Semi-classical analysis

Victor Guillemin and Shlomo Sternberg

April 25, 2012

Preface

0.1 Semi-classical analysis

There are a number of excellent texts available on the topic of this monograph, among them Dimassi and Sjostrand's "Spectral Asymptotics in the Semi-classical Analysis" [DiSj], Zworski's, "Lectures on Semi-classical Analysis" [Zwor], Martinez's "An introduction to Semi-classical and Microlocal Analysis and Microlocal Analysis " [Mart], Didier Robert's "Autour de l'Approximation Semi-classique", [Did] and Colin de Verdiere's, "Méthodes Semi-classiques et Théorie Spectral", [Col]. The focus of this monograph, however, is an aspect of this subject which is somewhat less systematically developed in the texts cited above than it will be here: In semi-classical analysis many of the basic results involve asymptotic expansions in which the terms can by computed by symbolic techniques and the focus of these notes will be the "symbol calculus" that this creates. In particular, the techniques involved in this symbolic calculus have their origins in symplectic geometry and the first seven chapters of this monograph will, to a large extent, be a discussion of this underlying symplectic geometry.

Another feature which, to some extent, differentiates this monograph from the texts above is an emphasis on the *global* aspects of this subject: We will spend a considerable amount of time showing that the objects we are studying are coordinate invariant and hence make sense on manifolds; and, in fact, we will try, in so far as possible, to give intrinsic coordinate free descriptions of these objects. In particular, although one can find an excellent account of the global symbol calculus of Fourier integral operators in Hörmander's seminal paper "Fourier integral operators I", the adaptation of this calculus to the semi-classical setting with all the i's dotted and t's crossed is not entirely trivial, and most of chapters 6 and 7 will be devoted to this task.

This emphasis on globality will also be reflected in our choice of topics in the later chapters of this book, for instance: wave and heat trace formulas for globally defined semi-classical differential operators on manifolds and equivariant versions of these results involving Lie group actions. (Also, apropos of Lie groups, we will devote most of Chapter 12 to discussing semi-classical aspects of the representation theory of these groups.)

We will give a more detailed description of these later chapters (and, in fact, of the whole book) in Section 4 of this preface. However before we do so we will attempt to describe in a few words what "semi-classical" analysis is concerned with and what role symplectic geometry plays in this subject.

0.2 The Bohr correspondence principle

One way to think of semi-classical analysis is as an investigation of the mathematical implications of the Bohr correspondence principle: the assertion that classical mechanics is the limit, as $\hbar$ tends to zero, of quantum mechanics.[1] To illustrate how this principle works, let's consider a physical system consisting of a single point particle, p, of mass, m, in $\mathbb{R}^n$ acted on by a conservative force $F = -\nabla V$, $V \in \mathcal{C}^\infty(\mathbb{R}^n)$. The total energy of this system (kinetic plus potential) is given by $H(x,\xi) = \frac{1}{2m}|\xi|^2 + V(x)$, where x is the position and ξ the momentum of p, and the motion of this system in phase space is described by the Hamilton–Jacobi equations

$$\begin{aligned} \frac{dx}{dt} &= \frac{\partial H}{\partial \xi}(x,\xi) \qquad (1) \\ \frac{d\xi}{dt} &= -\frac{\partial H}{\partial x}(x,\xi) \end{aligned}$$

The quantum mechanical description of this system on the other hand is given by the Schrödinger equation

$$ih\frac{\partial}{\partial t}\varphi = -\frac{h^2}{2m}\Delta\varphi + V\varphi \qquad (2)$$

whose L^2 normalized solution, $\int |\varphi|^2\,dx = 1$, gives one a probability measure $\mu_t = |\varphi(x,t)|^2\,dx$ that describes the "probable" position of the state described by ϕ at time t.

Of particular interest are the steady state solutions of (2). If we assume for simplicity that the eigenvalues $\lambda_k(h)$ of the Schrödinger operator are discrete and the corresponding L^2 normalized eigenfunctions are $\varphi_k(x)$ then the functions, $e^{-i\frac{t\lambda_k}{\hbar}}\varphi_k(x)$, are steady state solutions of (2) in the sense that the measures $\mu_k = |\varphi_k(x,t)|^2\,dx$ are independent of t. The $\lambda_k(\hbar)$'s are, by definition the energies of these steady state solutions, and the number of states with energies lying on the interval $a < \lambda < b$ is given by

$$N(a,b,\hbar) = \sharp\{a < \lambda_k(h) < b\}\,. \qquad (3)$$

On the other hand a crude semi-classical method for computing this number of states is to invoke the Heisenberg uncertainty principle

$$|\delta x_i \delta \xi_i| \geq 2\pi\hbar \qquad (4)$$

[1]Mathematicians are sometimes bothered by this formulation of the BCP since $\hbar$ is a fixed constant, i.e., is (approximately) 10^{-27} erg secs., (a conversion factor from the natural units of inverse seconds to the conventional unit of ergs) not a parameter that one can vary at will. However, unlike e and π, it is a *physical* constant: in the world of classical physics in which quantities are measured in ergs and secs, it is negligibly small, but in the world of subatomic physics it's not. Therefore the transition from quantum to semi-classical can legitimately be regarded as an "$\hbar$ tends to zero" limit.

and the Pauli exclusion principle (which can be interpreted as saying that no two of these states can occupy the same position in phase space) to conclude that the maximum number of classical states with energies on the interval $a < H < b$ is approximately equal to the maximal number of disjoint rectangles lying in the region, $a < H(x, \xi) < b$ and satisfying the volume constraint imposed by (4). For $\hbar$ small the number of such rectangles is approximately

$$\left(\frac{1}{2\pi\hbar}\right)^n \mathrm{vol}\,(a < H(x,\xi) < b) \tag{5}$$

so as $\hbar$ tends to zero

$$(2\pi\hbar)^n N(a,b,\hbar) \to \mathrm{vol}\,(a < H(x,\xi) < b)\,. \tag{6}$$

We will see in Chapter 10 of this monograph that the empirical derivation of this "Weyl law" can be made rigorous and is valid, not just for the Schrödinger operator, but for a large class of semi-classical and classical differential operators as well.

0.3 The symplectic category

We recall that a symplectic manifold is a pair (M, ω) where M is a $2n$-dimensional manifold and $\omega \in \Omega^2(M)$ a closed two-form satisfying $\omega_p^n \neq 0$ for all $p \in M$. Given a symplectic manifold (M, ω) we will denote by M^- the symplectic manifold, $(M, -\omega)$ and given two symplectic manifolds, M_i, $i = 1, 2$ we will denote by $M_1 \times M_2$ the product of these two manifolds equipped with the symplectic form

$$\omega = (pr_1)^*\omega_1 + (pr_2)^*\omega_2\,.$$

Finally, given a $2n$-dimensional symplectic manifold, (M, ω), we'll call an n-dimensional submanifold, Λ of M *Lagrangian* if the inclusion map, $\iota_\Lambda : \Lambda \to M$ satisfies $\iota_\Lambda^*\omega = 0$, i.e. ω vanishes when restricted to Λ. Of particular importance for us will be Lagrangian submanifolds of the product manifold, $M_1^- \times M_2$, and these we will call *canonical relations*.

With these notations in place we will define the *symplectic category* to be the category whose objects are symplectic manifolds and whose morphisms are canonical relations: i.e. given symplectic manifolds, M_1 and M_2, we will define a morphism of M_1 into M_2 to be a canonical relation, Γ, in $M_1^- \times M_2$. (We will use double arrow notation, $\Gamma : M_1 \twoheadrightarrow M_2$ for these morphisms to distinguish them from a more conventional class of morphisms, symplectic maps.)

To make these objects and morphisms into a category we have to specify a composition law for pairs of morphisms, $\Gamma_i : M_i \twoheadrightarrow M_{i+1}$ $i = 1, 2$ and this we do by the recipe

$$(m_1, m_3) \in \Gamma \Leftrightarrow (m_1, m_2) \in \Gamma_1 \ \text{ and } \ (m_2, m_3) \in \Gamma_2 \tag{7}$$

for some $m_2 \in M_2$. Unfortunately the "Γ" defined by this recipe is not always a canonical relation (or even a manifold) but it is if one imposes some transversality conditions on Γ_1 and Γ_2 which we'll spell out in detail in Chapter 4.

The fundamental notion in our approach to semi-classical analysis is a "quantization operation" for canonical relations. We're not yet in position to discuss this quantization operation in general. (This will be the topic of Chapters 8-11 of this monograph.) But we'll briefly discuss an important special case: Let X be a manifold and let $M = T^*X$ be the cotangent bundle of X equipped with its standard symplectic form (the two-form, ω, which, in standard cotangent coordinates, is given by, $\sum dx_i \wedge d\xi_i$). A Lagrangian manifold Λ of M is *horizontal* if the cotangent fibration, $\pi(x,\xi) = x$, maps Λ bijectively onto X. Assuming X is simply connected this condition amounts to the condition

$$\Lambda = \Lambda_\varphi \tag{8}$$

where φ is a real-valued $\mathcal{C}^\infty$ function on X and

$$\Lambda_\varphi = \{(x,\xi) \in T^*X\,,\ \xi = d\varphi_x\}\,. \tag{9}$$

Now let $M_i = T^*X_i$, $i = 1,2$ and let $\Gamma : M_1 \twoheadrightarrow M_2$ be a canonical relation. Then

$$\Gamma^\sharp = \{(x_1, -\xi_1, x_2, \xi_2)\,,\ (x_1, \xi_1, x_2, \xi_2) \in \Gamma\}$$

is a Lagrangian submanifold of the product manifold

$$M_1 \times M_2 = T^*(X_1 \times X_2)$$

and hence if $\Gamma^\sharp$ is horizontal it is defined as above by a real-valued $\mathcal{C}^\infty$ function $\varphi(x_1, x_2)$ on $X_1 \times X_2$. We will quantize Γ by associating with it the set of linear operators

$$T_h : \mathcal{C}_0^\infty(X_1) \to \mathcal{C}^\infty(X_2) \tag{10}$$

of the form

$$T_h f(x_2) = \int e^{i\frac{\varphi(x_1,x_2)}{\hbar}} a(x_1, x_2, h) f(x_1)\, dx_1 \tag{11}$$

where $a(x_1, x_2, \hbar)$ is in $\mathcal{C}^\infty(X_1 \times X_2 \times \mathbb{R})$ and $\hbar$ is positive parameter (our stand-in for Planck's constant). These "semi-classical Fourier integral operators" are the main topic of this monograph, and our goal in Chapters 8-11 will be to show that their properties under composition, taking of transposes, computing traces, etc. is governed symbolically by symplectic properties of their corresponding canonical relations. In fact we will show that the symbolic calculus that describes the leading asymptotics of these operators in the $h \to 0$ limit can be entirely described by constructing (as we will do in Chapter 7) an "enhanced symplectic category" consisting of pairs (Γ, σ) where Γ is a canonical relation and σ a section of the "pre-quantum line bundle" on Γ.

0.4 The plan of attack, part 1

Chapter 1 of this monograph will essentially be a fleshed out version of this preface. We will show how one can construct solutions of hyperbolic partial differential eqautions for short time intervals, modulo error terms of order $O(\hbar)$, by reducing this problem to a problem involving solutions of the Hamilton–Jacobi equation (1). Then, using an embryonic version of the symbol theory mentioned above we will show that these "solutions modulo $O(\hbar)$" can be converted into "solutions modulo $O(\hbar^\infty)$". We will also show that this method of solving (2) breaks down when the solution of the associated classical equation (1) develops caustics, thus setting the stage for the much more general approach to this problem described in Chapter 8 where methods for dealing with caustics and related complications are developed in detail.

In Chapter 1 we will also discuss in more detail the Weyl law (6). (At this stage we are not prepared to prove (6) in general but we will show how to prove it in two simple illuminating special cases.)

Chapter 2 will be short crash course in symplectic geometry in which we will review basic definitions and theorems. Most of this material can be found in standard references such as [AM], [Can] or [GSSyT], however the material at the end of this section on the Lagrangian Grassmannian and Maslov theory is not so readily accessible and for this reason we've treated it in more detail.

In Chapter 3 we will, as preparation for our "categorical" approach to symplectic geometry, discuss some prototypical examples of categories. The category of finite sets and relations, and the linear symplectic category (in which the objects are symplectic vector spaces and the morphisms are linear canonical relations). The first of these examples are mainly introduced for the propose of illustrating categorical concepts; however the second example will play an essential role in what follows. In particular the fact that the linear symplectic category is a true category: that the composition of *linear* canonical relations is always well defined, will be a key ingredient in our construction of a symbol calculus for semi-classical Fourier integral operators.

Chapter 4 will begin our account of the standard non-linear version of this category, the symplectic "category" itself.[2] Among the topics we will discuss are composition operations, a factorization theorem of Weinstein (which asserts that every canonical relation is the composition of an immersion and submersion), an imbedding result (which shows that the standard differential category of $\mathcal{C}^\infty$ manifolds, and $\mathcal{C}^\infty$ maps is a subcategory of the symplectic category) and other examples of such subcategories. In particular one important such subcategory is the *exact* symplectic category, whose objects are pairs, (M,α) where α is a one-form on M whose exterior derivative is symplectic. In this category the Lagrangian submanifolds, Λ, of M will also be required to be exact, i.e. to satisfy $\iota_\Lambda^*\alpha = d\varphi_\Lambda$ for some $\varphi_\Lambda \in \mathcal{C}^\infty(\Lambda)$. (In Chapter 8 when we associate oscillatory integrals with Lagrangian submanifolds, Λ, of T^*X the fixing of this φ_Λ will enable us to avoid the presence of troublesome undefined oscillatory factors in these integrals.)

We will also describe in detail a number of examples of canonical relations that will frequently be encountered later on. To give a brief description of

[2] Many of the ideas discussed in this chapter are directly or indirectly inspired by Alan Weinstein's 1981 *Bulletin* article "Symplectic geometry", not the least of these being the term, "category", for a collection of morphisms for which there are simple, easy-to-verify criteria for composability.

some of examples in this preface let's denote by "pt." the "point-object" in the symplectic category: the unique-up-to-symplectomorphism connected symplectic manifold of dimension zero, and regard a Lagrangian submanifold of a symplectic manifold, M, as being a morphism

$$\Lambda : \text{pt.} \twoheadrightarrow M\,.$$

In addition given a canonical relation $\Gamma : M_1 \twoheadrightarrow M_2$ let's denote by $\Gamma^t : M_2 \twoheadrightarrow M_1$ the transpose canonical relation; i.e. require that $(m_2, m_1) \in \Gamma^t$ if $(m_1, m_2) \in \Gamma$.

Example 1. Let X and Y be manifolds and $f : X \to Y$ a $\mathcal{C}^\infty$ map. Then

$$\Gamma_f : T^*X \twoheadrightarrow T^*Y \tag{12}$$

is the canonical relation defined by

$$(x, \xi, y, \eta) \in \Gamma_f \Leftrightarrow y = f(x) \text{ and } \xi = df_X^* \eta\,. \tag{13}$$

The correspondence that associates Γ_f to f gives us the imbedding of the differential category into the symplectic category that we mentioned above. Moreover we will see in Chapter 8 that Γ_f and its transpose have natural quantizations: Γ_f^t as the pull-back operation

$$f^* : \mathcal{C}^\infty(Y) \to \mathcal{C}^\infty(X) \tag{14}$$

and Γ_f as the transpose of this operation on distributions.

Example 2. If $\pi : Z \to X$ is a $\mathcal{C}^\infty$ fibration the distributional transpose of (14) maps $\mathcal{C}_0^\infty(Z)$ into $\mathcal{C}_0^\infty(X)$ and hence defines a fiber integration operation

$$\pi_* : \mathcal{C}_0^\infty(Z) \to \mathcal{C}_0^\infty(X) \tag{15}$$

about which we will have more to say when we preview the "quantum" chapters of this monograph in the next section.

Example 3. Let Z be a closed submanifold of $X_1 \times X_2$ and let π_i be the projection of Z onto X_i. Then by quantizing $\Gamma_{\pi_2} \times \Gamma_{\pi_1}^t$ we obtain a class of Fourier integral operators which play a fundamental role in integral geometry: generalized Radon transforms.

Example 4. The identity map of T^*X onto itself. We will show in Chapter 8 that the entity in the quantum world that corresponds to the identity map is the algebra of semi-classical pseudodifferential operators (about which we will have a lot more to say below!)

Example 5. The symplectic involution

$$\Gamma : T^*\mathbb{R}^n \to T^*\mathbb{R}^n \qquad (x, \xi) \to (\xi - x)\,. \tag{16}$$

This is the horizontal canonical relation in $(T^*\mathbb{R}^n)^- \times T^*\mathbb{R}^n$ associated with the Lagrangian manifold Λ_φ where $\varphi \in \mathcal{C}^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ is the function, $\varphi(x, y) = -x.y$.

If one quantizes Γ by the recipe (11) taking $a(x, y, h)$ to be the constant function $(2\pi h)^{-n}$ one gets the semi-classical Fourier transform

$$F_h f(x) = (2\pi\hbar)^{-n/2} \int e^{-i\frac{x \cdot y}{\hbar}} f(y)\, dy\,. \tag{17}$$

(See Chapter 5, §15 and Chapter 8, §9.)

This operator will play an important role in our "local" description of the algebra of semi-classical pseudodifferential operators when the manifold X in Example 3 is an open subset of $\mathbb{R}^n$.

Example 6. (Generating functions) Given a Lagrangian manifold, $\Lambda \subseteq T^*X$, a fiber bundle $\pi : Z \to X$ and a function $\varphi \in \mathcal{C}^\infty(Z)$ we will say that φ is a *generating function* for Λ with respect to the fibration, π, if Λ is the composition of the relations, $\Lambda_\varphi : pt \twoheadrightarrow T^*Z$ and $\Gamma_\pi : T^*Z \twoheadrightarrow T^*X$. In the same spirit, if $\Gamma : T^*X \twoheadrightarrow T^*Y$ is a canonical relation, $\pi : Z \to X \times Y$ is a fiber bundle and $\varphi \in \mathcal{C}^\infty(Z)$ we will say that φ is a generating function for Γ with respect to π if it is a generating function for the associated Lagrangian manifold, $\Gamma^\sharp$ in $T^*(X \times Y)$. These functions will play a key role in our definition of Fourier integral operators in Chapter 8, and in Chapter 5 we will give a detailed account of their properties. In particular we will show that locally every Lagrangian manifold is definable by a generating function and we will also prove a uniqueness result which says that locally any generating function can be obtained from any other by a sequence of clearly defined "Hörmander moves". We will also prove a number of functorial properties of generating functions: e.g. show that if

$$\Gamma_i : T^*X_i \twoheadrightarrow T^*X_{i+1} \qquad i = 1, 2$$

are canonical relations and (Z_i, π_i, φ_i) generating data for Γ_i, then if Γ_i and Γ_2 are composable, the φ_i's are composible as well in the sense that there is a simple procedure for constracting from the φ_i's a generating function for $\Gamma_2 \circ \Gamma_1$. Finally in the last part of Chapter 5 we will investigate the question, "Do global generating functions exist?" This question is one of the main unanswered open questions in present-day symplectic topology; so we will not be able to say much about it; however we will show that if one tries to construct a global generating function by patching together local generating functions one encouters a topological obstacle: the vanishing of a cohomology class in $H^1(\Lambda, \mathbb{Z})$. This cohomology class, the Maslov class, puts in its appearance in this theory in other contexts as well. In particular the line bundle on Λ associated with the mod 4 reduction of this cohomology class is a main ingredient in the leading symbol theory of semi-classical Fourier integral operators.

The other main ingredient in this symbol theory is *half-densities.* These will be discussed in Chapter 6, and in Chapter 7 we will show how to "enhance" the symplectic category by replacing canonical relations by pairs, (Γ, σ) where Γ is a canonical relation and σ a half-density on Γ, and by showing that the composition law for canonical relations that we discussed above extends to a composition law for these pairs. (In §7.8 we will add a further complication to this picture by replacing the σ's by $\sigma \otimes m$'s where m is a section of the Maslov bundle.)

0.5 The plan of attack, part 2

Section 4 was an overview of Chapters 1–7, the symplectic or "classical" half of this monograph, We'll turn next to the material in the next five chapters, the application of these results to semi-classical analysis. Let $(\Lambda, \varphi_\Lambda)$ be an exact Lagrangian submanifold of T^*X. If Λ is horizontal, i.e. of the form (8)–(9) one can associate with Λ the space of oscillatory functions

$$\mu \in I^k(X;\Lambda) \Leftrightarrow \mu = \hbar^k a(x,\hbar) e^{i\frac{\varphi(x)}{\hbar}} \tag{18}$$

where a is a $\mathcal{C}^\infty$ function on $X \times \mathbb{R}$ and φ_Λ is the pull-back of φ to Λ. More generally if Λ is defined by generating data, (Z, π, φ) and φ and φ_Λ are compatible in an appropriate sense (see Section 8.1) we will define the elements of $I^k(X;\Lambda)$ to be oscillatory fuctions of the form

$$\mu = \hbar^{k-d/2} \pi_* \left(a(z,\hbar) e^{i\frac{\varphi(Z)}{\hbar}} \right) \tag{19}$$

where d is the fiber dimension of Z, $a(z,\hbar)$ is a $\mathcal{C}^\infty$ function on $Z \times \mathbb{R}$ and π_* is the operator (15)[3]

More generally if $(\Lambda, \varphi_\Lambda)$ is an arbitrary exact Lagrnagian manifold in T^*X one can define $I^k(X;\Lambda)$ by patching together local versions of this space. (As we mentioned in §4, φ_Λ plays an important role in this patching process. The compatibility of φ_Λ with local generating data avoids the presence of a lot of undefined oscillatory factors in the elements of $I^k(X;\Lambda)$ that one obtains from this patching process.)

Some of our goals in Chapter 8 will be:

1. To show that the space $I^k(X;\Lambda)$ is well-defined. (Doing so will rely heavily on the uniqueness theorem for generating functions proved in Chapter 5).

2. To show that if $\mathbb{L}_\Lambda$ is the line bundle over Λ defined in §7.8 (the tensor product of the Maslov bundle and the half-density bundle) there is a canonical leading symbol map

$$\sigma : I^k(X;\Lambda)/I^{k+1}(X,\Lambda) \to \mathcal{C}^\infty(\mathbb{L}_\Lambda)\,. \tag{20}$$

3. To apply these results to canonical relations. In more detail, if $\Gamma : T^*X \twoheadrightarrow T^*Y$ is a canonical relation and $\Gamma^\sharp$ is, as in §3, the associated Lagrangian

[3]Strictly speaking to define π_* one needs to equip X and Z with densities, dx and dz, so as to make sense of the pairing

$$\int \pi_* \mu \mathrm{v}\, dx = \int u \pi^* \mathrm{v}\, dz\,.$$

However in §8 we will give a slightly different definition of π_* that avoids these choices: We will let Γ_π be an enhanced canonical relation in the sense of §4.7, i.e. equipped with a $\frac{1}{2}$-density symbol, and let μ and v be $\frac{1}{2}$-densities. Thus in this approach $I^k(X;\Lambda)$ becomes a space of *half-densities.*

submanifold of $T^*(X \times Y)$, then, given an element, μ, of $I^{k-n/2}(X \times Y, \Gamma^\sharp)$, $n = \dim Y$, we can define an operator

$$F_\mu : \mathcal{C}_0^\infty(X) \to \mathcal{C}^\infty(Y) \tag{21}$$

by the recipe

$$F_\mu f(y) = \int f(x)\mu(x, y, \hbar)\, dx\,; \tag{22}$$

and we will call this operator a *semi-classical Fourier integral operator* of order k. We will also define its *symbol* to be the leading symbol of μ and we will denote the space of these operators by $\mathcal{F}^k(\Gamma)$. One of our main goals in Chapter 8 will be to show that the assignment

$$\Gamma \to \mathcal{F}^k(\Gamma) \tag{23}$$

is a functor, i.e. to show that if $\Gamma_i : T^*X_i \twoheadrightarrow T^*X_{i+1}$, $i = 1, 2$, are canonical relations and Γ_1 and Γ_2 are transversally composable, then for $F_i \in \mathcal{F}^{k_i}(\Gamma_i)$, F_2F_1 is in $\mathcal{F}^{k_1+k_2}$ and the leading symbol of F_2F_1 can be computed from the leading symbols of F_2F_1 by the composition law for symbols that we defined in Chapter 7. (We will also prove an analogous result for cleanly composable canonical relations.)

