\tilde{x}_j \neq \tilde{x}_k$ for $j \neq k$. We set $q(c) = (\tilde{x}_0, \tilde{\xi}_0)$, where $\tilde{x}_0$ is uniquely determined by $\tilde{q}(c) = \tilde{x}_0$. By construction, q is a homogeneous cover of $\dot{\mathbb{R}}^n$-bundles over $\widetilde{M}$. q.e.d.

When $C \subset \dot{T}^*M_1 \times \dot{T}^*M_2^-$ is a unitarizable canonical relation, Proposition 3.2 leads to the covering diagram

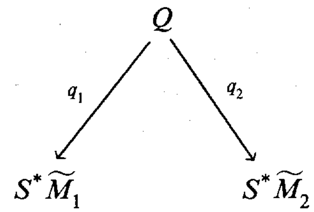
$$\begin{array}{ccc} & C & \\ {\scriptstyle q_1}\swarrow & & \searrow{\scriptstyle q_2} \\ \dot{T}^*\widetilde{M}_1 & & \dot{T}^*\widetilde{M}_2 \\ {\scriptstyle \tilde{p}_1}\downarrow & & \downarrow{\scriptstyle \tilde{p}_2} \\ \dot{T}^*M_1 & & \dot{T}^*M_2 \end{array} \tag{3.5}$$

with the q_i and $\tilde{p}_i$ as in Proposition 3.2. Each connected component of C gives rise to a similar diagram, so henceforth we fix one, say C_0. Via C_0 we then define the symplectic correspondence $\tilde{\chi} \stackrel{\text{def}}{=} q_2 \circ q_1^{-1} : \dot{T}^*\widetilde{M}_1 \to \dot{T}^*\widetilde{M}_2$ (where the q_i are restricted to C_0).

$\tilde{\chi}$ must be a diffeomorphism if $\dim M_i \geq 3$ (this need not imply the $\widetilde{M}_i$ are diffeomorphic [1]). $\tilde{\chi}$ may perhaps fail to be a diffeomorphism if $\dim M_i = 2$. However, the $\widetilde{M}_i$ must be surfaces of the same genus. This is obvious unless both surfaces have genus $g \geq 2$: For that case, we note that the center Z_i of $\pi_1(\dot{T}^*\widetilde{M}_i, (x_i, \xi_i))$ $(\simeq \pi_1(S^*\widetilde{M}, (x_i, \xi_i)))$ is generated by the class z_i of the fiber $(\dot{T}^*\widetilde{M}_i)_{(x_i, \xi_i)}$. Similarly, the center Z of $\pi_1(C, c)$ is generated by the class of the fiber of $C \to M_i$ (either projection). Since q is just unwinding the fibers of $\dot{T}^*\widetilde{M}_i \to \widetilde{M}_i$, the induced q_{i*} on π_1 takes Z to Z_i, and is an isomorphism from $\pi_1(C, c)/Z$ to $\pi_i(\dot{T}^*\widetilde{M}_i, (x_i, \xi_i))/Z_i$. But it is well known that this quotient is isomorphic to $\pi_1(\widetilde{M}_1, x_i)$.

Suppose now that C is a UIM between the geodesic flows G_i^t on $\dot{T}^*\widetilde{M}_i$. The metrics g_i on M_i lift to metrics $\tilde{g}_i$ on $\widetilde{M}_i$, and hence the G_i^t lift to geodesic flows $\widetilde{G}_i^t$ on $\dot{T}^*\widetilde{M}_i$. Obviously, $\tilde{\chi}$ conjugates the lifted flows. To put this conjugacy in a more familiar form, we slice the $\mathbb{R}^+$-action by defining $Q = C_0 \cap (S^*M_1 \times S^*M_2)$. Since the difference norm $f(x_1, \xi_1, x_2, \xi_2) = |\xi_1|_1 - |\xi_2|_2$ on $\dot{T}^*M_1 \times \dot{T}^*M_2$ vanishes on C_0 (Proposition 1.5), Q is just the hypersurface $\{|\xi_1|_1 = 1\}$ in C_0. The maps in (3.5) therefore restrict to Q to define a diagram