4. To apply these results to the identity map of T^*X onto T^*X. If Γ is this identity map then $\Gamma \circ \Gamma = \Gamma$ and this composition is a transversal composition, so the space of Fourier integral operators, $\mathcal{F}(\Gamma)$, is a filtered ring. This ring is the *ring* of *semi-classical pseudodifferential* operators and we will henceforth denote it by $\Psi(X)$. We will show that the symbol calculus for this ring is much simpler than the symbol calculus for arbitrary Γ; namely, we will show that $\mathbb{L}_\Gamma \cong \mathbb{C}$ and hence that the leading symbol of an element of Ψ^k/Ψ^{k+1} is just a $\mathcal{C}^\infty$ function on T^*X.

5. To observe that $I(X, \Lambda)$ is a *module* over $\Psi(X)$; More explicitly if $\Lambda : pt \twoheadrightarrow T^*X$ is a Lagrangian manifold and Γ is the identity map of T^*X onto itself, then $\Gamma \centerdot \Lambda = \Lambda$, and this composition is transversal. Hence, for $\mu \in I^k(X, \Lambda)$ and $P \in \Psi^\varphi(X)$, $P_\mu \in I^{k+\ell}(X, \Lambda)$. We will make use of this module structure to deal with some problems in PDE theory that we were unable to resolve in Chapter 1, in particular to construct solutions mod $O(\hbar^\infty)$ of the Schrödinger equation and other semi-classical differential equations in the presence of caustics.

6. To give a concrete description of the algebra of semi-classical pseudodifferential operators for $X = \mathbb{R}^n$, in particular to show that locally on $\mathbb{R}^n$ these operators are of the form

$$\hbar^{-k} Pf(x) = (2\pi\hbar)^{-\frac{n}{2}} \int a(x, \xi, h) e^{i\frac{x \cdot \xi}{\hbar}} F_h f(\xi)\, d\xi \tag{24}$$

where F_h is the semi-classical Fourier transform (17).

Finally one last (very important) goal of Chapter 8 will be to describe the role of "microlocaity" in semi-classical analysis. If P is the pseudodifferential operator (**??**) and (x, ξ) a point in $T^*\mathbb{R}^n$ we will say that P *vanishes* on an open neighborhoor, U, of (x, ξ) if the function $a(x, \xi, h)$ vanishes to infinite order in $\hbar$ on this open neighborhood. We will show that this definition is coordinate independent and hence that one can make sense of the notion "$P = 0$ on U" for X an arbitrary manifold, P an element of $\Psi(X)$ and U an open subset of T^*X. Moreover, from this notion one gets a number of useful related notions. For instance, for an open set, U, in T^*X one can define the ring of pseudodifferential operators, $\Psi(U)$, to be the quotient of $\Psi(X)$ by the ideal of operators which vanish on U, and one can define the *microsupport* of an operator, $P \in \Psi(X)$ by decreeing that $(x, \xi) \notin \mathrm{Supp}(P)$ if P vanishes on a neighborhood of (x, ξ). Moreover, owing to the fact that $I(X, \Lambda)$ is a module over $\Psi(X)$ one can define analogous notions for this module. (We refer to §8.6 of Chapter 6 for details.) In particular these "microlocalizations" of the basic objects in semi-classical analysis convert this into a subject which essentially lives on T^*X rather than X.

One last word about microlocality: In definition (19) we have been a bit sloppy in not specifying conditions on the support of $a(z, \hbar)$. For this expression to be well-defined we clearly have to assume that for every $p \in X$, $a(z, \hbar)$ is compactly supported on the fiber above, p, or at least, in lieu of this, impose some decay-at-infinity conditions on the restriction of a to these fibers. However sometimes one can get around such assumptions using microlocal cutoffs, i.e. define generalized elements, μ of $I^k(X,\ V)$ by requiring that such an element satisfy $P\mu \in I^k(X; \Lambda)$ for every compactly supported cutoff "function", $P \in \Psi(X)$. In Chapter 9 we will apply this idea to the ring of pseudodifferential operators itself. First, however, as an illustration of this idea, we will show that the algebra of *classical* pseudodifferential operators: operators with polyhomogeneous symbols (but with no $\hbar$ dependence) has such a characterization. Namely let $\Psi_0(X)$ be the ring of semi-classical pseudodifferential operators having compact micro-support and $\Psi_0(X)$, let $\Psi_{00}(X)$ be the elements of this ring for which the micro-support doesn't intersect the zero section. We will prove

Theorem 1. *A linear operator, $A : \mathcal{C}_0^\infty(X) \to \mathcal{C}^{-\infty}(X)$, with distributional kernel is a classical pseudodifferential operator with polyhomogeneous symbol if and only if $AP \in \Psi_{00}(X)$ for all $P \in \Psi_{00}(X)$, and is a differential operator if $AP \in \Psi_0(X)$ for all $P \in \Psi_0(X)$.*

We will then generalize this to the semi-classical setting by showing that semi-classical pseudodifferential operators with polyhomogeneous symbols are characterized by the properties:

(i) $A_\hbar$ depends smoothly on $\hbar$.

(ii) For fixed $\hbar$, A_h is polyhomogeneous.

(iii) $A_h P \in \Psi_0(X)$ for all $P \in \Psi_0(X)$.

The second half of Chapter 9 will be devoted to discussing the symbol calculus for this class of operators, for the most part focussing on operators on $\mathbb{R}^n$ of the form (24).[4]

If $a(x,\xi,\hbar)$ is polyhomogeneous of degree less than n in ξ then the Schwartz kernel of P can be written in the form

$$h^k(2\pi h)^{-h}\int a(x,\xi,h)e^{i\frac{x-y\cdot\xi}{h}}\,d\xi\,; \tag{25}$$

however, we will show that these are several alternative expressions for (25): $a(x,\xi,h)$ can be replaced by a function of the form $a(y,\xi,h)$, a function of the form $a(\frac{x+y}{2},\xi,h)$ or a function of the form $a(x,y,\xi,h)$ and we will show how all these symbols are related and derive formulas for the symbols of products of these operators. Then in the last section of Chapter 9 we will show that there is a local description in coordinates for the space $I(X;\Lambda)$ similar to (25) and give a concrete description in coordinates of the module structure of $I(X;\Lambda)$ as a module over $\Psi(X)$.

In Chapter 10 we will study the functional calculus associated with polyhomogeneous semi-classical pseudodifferential operators. We recall that if $\mathcal{H}$ is a Hilbert space and A a densely defined self-adjoint operator on $\mathcal{H}$ then by Stone's theorem A generates a one-parameter group of unitary operators

$$U(t)=e^{itA}$$

and one can make use of this fact to define functions of A by the recipe

$$f(A)=\frac{1}{2\pi}\int\hat{f}(t)e^{itA}\,dt$$

for f a compactly supported continuous function and $\hat{f}$ its Fourier transform. We will give an account of these results in Chapter 13 and also describe an adaptation of this theory to the setting of semi-classical pseudodifferential operators by Dimassi–Sjostrand. In Chapter 10, however, we will mainly be concerned with the "mod $O(\hbar^{-\infty})$" version of this functional calculus. More explicitly we will show that if $P:\mathcal{C}^\infty(\mathbb{R}^n)\to\mathcal{C}^\infty(\mathbb{R}^n)$ is a self-adjoint elliptic psedudifferential operator of order zero with leading symbol $P_0(x,\xi)$ and $f:\mathbb{R}\to\mathbb{R}$ a compactly supported $\mathcal{C}^\infty$ function then $f(P)$ is a semi-classical pseudodifferential operator with Schwartz kernel

$$(2\pi\hbar)^{-n}\int b_f(x,\xi,h)e^{i\frac{(x-y)\cdot\xi}{\hbar}}\,d\xi \tag{26}$$

where $b_f(x,\xi,h)$ admits an asymptotic expansion

$$\sum\hbar^k\sum_{\ell\leq 2k}b_{k,\ell}(x,\xi)\left(\frac{1}{i}\frac{d}{ds}\right)^\ell f(P_0(x,\xi)) \tag{27}$$

[4]We will, however, show that our results are valid under change of variables and hence make sense on manifolds.

in which the $b_{k,\ell}$'s are explicitly computable, and from this we will deduce the following generalization of the Weyl law that we described in Section 2 above.

Theorem 2. *Suppose that for some interval,* $[a, b]$, *the set* P_0^{-1} *([a,b]) is compact. Then the spectrum of* P *intersected with* (a, b) *consists of a finite number of discrete eigenvalues,* $\lambda_k(\hbar)$, $q \leq k \leq N(\hbar)$ *where*

$$N(\hbar) \sim (2\pi h)^{-n} \quad volume \quad (P_0^{-1}([a, b]))\,. \tag{28}$$

We will in fact derive this result from a much sharper result. Namely the formula (27) gives us for $f \in \mathcal{C}_0^\infty(\mathbb{R})$ an asymptotic expansion for

$$\text{Trace } f(P) = \int b_f(x, \xi, h)\, dx\, d\xi \tag{29}$$

in powers of $\hbar$ and hence an asymptotic expansion of the sum

$$\sum f(\lambda_k(h)) \qquad 1 \leq k \leq N(\hbar)\,. \tag{30}$$

The second half of Chapter 10 will basically be concerned with applications of this result. For P the Schrödinger operator we will compute the first few terms in this expansion in terms of the Schrödinger potential, V, and will prove in dimension one an inverse result of Colin de Verdiere which shows that modulo weak assymetry assumptions on V, V is spectrally determined. We will also show in dimension one that there is a simple formula linking the spectral measure $\mu(f) = \text{trace } f(P)$ and the quantum Birkhoff canonical form of P.

The results above are concerned with semi-classical pseudodifferential operators on $\mathbb{R}^n$; however we will show at the end of Chapter 10 that they can easily be generalized to manifolds and will show that these generalizations are closely related to classical heat trace results for elliptic differential operators on manifolds.

In Chapter 11 we will discuss results similar to these for Fourier integral operators. A succinct table of contents for Chapter 11 (which we won't bother to reproduce here) can be found at the very beginning of the chapter. However, in fifty words or less the main goal of the chapter will be to compute the trace of a Fourier integral operator $F : \mathcal{C}^\infty(X) \to \mathcal{C}^\infty(X)$ whose canonical relation is the graph of a symlpectomorphism, $f : T^*X \to T^*X$, and to apply this result to the wave trace

$$\text{trace} \exp i\frac{tP}{\hbar} \tag{31}$$

where P is an elliptic zeroth order semi-classical pseudodifferential operator.

The last chapter in this semi-classical segment of the monograph, Chapter 12, has to do with a topic that, as fas as we know, has not been much investigated in the mathematical literature (at least not from the semi-classical perspective). Up to this point our objects of study have been *exact* Lagrangian manifolds and *exact* canonical relations, but these belong to a slightly larger class of Lagrangian manifolds and canonical relations: If (M, α) is an exact symplectic manifold and

$\Lambda \subseteq M$ a Lagrangian submanifold we will say that Λ is *integral* if there exists a function $f : \Lambda \to S^1$ such that

$$\iota_\Lambda^* \alpha = \frac{1}{\sqrt{-1}} \frac{df}{f} \,. \tag{32}$$

To quantize Lagrangian manifolds of this type we will be forced to impose a quantization condition on $\hbar$ itself: to require that $\hbar^{-1}$ tend to infinity in $\mathbb{Z}^+$ rather than in $\mathbb{R}^+$. An example which illustrates why this constraint is needed is the Lagrangian manifold, Λ_φ, = graph $d\varphi$ in the cotangent bundle of the n-torus, $\mathbb{R}^n/2\pi\mathbb{Z}^n$ where $\varphi(x) = \sum k_i x_i$, $k \in \mathbb{Z}^n$. As a function on the torus this function is multi-valued, but $d\varphi$ and Λ_φ are well-defined, and

$$\iota_\Lambda^* \alpha = \pi_\Lambda^* \frac{df}{f}$$

where π_Λ is the projection of Λ onto the torus and $f = e^{i\varphi}$, so Λ_φ is integral.

Suppose now that we quantize Λ_φ by the recipe (18), i.e. by associating to it oscillatory functions of the form

$$a(x,\hbar) e^{i\frac{\varphi(x)}{\hbar}} \,. \tag{33}$$

It's clear that for these expressions to be well-defined we have to impose the constraint, $\hbar^{-1} \in \mathbb{Z}_+$ on $\hbar$.

In Chapter 12 we will discuss a number of interesting results having to do with quantization in this integral category. The most interesting perhaps is some "observational mathematics" concerning the classical character formulas of Weyl, Kirillov and Gross–Kostant–Ramond–Sternberg for representations of Lie groups: Let G be a compact simply-connected semi-simple Lie group and γ_α the irreducile representation of G with highest weight, α. By semi-classical techniques adapted to this integral symplectic category, one can compute symbolically the leading order asymptotics of the character, χ_n = trace $\gamma_{n\alpha}$ as n tends to infinity. However, somewhat surprisingly, the asymptotic answer is, in fact, the exact answer (and in particular valid for $n = 1$).

0.6 The plan of attack, part 3

The last four chapters of this monograph are basically appendices and have to do with results that were cited (but not proved or not explained in detail) in the earlier chapters. Most of these results are fairly standard and a well-exposed in other texts, so we haven't, in all instances, supplied detailed proofs. (In the instances where we've failed to do so, however, we've attempted to give some sense of how the proofs go.) We've also, to provide some perspective on these results, discussed a number of their applications besides those specifically alluded to in the text.

1. Chapter 13:

   Here we gather various facts from functional analysis that we use, or which motivate our constructions in Chapter 10. All the material we present here is standard, and is available in excellent modern texts such as Davies, Reed-Simon, Hislop-Sigal, Schecter, and in the classical text by Yosida. Our problem is that the results we gather here are scattered among these texts. So we had to steer a course between giving a complete and self-contained presentation of this material (which would involve writing a whole book) and giving a bare boned listing of the results.

   We also present some of the results relating semi-classical analysis to functional analysis on L_2 which allow us to provide the background material for the results of Chapters 9-11. Once again the material is standard and can be found in the texts by Dimassi-Sjöstrand, Evans-Zworski, and Martinez. And once again we steer a course between giving a complete and self-contained presentation of this material giving a bare boned listing of the results.

2. Chapter 14: The purpose of this chapter is to give a rapid review of the basics of calculus of differential forms on manifolds. We will give two proofs of Weil's formula for the Lie derivative of a differential form: the first of an algebraic nature and then a more general geometric formulation with a "functorial" proof that we learned from Bott. We then apply this formula to the "Moser trick" and give several applications of this method. (This Moser trick is, incidentally, the basic ingredient in the proof of the main results of Chapter 5.)

3. Chapter 15: The topic of this chapter is the lemma of stationary phase. This lemma played a key role in the proofs of two of the main results of this monograph: It was used in Chapter 8 to show that the quantization functor that associates F.I.O's to canonical relations is well-defined and in chapter 11 to compute the traces of these operators. In this chapter we will prove the standard version of stationary phase (for oscillatory integrals whose phase functions are just quadratic forms) and also "manifold" versions for oscillatory integrals whose phase functions are Morse or Bott-Morse. In addition we've included, for edificatioal purposes, a couple corollaries of stationary phase that are not explicitly used earlier on: the Van der Corput theorem (for estimating the number of lattice points contained in a convex region of n-space) and the Fresnel version in geometric optics of Huygens's principle.

4. In Chapter 15 we come back to the Weyl calculus of semi-classical psuedodifferential operators that we developed in Chapter 9 and describe another way of looking at it (also due to Hermann Weyl.) This approach involves the representation theory of the Heisenberg group and is based upon the following fundamental result in the representation theory of locally compact topological groups: If one is given a unitary representation

of a group of this type, this representation extends to a representation of the convolution algebra of compactly supported continuous functions on the group. Applying this observation to the Heisenberg group and the irreducible representation, ρ_h, with "Planck's constant h", one gets an algebra of operators on $L^2(R^n)$ which is canonically isomorphic to the Weyl algebra of Chapter 9, and we show that this way of looking at the Weyl algebra makes a lot of its properties much more transparent.

Contents

0.1 Semi-classical analysis i
0.2 The Bohr correspondence principle ii
0.3 The symplectic category iii
0.4 The plan of attack, part 1 iv
0.5 The plan of attack, part 2 viii
0.6 The plan of attack, part 3 xiii

1 Introduction **1**
1.1 The problem. 2
1.2 The eikonal equation. 2
1.2.1 The principal symbol. 2
1.2.2 Hyperbolicity. 3
1.2.3 The canonical one form on the cotangent bundle. 3
1.2.4 The canonical two form on the cotangent bundle. 4
1.2.5 Symplectic manifolds. 4
1.2.6 Hamiltonian vector fields. 5
1.2.7 Isotropic submanifolds. 5
1.2.8 Lagrangian submanifolds. 7
1.2.9 Lagrangian submanifolds of the cotangent bundle. 7
1.2.10 Local solution of the eikonal equation. 8
1.2.11 Caustics. 8
1.3 The transport equations. 8
1.3.1 A formula for the Lie derivative of a $\frac{1}{2}$-density. 10
1.3.2 The total symbol, locally. 12
1.3.3 The transpose of P. 12
1.3.4 The formula for the sub-principal symbol. 13
1.3.5 The local expression for the transport operator R. 14
1.3.6 Putting it together locally. 17
1.3.7 Differential operators on manifolds. 17
1.4 Semi-classical differential operators. 19
1.4.1 Schrödinger's equation and Weyl's law. 19
1.4.2 The harmonic oscillator. 20
1.5 The Schrödinger operator on a Riemannian manifold. 25
1.5.1 Weyl's law for a flat torus with $V = 0$. 25
1.6 The plan. 26

2 Symplectic geometry. 27
2.1 Symplectic vector spaces. 27
2.1.1 Special kinds of subspaces. 27
2.1.2 Normal forms. 28
2.1.3 Existence of Lagrangian subspaces. 28
2.1.4 Consistent Hermitian structures. 28
2.2 Lagrangian complements. 29
2.2.1 Choosing Lagrangian complements "consistently". 29
2.3 Equivariant symplectic vector spaces. 33
2.3.1 Invariant Hermitian structures. 33
2.3.2 The space of fixed vectors for a compact group of symplectic automorphisms is symplectic. 34
2.3.3 Toral symplectic actions. 34
2.4 Symplectic manifolds. 35
2.5 Darboux style theorems. 35
2.5.1 Compact manifolds. 36
2.5.2 Compact submanifolds. 37
2.5.3 The isotropic embedding theorem. 38
2.6 The space of Lagrangian subspaces of a symplectic vector space. 41
2.7 The set of Lagrangian subspaces transverse to a pair of Lagrangian subspaces 43
2.8 The Maslov line bundle 45
2.9 A look ahead - a simple example of Hamilton's idea. 46
2.9.1 A different kind of generating function. 46
2.9.2 Composition of symplectic transformations and addition of generating functions. 47

3 The language of category theory. 51
3.1 Categories. 51
3.2 Functors and morphisms. 52
3.2.1 Covariant functors. 52
3.2.2 Contravariant functors. 52
3.2.3 The functor to families. 52
3.2.4 Morphisms 53
3.2.5 Involutory functors and involutive functors. 53
3.3 Example: Sets, maps and relations. 54
3.3.1 The category of finite relations. 54
3.3.2 Categorical "points". 55
3.3.3 The universal associative law. 56
3.3.4 The transpose. 58
3.3.5 Some notation. 58
3.4 The linear symplectic category. 59
3.4.1 The space $\Gamma_2 \star \Gamma_1$. 59
3.4.2 The transpose. 60
3.4.3 The projection $\alpha : \Gamma_2 \star \Gamma_1 \to \Gamma_2 \circ \Gamma_1$. 61
3.4.4 The kernel and image of a linear canonical relation. 61

3.4.5 Proof that $\Gamma_2 \circ \Gamma_1$ is Lagrangian. 62
3.4.6 Details concerning the identity $\tilde{\Delta}_{XYZ} \circ (\Gamma_1 \times \Gamma_2) = \Gamma_2 \circ \Gamma_1$. 63
3.4.7 The category **LinSym** and the symplectic group. 64
3.4.8 Reductions in the linear symplectic category. 64
3.4.9 Composition with reductions or co-reductions. 66
3.5 The category of oriented linear canonical relations. 66

4 The Symplectic "Category". 69
4.1 Clean intersection. 70
4.2 Composable canonical relations. 72
4.3 Transverse composition. 73
4.4 Lagrangian submanifolds as canonical relations. 73
4.5 The involutive structure on $\mathcal{S}$. 74
4.6 Reductions in the symplectic "category". 74
4.6.1 Reductions in the symplectic "category" are reductions by coisotropics. 75
4.6.2 The decomposition of any morphism into a reduction and a coreduction. 75
4.6.3 Composition with reductions or co-reductions. 76
4.7 Canonical relations between cotangent bundles. 76
4.8 The canonical relation associated to a map. 77
4.9 Pushforward of Lagrangian submanifolds of the cotangent bundle. 78
4.9.1 Envelopes. 80
4.10 Pullback of Lagrangian submanifolds of the cotangent bundle. . . 82
4.11 The moment map. 83
4.11.1 The classical moment map. 83
4.11.2 Families of symplectomorphisms. 84
4.11.3 The moment map in general. 86
4.11.4 Proofs. 88
4.11.5 The derivative of Φ. 91
4.11.6 A converse. 91
4.11.7 Back to families of symplectomorphisms. 92
4.12 Double fibrations. 93
4.12.1 The moment image of a family of symplectomorphisms . . 94
4.12.2 The character Lagrangian. 95
4.12.3 The period–energy relation. 96
4.12.4 The period–energy relation for families of symplectomorphisms. 96
4.13 The category of exact symplectic manifolds and exact canonical relations. 98
4.13.1 Exact symplectic manifolds. 98
4.13.2 Exact Lagrangian submanifolds of an exact symplectic manifold. 99
4.13.3 The sub"category" of $\mathcal{S}$ whose objects are exact. 99
4.13.4 Functorial behavior of β_Γ. 100

4.13.5 Defining the "category" of exact symplectic manifolds and canonical relations. 101
4.13.6 Pushforward via a map in the "category" of exact canonical relations between cotangent bundles. 102