(3.6)



of compact covers. Equipping $S^*\widetilde{M}_i$ with its canonical contact form $\tilde{\alpha}^i (= \tilde{\xi}^i \, d\tilde{x}^i)$, (3.6) determines a contact correspondence, still denoted $\tilde{\chi}$, from $S^*\widetilde{M}_1 \to S^*\widetilde{M}_2$. From (3.1) we conclude:

(3.7) **Corollary.** *Suppose there exists a UIM C between the geodesic flows G_i^t on $\dot{T}^*M_1$. Then the following hold*:

(a) *If* $\dim M \geq 3$, *there must exist finite covers $p_i : \widetilde{M}_i \to M_i$ and a contact diffeomorphism $\tilde{\chi} : S^*\widetilde{M}_1 \to S^*\widetilde{M}_2$ so that $\tilde{\chi} \circ \widetilde{G}_1^t \circ \tilde{\chi}^{-1} = \widetilde{G}_2^t$.*

(b) *If* $\dim M = 2$, *there exist finite covers* $p_i\colon \widetilde{M}_i \to M_i$ *with* $\widetilde{M}_1 \approx \widetilde{M}_2$ (*diffeomorphic*). *Further, there exists a common connected cover* $q_i\colon Q \to S^*\widetilde{M}_i$ *so that* q_i *is a bundle map of* S^1*-bundles over* $\widetilde{M}_i$ *and so that the contact correspondence* $\tilde{\chi} \stackrel{\text{def}}{=} q_2 \circ q_1^{-1}\colon S^*\widetilde{M}_1 \to S^*\widetilde{M}_2$ *conjugates the flows* $\widetilde{G}_i^t$.

We can sharpen 3.7(b) if the metrics $\tilde{g}_i$ have the same area A_i. First, for simplicity we will henceforth denote $\widetilde{M}_1$ by M and will fix a diffeomorphism $\varphi\colon \widetilde{M}_1 \to \widetilde{M}_2$. We thus get two metrics, $\tilde{g}_1$ and $\varphi^*\tilde{g}_2$ on M, and hence two unit tangent bundles S_1^*M and S_2^*M (say). Replacing q_2 by $\varphi \circ q_2$, we also get a pair of covers $Q \to S_i^*M$ (which we will continue to denote by q_i; by abuse of notation, we will also denote $\varphi^*\tilde{g}_2$ by $\tilde{g}_2$).

(3.8) **Proposition.** *Suppose the metrics* $\tilde{g}_1$ *and* $\tilde{g}_2$ *on* M *have the same area. Then there is a contact diffeomorphism* $\Phi\colon Q \to Q$ *so that* $q_2 = q_1 \circ \Phi$. *Hence the flows* $\widetilde{G}_i^t$ *are conjugate via* $\tilde{\chi} = q_1 \circ \Phi \circ q_1^{-1}$.

Proof. First, $\deg(q_1) = \deg(q_2)$. Indeed, since C is homogeneous Lagrangian, the canonical 1-forms $\alpha^{(i)}$ on S_i^*M must pull back to the same 1-form $\alpha \stackrel{\text{def}}{=} q_i^*(\alpha^{(i)})$ on Q. Hence, $q_1^*(\alpha^{(1)} \wedge d\alpha^{(1)}) = q_2^*(\alpha^{(2)} \wedge d\alpha^{(2)})$. Since $\int_Q q_i^*(\alpha^{(i)} \wedge d\alpha^{(i)}) = 2\pi \deg(q_i) A_i$, equality of the A_i implies equality of the $\deg(q_i)$.

Next, both of the q_i are cyclic covers of S^1-bundles over M. Equality of the degrees $\deg(q_i)$ implies that the subgroups $q_{i*}(C, c)$ coincide. In the standard way, we path-lift the projection q_2 to an isomorphism $\Phi\colon Q \to Q$ of the covers. Since $q_2 = q_1 \circ \Phi$, Φ must be a contact diffeomorphism.

(3.9) **Corollary.** *Let* $\tilde{g}_1$ *and* $\tilde{g}_2$ *have the same area. Then the geodesic flows* $\widetilde{G}_i^t$ *are covered by contact flows* H_i^t *on* Q, *with* $H_2^t = \Phi \circ H_1^t \circ \Phi^{-1}$.