5 Generating functions. 103
5.1 Fibrations. 103
5.1.1 Transverse vs. clean generating functions. 105
5.2 The generating function in local coordinates. 106
5.3 Example - a generating function for a conormal bundle. 107
5.4 Example. The generating function of a geodesic flow. 108
5.5 The generating function for the transpose. 111
5.6 The generating function for a transverse composition. 112
5.7 Generating functions for clean composition of canonical relations between cotangent bundles. 115
5.8 Reducing the number of fiber variables. 116
5.9 The existence of generating functions. 120
5.10 The Legendre transformation. 123
5.11 The Hörmander-Morse lemma. 125
5.12 Changing the generating function. 132
5.13 The Maslov bundle. 132
5.13.1 The Čech description of locally flat line bundles. 133
5.13.2 The local description of the Maslov cocycle. 133
5.13.3 The global definition of the Maslov bundle. 135
5.13.4 The Maslov bundle of a canonical relation between cotangent bundles. 135
5.13.5 Functoriality of the Maslov bundle. 136
5.14 Identifying the two definitions of the Maslov bundle. 137
5.15 More examples of generating functions. 139
5.15.1 The image of a Lagrangian submanifold under geodesic flow. 139
5.15.2 The billiard map and its iterates. 139
5.15.3 The classical analogue of the Fourier transform. 141
5.15.4 Quadratic generating functions. 142

6 The calculus of $\frac{1}{2}$-densities. 145
6.1 The linear algebra of densities. 145
6.1.1 The definition of a density on a vector space. 145
6.1.2 Multiplication. 147
6.1.3 Complex conjugation. 147
6.1.4 Elementary consequences of the definition. 147
6.1.5 Pullback and pushforward under isomorphism. 149
6.1.6 Pairs of Lagrangian subspaces. 150
6.1.7 Spanning pairs of subspaces of a symplectic vector space. 150
6.1.8 Lefschetz symplectic linear transformations. 151
6.2 Densities on manifolds. 154

6.2.1 Multiplication of densities. 155
6.2.2 Support of a density. 155
6.3 Pull-back of a density under a diffeomorphism. 155
6.4 Densities of order 1. 156
6.5 The principal series representations of Diff(X). 157
6.6 The push-forward of a density of order one by a fibration. 158

7 The Enhanced Symplectic "Category". 161
7.1 The underlying linear algebra. 161
7.1.1 Transverse composition of $\frac{1}{2}$ densities. 163
7.2 Half densities and clean canonical compositions. 164
7.3 Rewriting the composition law. 165
7.4 Enhancing the category of smooth manifolds and maps. 166
7.4.1 Enhancing an immersion. 167
7.4.2 Enhancing a fibration. 167
7.4.3 The pushforward via an enhanced fibration. 167
7.5 Enhancing a map enhances the corresponding canonical relation. 168
7.6 The involutive structure of the enhanced symplectic "category". . 169
7.6.1 Computing the pairing $\langle(\Lambda_1,\rho_1),(\Lambda_2,\rho_2)\rangle$ 170
7.6.2 † and the adjoint under the pairing. 171
7.7 The symbolic distributional trace. 171
7.7.1 The $\frac{1}{2}$-density on Γ. 171
7.7.2 Example: The symbolic trace. 172
7.7.3 General transverse trace. 172
7.7.4 Example: Periodic Hamiltonian trajectories. 174
7.8 The Maslov enhanced symplectic "category". 176

8 Oscillatory $\frac{1}{2}$-densities. 179
8.1 Definition of $I^k(X,\Lambda)$ in terms of a generating function. 180
8.1.1 Local description of $I^k(X,\Lambda,\phi)$. 181
8.1.2 Independence of the generating function. 181
8.1.3 The global definition of $I^k(X,\Lambda)$. 183
8.2 Semi-classical Fourier integral operators. 183
8.2.1 Composition of semi-classical Fourier integral operators. . 184
8.3 The symbol of an element of $I^k(X,\Lambda)$. 185
8.3.1 A local description of $I^k(X,\Lambda)/I^{k+1}(X,\Lambda)$. 185
8.3.2 The local definition of the symbol. 187
8.3.3 The intrinsic line bundle and the intrinsic symbol map. . 187
8.4 Symbols of semi-classical Fourier integral operators. 188
8.4.1 The functoriality of the symbol. 189
8.5 The Keller-Maslov-Arnold description of the line bundle $\mathbb{L}$ 193
8.6 Microlocality. 196
8.6.1 The microsheaf. 200
8.6.2 Functoriality of the sheaf $\mathcal{E}^\ell$. 201
8.7 Semi-classical pseudo-differential operators. 203
8.7.1 The line bundle and the symbol. 203

8.7.2 The commutator and the bracket. 204
8.7.3 $I(X, \Lambda)$ as a module over $\Psi(X)$. 204
8.7.4 Microlocality. 205
8.7.5 The semi-classical transport operator. 206
8.8 The local theory. 209
8.8.1 The composition law for symbols. 210
8.9 The semi-classical Fourier transform. 212
8.9.1 The local structure of oscillatory $\frac{1}{2}$-densities. 214
8.9.2 The local expression of the module structure of $I(X, \Lambda)$ over $\Psi(X)$. 215
8.9.3 Egorov's theorem. 215
8.10 Semi-classical differential operators and semi-classical pseudo-differential operators. 216
8.10.1 Semi-classical differential operators act microlocally as semi-classical pseudo-differential operators. 218
8.10.2 Pull-back acts microlocally as a semi-classical Fourier integral operator. 220
8.11 Description of the space $I^k(X, \Lambda)$ in terms of a clean generating function. 220
8.12 The clean version of the symbol formula. 222
8.13 Clean composition of Fourier integral operators. 224
8.13.1 A more intrinsic description. 225
8.13.2 The composition formula for symbols of Fourier integral operators when the underlying canonical relations are cleanly composable. 226
8.14 An abstract version of stationary phase. 227

9 Pseudodifferential Operators. 231
9.1 Semi-classical pseudo-differential operators with compact micro-support. 231
9.2 Classical ΨDO's with polyhomogeneous symbols. 233
9.3 Semi-classical pseudo-differential operators. 238
9.4 The symbol calculus. 243
9.4.1 Composition. 246
9.4.2 Behavior under coordinate change. 247
9.5 The formal theory of symbols. 250
9.5.1 Multiplication properties of symbols. 251
9.6 The Weyl calculus. 253
9.7 The structure of $I(X, \Lambda)$ as a module over the ring of semi-classical pseudo-differential operators. 254

10 Trace invariants. 255
10.1 Functions of pseudo-differential operators. 255
10.2 The wave operator for semi-classical pseudo-differential operators. 258
10.3 The functional calculus modulo $O(\hbar^\infty)$. 260
10.4 The trace formula. 261

10.5 Spectral invariants for the Schrödinger operator. 262
10.6 An Inverse Spectral Result: Recovering the Potential Well 266
10.7 Semiclassical Spectral Invariants for Schrödinger Operators with Magnetic Fields 268
10.8 An Inverse Result for The Schrödinger Operator with A Magnetic Field 270
10.9 Counterexamples. 271
10.10The functional calculus on manifolds. 273

11 Fourier Integral operators. 275
11.1 Semi-classical Fourier integral operators. 275
11.2 The lemma of stationary phase. 277
11.3 The trace of a semiclassical Fourier integral operator. 278
11.3.1 Examples. 281
11.3.2 The period spectrum of a symplectomorphism. 282
11.4 The mapping torus of a symplectic mapping. 284
11.5 The Gutzwiller formula. 287
11.5.1 The phase function for the flowout. 290
11.5.2 Periodic trajectories of v_{p}. 291
11.5.3 The trace of the operator (11.29). 292
11.5.4 Density of states. 293
11.6 The Donnelly theorem. 294

12 Integrality in semi-classical analysis. 297
12.1 Introduction. 297
12.2 Line bundles and connections. 299
12.3 Integrality in DeRham theory. 304
12.4 Integrality in symplectic geometry. 306
12.5 Symplectic reduction and the moment map. 310
12.6 Coadjoint orbits. 315
12.7 Integrality in semi-classical analysis 318
12.8 The Weyl character formula. 319
12.9 The Kirillov character formula. 326
12.10The GKRS character formula. 329
12.11 The pseudodifferential operators on line bundles 330
12.12 Spectral properties of the operators, $A_\hbar$ 334
12.13 Equivariant spectral problems in semi-classical analysis 336

13 Spectral theory and Stone's theorem. 341
13.1 Unbounded operators, their domains, their spectra and their resolvents. 342
13.1.1 Linear operators and their graphs. 342
13.1.2 Closed linear transformations. 344
13.1.3 The resolvent, the resolvent set and the spectrum. 344
13.1.4 The resolvent identities. 346
13.1.5 The adjoint of a densely defined linear operator. 348

13.2 Self-adjoint operators on a Hilbert space. 349
13.2.1 The graph and the adjoint of an operator on a Hilbert space. 349
13.2.2 Self-adjoint operators. 349
13.2.3 Symmetric operators. 350
13.2.4 The spectrum of a self-adjoint operator is real. 352
13.3 Stone's theorem. 354
13.3.1 Equibounded continuous semi-groups. 355
13.3.2 The infinitesimal generator. 355
13.3.3 The resolvent of the infinitesimal generator. 359
13.3.4 Application to Stone's theorem. 360
13.3.5 The exponential series and sufficient conditions for it to converge. 361
13.3.6 The Hille Yosida theorem. 362
13.3.7 The case of a Banach space. 365
13.3.8 The other half of Stone's theorem. 366
13.4 The spectral theorem. 366
13.4.1 The functional calculus for functions in $\mathcal{S}$. 366
13.4.2 The multiplication version of the spectral theorem. 368
13.5 The Calderon-Vallaincourt theorem. 372
13.5.1 Existence of inverses. 373
13.6 The functional calculus for Weyl operators. 376
13.6.1 Trace class Weyl operators. 378
13.7 Kantorovitz's non-commutative Taylor's formula. 378
13.7.1 A Dynkin-Helffer-Sjöstrand formula for derivatives. 378
13.7.2 The exponential formula. 379
13.7.3 Kantorovitz's theorem. 380
13.7.4 Using the extended Dynkin-Helffer-Sjöstrand formula. . . 383
13.8 Appendix: The existence of almost holomorphic extensions. . . . 383

14 Differential calculus of forms, Weil's identity and the Moser trick. 387
14.1 Superalgebras. 387
14.2 Differential forms. 388
14.3 The d operator. 388
14.4 Derivations. 389
14.5 Pullback. 390
14.6 Chain rule. 391
14.7 Lie derivative. 391
14.8 Weil's formula. 392
14.9 Integration. 394
14.10Stokes theorem. 395
14.11Lie derivatives of vector fields. 395
14.12Jacobi's identity. 397
14.13A general version of Weil's formula. 397
14.14The Moser trick. 400

14.14.1 Volume forms. . . . 401
14.14.2 Variants of the Darboux theorem. . . . 402
14.14.3 The classical Morse lemma. . . . 402

15 The method of stationary phase 405
15.1 Gaussian integrals. . . . 405
15.1.1 The Fourier transform of a Gaussian. . . . 405
15.2 The integral $\int e^{-\lambda x^2/2} h(x)dx$. . . . 407
15.3 Gaussian integrals in n dimensions. . . . 408
15.4 Using the multiplication formula for the Fourier transform. . . . 409
15.5 A local version of stationary phase. . . . 410
15.6 The formula of stationary phase. . . . 411
15.6.1 Critical points. . . . 411
15.6.2 The formula. . . . 412
15.6.3 The clean version of the stationary phase formula. . . . 414
15.7 Group velocity. . . . 415
15.8 The Fourier inversion formula. . . . 417
15.9 Fresnel's version of Huygen's principle. . . . 417
15.9.1 The wave equation in one space dimension. . . . 417
15.9.2 Spherical waves in three dimensions. . . . 418
15.9.3 Helmholtz's formula 419
15.9.4 Asymptotic evaluation of Helmholtz's formula 420
15.9.5 Fresnel's hypotheses. . . . 421
15.10 The lattice point problem. . . . 421
15.10.1 The circle problem. . . . 422
15.10.2 The divisor problem. . . . 424
15.10.3 Using stationary phase. . . . 425
15.10.4 Recalling Poisson summation. . . . 426
15.11 Van der Corput's theorem. . . . 427

16 The Weyl Transform. 431
16.1 The Weyl transform in the physics literature. . . . 432
16.1.1 The Weyl transform and the Weyl ordering. . . . 433
16.2 Definition of the semi-classical Weyl transform. . . . 433
16.3 Group algebras and representations. . . . 434
16.3.1 The group algebra. . . . 434
16.3.2 Representing the group algebra. . . . 434
16.3.3 Application that we have in mind. . . . 435
16.4 The Heisenberg algebra and group. . . . 435
16.4.1 The Heisenberg algebra. . . . 435
16.4.2 The Heisenberg group. . . . 435
16.4.3 Special representations. . . . 436
16.5 The Stone-von-Neumann theorem. . . . 436
16.6 Constructing $\rho_\hbar$. . . . 437
16.7 The "twisted convolution". . . . 439
16.8 The group theoretical Weyl transform. . . . 440

16.9 Two two by two matrices. 440
16.10 Schrödinger representations. 441
16.11 The Weyl transform. 442
16.11.1 Repeat of the definition of the semi-classical Weyl transform. 442
16.11.2 Weyl_σ and the Schrödinger representation of the Heisenberg group. 442
16.12 Weyl transforms with symbols in $L^2(\mathbb{R}^{2n})$. 443
16.13 Weyl transforms associated to linear symbols and their exponentials. 444
16.13.1 The Weyl transform associated to ξ^α is $(\hbar D)^\alpha$. 444
16.13.2 The Weyl transform associated to $a = a(x)$ is multiplication by a. 444
16.13.3 The Weyl transform associated to a linear function. 445
16.13.4 The composition $L \circ B$. 445
16.14 The one parameter group generated by L. 446
16.15 Composition. 447
16.16 Hilbert-Schmidt Operators. 449
16.17 Proof of the irreducibility of $\rho_{\ell,\hbar}$. 450
16.18 Completion of the proof. 452

Chapter 1

Introduction

Let $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$ with coordinates $(x^1, \ldots, x^n, t)$. Let

$$P = P\left(x, t, \frac{\partial}{\partial x}, \frac{\partial}{\partial t}\right)$$

be a k-th order linear partial differential operator. Suppose that we want to solve the partial differential equation

$$Pu = 0$$

with initial conditions

$$u(x,0) = \delta_0(x), \quad \frac{\partial^i}{\partial t^i} u(x,0) = 0, \quad i = 1, \ldots, k-1,$$

where δ_0 is the Dirac delta function.

Let ρ be a C^∞ function of x of compact support which is identically one near the origin. We can write

$$\delta_0(x) = \frac{1}{(2\pi)^n} \rho(x) \int_{\mathbb{R}^n} e^{ix\cdot\xi} d\xi.$$

Let us introduce polar coordinates in ξ space:

$$\xi = \omega \cdot r, \quad \|\omega\| = 1, \quad r = \|\xi\|$$

so we can rewrite the above expression as

$$\delta_0(x) = \frac{1}{(2\pi)^n} \rho(x) \int_{\mathbb{R}_+} \int_{S^{n-1}} e^{i(x\cdot\omega)r} r^{n-1} dr d\omega$$

where $d\omega$ is the measure on the unit sphere S^{n-1}.

Passing the differential operator under the integrals shows that we are interested in solving the partial differential equation $Pu = 0$ with the initial conditions

$$u(x,0) = \rho(x) e^{i(x\cdot\omega)r} r^{n-1}, \quad \frac{\partial^i}{\partial t^i} u(x,0) = 0, \quad i = 1, \ldots, k-1.$$

1.1 The problem.

More generally, set

$$r = \hbar^{-1}$$

and let

$$\psi \in C^\infty(\mathbb{R}^n).$$

We look for solutions of the partial differential equation with initial conditions

$$Pu(x,t) = 0, \quad u(x,0) = \rho(x)e^{i\frac{\psi(x)}{\hbar}}\hbar^{-\ell}, \quad \frac{\partial^i}{\partial t^i}u(x,0) = 0, \quad i = 1,\dots,k-1. \tag{1.1}$$

Here ℓ can be any integer; in the preceding example we had $\ell = 1-n$.

1.2 The eikonal equation.

Look for solutions of (1.1) of the form

$$u(x,t) = a(x,t,\hbar)e^{i\phi(x,t)/\hbar} \tag{1.2}$$

where

$$a(x,t,\hbar) = \hbar^{-\ell}\sum_{i=0}^{\infty} a_i(x,t)\hbar^i. \tag{1.3}$$

1.2.1 The principal symbol.

Define the **principal symbol** $H(x,t,\xi,\tau)$ of the differential operator P by

$$\hbar^k e^{-i\frac{x.\xi+t\tau}{\hbar}}Pe^{i\frac{x.\xi+t\tau}{\hbar}} = H(x,t,\xi,\tau) + O(\hbar). \tag{1.4}$$

We think of H as a function on $T^*\mathbb{R}^{n+1}$.

If we apply P to $u(x,t) = a(x,t,\hbar)e^{i\phi(x,t)/\hbar}$, then the term of degree $\hbar^{-k}$ is obtained by applying all the differentiations to $e^{i\phi(x,t)/\hbar}$. In other words,

$$\hbar^k e^{-i\phi/\hbar}Pa(x,t)e^{i\phi/\hbar} = H\left(x,t,\frac{\partial\phi}{\partial x},\frac{\partial\phi}{\partial t}\right)a(x,t) + O(\hbar). \tag{1.5}$$

So as a first step we must solve the first order non-linear partial differential equation

$$H\left(x,t,\frac{\partial\phi}{\partial x},\frac{\partial\phi}{\partial t}\right) = 0 \tag{1.6}$$

for ϕ. Equation (1.6) is known as the **eikonal equation** and a solution ϕ to (1.6) is called an **eikonal** . The Greek word eikona $\epsilon\iota\kappa\omega\nu\alpha$ means image.

1.2.2 Hyperbolicity.

For all (x, t, ξ) the function

$$\tau \mapsto H(x, t, \xi, \tau)$$

is a polynomial of degree (at most) k in τ. We say that P is **hyperbolic** if this polynomial has k distinct real roots

$$\tau_i = \tau_i(x, t, \xi).$$

These are then smooth functions of (x, t, ξ).

We assume from now on that P is hyperbolic. For each $i = 1, \ldots, k$ let

$$\Sigma_i \subset T^*\mathbb{R}^{n+1}$$

be defined by

$$\Sigma_i = \{(x, 0, \xi, \tau) | \xi = d_x\psi,\ \tau = \tau_i(x, 0, \xi)\} \tag{1.7}$$

where ψ is the function occurring in the initial conditions in (1.1). The classical method for solving (1.6) is to reduce it to solving a system of ordinary differential equations with initial conditions given by (1.7). We recall the method:

1.2.3 The canonical one form on the cotangent bundle.

If X is a differentiable manifold, then its cotangent bundle T^*X carries a **canonical one form** $\alpha = \alpha_X$ defined as follows: Let

$$\pi : T^*X \to X$$

be the projection sending any covector $p \in T_x^*X$ to its base point x. If $v \in T_p(T^*X)$ is a tangent vector to T^*X at p, then

$$d\pi_p v$$

is a tangent vector to X at x. In other words, $d\pi_p v \in T_xX$. But $p \in T_x^*X$ is a linear function on T_xX, and so we can evaluate p on $d\pi_p v$. The canonical linear differential form α is defined by

$$\langle \alpha_p, v \rangle := \langle p, d\pi_p v \rangle \quad \text{if} \quad v \in T_p(T^*X). \tag{1.8}$$

For example, if our manifold is $\mathbb{R}^{n+1}$ as above, so that we have coordinates (x, t, ξ, τ) on $T^*\mathbb{R}^{n+1}$ the canonical one form is given in these coordinates by

$$\alpha = \xi \cdot dx + \tau dt = \xi_1 dx^1 + \cdots \xi_n dx^n + \tau dt. \tag{1.9}$$

1.2.4 The canonical two form on the cotangent bundle.

This is defined as

$$\omega_X = -d\alpha_X. \tag{1.10}$$

Let $q^1, \ldots, q^n$ be local coordinates on X. Then $dq^1, \ldots, dq^n$ are differential forms which give a basis of T_x^*X at each x in the coordinate neighborhood U. In other words, the most general element of T_x^*X can be written as $p_1(dq^1)_x + \cdots + p_n(dq^n)_x$. Thus $q^1, \ldots, q^n, p_1, \ldots, p_n$ are local coordinates on

$$\pi^{-1}U \subset T^*X.$$

In terms of these coordinates the canonical one-form is given by

$$\alpha = p \cdot dq = p_1 dq^1 + \cdots + p_n dq^n$$

Hence the canonical two-form has the local expression

$$\omega = dq \wedge \cdot dp = dq^1 \wedge dp_1 + \cdots + dq^n \wedge dp_n. \tag{1.11}$$

The form ω is closed and is of maximal rank, i.e., ω defines an isomorphism between the tangent space and the cotangent space at every point of T^*X.

1.2.5 Symplectic manifolds.

A two form which is closed and is of maximal rank is called **symplectic**. A manifold M equipped with a symplectic form is called a **symplectic manifold**. We shall study some of the basic geometry of symplectic manifolds in Chapter 2. But here are some elementary notions which follow directly from the definitions: A diffeomorphism $f : M \to M$ is called a **symplectomorphism** if $f^*\omega = \omega$. More generally if (M, ω) and (M', ω') are symplectic manifolds then a diffeomorphism

$$f : M \to M'$$

is called a symplectomorphism if

$$f^*\omega' = \omega.$$

If v is a vector field on M, then the general formula for the Lie derivative of a differential form Ω with respect to v is given by

$$D_v\Omega = i(v)d\Omega + di(v)\Omega.$$

This is known as Weil's identity. See (14.2) in Chapter 14 below. If we take Ω to be a symplectic form ω, so that $d\omega = 0$, this becomes

$$D_v\omega = di(v)\omega.$$

So the flow $t \mapsto \exp tv$ generated by v consists (locally) of symplectomorphisms if and only if

$$di(v)\omega = 0.$$

1.2.6 Hamiltonian vector fields.

In particular, if H is a function on a symplectic manifold M, then the **Hamiltonian vector field** v_H associated to H and defined by

$$i(v_H)\omega = dH \tag{1.12}$$

satisfies

$$(\exp tv_H)^*\omega = \omega.$$

Also

$$D_{v_H}H = i(v_H)dH = i(v_H)i(v_H)\omega = \omega(v_H, v_H) = 0.$$

Thus

$$(\exp tv_H)^*H = H. \tag{1.13}$$

So the flow $\exp tv_H$ preserves the level sets of H. In particular, it carries the zero level set - the set $H = 0$ - into itself.

1.2.7 Isotropic submanifolds.

A submanifold Y of a symplectic manifold is called **isotropic** if the restriction of the symplectic form ω to Y is zero. So if

$$\iota_Y : \ Y \to M$$

denotes the injection of Y as a submanifold of M, then the condition for Y to be isotropic is

$$\iota_Y^*\omega = 0$$

where ω is the symplectic form of M.

For example, consider the submanifold Σ_i of $T^*(\mathbb{R}^{n+1})$ defined by (1.7). According to (1.9), the restriction of $\alpha_{\mathbb{R}^{n+1}}$ to Σ_i is given by

$$\frac{\partial\psi}{\partial x_1}dx_1 + \cdots \frac{\partial\psi}{\partial x_n}dx_n = d_x\psi$$

since $t \equiv 0$ on Σ_i. So

$$\iota_{\Sigma_i}^*\omega_{\mathbb{R}^{n+1}} = -d_x d_x\psi = 0$$

and hence Σ_i is isotropic.

Let H be a smooth function on a symplectic manifold M and let Y be an isotropic submanifold of M contained in a level set of H. For example, suppose that

$$H_{|Y} \equiv 0. \tag{1.14}$$

Consider the submanifold of M swept out by Y under the flow $\exp tv_H$. More precisely suppose that

- v_H is transverse to Y in the sense that for every $y \in Y$, the tangent vector $v_H(y)$ does *not* belong to T_yY and

- there exists an open interval I about 0 in $\mathbb{R}$ such that $\exp tv_H(y)$ is defined for all $t \in I$ and $y \in Y$.

We then get a map

$$j : Y \times I \to M, \quad j(y,t) := \exp tv_H(y)$$

which allows us to realize $Y \times I$ as a submanifold Z of M. The tangent space to Z at a point (y,t) is spanned by

$$(\exp tv_H)_* TY_y \quad \text{and} \quad v_H(\exp tv_H y)$$

and so the dimension of Z is $\dim Y + 1$.

Proposition 1.2.1. *With the above notation and hypotheses, Z is an isotropic submanifold of M.*

Proof. We need to check that the form ω vanishes when evaluated on

1. two vectors belonging to $(\exp tv_H)_* TY_y$ and
2. $v_H(\exp tv_H y)$ and a vector belonging to $(\exp tv_H)_* TY_y$.

For the first case observe that if $w_1, w_2 \in T_yY$ then

$$\omega((\exp tv_H)_* w_1, (\exp tv_H)_* w_2) = (\exp tv_H)^* \omega(w_1, w_2) = 0$$

since

$$(\exp tv_H)^* \omega = \omega$$

and Y is isotropic. For the second case observe that $i(v_H)\omega = dH$ and so for $w \in T_yY$ we have

$$\omega(v_H(\exp tv_H y), (\exp tv_H)_* w) = dH(w) = 0$$

since H is constant on Y. □

If we consider the function H arising as the symbol of a hyperbolic equation, i.e. the function H given by (1.4), then H is a homogeneous polynomial in ξ and τ of the form $b(x,t,\xi)\prod_i(\tau - \tau_i)$, with $b \neq 0$ so

$$\frac{\partial H}{\partial \tau} \neq 0 \quad \text{along} \quad \Sigma_i.$$

But the coefficient of $\partial/\partial t$ in v_H is $\partial H/\partial \tau$. Now $t \equiv 0$ along Σ_i so v_H is transverse to Σ_i. Our transversality condition is satisfied. We can arrange that the second of our conditions, the existence of solutions for an interval I can be satisfied locally. (In fact, suitable compactness conditions that are frequently satisfied will guarantee the existence of global solutions.)

Thus, at least locally, the submanifold of $T^*\mathbb{R}^{n+1}$ swept out from Σ_i by $\exp tv_H$ is an $n+1$ dimensional isotropic submanifold.

1.2.8 Lagrangian submanifolds.

A submanifold of a symplectic manifold which is isotropic and whose dimension is one half the dimension of M is called **Lagrangian**. We shall study Lagrangian submanifolds in detail in Chapter 2. Here we shall show how they are related to our problem of solving the eikonal equation (1.6).

The submanifold Σ_i of $T^*\mathbb{R}^{n+1}$ is isotropic and of dimension n. It is transversal to v_H. Therefore the submanifold Λ_i swept out by Σ_i under $\exp tv_H$ is Lagrangian. Also, near $t = 0$ the projection

$$\pi : T^*\mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$$

when restricted to Λ_i is (locally) a diffeomorphism. It is (locally) **horizontal** in the sense of the next section.

1.2.9 Lagrangian submanifolds of the cotangent bundle.

To say that a submanifold $\Lambda \subset T^*X$ is Lagrangian means that Λ has the same dimension as X and that the restriction to Λ of the canonical one form α_X is closed.

Suppose that Z is a submanifold of T^*X and that the restriction of $\pi : T^*X \to X$ to Z is a diffeomorphism. This means that Z is the image of a section

$$s : \; X \to T^*X.$$

Giving such a section is the same as assigning a covector at each point of X, in other words it is a linear differential form. For the purposes of the discussion we temporarily introduce a redundant notation and call the section s by the name β_s when we want to think of it as a linear differential form. We claim that

$$s^*\alpha_X = \beta_s.$$

Indeed, if $w \in T_xX$ then $d\pi_{s(x)} \circ ds_x(w) = w$ and hence

$$s^*\alpha_X(w) = \langle (\alpha_X)_{s(x)}, ds_x(w)\rangle =$$

$$= \langle s(x), d\pi_{s(x)} ds_x(w)\rangle = \langle s(x), w\rangle = \beta_s(x)(w).$$

Thus the submanifold Z is Lagrangian if and only if $d\beta_s = 0$. Let us suppose that X is connected and simply connected. Then $d\beta = 0$ implies that $\beta = d\phi$ where ϕ is determined up to an additive constant.

With some slight abuse of language, let us call a Lagrangian submanifold Λ of T^*X **horizontal** if the restriction of $\pi : T^*X \to X$ to Λ is a diffeomorphism. We have proved

Proposition 1.2.2. *Suppose that X is connected and simply connected. Then every horizontal Lagrangian submanifold of T^*X is given by a section $\gamma_\phi : X \to T^*X$ where γ_ϕ is of the form*

$$\gamma_\phi(x) = d\phi(x)$$

where ϕ is a smooth function determined up to an additive constant.

1.2.10 Local solution of the eikonal equation.

We have now found a local solution of the eikonal equation! Starting with the initial conditions Σ_i given by (1.7) at $t = 0$, we obtain the Lagrangian submanifold Λ_i. Locally (in x and in t near zero) the manifold Λ_i is given as the image of γ_{ϕ_i} for some function ϕ_i. The fact that Λ_i is contained in the set $H = 0$ then implies that ϕ_i is a solution of (1.6).

1.2.11 Caustics.

What can go wrong globally? One problem that might arise is with integrating the vector field v_H. As is well known, the existence theorem for non-linear ordinary differential equations is only local - solutions might "blow up" in a finite interval of time. In many applications this is not a problem because of compactness or boundedness conditions. A more serious problem - one which will be a major concern of this book - is the possibility that after some time the Lagrangian manifold is no longer horizontal.

If $\Lambda \subset T^*X$ is a Lagrangian submanifold, we say that a point $m \in \Lambda$ is a **caustic** if

$$d\pi_m T_m\Lambda \to T_xX. \quad x = \pi(m)$$

is *not* surjective. A key ingredient in what we will need to do is to describe how to choose convenient parametrizations of Lagrangian manifolds near caustics. The first person to deal with this problem (through the introduction of so-called "angle characteristics") was Hamilton (1805-1865) in a paper he communicated to Dr. Brinkley in 1823, by whom, under the title "Caustics" it was presented in 1824 to the Royal Irish Academy.

We shall deal with caustics in a more general manner, after we have introduced some categorical language.

1.3 The transport equations.

Let us return to our project of looking for solutions of the form (1.2) to the partial differential equation and initial conditions (1.1). Our first step was to find the Lagrangian manifold $\Lambda = \Lambda_\phi$ which gave us, locally, a solution of the eikonal equation (1.6). This determines the "phase function" ϕ up to an overall additive constant, and also guarantees that no matter what a_i's enter into the expression for u given by (1.2) and (1.3), we have

$$Pu = O(\hbar^{-k-\ell+1}).$$

The next step is obviously to try to choose a_0 in (1.3) such that

$$P\left(a_0e^{i\phi(x,t)/\hbar}\right) = O(\hbar^{-k+2}).$$

In other words, we want to choose a_0 so that there are no terms of order $\hbar^{-k+1}$ in $P\left(a_0e^{i\phi(x,t)/\hbar}\right)$. Such a term can arise from three sources:

1. We can take the terms of degree $k-1$ in P and apply all the differentiations to $e^{i\phi/\hbar}$ with none to a or to ϕ. We will obtain an expression C similar to the principal symbol but using the operator Q obtained from P by eliminating all terms of degree k. This expression C will then multiply a_0.

2. We can take the terms of degree k in P, apply all but one differentiation to $e^{i\phi/\hbar}$ and the remaining differentiation to a partial derivative of ϕ. The resulting expression B will involve the second partial derivatives of ϕ. This expression will also multiply a_0.

3. We can take the terms of degree k in P, apply all but one differentiation to $e^{i\phi/\hbar}$ and the remaining differentiation to a_0. So we get a first order differential operator
$$\sum_{i=1}^{n+1} A_i \frac{\partial}{\partial x_i}$$
applied to a_0. In the above formula we have set $t = x_{n+1}$ so as to write the differential operator in more symmetric form.

So the coefficient of $\hbar^{-k+1}$ in $P\left(a_0 e^{i\phi(x,t)/\hbar}\right)$ is
$$\left(Ra_0\right) e^{i\phi(x,t)/\hbar}$$
where R is the first order differential operator
$$R = \sum A_i \frac{\partial}{\partial x_i} + B + C.$$
We will derive the explicit expressions for the A_i, B and C below.

The strategy is then to look for solutions of the first order homogenous linear partial differential equation
$$Ra_0 = 0.$$
This is known as the **first order transport equation**.

Having found a_0, we next look for a_1 so that
$$P\left((a_0 + a_1\hbar)e^{i\phi/\hbar}\right) = O(h^{-k+3}).$$
From the above discussion it is clear that this amounts to solving an inhomogeneous linear partial differential equation of the form
$$Ra_1 = b_0$$
where b_0 is the coefficient of $\hbar^{-k+2}e^{i\phi/\hbar}$ in $P(a_0 e^{i\phi/\hbar})$ and where R is the *same operator as above.* Assuming that we can solve all these equations, we see that we have a recursive procedure involving the operator R for solving (1.1) to all orders, at least locally - up until we hit a caustic!

We will find that when we regard P as acting on $\frac{1}{2}$-densities (rather than on functions) then the operator R has an invariant (and beautiful) expression as a differential operator acting on $\frac{1}{2}$-densities on Λ, see equation (1.21) below. In fact, the differentiation part of the differential operator will be given by the vector field v_H which we know to be tangent to Λ. The differential operator on Λ will be defined even at caustics. This fact will be central in our study of global asymptotic solutions of hyperbolic equations.

In the next section we shall assume only the most elementary facts about $\frac{1}{2}$-densities - the fact that the product of two $\frac{1}{2}$-densities is a density and hence can be integrated if this product has compact support. Also that the concept of the Lie derivative of a $\frac{1}{2}$-density with respect to a vector field makes sense. If the reader is unfamiliar with these facts they can be found with many more details in Chapter 6.

1.3.1 A formula for the Lie derivative of a $\frac{1}{2}$-density.

We want to consider the following situation: H is a function on T^*X and Λ is a Lagrangian submanifold of T^*X on which $H = 0$. This implies that the corresponding Hamiltonian vector field is tangent to Λ. Indeed, for any $w \in T_z\Lambda,\ z \in \Lambda$ we have

$$\omega_X(v_H, w) = dH(w) = 0$$

since H is constant on Λ. Since Λ is Lagrangian, this implies that $v_H(z) \in T_z(\Lambda)$.

If τ is a smooth $\frac{1}{2}$-density on Λ, we can consider its Lie derivative with respect to the vector field v_H restricted to Λ. We want an explicit formula for this Lie derivative in terms of local coordinates on X on a neighborhood over which Λ is horizontal.

Let

$$\iota : \Lambda \to T^*X$$

denote the embedding of Λ as submanifold of X so we are assuming that

$$\pi \circ \iota : \Lambda \to X$$

is a diffeomorphism. (We have replaced X by the appropriate neighborhood over which Λ is horizontal and on which we have coordinates $x^1, \ldots, x^m$.) We let $dx^{\frac{1}{2}}$ denote the standard $\frac{1}{2}$-density relative to these coordinates. Let a be a function on X, so that

$$\tau := (\pi \circ \iota)^* \left(a dx^{\frac{1}{2}}\right)$$

is a $\frac{1}{2}$-density on Λ, and the most general $\frac{1}{2}$-density on Λ can be written in this form. Our goal in this section is to compute the Lie derivative $D_{v_H}\tau$ and express it in a similar form. We will prove:

Proposition 1.3.1. *If $\Lambda = \Lambda_\phi = \gamma_\phi(X)$ then*

$$D_{v_H|\Lambda}(\pi \circ \iota)^* \left(a dx^{\frac{1}{2}}\right) = b(\pi \circ \iota)^* \left(dx^{\frac{1}{2}}\right)$$

where

$$b = D_{v_H|\Lambda}((\pi\circ\iota)^* a) + \iota^* \left[\frac{1}{2}\sum_{i,j}\frac{\partial^2 H}{\partial\xi_i\partial\xi_j}\frac{\partial^2\phi}{\partial x^i\partial x^j} + \frac{1}{2}\sum_i \frac{\partial^2 H}{\partial\xi_i\partial x^i}\right]((\pi\circ\iota)^* a). \tag{1.15}$$

Proof. Since $D_v(f\tau) = (D_v f)\tau + f D_v\tau$ for any vector field v, function f and any $\frac{1}{2}$-density τ, it suffices to prove (1.15) for the case the $a \equiv 1$ in which case the first term disappears. By Leibnitz's rule,

$$D_{v_H}(\pi\circ\iota)^*\left(dx^{\frac{1}{2}}\right) = \frac{1}{2}c(\pi\circ\iota)^*\left(dx^{\frac{1}{2}}\right)$$

where

$$D_{v_H}(\pi\circ\iota)^*|dx| = c(\pi\circ\iota)^*|dx|.$$

Here we are computing the Lie derivative of the density $(\pi\circ\iota)^*|dx|$, but we get the same function c if we compute the Lie derivative of the m-form

$$D_{v_H}(\pi\circ\iota)^*(dx^1\wedge\cdots\wedge dx^m) = c(\pi\circ\iota)^*(dx^1\wedge\cdots\wedge dx^m).$$

Now $\pi^*(dx^1\wedge\cdots\wedge dx^m)$ is a well defined m-form on T^*X and

$$D_{v_H|\Lambda}(\pi\circ\iota)^*(dx^1\wedge\cdots\wedge dx^m) = \iota^* D_{v_H}\pi^*(dx^1\wedge\cdots\wedge dx^m).$$

We may write dx^j instead of $\pi^* dx^j$ with no risk of confusion and we get

$$\begin{aligned}
D_{v_H}(dx^1\wedge\cdots\wedge dx^m) &= \sum_j dx^1\wedge\cdots\wedge d(i(v_H)dx^j)\wedge\cdots\wedge dx^m \\
&= \sum_j dx^1\wedge\cdots\wedge d\frac{\partial H}{\partial\xi_j}\wedge\cdots\wedge dx^m \\
&= \sum_j \frac{\partial^2 H}{\partial\xi_j\partial x^j}dx^1\wedge\cdots\wedge dx^m + \\
&\qquad \sum_{jk} dx^1\wedge\cdots\wedge\frac{\partial^2 H}{\partial\xi_j\partial\xi_k}d\xi_k\wedge\cdots\wedge dx^m.
\end{aligned}$$

We must apply ι^* which means that we must substitute $d\xi_k = d\left(\frac{\partial\phi}{\partial x^k}\right)$ into the last expression. We get

$$c = \sum_{i,j}\frac{\partial^2 H}{\partial\xi_i\partial\xi_j}\frac{\partial^2\phi}{\partial x^i\partial x^j} + \sum_i\frac{\partial^2 H}{\partial\xi_i\partial x^i}$$

proving (1.15). □

1.3.2 The total symbol, locally.

Let U be an open subset of $\mathbb{R}^m$ and $x_1, \ldots x_m$ the standard coordinates. We will let D_j denote the differential operator

$$D_j = \frac{1}{i}\frac{\partial}{\partial x_j} = \frac{1}{\sqrt{-1}}\frac{\partial}{\partial x_j}.$$

For any multi-index $\alpha = (\alpha_1, \ldots, \alpha_m)$ where the α_j are non-negative integers, we let

$$D^\alpha := D_1^{\alpha_1} \cdots D_m^{\alpha_m}$$

and

$$|\alpha| := \alpha_1 + \cdots + \alpha_m.$$

So the most general k-th order linear differential operator P can be written as

$$P = P(x, D) = \sum_{|\alpha| \leq k} a_\alpha(x) D^\alpha.$$

The **total symbol** of P is defined as

$$e^{-i\frac{x\cdot\xi}{\hbar}} P e^{i\frac{x\cdot\xi}{\hbar}} = \sum_{j=0}^{k} \hbar^{-j} p_j(x,\xi)$$

so that

$$p_j(x,\xi) = \sum_{|\alpha|=j} a_\alpha(x)\xi^\alpha. \tag{1.16}$$

So p_k is exactly the principal symbol as defined in (1.4).

Since we will be dealing with operators of varying orders, we will denote the principal symbol of P by

$$\sigma(P).$$

We should emphasize that the definition of the total symbol is heavily coordinate dependent: If we make a non-linear change of coordinates, the expression for the total symbol in the new coordinates will not look like the expression in the old coordinates. However the principal symbol *does* have an invariant expression as a function on the cotangent bundle which is a polynomial in the fiber variables.

1.3.3 The transpose of P.

We continue our study of linear differential operators on an open subset $U \subset \mathbb{R}^n$. If f and g are two smooth functions of compact support on U then

$$\int_U (Pf) g dx = \int_U f P^t g dx$$

where, by integration by parts,

$$P^t g = \sum (-1)^{|\alpha|} D^\alpha (a_\alpha g).$$

(Notice that in this definition, following convention, we are using g and not $\overline{g}$ in the definition of P^t.) Now

$$D^\alpha(a_\alpha g) = a_\alpha D^\alpha g + \cdots$$

where the $\cdots$ denote terms with fewer differentiations in g. In particular, the principal symbol of P^t is

$$p_k^t(x,\xi) = (-1)^k p_k(x,\xi). \tag{1.17}$$

Hence the operator

$$Q := \frac{1}{2}(P - (-1)^k P^t) \tag{1.18}$$

is of order $k-1$ The **sub-principal symbol** is defined as the principal symbol of Q (considered as an operator of degree $(k-1)$). So

$$\sigma_{sub}(P) := \sigma(Q)$$

where Q is given by (1.18).

1.3.4 The formula for the sub-principal symbol.

We claim that

$$\sigma_{sub}(P)(x,\xi) = p_{k-1}(x,\xi) + \frac{\sqrt{-1}}{2} \sum_i \frac{\partial^2}{\partial x_i \partial \xi_i} p_k(x,\xi). \tag{1.19}$$

Proof. If $p_k(x,\xi) \equiv 0$, i.e. if P is actually an operator of degree $k-1$, then it follows from (1.17) (applied to $k-1$) and (1.18) that the principal symbol of Q is p_{k-1} which is the first term on the right in (1.19). So it suffices to prove (1.19) for operators which are strictly of order k. By linearity, it suffices to prove (1.19) for operators of the form

$$a_\alpha(x) D^\alpha.$$

By polarization it suffices to prove (1.19) for operators of the form

$$a(x) D^k, \quad D = \sum_{j=1}^k c_j D_j, \qquad c_i \in \mathbb{R}$$

and then, by making a linear change of coordinates, for an operator of the form

$$a(x) D_1^k.$$

For this operator
$$p_k(x,\xi) = a(x)\xi_1^k.$$
By Leibnitz's rule,
$$\begin{aligned}
P^t f &= (-1)^k D_1^k(af)\\
&= (-1)^k \sum_j \binom{k}{j} D_1^j a D_1^{k-j} f\\
&= (-1)^k \left(a D_1^k f + \frac{k}{i}\left(\frac{\partial a}{\partial x_1}\right) D_1^{k-1} f + \cdots \right) \quad \text{so}\\
Q &= \frac{1}{2}(P - (-1)^k P^t)\\
&= -\frac{k}{2i}\left(\frac{\partial a}{\partial x_1} D_1^{k-1} + \cdots\right)
\end{aligned}$$
and therefore
$$\begin{aligned}
\sigma(Q) &= \frac{ik}{2}\frac{\partial a}{\partial x_1}\xi_1^{k-1}\\
&= \frac{i}{2}\frac{\partial}{\partial x_1}\frac{\partial}{\partial \xi_1}(a\xi_1^k)\\
&= \frac{i}{2}\sum_j \frac{\partial^2 p_k}{\partial x_j \partial \xi_j}(x,\xi)
\end{aligned}$$
since p_k does not depend on ξ_j for $j > 1$, in this case. □

1.3.5 The local expression for the transport operator R.

We claim that
$$\hbar^k e^{-i\phi/\hbar} P(u e^{i\phi/\hbar}) = p_k(x, d\phi)u + \hbar R u + \cdots$$
where R is the first order differential operator
$$Ru =$$
$$\sum_j \frac{\partial p_k}{\partial \xi_j}(x,d\phi) D_j u + \left[\frac{1}{2\sqrt{-1}} \sum_{ij} \frac{\partial^2 p_k}{\partial \xi_i \partial \xi_j}(x,d\phi) \frac{\partial^2 \phi}{\partial x_i \partial x_j} + p_{k-1}(x,d\phi)\right] u. \tag{1.20}$$

Proof. The term coming from p_{k-1} is clearly the result of applying
$$\sum_{|\alpha|=k-1} a_\alpha D^\alpha.$$

So we only need to deal with a homogeneous operator of order k. Since the coefficients a_α are not going to make any difference in this formula, we need only prove it for the differential operator

$$P(x, D) = D^\alpha$$

which we will do by induction on $|\alpha|$.

For $|\alpha| = 1$ we have an operator of the form D_j and Leibnitz's rule gives

$$\hbar e^{-i\phi/\hbar} D_j(ue^{i\phi/\hbar}) = \frac{\partial \phi}{\partial x_j} u + \hbar D_j u$$

which is exactly (1.20) as $p_1(\xi) = \xi_j$, and so the second and third terms in (1.20) do not occur.