Proof. The Hamilton vector fields of the norm functions $|\eth|_i$ of $\tilde{g}_i$ lift under the q_i to contact vector fields Ξ_i on Q. Their flows H_i^t cover the $\widetilde{G}_i^t$ and are conjugate via Φ.

4. Proofs of Theorem 4.1 and 4.2

Proof of Theorem 4.1. We are given an FIO U conjugating the Laplacians, and hence a UIM C intertwining the geodesic flows G_i^t (see Proposition 1.5). The surfaces M_i therefore have a common cover M (see Corollary 3.7(b)). Further, the induced metrics $\tilde{g}_i$ on M must have the same area (the M_i, being isospectral, had the same genus and area).

Hence, the geodesic flows G_i^t on S_i^*M are conjugate via a contact corresponding $q_1 \circ \Phi \circ q_1^{-1}$, where $q_1 : Q \to S_1^*M$ is a finite cover which only unwraps the circles $S_1^*M_m$, and Φ is a contact diffeomorphism of Q (see Proposition 3.8). Alternatively, the G_i^t are covered by conjugate contact flows H_i^t on Q (see Corollary 3.9).

Now suppose that M has genus $g \geq 2$ and that g_1 is a metric of nonpositive curvature. In view of [4, Theorem A, §0], we most show that $(M_1, \tilde{g}_2)$ has no conjugate points and that Φ induces a bijection $\Phi_* : \hat{\pi}_1(M) \to \hat{\pi}_1(M)$ which presents lengths of closed geodesics.

Both steps are relatively straightforward from [3, I, Lemma 3.2]. We first observe that Φ induces an isomorphism Φ_* on $\pi_1(M)$. Indeed, as above, the fiber of $Q \to M$ (either projection) generates the center Z of $\pi_1(Q)$. The isomorphism induced by Φ on $\pi_1(Q)$ must take Z to Z, and hence it determines a quotient isomorphism on $\pi_1(M)$. It follows that Φ_* induces a bijection on $\hat{\pi}_1(M)$. We claim that it is length preserving on closed geodesics. Indeed, let γ be a closed geodesic of length $L(\gamma)$ for $(M_1, \tilde{g}_1)$. Lift it to S_1^*M as an orbit $(\gamma, \dot{\gamma})$ of G_i^t. Now, $S_1^*M|_\gamma$ (the unit cotangent bundle along γ) is a trivial S^1-bundle over γ. So is $Q|_\gamma$ (the inverse image of $S_1^*M|_\gamma$ under q_1). Further $q_1 : Q|_\gamma \to S_1^*M|_\gamma$ is just the standard d-fold cover on the second factor of $\gamma \times S^1 \to \gamma \times S^1$. Hence, $q_1^{-1}(\gamma, \dot{\gamma})$ is a set of d orbits of H_1^t of period $L(\gamma)$. Under Φ, this goes over to a set of d orbits of H_2^+ of period $L(\gamma)$, which project to M as d (freely homotopic) closed geodesics of length $L(\gamma)$. The reverse argument also holds, so Φ_* is a length preserving bijection of free homotopy classes of closed geodesics.

Now, it is well known that on a manifold of nonpositive curvature, freely homotopic closed geodesics have the same length ([3, I]). Hence freely homotopic closed geodesics of $(M, \tilde{g}_2)$ must have the same length. It follows that Φ_* identifies the marked length spectra of $(M, \tilde{g}_1)$ and $(M, \tilde{g}_2)$.

It remains to show that $(M, \tilde{g}_2)$ has no conjugate points. This follows as long as the lift $\tilde{\gamma}$ of each geodesic γ of $(M, \tilde{g}_2)$ to the universal is minimizing [13, II, Theorem 5.7]. As in [3, Lemma 3.2] we argue that $\tilde{\gamma}$ is minimizing for any closed γ because γ is the shortest loop in its free homotopy class. Further, closed geodesics for $(M, \tilde{g}_2)$ must be dense in S_2^*M. Indeed, those for $(M, \tilde{g}_1)$ are well known to be dense in S_1^*M and under $q_1 \circ \Phi \circ q_1^{-1}$ the same must hold for $\tilde{g}_2$ [loc. cit]. Hence, $\tilde{\gamma}$ minimizing for closed geodesics γ implies $\tilde{\gamma}$ minimizing for all γ.