Suppose we have verified (1.20) for D^α and we want to check it for

$$D_r D^\alpha = D^{\alpha+\delta_r}.$$

So

$$\hbar^{|\alpha|+1} e^{-i\phi/\hbar}\left(D_r D^\alpha(ue^{i\phi/\hbar})\right) = \hbar e^{-i\phi/\hbar} D_r[(d\phi)^\alpha u e^{i\phi/\hbar} + \hbar(R_\alpha u)e^{i\phi/\hbar}] + \cdots$$

where R_α denotes the operator in (1.20) corresponding to D^α. A term involving the zero'th power of $\hbar$ can only come from applying the D_r to the exponential in the first expression and this will yield

$$(d\phi)^{\alpha+\delta_r} u$$

which $p_{|\alpha|+1}(d\phi)u$ as desired. In applying D_r to the second term in the square brackets and multiplying by $\hbar e^{-i\phi/\hbar}$ we get

$$\hbar^2 D_r(R_\alpha u) + \hbar \frac{\partial \phi}{\partial x_r} R_\alpha u$$

and we ignore the first term as we are ignoring all powers of $\hbar$ higher than the first. So all have to do is collect coefficients:

We have

$$D_r((d\phi^\alpha)u) = (d\phi)^\alpha D_r u + \frac{1}{\sqrt{-1}}\left[\alpha_1 (d\phi)^{\alpha-\delta_1} \frac{\partial^2 \phi}{\partial x_1 \partial x_r} + \cdots + \alpha_m (d\phi)^{\alpha-\delta_m} \frac{\partial^2 \phi}{\partial x_m \partial x_r}\right] u.$$

Also

$$\frac{\partial \phi}{\partial x_r} R_\alpha u =$$

$$\sum \alpha_i (d\phi)^{\alpha-\delta_i+\delta_r} D_i u + \frac{1}{2\sqrt{-1}} \sum_{ij} \alpha_i(\alpha_j - \delta_{ij})(d\phi)^{\alpha-\delta_i-\delta_j+\delta_r} \frac{\partial^2 \phi}{\partial x_i \partial x_j} u.$$

The coefficient of $D_j u, \quad j \neq r$ is

$$\alpha_j (d\phi)^{(\alpha+\delta_r-\delta_j)}$$

as desired. The coefficient of $D_r u$ is

$$(d\phi)^\alpha + \alpha_r (d\phi)^\alpha = (\alpha_r + 1)(d\phi)^{(\alpha+\delta_r)-\delta_r}$$

as desired.

Let us now check the coefficient of $\frac{\partial^2\phi}{\partial x_i \partial x_j}$. If $i \neq r$ and $j \neq r$ then the desired result is immediate.

If $j = r$, there are two sub-cases to consider: 1) $j = r, j \neq i$ and 2) $i = j = r$.

If $j = r$, $j \neq i$ remember that the sum in R_α is over *all* i and j, so the coefficient of $\frac{\partial^2\phi}{\partial x_i \partial x_j}$ in

$$\sqrt{-1}\frac{\partial\phi}{\partial x_r} R_\alpha u$$

is

$$\frac{1}{2}\left(\alpha_i\alpha_j + \alpha_j\alpha_i\right)(d\phi)^{\alpha-\delta_i} = \alpha_i\alpha_j (d\phi)^{\alpha-\delta_i}$$

to which we add

$$\alpha_i (d\phi)^{\alpha-\delta_i}$$

to get

$$\alpha_i(\alpha_j + 1)(d\phi)^{\alpha-\delta_i} = (\alpha+\delta_r)_i(\alpha+\delta_r)_j (d\phi)^{\alpha-\delta_i}$$

as desired.

If $i = j = r$ then the coefficient of $\frac{\partial^2\phi}{(\partial x_i)^2}$ in

$$\sqrt{-1}\frac{\partial\phi}{\partial x_r} R_\alpha u$$

is

$$\frac{1}{2}\alpha_i(\alpha_i - 1)(d\phi)^{\alpha-\delta_i}$$

to which we add

$$\alpha_i (d\phi)^{\alpha-\delta_i}$$

giving

$$\frac{1}{2}\alpha_i(\alpha_i + 1)(d\phi)^{\alpha-\delta_i}$$

as desired.

This completes the proof of (1.20).

1.3.6 Putting it together locally.

We have the following three formulas, some of them rewritten with H instead of p_k so as to conform with our earlier notation: The formula for the transport operator R given by (1.20):

$$\sum_j \frac{\partial H}{\partial \xi_j}(x,d\phi)D_j a + \left[\frac{1}{2\sqrt{-1}}\sum_{ij}\frac{\partial^2 H}{\partial\xi_i\partial\xi_j}(x,d\phi)\frac{\partial^2\phi}{\partial x_i\partial x_j} + p_{k-1}(x,d\phi)\right]a,$$

and the formula for the Lie derivative with respect to v_H of the pull back $(\pi\circ\iota)^*(adx^{\frac{1}{2}})$ given by $(\pi\circ\iota)*bdx^{\frac{1}{2}}$ where b is

$$\sum_j \frac{\partial H}{\partial \xi_j}(x,d\phi)\frac{\partial a}{\partial x_j} + \left[\frac{1}{2}\sum_{i,j}\frac{\partial^2 H}{\partial\xi_i\partial\xi_j}(x,d\phi)\frac{\partial^2\phi}{\partial x^i\partial x^j} + \frac{1}{2}\sum_i\frac{\partial^2 H}{\partial\xi_i\partial x^i}\right]a.$$

This is equation (1.15). Our third formula is the formula for the sub-principal symbol, equation (1.19), which says that

$$\sigma_{sub}(P)(x,\xi)a = \left[p_{k-1}(x,\xi) + \frac{\sqrt{-1}}{2}\sum_i\frac{\partial^2 H}{\partial x_i\partial\xi_i}(x,\xi)\right]a.$$

As first order partial differential operators on a, if we multiply the first expression above by $\sqrt{-1}$ we get the second plus $\sqrt{-1}$ times the third! So we can write the transport operator as

$$(\pi\circ\iota)^*[(Ra)dx^{\frac{1}{2}}] = \frac{1}{i}\left[D_{v_H} + i\sigma_{sub}(P)(x,d\phi)\right](\pi\circ\iota)^*(adx^{\frac{1}{2}}). \tag{1.21}$$

The operator inside the brackets on the right hand side of this equation is a perfectly good differential operator on $\frac{1}{2}$-densities on Λ. We thus have two questions to answer: Does this differential operator have invariant significance when Λ is horizontal - but in terms of a general coordinate transformation? Since the first term in the brackets comes from H and the symplectic form on the cotangent bundle, our question is one of attaching some invariant significance to the sub-principal symbol. We will deal briefly with this question in the next section and at more length in Chapter 6.

The second question is how to deal with the whole method - the eikonal equation, the transport equations, the meaning of the series in $\hbar$ etc. when we pass through a caustic. The answer to this question will occupy us for the whole book.

1.3.7 Differential operators on manifolds.

Differential operators on functions.

Let X be an m-dimensional manifold. An operator

$$P:\ C^\infty(X)\to C^\infty(X)$$

is called a differential operator of order k if, for every coordinate patch $(U, x_1, \ldots, x_m)$ the restriction of P to $C_0^\infty(U)$ is of the form

$$P = \sum_{|\alpha| \le k} a_\alpha D^\alpha, \qquad a_\alpha \in C^\infty(U).$$

As mentioned above, the total symbol of P is no longer well defined, but the principal symbol *is* well defined as a function on T^*X. Indeed, it is defined as in Section 1.2.1: The value of the principal symbol H at a point $(x, d\phi(x))$ is determined by

$$H(x, d\phi(x))u(x) = \hbar^k e^{-i\frac{\phi}{\hbar}}(P(ue^{i\frac{\phi}{\hbar}})(x) + O(\hbar).$$

What about the transpose and the sub-principal symbol?

Differential operators on sections of vector bundles.

Let $E \to X$ and $F \to X$ be vector bundles. Let E be of dimension p and F be of dimension q. We can find open covers of X by coordinate patches $(U, x_1, \ldots, x_m)$ over which E and F are trivial. So we can find smooth sections $r_1, \ldots, r_p$ of E such that every smooth section of E over U can be written as

$$f_1 r_1 + \cdots f_p r_p$$

where the f_i are smooth functions on U and smooth sections $s_1, \ldots, s_q$ of F such that every smooth section of F over U can be written as

$$g_1 s_1 + \cdots + g_q s_q$$

over U where the g_j are smooth functions. An operator

$$P : \ C^\infty(X, E) \to C^\infty(X, F)$$

is called a differential operator of order k if, for every such U the restriction of P to smooth sections of compact support supported in U is given by

$$P\left(f_1 r_1 + \cdots f_p r_p\right) = \sum_{j=1}^{q} \sum_{i=1}^{p} P_{ij} f_i s_j$$

where the P_{ij} are differential operators of order k.

In particular if E and F are line bundles so that $p = q = 1$ it makes sense to talk of differential operators of order k from smooth sections of E to smooth sections of F. In a local coordinate system with trivializations r of E and s of F a differential operator locally is given by

$$fr \mapsto (Pf)s.$$

If $E = F$ and $r = s$ it is easy to check that the principal symbol of P is independent of the trivialization. (More generally the matrix of principal symbols in the vector bundle case is well defined up to appropriate pre and post multiplication by change of bases matrices, i.e. is well defined as a section of $\mathrm{Hom}(E, F)$ pulled up to the cotangent bundle. See Chaper II of [GSGA] for the general discussion.)

In particular it makes sense to talk about a differential operator of degree k on the space of smooth $\frac{1}{2}$-densities and the principal symbol of such an operator.

The transpose and sub-principal symbol of a differential operator on $\frac{1}{2}$-densities.

If μ and ν are $\frac{1}{2}$-densities on a manifold X, their product $\mu \cdot \nu$ is a density (of order one). If this product has compact support, for example if μ or ν has compact support, then the integral

$$\int_X \mu \cdot \nu$$

is well defined. See Chapter 6 for details. So if P is a differential operator of degree k on $\frac{1}{2}$-densities, its transpose P^t is defined via

$$\int_X (P\mu) \cdot \nu = \int_X \mu \cdot (P^t \nu)$$

for all μ and ν one of which has compact support. Locally, in terms of a coordinate neighborhood $(U, x_1, \ldots, x_m)$, every $\frac{1}{2}$-density can be written as $f dx^{\frac{1}{2}}$ and then the local expression for P^t is given as in Section 1.3.3. We then define the operator Q as in equation (1.18) and the sub-principal symbol as the principal symbol of Q as an operator of degree $k-1$ just as in Section 1.3.3.

We have now answered our first question - that of giving a coordinate-free interpretation to the transport equation: Equation (1.21) makes good invariant sense if we agree that our differential operator is acting on $\frac{1}{2}$-densities rather than functions.

1.4 Semi-classical differential operators.

Until now, we have been considering asymptotic solutions to (hyperbolic) partial differential equations. The parameter $\hbar$ entered into the (approximate) solution, but was not part of the problem. In physics, $\hbar$ is a constant which enters into the formulation of the problem. This is most clearly seen in the study of Schrödinger's equation.

1.4.1 Schrödinger's equation and Weyl's law.

Consider the Schrödinger operator in n-dimensions:

$$P(\hbar): \; u \mapsto \left(-\hbar^2 \left(\frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}\right) + V\right) u.$$

In physics $\hbar$ is a constant closely related to Planck's constant. But we want to think of $\hbar$ as a small parameter. Weyl's law says that, under appropriate growth hypotheses on V, the operators $P(\hbar)$ have discrete spectrum (cf. Chapter 13, especially Section **??**) and that for any pair of real numbers $a < b$ the number of eigenvalues $E(\hbar)$ of $P(\hbar)$ between a and b can estimated by a certain volume in phase space:

$$\begin{aligned}&\#\left\{E(\hbar) : a \leq E(\hbar) \leq b\right\} \\ &= \frac{1}{(2\pi\hbar)^n}\left[\mathrm{Vol}\left(a \leq \|\xi\|^2 + V(x) \leq b\right) + o(1)\right]. \qquad (1.22)\end{aligned}$$

Physicists know Weyl's law as the "formula for the density of states".

We will give a proof of (1.22) in Chapter 9. For the moment, let us do two special cases where we can compute the spectrum explicitly, and so verify Weyl's law.

1.4.2 The harmonic oscillator.

Here V is assumed to be a positive definite quadratic function of x. The following exposition is taken from Evans and Zworski (a preliminary version of [Zwor].

$n = 1, \ \hbar = 1$**.**

This is taught in all elementary quantum mechanics courses. The operator $P = P(1)$ is

$$Pu = \left(-\frac{d^2}{dx^2} + x^2\right)u.$$

We have

$$\frac{d}{dx}e^{-x^2/2} = -xe^{-x^2/2} \quad \text{so} \quad \frac{d^2}{dx^2}e^{-x^2/2} = -e^{-x^2/2} + x^2e^{-x^2/2}$$

and hence $e^{-x^2/2}$ is an eigenvector of P with eigenvalue 1. The remaining eigenvalues are found by the method of "spectrum generating algebras": Define the **creation operator**

$$A_+ := D + ix.$$

Here

$$D = \frac{1}{i}\frac{d}{dx}$$

and ix denotes the operator of multiplication by ix. Notice that D is formally self-adjoint in the sense that integration by parts shows that

$$\int_{\mathbb{R}}(Df)\overline{g}dx = \int_{\mathbb{R}} f\overline{Dg}dx$$

for all smooth functions vanishing at infinity. Even more directly the operator of multiplication by ix is skew adjoint so we can write

$$A_+^* = A_- := D - ix$$

in the formal sense. The operator A_- is called the **annihilation** operator.

Also

$$\begin{aligned} A_+A_-u &= -u_{xx} - (xu)_x + xu_x + x^2u \\ &= -u_{xx} - u + x^2u \\ &= Pu - u \\ \text{and} \\ A_-A_+u &= -u_{xx} + (xu)_x - xu_x + x^2u \\ &= Pu + u. \end{aligned}$$

So we have proved

$$P = A_+A_- + I = A_-A_+ - I. \tag{1.23}$$

Notice that

$$A_-\left(e^{-x^2/2}\right) = ixe^{-x^2/2} - ixe^{-x^2/2} = 0$$

so the first equation above shows again that

$$v_0(x) := e^{-x^2/2}$$

is an eigenvector of P with eigenvalue 1. Let $v_1 := A_+v_0$. Then

$$Pv_1 = (A_+A_- + I)A_+v_0 = A_+(A_-A_+ - I)v_0 + 2A_+v_0 = A_+Pv_0 + 2A_+v_0 = 3v_1.$$

So v_1 is an eigenvector of P with eigenvalue 3. Proceeding inductively, we see that if we define

$$v_n := A_+^n v_0$$

then v_n is an eigenvector of P with eigenvalue $2n+1$.

Also,

$$[A_-, A_+] = A_-A_+ - A_+A_- = P + I - (P - I) = 2I.$$

This allows us to conclude the $(v_n, v_m) = 0$ if $m \neq n$. Indeed, we may suppose that $m > n$. Then $(v_n, v_m) = (A_+^n, v_0, A_+^m v_0) = (A_-^m A_+^n v_0, v_0)$ since $A_- = A_+^*$. If $n = 0$ this is 0 since A_-v_0=0. If $n > 0$ then

$$A_-^m A_+^n = A_-^{m-1} A_- A_+ A_+^{n-1} = A_-^{m-1}(A_+A_- + 2I)A_+^{n-1}.$$

By repeated use of this argument, we end up with a sum of expressions all being left multiples of A_- and hence give 0 when applied to v_0.

We let

$$u_n := \frac{1}{\|v_n\|} v_n$$

so that the u_n form an orthonormal set of eigenvectors. By construction, the v_n, and hence the u_n, are polynomials of degree (at most) n times v_0. So we have

$$u_n(x) = H_n(x)e^{-x^2/2}$$

and the H_n are called the **Hermite** polynomials of degree n. Since the u_n are linearly independent and of degree at most n, the coefficient of x^n in H_n can not vanish.

Finally, the u_n form a basis of $L^2(\mathbb{R})$. To prove this, we must show that if $g \in L^2(\mathbb{R})$ is orthogonal to all the u_n then $g = 0$. To see that this is so, if, $(g, u_n) = 0$ for all n, then $(g, pe^{-x^2/2}) = 0$ for all polynomials. Take p to be the n-th Taylor expansion of e^{ix}. These are all majorized by $e^{|x|}$ and $e^{|x|}e^{-x^2/2} \in L^2(\mathbb{R})$. So from the Lebesgue dominated convergence theorem we see that $(g, e^{ix}e^{-x^2/2}) = 0$ which says that the Fourier transform of $ge^{-x^2/2}$ vanishes. This implies that $ge^{-x^2/2} \equiv 0$. Since $e^{-x^2/2}$ does not vanish anywhere, this implies that $g = 0$.

$\hbar = 1,\ n$ **arbitrary.**

We may identify $L_2(\mathbb{R}^n)$ with the (completed) tensor product

$$L^2(\mathbb{R})\hat{\otimes}\cdots\hat{\otimes}L^2(\mathbb{R}) \quad n-\text{factors}$$

where $\hat{\otimes}$ denotes the completed tensor product.

Then the n-dimensional Schrödinger harmonic oscillator has the form

$$P\hat{\otimes}I\hat{\otimes}\cdots\hat{\otimes}I + I\hat{\otimes}P\hat{\otimes}\cdots\hat{\otimes}I + \cdots + I\hat{\otimes}\cdots\hat{\otimes}P$$

where P is the one dimensional operator. So the tensor products of the u's form an orthonormal basis of $L^2(\mathbb{R}^n)$ consisting of eigenvectors. Explicitly, let $\alpha = (\alpha_1, \ldots, \alpha_n)$ be a multi-index of non-negative integers and

$$u_\alpha(x_1, \ldots, x_n) := \prod_{j=1}^n H_{\alpha_j}(x_j)e^{-\frac{1}{2}(x_1^2+\cdots+x_n^2)}.$$

Then the u_α are eigenvectors of the operator

$$u \mapsto -\left(\frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}\right)u + \|x\|^2 u$$

with eigenvalues

$$2|\alpha| + n$$

where

$$|\alpha| := \alpha_1 + \cdots + \alpha_n.$$

Furthermore the u_α form an orthonormal basis of $L^2(\mathbb{R}^n)$.

$n = 1,\ \hbar$ **arbitrary.**

Consider the "rescaling operator"

$$S_\hbar : \quad u(x) \mapsto \hbar^{-\frac{1}{4}} u\left(\frac{x}{\hbar^{\frac{1}{2}}}\right).$$

This is a unitary operator on $L^2(\mathbb{R})$ and on smooth functions we have

$$\frac{d}{dx} \circ S_\hbar = \hbar^{-\frac{1}{2}} S_\hbar \circ \frac{d}{dx}$$

and

$$x^2 S_\hbar u = \hbar S_\hbar (x^2 u).$$

So

$$\left(-\hbar^2 \frac{d^2}{dx^2} + x^2\right) S_\hbar u = \hbar S_\hbar \left(-\frac{d^2}{dx^2} + x^2\right) u.$$

This shows that if we let

$$u_{j,\hbar}(x) = S_\hbar(u_j)$$

Then the $u_{j,\hbar}$ form an orthonormal basis of $L^2(\mathbb{R})$ and are eigenvectors of $P(\hbar)$ with eigenvalues $\hbar(2j+1)$.

n and $\hbar$ arbitrary.

We combine the methods of the two previous sections and conclude that

$$u_{\alpha,\hbar}(x) := \hbar^{-n/4} \prod_1^n H_{\alpha_j}\left(\frac{x_j}{\hbar^{\frac{1}{2}}}\right) e^{-\frac{\|x\|^2}{2\hbar}}$$

are eigenvectors of $P(\hbar)$ with eigenvalues

$$E_\alpha(\hbar) = (2|\alpha| + n)\hbar, \tag{1.24}$$

and the $u_{\alpha,\hbar}$ form an orthonormal basis of $L^2(\mathbb{R}^n)$.

Verifying Weyl's law.

In verifying Weys' law we may take $a = 0$ so

$$\begin{aligned} \#\left\{E(\hbar)|0 \leq E(\hbar) \leq b\right\} &= \left\{\alpha|0 \leq 2|\alpha| + n \leq \frac{b}{\hbar}\right\} \\ &= \left\{\alpha|\alpha_1 + \cdots + \alpha_n \leq \frac{b - n\hbar}{2\hbar}\right\}, \end{aligned}$$

the number of lattice points in the simplex

$$x_1 \geq 0, \ldots x_n \geq 0, x_1 + \cdots x_n \leq \frac{b - n\hbar}{\hbar}.$$

This number is (up to lower order terms) the volume of this simplex. Also, up to lower order terms we can ignore the $n\hbar$ in the numerator. Now the volume of the simplex is $1/n! \times$ the volume of the cube. So

$$\#\left\{E(\hbar)|0 \leq E(\hbar) \leq b\right\} = \frac{1}{n!}\left(\frac{b}{2\hbar}\right)^n + o\left(\frac{1}{\hbar^n}\right).$$

This gives the left hand side of Weyl's formula. As to the right hand side,

$$\mathrm{Vol}\left(\left\{\|x\|^2 + \|\xi\|^2 \le b\right\}\right)$$

is the volume of the ball of radius b in $2n$-dimensional space which is $\pi^n b^n/n!$, as we recall below. This proves Weyl's law for the harmonic oscillator.

Recall about the volume of spheres in $\mathbb{R}^k$.

Let A_{k-1} denote the volume of the $k-1$ dimensional unit sphere and V_k the volume of the k-dimensional unit ball, so

$$V_k = A_{k-1}\int_0^1 r^{k-1}dr = \frac{1}{k}A_{k-1}.$$

The integral $\int_{-\infty}^{\infty} e^{-x^2}dx$ is evaluated as $\sqrt{\pi}$ by the trick

$$\left(\int_{-\infty}^{\infty} e^{-x^2}dx\right)^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)}dxdy = 2\pi\int_0^{\infty} re^{-r^2}dr = \pi.$$

So

$$\pi^{k/2} = \left(\int_{-\infty}^{\infty} e^{-x^2}dx\right)^k = A_{k-1}\int_0^{\infty} r^{k-1}e^{-r^2}dr.$$

The usual definition of the Gamma function is

$$\Gamma(y) = \int_0^{\infty} t^{y-1}e^{-t}dt.$$

If we set $t = r^2$ this becomes

$$\Gamma(y) = 2\int_0^{\infty} e^{-r^2}r^{2y-1}dr.$$

So if we plug this back into the preceding formula we see that

$$A_{k-1} = \frac{2\pi^{k/2}}{\Gamma(\frac{k}{2})}.$$

Taking $k = 2n$ this gives

$$A_{2n-1} = \frac{2\pi^n}{(n-1)!}$$

and hence

$$V_{2n} = \frac{\pi^n}{n!}.$$

1.5 The Schrödinger operator on a Riemannian manifold.

As a generalization of the Schrödinger operator we studied above, we can consider the operator

$$\hbar^2\Delta + V$$

where Δ is the Laplacian of a Riemann manifold M. For example, if M is compact, then standard elliptical engineering tells us that this operator has discrete spectrum. Then once again Weyl's law is true, and the problem of estimating the "remainder" is of great interest.

We saw that Weyl's law in the case of a harmonic oscillator on Euclidean space involved counting the number of lattice points in a simplex. The problem of counting the number of lattice points in a polytope has attracted a lot of attention in recent years.

1.5.1 Weyl's law for a flat torus with $V = 0$.

Let us illustrate Weyl's law for the Schrödinger operator on a Riemannian manifold by examining what you might think is an "'easy case". Let M be the torus $M = (\mathbb{R}/(2\pi\mathbb{Z}) \times (\mathbb{R}/(2\pi\mathbb{Z})$ and take the flat (Euclidean) metric so that the Laplacian is

$$\Delta = -\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)$$

and take $V \equiv 0$! For simplicity in notation I will work with $\hbar = 1$. The eigenvectors of Δ are the functions $\phi_{m,n}$ where

$$\phi_{m,n}(x, y) = e^{imx+ny}$$

as m, n range over the integers and the corresponding eigenvalues are $m^2 + n^2$. So the number of eigenvalues $\leq r^2$ is the number of lattice points in the disk of radius r centered at the origin.

The corresponding region in phase space (with a slight change in notation) is the set of all (x, y, ξ, η) such that $\xi^2 + \eta^2 \leq r^2$. Since this condition does not involve x or y, this four dimensional volume is $(2\pi)^2\times$ the area of the disk of radius r. So we have verified Weyl's law.

But the problem of estimating the remainder is one of the great unsolved problems of mathematics: Gauss' problem in estimating the error term in counting the number of lattice points in a disk of radius r.

In the 1920's van der Corput made a major advance in this problem by introducing the method of stationary phase for this purpose, as we will expain in Chapter 15.

1.6 The plan.

We need to set up some language and prove various facts before we can return to our program of extending our method - the eikonal equation and the transport equations - so that they work past caustics.

In Chapter 2 we develop some necessary facts from symplectic geometry. In Chapter 3 we review some of the language of category theory. We also present a "baby" version of what we want to do later. We establish some facts about the category of finite sets and relations which will motivate similar constructions when we get to the symplectic "category" and its enhancement. We describe this symplectic "category" in Chapter 4. The objects in this "category" are symplectic manifolds and the morphisms are canonical relations. The quotation marks around the word "category" indicates that not all morphisms are composible.

In Chapter 5 we use this categorical language to explain how to find a local description of a Lagrangian submanifold of the cotangent bundle via "generating functions", a description which is valid even at caustics. The basic idea here goes back to Hamilton. But since this description depends on a choice, we must explain how to pass from one generating function to another. The main result here is the Hormander-Morse lemma which tells us that passage from one generating function to another can be accomplished by a series of "moves". The key analytic tool for proving this lemma is the method of stationary phase which we explain in Chapter 15. In Chapter 6 we study the calculus of $\frac{1}{2}$-densities, and in Chapter 7 we use half-densities to enhance the symplectic "category". In Chapter 8 we get to the main objects of study, which are oscillatory $\frac{1}{2}$-densities and develop their symbol calculus from an abstract and functorial point of view. In Chapter 9 we show how to turn these abstract considerations into local computations. In Chapter 14 we review the basic facts about the calculus of differential forms. In particular we review the Weil formula for the Lie derivative and the Moser trick for proving equivalence. In Chapter 13 we summarize, for the reader's convenience, various standard facts about the spectral theorem for self-adjoint operators on a Hilbert space.

Chapter 2

Symplectic geometry.

2.1 Symplectic vector spaces.

Let V be a (usually finite dimensional) vector space over the real numbers. A symplectic structure on V consists of an antisymmetric bilinear form

$$\omega : V \times V \to \mathbb{R}$$

which is non-degenerate. So we can think of ω as an element of $\wedge^2 V^*$ when V is finite dimensional, as we shall assume until further notice. A vector space equipped with a symplectic structure is called a symplectic vector space.

A basic example is $\mathbb{R}^2$ with

$$\omega_{\mathbb{R}^2}\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}\right) := \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

We will call this the standard symplectic structure on $\mathbb{R}^2$.

So if $u, v \in \mathbb{R}^2$ then $\omega_{\mathbb{R}^2}(u, v)$ is the oriented area of the parallelogram spanned by u and v.

2.1.1 Special kinds of subspaces.

If W is a subspace of symplectic vector space V then $W^\perp$ denotes the symplectic orthocomplement of W:

$$W^\perp := \{v \in V \mid \omega(v, w) = 0, \ \forall w \in W\}.$$

A subspace is called

1. **symplectic** if $W \cap W^\perp = \{0\}$,
2. **isotropic** if $W \subset W^\perp$,
3. **coisotropic** if $W^\perp \subset W$, and

4. **Lagrangian** if $W = W^\perp$.

Since $(W^\perp)^\perp = W$ by the non-degeneracy of ω, it follows that W is symplectic if and only if $W^\perp$ is. Also, the restriction of ω to any symplectic subspace W is non-degenerate, making W into a symplectic vector space. Conversely, to say that the restriction of ω to W is non-degenerate means precisely that $W \cap W^\perp = \{0\}$.

2.1.2 Normal forms.

For any non-zero $e \in V$ we can find an $f \in V$ such that $\omega(e, f) = 1$ and so the subspace W spanned by e and f is a two dimensional symplectic subspace. Furthermore the map

$$e \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad f \mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

gives a symplectic isomorphism of W with $\mathbb{R}^2$ with its standard symplectic structure. We can apply this same construction to $W^\perp$ if $W^\perp \neq 0$. Hence, by induction, we can decompose any symplectic vector space into a direct sum of two dimensional symplectic subspaces:

$$V = W_1 \oplus \cdots W_d$$

where dim $V = 2d$ (proving that every symplectic vector space is even dimensional) and where the W_i are pairwise (symplectically) orthogonal and where each W_i is spanned by e_i, f_i with $\omega(e_i, f_i) = 1$. In particular this shows that all $2d$ dimensional symplectic vector spaces are isomorphic, and isomorphic to a direct sum of d copies of $\mathbb{R}^2$ with its standard symplectic structure.

2.1.3 Existence of Lagrangian subspaces.

Let us collect the $e_1, \ldots, e_d$ in the above construction and let L be the subspace they span. It is clearly isotropic. Also, $e_1, \ldots, e_n, f_1, \ldots, f_d$ form a basis of V. If $v \in V$ has the expansion

$$v = a_1 e_1 + \cdots a_d e_d + b_1 f_1 + \cdots + b_d f_d$$

in terms of this basis, then $\omega(e_i, v) = b_i$. So $v \in L^\perp \Rightarrow v \in L$. Thus L is Lagrangian. So is the subspace M spanned by the f's.

Conversely, if L is a Lagrangian subspace of V and if M is a complementary Lagrangian subspace, then ω induces a non-degenerate linear pairing of L with M and hence any basis $e_1, \cdots e_d$ picks out a dual basis $f_1, \cdots .f_d$ of M giving a basis of V of the above form.

2.1.4 Consistent Hermitian structures.

In terms of the basis $e_1, \ldots, e_n, f_1, \ldots, f_d$ introduced above, consider the linear map

$$J: \quad e_i \mapsto -f_i, \quad f_i \mapsto e_i.$$

It satisfies

$$\begin{aligned} J^2 &= -I, & (2.1) \\ \omega(Ju, Jv) &= \omega(u, v), \quad \text{and} & (2.2) \\ \omega(Ju, v) &= \omega(Jv, u). & (2.3) \end{aligned}$$

Notice that any J which satisfies two of the three conditions above automatically satisfies the third. Condition (2.1) says that J makes V into a d-dimensional complex vector space. Condition (2.2) says that J is a symplectic transformation, i.e acts so as to preserve the symplectic form ω. Condition (2.3) says that $\omega(Ju, v)$ is a real symmetric bilinear form.

All three conditions (really any two out of the three) say that $(\ ,\ ) = (\ ,\ )_{\omega,J}$ defined by

$$(u, v) = \omega(Ju, v) + i\omega(u, v)$$

is a semi-Hermitian form whose imaginary part is ω. For the J chosen above this form is actually Hermitian, that is the real part of $(\ ,\ )$ is positive definite.

2.2 Lagrangian complements.

The results of this section will be used extensively, especially in Chapter 5.

Let V be a symplectic vector space.

Proposition 2.2.1. *Given any finite collection of Lagrangian subspaces $M_1, \ldots, M_k$ of V one can find a Lagrangian subspace L such that*

$$L \cap M_j = \{0\}, \quad i = 1, \ldots k.$$

Proof. We can always find an isotropic subspace L with $L \cap M_j = \{0\}, \quad i = 1, \ldots k$, for example a line which does not belong to any of these subspaces. Suppose that L is an isotropic subspace with $L \cap M_j = \{0\}$, $\forall j$ and is not properly contained in a larger isotropic subspace with this property. We claim that L is Lagrangian. Indeed, if not, $L^\perp$ is a coisotropic subspace which strictly contains L. Let $\pi : L^\perp \to L^\perp/L$ be the quotient map. Each of the spaces $\pi(L^\perp \cap M_j)$ is an isotropic subspace of the symplectic vector space $L^\perp/L$ and so each of these spaces has positive codimension. So we can choose a line ℓ in $L^\perp/L$ which does not intersect any of the $\pi(L^\perp \cap M_j)$. Then $L' := \pi^{-1}(\ell)$ is an isotropic subspace of $L^\perp \subset V$ with $L' \cap M_j = \{0\}$, $\forall\ j$ and strictly containing L, a contradiction. □

In words, given a finite collection of Lagrangian subspaces, we can find a Lagrangian subspace which is transversal to all of them.

2.2.1 Choosing Lagrangian complements "consistently".

The results of this section are purely within the framework of symplectic linear algebra. Hence their logical place is here. However their main interest is that they serve as lemmas for more geometrical theorems, for example the Weinstein isotropic embedding theorem. The results here all have to do with making choices in a "consistent" way, so as to guarantee, for example, that the choices can be made to be invariant under the action of a group.

For any a Lagrangian subspace $L \subset V$ we will need to be able to choose a complementary Lagrangian subspace L', and do so in a consistent manner, depending, perhaps, on some auxiliary data. Here is one such way, depending on the datum of a symmetric positive definite bilinear form B on V. (Here B has nothing to do with with the symplectic form.)

Let L^B be the orthogonal complement of L relative to the form B. So

$$\dim L^B = \dim L = \frac{1}{2} \dim V$$

and any subspace $W \subset V$ with

$$\dim W = \frac{1}{2} \dim V \quad \text{and } W \cap L = \{0\}$$

can be written as

$$\text{graph}(A)$$

where $A : L^B \to L$ is a linear map. That is, under the vector space identification

$$V = L^B \oplus L$$

the elements of W are all of the form

$$w + Aw, \quad w \in L^B.$$

We have

$$\omega(u + Au, w + Aw) = \omega(u, w) + \omega(Au, w) + \omega(u, Aw)$$

since $\omega(Au, Aw) = 0$ as L is Lagrangian. Let C be the bilinear form on L^B given by

$$C(u, w) := \omega(Au, w).$$

Thus W is Lagrangian if and only if

$$C(u, w) - C(w, u) = -\omega(u, w).$$

Now

$$\text{Hom}(L^B, L) \sim L \otimes L^{B*} \sim L^{B*} \otimes L^{B*}$$

under the identification of L with L^{B*} given by ω. Thus the assignment $A \leftrightarrow C$ is a bijection, and hence the space of all Lagrangian subspaces complementary to L is in one to one correspondence with the space of all bilinear forms C on L^B which satisfy $C(u, w) - C(w, u) = -\omega(u, w)$ for all $u, w \in L^B$. An obvious choice is to take C to be $-\frac{1}{2}\omega$ restricted to L^B. In short,

Proposition 2.2.2. *Given a positive definite symmetric form on a symplectic vector space V, there is a consistent way of assigning a Lagrangian complement L' to every Lagrangian subspace L.*

Here the word "consistent" means that the choice depends only on B. This has the following implication: Suppose that T is a linear automorphism of V which preserves both the symplectic form ω and the positive definite symmetric form B. In other words, suppose that

$$\omega(Tu, Tv) = \omega(u, v) \quad \text{and} \quad B(Tu, Tv) = B(u, v) \quad \forall\ u, v \in V.$$

Then if $L \mapsto L'$ is the correspondence given by the proposition, then

$$TL \mapsto TL'.$$

More generally, if $T : V \to W$ is a symplectic isomorphism which is an isometry for a choice of positive definite symmetric bilinear forms on each, the above equation holds.

Given L and B (and hence L') we determined the complex structure J by

$$J : L \to L', \ \ \omega(u, Jv) = B(u, v) \quad u, v \in L$$

and then

$$J := -J^{-1} : L' \to L$$

and extending by linearity to all of V so that

$$J^2 = -I.$$

Then for $u, v \in L$ we have

$$\omega(u, Jv) = B(u, v) = B(v, u) = \omega(v, Ju)$$

while

$$\omega(u, JJv) = -\omega(u, v) = 0 = \omega(Jv, Ju)$$

and

$$\omega(Ju, JJv) = -\omega(Ju, v) = -\omega(Jv, u) = \omega(Jv, JJu)$$

so (2.3) holds for all $u, v \in V$. We should write $J_{B,L}$ for this complex structure, or J_L when B is understood

Suppose that T preserves ω and B as above. We claim that

$$J_{TL} \circ T = T \circ J_L \tag{2.4}$$

so that T is complex linear for the complex structures J_L and J_{TL}. Indeed, for $u, v \in L$ we have

$$\omega(Tu, J_{TL}Tv) = B(Tu, Tv)$$

by the definition of J_{TL}. Since B is invariant under T the right hand side equals $B(u, v) = \omega(u, J_L v) = \omega(Tu, TJ_L v)$ since ω is invariant under T. Thus

$$\omega(Tu, J_{TL}Tv) = \omega(Tu, TJ_L v)$$

showing that

$$TJ_L = J_{TL}T$$

when applied to elements of L. This also holds for elements of L'. Indeed every element of L' is of the form $J_L u$ where $u \in L$ and $TJ_L u \in TL'$ so

$$J_{TL}TJ_L u = -J_{TL}^{-1}TJ_L u = -Tu = TJ_L(J_L u). \quad \square$$

Let I be an isotropic subspace of V and let $I^\perp$ be its symplectic orthogonal subspace so that $I \subset I^\perp$. Let

$$I_B = (I^\perp)^B$$

be the B-orthogonal complement to $I^\perp$. Thus

$$\dim I_B = \dim I$$

and since $I_B \cap I^\perp = \{0\}$, the spaces I_B and I are non-singularly paired under ω. In other words, the restriction of ω to $I_B \oplus I$ is symplectic. The proof of the preceding proposition gives a Lagrangian complement (inside $I_B \oplus I$) to I which, as a subspace of V has zero intersection with $I^\perp$. We have thus proved:

Proposition 2.2.3. *Given a positive definite symmetric form on a symplectic vector space V, there is a consistent way of assigning an isotropic complement I' to every co-isotropic subspace $I^\perp$.*

We can use the preceding proposition to prove the following:

Proposition 2.2.4. *Let V_1 and V_2 be symplectic vector spaces of the same dimension, with $I_1 \subset V_1$ and $I_2 \subset V_2$ isotropic subspaces, also of the same dimension. Suppose we are given*

- *a linear isomorphism $\lambda : I_1 \to I_2$ and*
- *a symplectic isomorphism $\ell : I_1^\perp / I_1 \to I_2^\perp / I_2$.*

Then there is a symplectic isomorphism

$$\gamma : V_1 \to V_2$$

such that

1. *$\gamma : I_1^\perp \to I_2^\perp$ and (hence) $\gamma : I_1 \to I_2$,*
2. *The map induced by γ on $I_1^\perp / I_1$ is ℓ and*
3. *The restriction of γ to I_1 is λ.*

Furthermore, in the presence of positive definite symmetric bilinear forms B_1 on V_1 and B_2 on V_2 the choice of γ can be made in a "canonical" fashion.

Indeed, choose isotropic complements I_{1B} to $I_1^\perp$ and I_{2B} to $I_2^\perp$ as given by the preceding proposition, and also choose B orthogonal complements Y_1 to I_1 inside $I_1^\perp$ and Y_2 to I_2 inside $I_2^\perp$. Then Y_i $(i = 1, 2)$ is a symplectic subspace of V_i which can be identified as a symplectic vector space with $I_i^\perp/I_i$. We thus have

$$V_1 = (I_1 \oplus I_{1B}) \oplus Y_1$$

as a direct sum decomposition into the sum of the two symplectic subspaces $(I_1 \oplus I_{1B})$ and Y_1 with a similar decomposition for V_2. Thus ℓ gives a symplectic isomorphism of $Y_1 \to Y_2$. Also

$$\lambda \oplus (\lambda^*)^{-1} : I_1 \oplus I_{1B} \to I_2 \oplus I_{2B}$$

is a symplectic isomorphism which restricts to λ on I_1. □

2.3 Equivariant symplectic vector spaces.

Let V be a symplectic vector space. We let $Sp(V)$ denote the group of all all symplectic automorphisms of V, i.e all maps T which satisfy $\omega(Tu, Tv) = \omega(u, v)\ \forall\ u, v \in V$.

A representation $\tau : G \to \mathrm{Aut}(V)$ of a group G is called symplectic if in fact $\tau : G \to Sp(V)$. Our first task will be to show that if G is compact, and τ is symplectic, then we can find a J satisfying (2.1) and (2.2), which commutes with all the $\tau(a),\ a \in G$ and such that the associated Hermitian form is positive definite.

2.3.1 Invariant Hermitian structures.

Once again, let us start with a positive definite symmetric bilinear form B. By averaging over the group we may assume that B is G invariant. (Here is where we use the compactness of G.) Then there is a unique linear operator K such that

$$B(Ku, v) = \omega(u, v) \quad \forall\ u, v \in V.$$

Since both B and ω are G-invariant, we conclude that K commutes with all the $\tau(a),\ a \in G$. Since $\omega(v, u) = -\omega(u, v)$ we conclude that K is skew adjoint relative to B, i.e. that

$$B(Ku, v) = -B(u, Kv).$$

Also K is non-singular. Then K^2 is symmetric and non-singular, and so V can be decomposed into a direct sum of eigenspaces of K^2 corresponding to distinct eigenvalues, all non-zero. These subspaces are mutually orthogonal under B and invariant under G. If $K^2 u = \mu u$ then

$$\mu B(u, u) = B(K^2 u, u) = -B(Ku, Ku) < 0$$

so all these eigenvalues are negative; we can write each μ as $\mu = -\lambda^2,\ \lambda > 0$. Furthermore, if $K^2 u = -\lambda^2 u$ then

$$K^2(Ku) = KK^2 u = -\lambda^2 Ku$$

so each of these eigenspaces is invariant under K. Also, any two subspaces corresponding to different values of λ^2 are orthogonal under ω. So we need only define J on each such subspace so as to commute with all the $\tau(a)$ and so as to satisfy (2.1) and (2.2), and then extend linearly. On each such subspace set

$$J := \lambda K^{-1}.$$

Then (on this subspace)

$$J^2 = \lambda^2 K^{-2} = -I$$

and

$$\omega(Ju, v) = \lambda\omega(K^{-1}u, v) = \lambda B(u, v)$$

is symmetric in u and v. Furthermore $\omega(Ju, u) = \lambda B(u, u) > 0$. □

Notice that if τ is irreducible, then the Hermitian form $(\ , \ ) = \omega(J\cdot, \cdot) + i\omega(\cdot, \cdot)$ is uniquely determined by the property that its imaginary part is ω.

2.3.2 The space of fixed vectors for a compact group of symplectic automorphisms is symplectic.

If we choose J as above, if $\tau(a)u = u$ then $\tau(a)Ju = Ju$. So the space of fixed vectors is a complex subspace for the complex structure determined by J. But the restriction of a positive definite Hermitian form to any (complex) subspace is again positive definite, in particular non-singular. Hence its imaginary part, the symplectic form ω, is also non-singular. □

This result need not be true if the group is not compact. For example, the one parameter group of shear transformations

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

in the plane is symplectic as all of these matrices have determinant one. But the space of fixed vectors is the x-axis.

2.3.3 Toral symplectic actions.

Suppose that $G = \mathbf{T}^n$ is an n-dimensional torus, and that $\mathbf{g}$ denotes its Lie algebra. Then $\exp: \mathbf{g} \to G$ is a surjective homomorphism, whose kernel $\mathbb{Z}_G$ is a lattice.

If $\tau : G \to U(V)$ as above, we can decompose V into a direct sum of one dimensional complex subspaces

$$V = V_1 \oplus \cdots \oplus V_d$$

where the restriction of τ to each subspace is given by

$$\tau_{|V_k}(\exp\,\xi)v = e^{2\pi i \alpha_k(\xi)}v$$

where

$$\alpha_k \in \mathbb{Z}_G^*,$$

the dual lattice.

2.4 Symplectic manifolds.

Recall that a manifold M is called **symplectic** if it comes equipped with a closed non-degenerate two form ω. A diffeomorphism is called symplectic if it preserves ω. We shall usually shorten the phrase "symplectic diffeomorphism" to **symplectomorphism**

A vector field v is called symplectic if

$$D_v\omega = 0.$$

Since $D_v\omega = d\iota(v)\omega + \iota(v)d\omega = d\iota(v)\omega$ as $d\omega = 0$, a vector field v is symplectic if and only if $\iota(v)\omega$ is closed.

Recall that a vector field v is called **Hamiltonian** if $\iota(v)\omega$ is exact. If θ is a closed one form, and v a vector field, then $D_v\theta = d\iota(v)\theta$ is exact. Hence if v_1 and v_2 are symplectic vector fields

$$D_{v_1}\iota(v_2)\omega = \iota([v_1, v_2])\omega$$

so $[v_1, v_2]$ is Hamiltonian with

$$\iota([v_1, v_2])\omega = d\omega(v_2, v_1).$$

2.5 Darboux style theorems.

These are theorems which state that two symplectic structures on a manifold are the same or give a normal form near a submanifold etc. We will prove them using the Moser-Weinstein method. This method hinges on the basic formula of differential calculus: If $f_t : X \to Y$ is a smooth family of maps and ω_t is a one parameter family of differential forms on Y then

$$\frac{d}{dt}f_t^*\omega_t = f_t^*\frac{d}{dt}\omega_t + Q_t d\omega_t + dQ_t\omega_t \tag{2.5}$$

where

$$Q_t : \Omega^k(Y) \to \Omega^{k-1}(X)$$

is given by

$$Q_t\tau(w_1, \ldots, w_{k-1}) := \tau(v_t, df_t(w_1), \ldots, df_t(w_{k-1}))$$

where

$$v_t : X \to T(Y), \quad v_t(x) := \frac{d}{dt}f_t(x).$$

If ω_t does not depend explicitly on t then the first term on the right of (2.5) vanishes, and integrating (2.5) with respect to t from 0 to 1 gives

$$f_1^* - f_0^* = dQ + Qd, \quad Q := \int_0^1 Q_t dt. \tag{2.6}$$

We give a review of all of this in Chapter 14. We urge the reader who is unfamiliar with these ideas to pause here and read Chapter 14.

Here is the first Darboux type theorem:

2.5.1 Compact manifolds.

Theorem 2.5.1. *Let M be a compact manifold, ω_0 and ω_1 two symplectic forms on M in the same cohomology class so that*

$$\omega_1 - \omega_0 = d\alpha$$

for some one form α. Suppose in addition that

$$\omega_t := (1-t)\omega_0 + t\omega_1$$

is symplectic for all $0 \le t \le 1$. Then there exists a diffeomorphism $f : M \to M$ such that

$$f^*\omega_1 = \omega_0.$$

Proof. Solve the equation

$$\iota(v_t)\omega_t = -\alpha$$

which has a unique solution v_t since ω_t is symplectic. Then solve the time dependent differential equation

$$\frac{df_t}{dt} = v_t(f_t), \quad f_0 = \mathrm{id}$$

which is possible since M is compact. Since

$$\frac{d\omega_t}{dt} = d\alpha,$$

the fundamental formula (2.5) gives

$$\frac{df_t^*\omega_t}{dt} = f_t^*\left[d\alpha + 0 - d\alpha\right] = 0$$

so

$$f_t^*\omega_t \equiv \omega_0.$$

In particular, set $t = 1$. □

This style of argument was introduced by Moser and applied to Darboux type theorems by Weinstein.