Proof of Theorem 4.2. By Theorem 4.1, (M_1, g_1) and (M_2, g_2) have a common finite negatively curved riemannian cover (M, g). Let $p_i: M \to M_i$ denote the covering maps. Also let $p: \widetilde{M} \to M$ denote the universal covering of M, let $\mathrm{Isom}(\widetilde{M})$ denote the isometry group of the metric $p^*(g)$, and let Γ_i denote the deck transformation groups of the covers $p_i \circ p: \widetilde{M} \to M_i$. Obviously, $\Gamma_i \subset \mathrm{Isom}(\widetilde{M})$. Since the M_i must have the same genus (by isospectrality), Γ_1 is isomorphic to Γ_2. We will now show that if $\mathrm{Lsp}(M_1, g_1)$ is also simple, then $\Gamma_1 = \Gamma_2$.

First, we recall that isospectral manifolds of negative curvature have the same length spectrum [2]. Indeed, the wave trace formula of [6] gives:

$$\mathrm{Tr}\cos t\sqrt{\Delta} = \sum_{\{\gamma\}} \frac{L_\gamma^{\#} e^{i\pi/4 m_\gamma}}{|I - P_\gamma|^{1/2}} \delta(t - L_\gamma) + \text{smoother}. \tag{4.3}$$

Here, Δ can be the Laplacian on any (M, g) whose closed geodesics are nondegenerate, $\{\gamma\}$ runs over the closed geodesics, $L_\gamma^{\#}$ is the primitive length of γ (once around), m_γ is the Morse index of γ, P_γ is its linear Poincaré map, and $|I - P_\gamma|$ is short for $|\det(I - P_\gamma)|$. Since $m_\gamma = 0$ for all γ, if (M, g) has negative curvature, all terms in (4.3) are positive. Hence, $\mathrm{Lsp}(M, g) = \mathrm{sing\,supp}\,\mathrm{Tr}\cos t\sqrt{\Delta}$. In particular, $\mathrm{Lsp}(M_1, g_1) = \mathrm{Lsp}(M_2, g_2)$.

Assuming $\mathrm{Lsp}(M_1, g_1)$ is simple, we claim that $\mathrm{Lsp}(M_2, g_2)$ is also simple. To see this, we first observe that (4.3) implies

$$\frac{L_\alpha^{\#}}{|I - P_\alpha|^{1/2}} = \sum_{\beta\,:\,L_\beta = L_\alpha} \frac{L_\beta}{|I - P_\beta|^{1/2}} \qquad (L_\alpha \in \mathrm{Lsp}(M_1, g_1)), \tag{4.4}$$

where α is a closed geodesic of (M_1, g_1), and β is one of (M_2, g_2)). Suppose now that α is a primitive closed geodesic, i.e., not an iterate, so that $L_\alpha = L_\alpha^{\#}$ lies in the primitive length spectrum $\mathrm{PLsp}(M_1, g_1)$ (lengths of primitive geodesics). Then L_α is also in $\mathrm{PLsp}(M_2, g_2)$. Indeed, if L_α were not primitive for (M_2, g_2), it would equal kL_β for some primitive β. Then L_β would occur as a length L_{α_0} in $\mathrm{Lsp}(M_1, g_1)$, with $L_\alpha = kL_{\alpha_0}$. By simplicity, $\alpha = \alpha_0^k$, a contradiction. Hence $L_\beta^{\#} = L_\beta = L_\alpha$ for each term in (4.4), and we conclude

$$|I - P_\alpha|^{-1/2} = \sum_{\beta\,:\,L_\beta = L_\alpha} |I - P_\beta|^{-1/2} \qquad (L_\alpha \in \mathrm{PLsp}(M_1, g_1). \tag{4.5}$$

Next, we claim that $|I - P_\alpha| = |I - P_\beta|$ for each β in (4.5); hence, only one term can occur. To see this, we first note that under the isometric correspondence $p_1 \circ p_2^{-1}\colon M_2 \to M_1$, each β in (4.5) must go over to α. Indeed, $p_1 \circ p_2^{-1}(\beta)$ must be a union of closed geodesics $\{\alpha, \cdots, \alpha_r\}$ of (M_1, g_1). Clearly, each L_{α_j} is a rational multiple of L_β. Hence, $L_{\alpha_j} = m_j L_{\alpha_1}/n_j$ for some m_j, $n_j \in \mathbb{N}$. By simplicity, $\alpha_j^{n_j} = \alpha_1^{m_j}$, so all α_j must be iterates of a simple primitive α_0. But α_0 must be α since their lengths are rationally related, and both are primitive. Hence $p_1 \circ p_2^{-1}(\beta) = \alpha$ as subsets of M_1.

Now let $T_\varepsilon(\beta)$ be the tube of radius ε around β, and let $T_\varepsilon(\alpha)$ be the tube around α. We claim $T_\varepsilon(\alpha)$ is isometric to $T_\varepsilon(\beta)$ for all β in (4.3) and for small enough ε. Indeed, under p_2^{-1}, β splits into closed geodesics $\{\tilde{\beta}_1, \cdots, \tilde{\beta}_r\}$, and $T_\varepsilon(\beta)$ splits into $\{T_\varepsilon(\tilde{\beta}_j)\}$. Pick one component, say $T_\varepsilon(\tilde{\beta}_1)$, and consider the covering diagram:

(4.6)
$$\begin{array}{ccc} & T_\varepsilon(\tilde{\beta}) & \\ {}^{p_1}\swarrow & & \searrow^{p_2} \\ T_\varepsilon(\alpha) & & T_\varepsilon(\beta) \end{array}$$

These covers are cyclic and riemannian. Since $L_\alpha = L_\beta$, they must have the same degrees. Hence, the deck transformation groups of the p_i are equal. It follows that $T_\varepsilon(\alpha)$ is isometric to $T_\varepsilon(\beta)$.

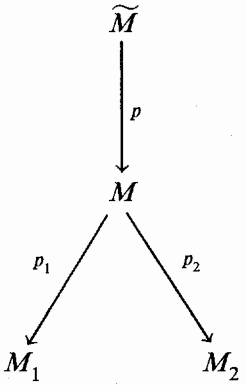
This implies $|I - P_\alpha| = |I - P_\beta|$. Indeed, lift α (resp. β) to the corresponding orbit $(\alpha, \dot{\alpha})$ (resp. $(\beta, \dot{\beta})$) of G_1^t on S^*M_1 (resp. G_2^t on S^*M_2). Also, let $T_\varepsilon(\alpha, \dot{\alpha})$ be the tube of radius ε around $(\alpha, \dot{\alpha})$ with respect to the natural metric on S^*M_1 induced from the metric g_1 on M_1 and the riemannian connection and (similarly for $T_\varepsilon(\beta, \dot{\beta})$). The isometry from $T_\varepsilon(\alpha)$ to $T_\varepsilon(\beta)$ has a natural lift to a contact diffeomorphism from $T_\varepsilon(\alpha, \dot{\alpha})$ to $T_\varepsilon(\beta, \dot{\beta})$ which takes the generator Ξ_1 of G_1^t to Ξ_2 for G_2^t. Hence the flows G_i^t have contact equivalent germs along the orbits $(\alpha, \dot{\alpha})$ (resp. $(\beta, \dot{\beta})$). In particular, the linear Poincaré cusps P_α and P_β are linearly symplectically equivalent; and so $|I - P_\alpha| = |I - P_\beta|$. It also follows that $|I - P_{\alpha^k}| = |I - P_{\beta^k}|$ for any $k = 1, 2, \cdots$. Hence, only one term can occur on the right side of (4.5) even if α is not primitive. We conclude that $\mathrm{Lsp}(M_2, g_2)$ is simple.

Just as with $p_1 \circ p_2^{-1}$ above, we now argue that $p_2 \circ p_1^{-1}(\alpha)$ consists of a single closed geodesic β of (M_2, g_2) with $L_\alpha = L_\beta$. Therefore,

$p_2 \circ p_1^{-1}$ induces a length preserving bijection between the closed geodesics of (M_1, g_1) and those of (M_2, g_2). We will see that this forces $\Gamma_1 = \Gamma_2$.