Here is a modification of the above:

Theorem 2.5.2. *Let M be a compact manifold, and ω_t, $0 \le t \le 1$ a family of symplectic forms on M in the same cohomology class.*

Then there exists a diffeomorphism $f : M \to M$ such that

$$f^*\omega_1 = \omega_0.$$

Proof. Break the interval $[0,1]$ into subintervals by choosing $t_0 = 0 < t_1 < t_2 < \cdots < t_N = 1$ and such that on each subinterval the "chord" $(1-s)\omega_{t_i} + s\omega_{t_{i+1}}$ is close enough to the curve $\omega_{(1-s)t_i + st_{i+1}}$ so that the forms $(1-s)\omega_{t_i} + s\omega_{t_{i+1}}$ are symplectic. Then successively apply the preceding theorem. □

2.5.2 Compact submanifolds.

The next version allows M to be non-compact but has to do with with behavior near a compact submanifold. We will want to use the following proposition:

Proposition 2.5.1. *Let X be a compact submanifold of a manifold M and let*

$$i : X \to M$$

denote the inclusion map. Let $\gamma \in \Omega^k(M)$ be a k-form on M which satisfies

$$\begin{aligned} d\gamma &= 0 \\ i^*\gamma &= 0. \end{aligned}$$

Then there exists a neighborhood U of X and a $k-1$ form β defined on U such that

$$\begin{aligned} d\beta &= \gamma \\ \beta_{|X} &= 0. \end{aligned}$$

(This last equation means that at every point $p \in X$ we have

$$\beta_p(w_1, \ldots, w_{k-1}) = 0$$

for all tangent vectors, not necessarily those tangent to X. So it is a much stronger condition than $i^*\beta = 0$.)

Proof. By choice of a Riemann metric and its exponential map, we may find a neighborhood of W of X in M and a smooth retract of W onto X, that is a one parameter family of smooth maps

$$r_t : W \to W$$

and a smooth map $\pi : W \to X$ with

$$r_1 = \mathrm{id}, \quad r_0 = i \circ \pi, \ \pi : W \to X, \ \ r_t \circ i \equiv i.$$

Write

$$\frac{dr_t}{dt} = w_t \circ r_t$$

and notice that $w_t \equiv 0$ at all points of X. Hence the form

$$\beta := Q\gamma$$

has all the desired properties where Q is as in (2.6). □

Theorem 2.5.3. *Let X, M and i be as above, and let ω_0 and ω_1 be symplectic forms on M such that*

$$i^*\omega_1 = i^*\omega_0$$

and such that

$$(1-t)\omega_0 + t\omega_1$$

is symplectic for $0 \le t \le 1$. *Then there exists a neighborhood* U *of* M *and a smooth map*

$$f : U \to M$$

such that

$$f_{|X} = id \quad \text{and} \quad f^*\omega_0 = \omega_1.$$

Proof. Use the proposition to find a neighborhood W of X and a one form α defined on W and vanishing on X such that

$$\omega_1 - \omega_0 = d\alpha$$

on W. Let v_t be the solution of

$$\iota(v_t)\omega_t = -\alpha$$

where $\omega_t = (1-t)\omega_0 + t\omega_1$. Since v_t vanishes identically on X, we can find a smaller neighborhood of X if necessary on which we can integrate v_t for $0 \le t \le 1$ and then apply the Moser argument as above. □

A variant of the above is to assume that we have a curve of symplectic forms ω_t with $i^*\omega_t$ independent of t.

Finally, a very useful variant is Weinstein's

Theorem 2.5.4. X, M, i *as above, and* ω_0 *and* ω_1 *two symplectic forms on* M *such that* $\omega_{1|X} = \omega_{0|X}$. *Then there exists a neighborhood* U *of* M *and a smooth map*

$$f : U \to M$$

such that

$$f_{|X} = id \quad \text{and} \quad f^*\omega_0 = \omega_1.$$

Here we can find a neighborhood of X such that

$$(1-t)\omega_0 + t\omega_1$$

is symplectic for $0 \le t \le 1$ since X is compact. □

One application of the above is to take X to be a point. The theorem then asserts that all symplectic structures of the same dimension are locally symplectomorphic. This is the original theorem of Darboux.

2.5.3 The isotropic embedding theorem.

Another important application of the preceding theorem is Weinstein's isotropic embedding theorem: Let (M, ω) be a symplectic manifold, X a compact manifold, and $i : X \to M$ an isotropic embedding, which means that $di_x(TX)_x$ is an isotropic subspace of $TM_{i(x)}$ for all $x \in X$. Thus

$$di_x(TX)_x \subset (di_x(TX)_x)^\perp$$

where $(di_x(TX)_x)^\perp$ denotes the orthogonal complement of $di_x(TX)_x$ in $TM_{i(x)}$ relative to $\omega_{i(x)}$. Hence

$$(di_x(TX)_x)^\perp/di_x(TX)_x$$

is a symplectic vector space, and these fit together into a symplectic vector bundle (i.e. a vector bundle with a symplectic structure on each fiber). We will call this the symplectic normal bundle of the embedding, and denote it by

$$SN_i(X)$$

or simply by $SN(X)$ when i is taken for granted.

Suppose that U is a neighborhood of $i(X)$ and $g : U \to N$ is a symplectomorphism of U into a second symplectic manifold N. Then $j = g \circ i$ is an isotropic embedding of X into N and f induces an isomorphism

$$g_* : NS_i(X) \to NS_j(X)$$

of symplectic vector bundles. Weinstein's isotropic embedding theorem asserts conversely, any isomorphism between symplectic normal bundles is in fact induced by a symplectomorphism of a neighborhood of the image:

Theorem 2.5.5. *Let (M, ω_M, X, i) and (N, ω_N, X, j) be the data for isotropic embeddings of a compact manifold X. Suppose that*

$$\ell : SN_i(X) \to SN_j(X)$$

is an isomorphism of symplectic vector bundles. Then there is a neighborhood U of $i(X)$ in M and a symplectomorphism g of U onto a neighborhood of $j(X)$ in N such that

$$g_* = \ell.$$

For the proof, we will need the following extension lemma:

Proposition 2.5.2. *Let*

$$i : X \to M, \quad j : Y \to N$$

be embeddings of compact manifolds X and Y into manifolds M and N. suppose we are given the following data:

- *A smooth map $f : X \to Y$ and, for each $x \in X$,*
- *A linear map $A_xTM_{i(x)} \to TN_{j(f(x))}$ such that the restriction of A_x to $TX_x \subset TM_{i(x)}$ coincides with df_x.*

Then there exists a neighborhood W of X and a smooth map $g : W \to N$ such that

$$g \circ i = f \circ i$$

and

$$dg_x = A_x \quad \forall\ x \in X.$$

Proof. If we choose a Riemann metric on M, we may identify (via the exponential map) a neighborhood of $i(X)$ in M with a section of the zero section of X in its (ordinary) normal bundle. So we may assume that $M = \mathcal{N}_i X$ is this normal bundle. Also choose a Riemann metric on N, and let

$$\exp : \mathcal{N}_j(Y) \to N$$

be the exponential map of this normal bundle relative to this Riemann metric. For $x \in X$ and $v \in N_i(i(x))$ set

$$g(x, v) := \exp_{j(x)}(A_x v).$$

Then the restriction of g to X coincides with f, so that, in particular, the restriction of dg_x to the tangent space to T_x agrees with the restriction of A_x to this subspace, and also the restriction of dg_x to the normal space to the zero section at x agrees A_x so g fits the bill. □

Proof of the theorem. We are given linear maps $\ell_x : (I_x^\perp / I_x) \to J_x^\perp / J_x$ where $I_x = di_x(TX)_x$ is an isotropic subspace of $V_x := TM_{i(x)}$ with a similar notation involving j. We also have the identity map of

$$I_x = TX_x = J_x.$$

So we may apply Proposition 2.2.4 to conclude the existence, for each x of a unique symplectic linear map

$$A_x : \ TM_{i(x)} \to TN_{j(x)}$$

for each $x \in X$. We may then extend this to an actual diffeomorphism, call it h on a neighborhood of $i(X)$, and since the linear maps A_x are symplectic, the forms

$$h^* \omega_N \quad \text{and } \omega_M$$

agree at all points of X. We then apply Theorem 2.5.4 to get a map k such that $k^*(h^* \omega_N) = \omega_M$ and then $g = h \circ k$ does the job. □

Notice that the constructions were all determined by the choice of a Riemann metric on M and of a Riemann metric on N. So if these metrics are invariant under a group G, the corresponding g will be a G-morphism. If G is compact, such invariant metrics can be constructed by averaging over the group.

An important special case of the isotropic embedding theorem is where the embedding is not merely isotropic, but is Lagrangian. Then the symplectic normal bundle is trivial, and the theorem asserts that all Lagrangian embeddings of a compact manifold are locally equivalent, for example equivalent to the embedding of the manifold as the zero section of its cotangent bundle.

2.6 The space of Lagrangian subspaces of a symplectic vector space.

Let $V = (V, \omega)$ be a symplectic vector space of dimension $2n$. We let $\mathcal{L}(V)$ denote the space of all Lagrangian subspaces of V. It is called the **Lagrangian Grassmannian**.

If $M \in \mathcal{L}(V)$ is a fixed Lagrangian subspace, we let $\mathcal{L}(V, M)$ denote the subset of $\mathcal{L}(V)$ consisting of those Lagrangian subspaces which are transversal to M.

Let $L \in \mathcal{L}(V, M)$ be one such subspace. The non-degenerate pairing between L and M identifies M with the dual space L^* of L and L with the dual space M^* of M. The vector space decomposition

$$V = M \oplus L = M \oplus M^*$$

tells us that any $N \in \mathcal{L}(V, M)$ projects bijectively onto L under this decomposition. In particular, this means that N is the graph of a linear map

$$T_N : L \to M = L^*.$$

So

$$N = \{(T_N\xi, \xi), \quad \xi \in L = M^*\}.$$

Giving a map from a vector space to its dual is the same as giving a bilinear form on the original vector space. In other words, N determines, and is determined by, the bilinear form β_N on $L = M^*$ where

$$\beta_N(\xi, \xi') = \frac{1}{2}\langle T_N\xi', \xi\rangle = \frac{1}{2}\omega(T_N\xi', \xi).$$

This is true for any n-dimensional subspace transversal to M. What is the condition on β_N for N to be Lagrangian? Well, if $w = (T_N\xi, \xi)$ and $w' = (T_N\xi', \xi'$ are two elements of N then

$$\omega(w, w') = \omega(T_N\xi, \xi') - \omega(T_N\xi', \xi)$$

since L and M are Lagrangian. So the condition is that β_N be symmetric. We have proved:

Proposition 2.6.1. *If $M \in \mathcal{L}(V)$ and we choose $L \in \mathcal{L}(V, M)$ then we get an identification of $\mathcal{L}(V, M)$ with $S^2(L)$, the space of symmetric bilinear forms on L.*

So every choice of a a pair of transverse Lagrangian subspaces L and M gives a coordinate chart on $\mathcal{L}(V)$ which is identified with $S^2(L)$. In particular, $\mathcal{L}(V)$ is a smooth manifold and

$$\dim \mathcal{L}(V) = \frac{n(n+1)}{2}$$

where $n = \frac{1}{2}\dim V$.

Description in terms of a basis.

Suppose that we choose a basis $e_1, \ldots, e_n$ of L and so get a dual basis $f_1, \ldots, f_n$ of M. If $N \in \mathcal{L}(V, M)$ then we get a basis $g_1, \ldots, g_n$ of N where

$$g_i = e_i + \sum_j S_{ij} f_j$$

where

$$S_{ij} = \beta_N(e_i, e_j).$$

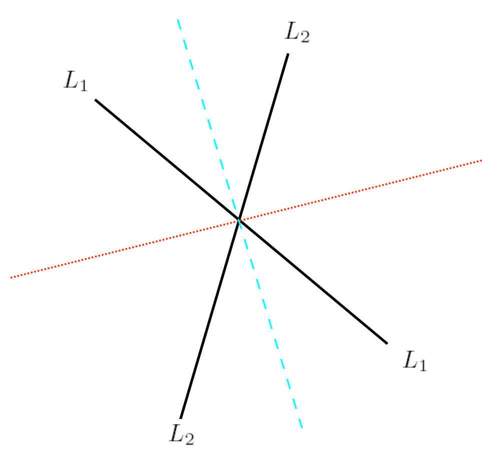
For later use we record the following fact: Let N and N' be two elements of $\mathcal{L}(V, M)$. The symplectic form ω induces a (possibly singular) bilinear form on $N \times N'$. In terms of the bases given above for N and N' we have

$$\omega(g_i, g_j') = S_{ij}' - S_{ij}. \tag{2.7}$$

$Sp(V)$ acts transitively on the space of pairs of transverse Lagrangian subspaces but not on the space of triples of Lagrangian subspaces.

Suppose that L_1 and L_2 are elements of $\mathcal{L}(V)$. An obvious invariant is the dimension of their intersection. Suppose that they are transverse, i.e. that $L_1 \cap L_2 = \{0\}$. We have seen that a basis $e_1, \ldots, e_n$ of L_1 determines a (dual) basis $f_1, \ldots, f_n$ of L_2 and together $e_1, \ldots, e_n, f_1, \ldots, f_n$ form a symplectic basis of V. Since $Sp(V)$ acts transitively on the set of symplectic bases, we see that it acts transitively on the space of pairs of transverse Lagrangian subspaces.

But $Sp(V)$ does *not* act transitively on the space of all (ordered, pairwise mutually transverse) triplets of Lagrangian subspaces. We can see this already in the plane: Every line through the origin is a Lagrangian subspace. If we fix two lines, the set of lines transverse to both is divided into two components corresponding to the two pairs of opposite cones complementary to the first two lines:



We can see this more analytically as follows: By an application of $Sl(2,\mathbb{R}) = Sp(\mathbb{R}^2)$ we can arrange that L_1 is the x-axis and L_2 is the y-axis. The subgroup of $Sl(2,\mathbb{R})$ which preserves both axes consists of the diagonal matrices (with determinant one), i.e. of all matrices of the form

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}.$$

If $\lambda > 0$ such a matrix preserves all quadrants, while if $\lambda < 0$ such a matrix interchanges the first and third and the second and fourth quadrants.

In any event, such a matrix carries a line passing through the first and third quadrant into another such line and the group of such matrices acts transitively on the set of all such lines. Similarly for lines passing through the second and fourth quadrant.

2.7 The set of Lagrangian subspaces transverse to a pair of Lagrangian subspaces

The situation depicted in the figure above has an n-dimensional analogue. Let M_1 and M_2 be Lagrangian subspaces of a symplectic vector space V. For the moment we will assume that they are transverse to each other, i.e., $M_1 \cap M_2 = \{0\}$. Let

$$\mathcal{L}(V, M_1, M_2) = \mathcal{L}(V, M_1) \cap \mathcal{L}(V, M_2)$$

be the set of Lagrangian subspaces, L of V which are transverse *both* to M_1 and to M_2. Since M_1 and M_2 are transverse, $V = M_1 \oplus M_2$, so L is the graph of a bijective mapping: $T_L : M_1 \to M_2$, and as we saw in the preceding section, this mapping defines a bilinear form, $\beta_L \in S^2(M_1)$ by the recipe

$$\beta_L(v, w) = \frac{1}{2}\omega(v, Lw)\,.$$

Moreover since T_L is bijective this bilinear form is non-degenerate. Thus, denoting by $S^2(M_1)_{\text{non-deg}}$ the set of non-degenerate symmetric bilinear forms on M_1, the bijective map

$$\mathcal{L}(V, M_1) \to S^2(M_1)$$

that we defined in §2.6 gives, by restriction, a bijective map

$$\mathcal{L}(V, M_1, M_2) \to S^2(M_1)_{\text{non-deg}}\,. \tag{2.8}$$

The connected components of $S^2(M_1)_{\text{non-deg}}$ are characterized by the signature invariant

$$\beta \in S^2(M_1)_{\text{non-sing}} \to \operatorname{sgn}\beta\,,$$

so, via the identification (2.8) the same is true of $\mathcal{L}(V, M_1, M_2)$: its connected components are characterized by the invariant $L \to \operatorname{sgn}\beta_L$. For instance in the two-dimensional case depicted in the figure above, $\operatorname{sgn}\beta_L$ is equal to 1 on one of the two components of $\mathcal{L}(V, M_1, M_2)$ and -1 on the other. Let

$$\sigma(M_1, M_2, L) =: \operatorname{sgn}\beta_L \tag{2.9}$$

This is by definition a *symplectic invariant* of the triple, M_1, M_2, L, so this shows that just as in two dimensions the group $Sp(V)$ does *not* act transitively on triples of mutually transversal Lagrangian subspaces.

Explicit computation of $\operatorname{sgn}\beta_L$**.**

We now describe how to compute this invariant explicitly in some special cases. Let $x_1, \ldots, x_n$, $\xi_1, \ldots, \xi_n$ be a system of Darboux coordinates on V such that M_1 and M_2 are the spaces, $\xi = 0$ and $x = 0$. Then L is the graph of a bijective linear map $\xi = Bx$ with $B^\dagger = B$ and hence

$$\sigma(M_1, M_2, L) = \operatorname{sgn}(B)\,. \tag{2.10}$$

Next we consider a slightly more complicated scenario. Let M_2 be, as above, the space, $x = 0$, but let M_1 be a Lagrangian subspace of V which is transverse to $\xi = 0$ and $x = 0$, i.e., a space of the form $x = A\xi$ where $A^\dagger = A$ and A is non-singular. In this case the symplectomorphism

$$(x, \xi) \to (x, \xi - A^{-1}x)$$

maps M_1 onto $\xi = 0$ and maps the space

$$L : \xi = Bx$$

onto the space

$$L_1 : \xi = (B - A^{-1})x.$$

and hence by the previous computation

$$\sigma(M_1, M_0, L) = \operatorname{sgn}(B - A^{-1})\,. \tag{2.11}$$

Notice however that the matrix

$$\begin{bmatrix} A & I \\ I & B \end{bmatrix}$$

can be written as the product

$$\begin{bmatrix} I & 0 \\ A^{-1} & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & B - A^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ A^{-1} & I \end{bmatrix}^\dagger \tag{2.12}$$

so

$$\operatorname{sgn} A + \operatorname{sgn}(B - A^{-1}) = \operatorname{sgn} \begin{bmatrix} A & I \\ I & B \end{bmatrix}\,. \tag{2.13}$$

Hence

$$\sigma(M_1, M_2, L) = \operatorname{sgn} \begin{bmatrix} A & I \\ I & B \end{bmatrix} - \operatorname{sgn} A\,. \tag{2.14}$$

In particular if L_1 and L_2 are Lagrangian subspaces of V which are transverse to M_1 and M_2 the difference,

$$\sigma(M_1, M_2, L_1) - \sigma(M_1, M_2, L_2)$$

is equal to

$$\operatorname{sgn} \begin{bmatrix} A & I \\ I & B_1 \end{bmatrix} - \operatorname{sgn} \begin{bmatrix} A & I \\ I & B_2 \end{bmatrix}.$$

In other words the quantity

$$\sigma(M_1, M_2, L_1, L_2) = \sigma(M_1, M_2, L_1) - \sigma(M_1, M_2, L_2)$$

is a symplectic invariant of M_1, M_2, L_1, L_2 which satisfies

$$\sigma(M_1, M_2, L_2, L_2) = \operatorname{sgn} \begin{bmatrix} A & I \\ I & B_1 \end{bmatrix} - \operatorname{sgn} \begin{bmatrix} A & I \\ I & B_2 \end{bmatrix}. \tag{2.15}$$

In the derivation of this identity we've assumed that M_1 and M_2 are transverse, however, the right hand side is well-defined provided the matrices

$$\begin{bmatrix} A & I \\ I & B_i \end{bmatrix} \qquad i = 1, 2$$

are non-singular, i.e., provided that L_1 and L_2 are transverse to the M_i. Hence to summarize, we've proved

Theorem 2.7.1. *Given Lagrangian subspaces M_1, M_2, L_1, L_2 of V such that the L_i's are transverse to the M_i's the formula (2.15) defines a symplectic invariant $\sigma(M_1, M_2, L_1, L_2)$ of M_1, M_2, L_1, L_2 and if M_1 and M_2 are transverse*

$$\sigma(M_1, M_2, L_1, L_2) = \sigma(M_1, M_2, L_1) - \sigma(M_1, M_2, L_2)\,. \tag{2.16}$$

2.8 The Maslov line bundle

We will use the results of the previous two sections to define an object which will play an important role in the analytical applications of the results of this chapter that we will discuss in Chapters 8 and 9.

Let X be an n-dimensional manifold and let $W = T^*X$ be its cotangent bundle. Given a Lagrangian submanifold, Λ, of W one has, at every point $p = (x, \xi)$, two Lagrangian subspaces of the symplectic vector space $V = T_pW$, namely the tangent space, M_1 to Λ at p and the tangent space M_2 at p to the cotangent fiber T^*_xX.

Let $\mathcal{O}_p = \mathcal{L}(V, M_1, M_2)$ and let $\mathbb{L}_p$ be the space of all functions

$$f : \mathcal{O}_p \to \mathbb{C}$$

which satisfy for $L_1, L_2 \in \mathcal{O}_p$

$$f(L_2) = e^{\frac{i\pi}{4}\sigma(M_1,M_2,L_2,L_1)} f(L^1)\,. \tag{2.17}$$

It is clear from (2.15) that this space is non-zero and from (2.16) that it is one-dimensional, i.e., is a complex line. Thus the assignment, $\Lambda \ni p \to \mathbb{L}_p$, defines a line bundle over Λ. We will denote this bundle by $\mathbb{L}_{\text{Maslov}}$ and refer to it henceforth as the *Maslov* line bundle of Λ. (The definition of it that we've just given, however, is due to Hörmander. An alternative definition, also due to Hörmander, will be described in §5.1.3. For the tie-in between these two definitions and the original definition of the Maslov bundle by Arnold, Keller, Maslov, see [[**?**]], Integrable Operators I, §3.3.)

2.9 A look ahead - a simple example of Hamilton's idea.

2.9.1 A different kind of generating function.

Let us go back to the situation described in Section 2.7. We have a symplectic vector space $V = M \oplus M^* = T^*M$ and we have a Lagrangian subspace $N \subset V$ which is transversal to M. This determines a linear map $T_N : M^* \to M$ and a symmetric bilinear form β_N on M^*. Suppose that we choose a basis of M and so identify M with $\mathbb{R}^n$ and so M^* with $\mathbb{R}^{n*}$. Then $T = T_N$ becomes a symmetric matrix and if we define

$$\gamma_N(\xi) := \frac{1}{2}\beta_N(\xi,\xi) = \frac{1}{2}T\xi\cdot\xi$$

then

$$T\xi = T_N\xi = \frac{\partial\gamma_N}{\partial\xi}.$$

Consider the function $\phi = \phi_N$ on $M \oplus M^*$ given by

$$\phi(x,\xi) = x\cdot\xi - \gamma_N(\xi), \quad x \in M,\ \xi \in M^*. \tag{2.18}$$

Then the equation

$$\frac{\partial\phi}{\partial\xi} = 0 \tag{2.19}$$

is equivalent to

$$x = T_N\xi.$$

Of course, we have

$$\xi = \frac{\partial\phi}{\partial x}$$

and at points where (2.19) holds, we have

$$\frac{\partial\phi}{\partial x} = d\phi,$$

the total derivative of ϕ in the obvious notation. So

Proposition 2.9.1. *Let M be a vector space and $V = T^*M = M \oplus M^*$ its cotangent bundle with its standard symplectic structure. Let N be a Lagrangian subspace of T^*M which is transversal to M. Then*

$$N = \{(x, d\phi(x,\xi))\}$$

where ϕ is the function on $M \times M^*$ given by (2.18) and where (x,ξ) satisfies (2.19).