Consider the following diagram of riemannian covers:

(4.7)



Let $\gamma \in \Gamma_1$ and let $A(\gamma)$ be its axis; i.e., the unique geodesic fixed by γ. Let $a(\gamma) = p_1 \circ p(A(\gamma))$, so that $a(\gamma)$ is a closed geodesic of M_1. Under $p_2 \circ p_1^{-1}$, $a(\gamma)$ goes to a single closed geodesic b of the same length. It follows first that $(p_1 \circ p)^{-1}(a(\gamma)) = (p_2 \circ p)^{-1}(b)$. But each component of $(p_2 \circ p)^{-1}(b)$ is the axis of some $\delta \in \Gamma_2$. Hence, for all $\gamma \in \Gamma_1$ there exists $\delta \in \Gamma_2$ with $A(\gamma) = A(\delta)$. Further, such a δ exists with the same displacement, say $d(\delta)$, as γ. Here, the displacement $d(\phi)$ of an isometry ϕ is given by

$$d(\phi) = \inf d(x, \phi(x)) \qquad (d = \text{distance}).$$

Indeed, $\{\delta' \in \Gamma_2 \colon A(\delta') = A(\delta)\}$ is just the centralizer $(\Gamma_2)_\delta$ of δ in Γ_2, and $(\Gamma_2)_\delta$ is a cyclic group, generated by a primitive hyperbolic element δ_0. Now, the quotient of $A(\delta)$ by $(\Gamma_2)_\delta$ is the closed geodesic b. So $d(\delta_0) = L_b$. Similarly the quotient of $A(\delta)$ by $(\Gamma_1)_\gamma$ is $a(\gamma)$. Since $L_b = L_{a(\gamma)}$, the generator γ_0 of $(\Gamma_1)_\gamma$ satisfies $d(\gamma_0) = d(\delta_0)$. It follows that for any $\gamma \in \Gamma_1$ there exists $\delta \in \Gamma_2$ with $A(\gamma) = A(\delta)$ and $d(\gamma) = d(\delta)$. But an orientation-preserving hyperbolic isometry in two dimensions is determined by its axis and displacement. Indeed, $\gamma\delta^{-1}$ would fix all points on $A(\delta)$ and therefore on all orthogonal horocircles. Hence it would fix all of $\widetilde{M}$.

It follows that $\Gamma_1 \subseteq \Gamma_2$. The reverse argument shows $\Gamma_2 \subseteq \Gamma_1$ as well.

5. The Sunada examples

(5.1) **Proposition.** *The Sunada isospectral pairs* $\{(M_1, g_1), (M_2, g_2)\}$ *are Fourier-isospectral.*

Proof (with A. Uribe). As discussed in §0, the M_i are assumed to fit into a diagram like (0.1), with $L^2(G/H_1) \simeq L^2(G/H_2)$ (isomorphic G-modules).

As is well known, the space of intertwining operators $A: L^2(G/H_1) \to L^2(G/H_2)$ is isomorphic to the space of convolution kernels $A(x^{-1}y)$ with $A \in \mathbb{C}[H_2\backslash G/H_1]$ (the double coset space [14, p. 365]). For each such A, define $U_A: L^2(M_1) \to L^2(M_2)$ by

$$U_A = \frac{1}{\#H_1} \sum_{g\in G} A(g)\pi_{2*}T_g\pi_1^*. \tag{5.2}$$

Here, $\pi_i: M \to M_i$ are the riemannian covers in (0.1), and T_g is translation by g. Since π_i and T_g are local isometries, U_A intertwines the Laplacians Δ_i, and U_A is clearly an FIO (cf. §0).

We now observe that A unitary implies U_A unitary. To simplify, we will view $L^2(M_i)$ as the space $L^2(M)^{H_i}$ of H_i-invariant elements of $L^2(M)$, and U_A as an operator from $L^2(M)^{H_1} \to L^2(M)^{H_2}$. Then π_{i*} becomes $\sum_{h\in H_i} T_h$, and π_i^* becomes the inclusion $L^2(M)^{H_i} \to L^2(M)$. We get

$$U_A^*U_A = \frac{1}{\#H_1\cdot\#H_2} \sum_{\substack{g_1,\, g_2\in G\\ h_1\in H_1,\, h_2\in H_2}} A(g_1)\overline{A}(g_2^{-1})T_{h_1g_1h_2}T_{g_2^{-1}}. \tag{5.3}$$

Set $\overline{g}_1 = h_1g_1h_2$ and change variables. Since $A(h_2^{-1}\overline{g}_1h_1^{-1}) = A(g_1)$, the sum in (5.3) simplifies (after another change) to

$$\sum_{g_1,\, g_2} A(g_1)\overline{A}(g_1g_2^{-1})T_{g_2} = \sum_{g_2} A^*A(g_2)T_{g_2} = \sum_{g_2} \delta_{H_1}(g_2)T_{g_2}$$

(by unitarity of A). Here, δ_{H_1} is $(\#H_1^{-1})$ times the characteristic function of H_1. $U_A^*U_A$ is thus the identity operator on $L^2(M)^{H_1}$.

Acknowledgment

We would like to thank C. Croke, D. DeTurck, A. Greenleaf, J. Morava, J. Smith, A. Uribe, M. Zworski, and the referee(s) for helpful comments.

References

[1] S. Cappel, private communication.

[2] Y. Colin de Verdiere, *Spectre du Laplacian et Longueurs des geodesiques periodiques.* I, Compositio Math. 27 (1973) 83–106.

[3] C. Croke, *Rigidity for surfaces of non-positive curvature*, Comment. Math. Helv. **65** (1990) 150–169.

[4] C. Croke, A. Fathi & J. Feldman, *The marked length-spectrum of a surface of non-positive curvature*, preprint, 1990.

[5] D. DeTurck & C. Gordon, *Isospectral deformations.* I: *Riemannian structures on two-step nilspaces*, Comm. Pure Appl. Math. **40** (1987) 367–387.

[6] H. Duistermaat & V. Guillemin, *The spectrum of positive elliptic operators and periodic bicharacteristics*, Invent. Math. **29** (1975) 39–79.

[7] J. P. Francois & V. Guillemin, *On the period spectrum of a symplectic mapping*, preprint, 1990.

[8] C. Gordon & E. Wilson, *Isospectral deformations of compact solvmanifolds*, J. Differential Geometry **19** (1984) 241–256.

[9] V. Guillemin & S. Sternberg, *Geometric asymptotics*, Amer. Math. Soc. Surveys, Vol. 14, Amer. Math. Soc., Providence, RI, 1977.

[10] V. Guillemin & A. Uribe, *Reduction and the trace formula*, J. Differential Geometry, to appear.

[11] L. Hörmander, *The analysis of linear partial differential operators*, Grundlehren Math. Wiss., Vols. 274–275, Springer, Berlin, 1985.

[12] A. Ikeda, *On lens spaces which are isospectral but not isometric*, Ann. Sci. École Norm. Sup. (4) **13** (1981) 303–316.

[13] S. Kobayashi & K. Nomizu, *Foundations of differential geometry*, Vol. II, Interscience tracts, No. 15, Wiley, New York, 1969.

[14] G. Mackey, *Unitary group representations*, Benjamin/Cummings, Reading, MA, 1978.

[15] F. Marhuenga, *Microlocal analysis of some isospectral problems*, Thesis, University of Rochester, 1990.

[16] J. Milnor, *Eigenvalues of the Laplacian on certain manifolds*, Proc. Nat. Acad. Sci. U.S.A. **51** (1964) 542.

[17] J. P. Otal, *Le spectre marqué des longueurs des surfaces à courbure negative*, Ann. of Math. (2) **131** (1990) 151–162.

[18] T. Sunada, *Riemannian coverings and isospectral manifolds*, Ann. of Math. (2) **121** (1985) 169–186.

[19] M. Vigneras, *Varietés riemanniennes isospectrales et non-isometriques*, Ann. of Math. (2) **112** (1980) 21–32.

[20] A. Weinstein, *Fourier integral operators, quantization, and the spectra of riemannian manifolds*, Colloq. Internat. de Geom. Gympos. et Phys. Math. (Aix 1974), CNRS.

[21] ——, *Symplectic geometry*, Bull. Amer. Math. Soc. (N.S.) **5** (1981) 1–13.

THE JOHNS HOPKINS UNIVERSITY