The function ϕ is an example of the type of (generalized) generating function that we will study in detail in Chapter 5. Notice that in contrast to the generating functions of Chapter I, ϕ is not a function of x alone, but depends on an auxiliary variable (in this case ξ). But this type of generating function can describe a Lagrangian subspace which is not horizontal. At the extreme, the subspace M^* is described by the case $\beta_T \equiv 0$.

We will show in Chapter 5 that every Lagrangian submanifold of any cotangent bundle can locally be described by a generating function, when we allow dependence on auxiliary variables.

2.9.2 Composition of symplectic transformations and addition of generating functions.

Let $V = (V, \omega)$ be a symplectic vector space. We let $V^- = (V, -\omega)$. In other words, V is the same vector space as V but with the symplectic form $-\omega$.

We may consider the direct sum $V^- \oplus V$ (with the symplectic form $\Omega = (-\omega, \omega)$. If $T \in Sp(V)$, then its graph $\Gamma :=$ graph $T = \{(v, Tv), \quad v \in V\}$ is a Lagrangian subspace of $V^- \oplus V$. Indeed, if $v, w \in V$ then

$$\Omega((v,Tv),(w,Tw)) = \omega(Tv,Tw) - \omega(v,w) = 0.$$

Suppose that $V = X \oplus X^*$ where X is a vector space and where V is given the usual symplectic form:

$$\omega\left(\begin{pmatrix} x \\ \xi \end{pmatrix}, \begin{pmatrix} x' \\ \xi' \end{pmatrix}\right) = \langle \xi', x\rangle - \langle \xi, x'\rangle.$$

The map $\varsigma : \ V \to V$

$$\varsigma\left(\begin{pmatrix} x \\ \xi \end{pmatrix}\right) = \begin{pmatrix} x \\ -\xi \end{pmatrix}$$

is a symplectic isomorphism of V with V^-. So $\varsigma \oplus \mathrm{id}$ gives a symplectic isomorphism of $V^- \oplus V$ with $V \oplus V$.

A generating function (either in the sense of Chapter I or in the sense of Section 2.9.1 for $(\iota \oplus \mathrm{id})(\Gamma)$ will also (by abuse of language) be called a generating function for Γ or for T.

Let us consider the simplest case, where $X = \mathbb{R}$. Then

$$V \oplus V = \mathbb{R} \oplus \mathbb{R}^* \oplus \mathbb{R} \oplus \mathbb{R}^* = T^*(\mathbb{R} \oplus \mathbb{R}).$$

Let (x, y) be coordinates on $\mathbb{R} \oplus \mathbb{R}$ and consider a generating function (of the type of Chapter I) of the form

$$\phi(x,y) = \frac{1}{2}(ax^2 + 2bxy + cy^2),$$

where

$$b \neq 0.$$

Taking into account the transformation ς, the corresponding Lagrangian subspace of $V^- \oplus V$ is given by the equations

$$\xi = -(ax + by), \qquad \eta = bx + cy.$$

Solving these equations for y, η in terms of x, ξ gives

$$y = -\frac{1}{b}(ax + \xi), \quad \eta = \left(b - \frac{c}{b}\right) x - \frac{c}{b}\xi.$$

In other words, the matrix (of) T is given by

$$\begin{pmatrix} -\frac{a}{b} & -\frac{1}{b} \\ \\ b - \frac{ca}{b} & -\frac{c}{b} \end{pmatrix}.$$

(Notice that by inspection the determinant of this matrix is 1, which is that condition that T be symplectic.)

Notice also that the upper right hand corner of this matrix is not zero. Conversely, starting with a matrix

$$T = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

of determinant one, with $\beta \neq 0$ we can solve the equation

$$\begin{pmatrix} -\frac{a}{b} & -\frac{1}{b} \\ \\ b - \frac{ca}{b} & -\frac{c}{b} \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

for a, b, c in terms of $\alpha, \beta, \gamma, \delta$. So the most general two by two matrix of determinant one with the upper right hand corner $\neq 0$ is represented by a generating function of the above form.

Suppose we have two functions

$$\phi_1(x,y) = \frac{1}{2}[ax^2 + 2bxy + cy^2], \quad \phi_2(y,z) = \frac{1}{2}[Ay^2 + 2Byz + Cz^2],$$

with $b \neq 0$ and $B \neq 0$, and consider their sum:

$$\phi(x,z,y) = \phi_1(x,y) + \phi_2(y,z).$$

Here we are considering y as an "auxiliary variable" in the sense of Section 2.9.1, so we want to impose the constraint

$$\frac{\partial \phi}{\partial y} = 0, \tag{2.20}$$

and on this constrained set let

$$\xi = -\frac{\partial \phi}{\partial x}, \quad \zeta = \frac{\partial \phi}{\partial z}, \tag{2.21}$$

and use these equations to express $\begin{pmatrix} z \\ \zeta \end{pmatrix}$ in terms of $\begin{pmatrix} x \\ \xi \end{pmatrix}$.

Equation (2.20) gives

$$(A + c)y + bx + Bz = 0. \tag{2.22}$$

There are now two alternatives:

- If $A + c \neq 0$ we can solve (2.22) for y in terms of x and z. This then gives a generating function of the above type (i.e. quadratic in x and z). It is easy to check that the matrix obtained from this generating function is indeed the product of the corresponding matrices. This is an illustration of Hamilton's principle that the composition of two symplectic transformations is given by the sum of their generating functions. This will be explained in detail in Chapter 5, in Sections 5.6 and 5.7. Notice also that because $\partial^2\phi/\partial y^2 = A+c \neq 0$, the effect of (2.20) was to allow us to eliminate y. The general setting of this phenomenon will be explained in Section 5.8.

- If $A+c = 0$, then (2.22) imposes no condition on y but does give $bx+Bz = 0$, i.e

  $$z = -\frac{b}{B}x$$

  which means precisely that the upper right hand corner of the corresponding matrix vanishes. Since y is now a "free variable", and $b \neq 0$ we can solve the first of equations (2.21) for y in terms of x and ξ giving

  $$y = -\frac{1}{b}(\xi + ax)$$

  and substitute this into the second of the equations (2.21) to solve for ζ in terms of x and ξ. We see that the corresponding matrix is

  $$\begin{pmatrix} -\frac{b}{B} & 0 \\ -\frac{aB}{b} - \frac{Cb}{B} & -\frac{B}{b} \end{pmatrix}.$$

  Again, this is indeed the product of the correpsonding matrices.

Chapter 3

The language of category theory.

3.1 Categories.

We briefly recall the basic definitions:

A **category** $\mathcal{C}$ consists of the following data:

(i) A family, $Ob(\mathcal{C})$, whose elements are called the **objects** of $\mathcal{C}$,

(ii) For every pair (X, Y) of $Ob(\mathcal{C})$ a family, $\mathrm{Morph}(X, Y)$, whose elements are called the **morphisms** or **arrows** from X to Y,

(iii) For every triple (X, Y, Z) of $Ob(\mathcal{C})$ a map from $\mathrm{Morph}(X, Y) \times \mathrm{Morph}(Y, Z)$ to $\mathrm{Morph}(X, Z)$ called the **composition map** and denoted $(f, g) \rightsquigarrow g \circ f$.

These data are subject to the following conditions:

(iv) The composition of morphisms is *associative*

(v) For each $X \in Ob(\mathcal{C})$ there is an $id_X \in \mathrm{Morph}(X, X)$ such that

$$f \circ id_X = f, \ \forall f \in \mathrm{Morph}(X, Y)$$

(for any Y) and

$$id_X \circ f = f, \ \forall f \in \mathrm{Morph}(Y, X)$$

(for any Y).

It follows from the definitions that id_X is unique.

3.2 Functors and morphisms.

3.2.1 Covariant functors.

If $\mathcal{C}$ and $\mathcal{D}$ are categories, a **functor** F from $\mathcal{C}$ to $\mathcal{D}$ consists of the following data:

(vi) a map $F : Ob(\mathcal{C}) \to Ob(\mathcal{D})$

and

(vii) for each pair (X, Y) of $Ob(\mathcal{C})$ a map

$$F : \mathrm{Hom}(X, Y) \to \mathrm{Hom}(F(X), F(Y))$$

subject to the rules

(viii)

$$F(id_X) = id_{F(X)}$$

and

(ix)

$$F(g \circ f) = F(g) \circ F(f).$$

This is what is usually called a **covariant functor**.

3.2.2 Contravariant functors.

A **contravariant functor** would have $F : \mathrm{Hom}(X, Y) \to \mathrm{Hom}(F(Y), F(X))$ in (vii) and $F(f) \circ F(g)$ on the right hand side of (ix).)

3.2.3 The functor to families.

Here is an important example, valid for any category $\mathcal{C}$. Let us fix an $X \in Ob(\mathcal{C})$. We get a functor

$$F_X : \mathcal{C} \to \mathbf{Set}$$

(where **Set** denotes the category whose objects are all families, and morphisms are all maps) by the rule which assigns to each $Y \in Ob(\mathcal{C})$ the family $F_X(Y) = \mathrm{Hom}(X, Y)$ and to each $f \in \mathrm{Hom}(Y, Z)$ the map $F_X(f)$ consisting of composition (on the left) by f. In other words, $F_X(f) : \mathrm{Hom}(X, Y) \to \mathrm{Hom}(X, Z)$ is given by

$$g \in \mathrm{Hom}(X, Y) \mapsto f \circ g \in \mathrm{Hom}(X, Z).$$

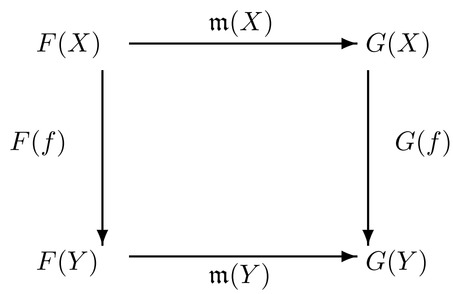


Figure 3.1:

3.2.4 Morphisms

Let F and G be two functors from $\mathcal{C}$ to $\mathcal{D}$. A **morphism**, $\mathfrak{m}$, from F to G (older name: "natural transformation") consists of the following data:

(x) for each $X \in Ob(\mathcal{C})$ an element $\mathfrak{m}(X) \in \mathrm{Hom}_{\mathcal{D}}(F(X), G(X))$ subject to the "naturality condition"

(xi) for any $f \in \mathrm{Hom}_{\mathcal{C}}(X, Y)$ the diagram in Figure 3.1 commutes. In other words

$$\mathfrak{m}(Y) \circ F(f) = G(f) \circ \mathfrak{m}(X) \quad \forall\ f \in f \in \mathrm{Hom}_{\mathcal{C}}(X, Y).$$

3.2.5 Involutory functors and involutive functors.

Consider the category $\mathcal{V}$ whose objects are finite dimensional vector spaces (over some given field $\mathbb{K}$) and whose morphisms are linear transformations. We can consider the "transpose functor" $F : \mathcal{V} \to \mathcal{V}$ which assigns to every vector space V its dual space

$$V^* = \mathrm{Hom}(V, \mathbb{K})$$

and which assigns to every linear transformation $\ell : V \to W$ its transpose

$$\ell^* : W^* \to V^*.$$

In other words,

$$F(V) = V^*, \quad F(\ell) = \ell^*.$$

This is a contravariant functor which has the property that F^2 is naturally equivalent to the identity functor. There does not seem to be a standard name for this type of functor. We will call it an **involutory** functor.

A special type of involutory functor is one in which $F(X) = X$ for all objects X and $F^2 = \mathrm{id}$ (not merely naturally equivalent to the identity). We shall call such a functor a **involutive** functor. We will refer to a category with an involutive functor as an **involutive category**, or say that we have a category with an involutive structure.

For example, let $\mathcal{H}$ denote the category whose objects are Hilbert spaces and whose morphisms are bounded linear transformations. We take $F(X) = X$ on objects and $F(L) = L^\dagger$ on bounded linear transformations where $L^\dagger$ denotes the adjoint of L in the Hilbert space sense.

3.3 Example: Sets, maps and relations.

The category **Set** is the category whose objects are ("all") families and and whose morphisms are ("all") maps between families. For reasons of logic, the word "all" must be suitably restricted to avoid contradiction.

We will take the extreme step in this section of restricting our attention to the class of finite sets. Our main point is to examine a category whose objects are finite sets, but whose morphisms are much more general than maps. Some of the arguments and constructions that we use in the study of this example will be models for arguments we will use later on, in the context of the symplectic "category".

3.3.1 The category of finite relations.

We will consider the category whose objects are finite sets. But we enlarge the set of morphisms by defining

$$\mathrm{Morph}(X, Y) = \text{the collection of all subsets of } X \times Y.$$

A subset of $X \times Y$ is called a **relation**. We must describe the map

$$\mathrm{Morph}(X, Y) \times \mathrm{Morph}(Y, Z) \to \mathrm{Morph}(X, Z)$$

and show that this composition law satisfies the axioms of a category. So let

$$\Gamma_1 \in \mathrm{Morph}(X, Y) \quad \text{and } \Gamma_2 \in \mathrm{Morph}(Y, Z).$$

Define

$$\Gamma_2 \circ \Gamma_1 \subset X \times Z$$

by

$$(x, z) \in \Gamma_2 \circ \Gamma_1 \;\Leftrightarrow\; \exists\, y \in Y \text{ such that } (x, y) \in \Gamma_1 \text{ and } (y, z) \in \Gamma_2. \tag{3.1}$$

Notice that if $f : X \to Y$ and $g : Y \to Z$ are maps, then

$$\mathrm{graph}(f) = \{(x, f(x)\} \in \mathrm{Morph}(X, Y) \quad \text{and} \quad \mathrm{graph}(g) \in \mathrm{Morph}(Y, Z)$$

with

$$\mathrm{graph}(g) \circ \mathrm{graph}(f) = \mathrm{graph}(g \circ f).$$

So we have indeed enlarged the category of finite sets and maps.

We still must check the axioms. Let $\Delta_X \subset X \times X$ denote the diagonal:

$$\Delta_X = \{(x, x),\ x \in X\},$$

so

$$\Delta_X \in \mathrm{Morph}(X, X).$$

If $\Gamma \in \mathrm{Morph}(X, Y)$ then

$$\Gamma \circ \Delta_X = \Gamma \quad \text{and } \Delta_Y \circ \Gamma = \Gamma.$$

So Δ_X satisfies the conditions for id_X.

Let us now check the associative law. Suppose that $\Gamma_1 \in \mathrm{Morph}(X, Y), \Gamma_2 \in \mathrm{Morph}(Y, Z)$ and $\Gamma_3 \in \mathrm{Morph}(Z, W)$. Then both $\Gamma_3 \circ (\Gamma_2 \circ \Gamma_1)$ and $(\Gamma_3 \circ \Gamma_2) \circ \Gamma_1$ consist of all $(x, w) \in X \times W$ such that there exist $y \in Y$ and $z \in Z$ with

$$(x, y) \in \Gamma_1,\ (y, z) \in \Gamma_2, \quad \text{and } (z, w) \in \Gamma_3.$$

This proves the associative law.

Let us call this category **FinRel**.

3.3.2 Categorical "points".

Let us pick a distinguished one element set and call it "pt.". Giving a *map* from pt. to any set X is the same as picking a point of X. So in the category **Set** of sets and *maps*, the points of X are the same as the morphisms from our distinguished object pt. to X.

In a more general category, where the objects are not necessarily sets, we can not talk about the points of an object X. However if we have a distinguished object pt., then we can *define* a **"point"** of any object X to be an element of $\mathrm{Morph}(\mathrm{pt.}, X)$. For example, later on, when we study the symplectic "category" whose objects are symplectic manifolds, we will find that the "points" in a symplectic manifold are its Lagrangian submanifolds. This idea has been emphasized by Weinstein. As he points out, this can be considered as a manifestation of the Heisenberg uncertainty principle in symplectic geometry.

In the category **FinRel**, the category of finite sets and relations, an element of $\mathrm{Morph}(\mathrm{pt.}, X)$, i.e a subset of $\mathrm{pt.} \times X$, is the same as a subset of X (by projection onto the second factor). So in this category, the "points" of X are the subsets of X. Many of the constructions we do here can be considered as warm ups to similar constructions in the symplectic "category".

Suppose we have a category with a distinguished object pt.. A morphism $\Gamma \in \mathrm{Morph}(X, Y)$ yields a map from "points" of X to "points" of Y. Namely, a "point" of X is an element $p \in \mathrm{Morph}(\mathrm{pt.}, X)$ so if $f \in \mathrm{Morph}(X, Y)$ we can form

$$f \circ p \in \mathrm{Morph}(\mathrm{pt.}, Y)$$

which is a "point" of Y. So f maps "points" of X to "points" of Y.

We will sometimes use the more suggestive language $f(p)$ instead of $f \circ p$.

3.3.3 The universal associative law.

Consider three objects X, Y, Z. Inside

$$X \times X \times Y \times Y \times Z \times Z$$

we have the subset

$$\Delta^3 = \Delta^3_{XYZ} = \Delta_X \times \Delta_Y \times \Delta_Z$$

consisting of all points of the form

$$(xxyyzz).$$

Let us move the first X factor past the others until it lies to immediate left of the right Z factor, so consider the subset

$$\tilde{\Delta}^3 = \tilde{\Delta}^3_{XYZ} \subset X \times Y \times Y \times Z \times X \times Z, \quad \tilde{\Delta}^3_{XYZ} = \{(x, y, y, z, x, z)\}.$$

By introducing parentheses around the first four and last two factors we can write

$$\tilde{\Delta}^3_{XYZ} \subset (X \times Y \times Y \times Z) \times (X \times Z).$$

In other words,

$$\tilde{\Delta}^3_{XYZ} \in \mathrm{Morph}(X \times Y \times Y \times Z, X \times Z).$$

Let $\Gamma_1 \in \mathrm{Morph}(X, Y)$ and $\Gamma_2 \in \mathrm{Morph}(Y, Z)$. Then

$$\Gamma_1 \times \Gamma_2 \subset X \times Y \times Y \times Z$$

is a "point" of $X \times Y \times Y \times Z$. We identify this "point" with an element of

$$\mathrm{Morph}(\mathrm{pt.}, X \times Y \times Y \times Z)$$

so that we can form

$$\tilde{\Delta}^3_{XYZ} \circ (\Gamma_1 \times \Gamma_2)$$

which consists of all (x, z) such that

$$\exists (x_1, y_1, y_2, z_1, x, z) \text{ with}$$

$$\begin{array}{rcl} (x_1, y_1) & \in & \Gamma_1, \\ (y_2, z_1) & \in & \Gamma_2, \\ x_1 & = & x, \\ y_1 & = & y_2, \\ z_1 & = & z. \end{array}$$

Thus

$$\tilde{\Delta}^3_{XYZ} \circ (\Gamma_1 \times \Gamma_2) = \Gamma_2 \circ \Gamma_1. \tag{3.2}$$

Suppose we have four sets X, Y, Z, W. We can form

$$\tilde{\Delta}^3_{XYZ} \bowtie \Delta^2_{ZW} \subset X \times Y \times Y \times Z \times Z \times W \times X \times Z \times Z \times W$$

consisting of all points of the form

$$(xyyzz'wxzz'w).$$

By inserting parentheses about the first six and last four positions we can regard $\tilde{\Delta}^3_{XYZ} \bowtie \Delta^2_{ZW}$ as an element of

$$\mathrm{Morph}((X \times Y \times Y \times Z \times Z \times W), (X \times Z \times Z \times W)).$$

If we compose $\tilde{\Delta}^3_{XYZ} \bowtie \Delta^2_{ZW}$ with

$$\Gamma_1 \times \Gamma_2 \times \Gamma_3 \in \mathrm{Morph}(\mathrm{pt.}, X \times Y \times Y \times Z \times Z \times W)$$

we obtain

$$(\Gamma_2 \circ \Gamma_1) \times \Gamma_3 \subset (X \times Z) \times (Z \times W).$$

Now let us consider

$$\tilde{\Delta}^3_{XZW} \circ \left(\tilde{\Delta}^3_{XYZ} \bowtie \Delta^2_{ZW}\right).$$

It consists of all pairs $(xyyzz'w), (xw)$ such that $(xzz'w) = (xzzw)$ i.e. such that $z = z'$. Removing the parentheses we obtain

$$\tilde{\Delta}^4_{XYZW} \subset X \times Y \times Y \times Z \times Z \times W \times X \times W,$$

given by

$$\tilde{\Delta}^4_{XYZW} = \{(xyyzzwxw)\}.$$

So putting in some parentheses shows that we can regard $\tilde{\Delta}^4_{XYZW}$ as an element of

$$\mathrm{Morph}(X \times Y \times Y \times Z \times Z \times W, X \times W).$$

If $\Gamma_1 \in \mathrm{Morph}(X, Y)$, $\Gamma_2 \in \mathrm{Morph}(Y, Z)$, and $\Gamma_3 \in \mathrm{Morph}(Z, W)$ then we can compose $\tilde{\Delta}^4_{XYZW}$ with $\Gamma_1 \times \Gamma_2 \times \Gamma_3$ to obtain an element of $\mathrm{Morph}(X, W)$. Thus the equation

$$\tilde{\Delta}^4_{XYZW} = \tilde{\Delta}^3_{XZW} \circ \left(\tilde{\Delta}^3_{XYZ} \bowtie \Delta^2_{ZW}\right) \tag{3.3}$$

is a sort of universal associative law in the sense that if we compose (3.3) with $\Gamma_1 \times \Gamma_2 \times \Gamma_3$ regarded as an element of $\mathrm{Morph}(\mathrm{pt.}., X \times Y \times Y \times Z \times Z \times W)$ we obtain the equation

$$\Gamma_3 \circ (\Gamma_2 \circ \Gamma_1) = \tilde{\Delta}^4_{XYZW}(\Gamma_1 \times \Gamma_2 \times \Gamma_3).$$

Similar to (3.3) we have an equation of the form

$$\tilde{\Delta}^3_{XZW} \circ \left(\Delta_{XY} \bowtie \tilde{\Delta}^3_{YZW}\right) = \tilde{\Delta}^4_{XYZW} \tag{3.4}$$

which implies that

$$(\Gamma_3 \circ \Gamma_2) \circ \Gamma_2 = \tilde{\Delta}^4_{XYZW}(\Gamma_1 \times \Gamma_2 \times \Gamma_3).$$

From this point of view the associative law is a consequence of equations (3.3) and (3.4) and of the fact that

$$(\Gamma_1 \times \Gamma_2) \times \Gamma_3 = \Gamma_1 \times (\Gamma_2 \times \Gamma_3) = \Gamma_1 \times \Gamma_2 \times \Gamma_3.$$

3.3.4 The transpose.

In our category **FinRel**, if $\Gamma \in \mathrm{Morph}(X, Y)$ define $\Gamma^\dagger \in \mathrm{Morph}(Y, X)$ by

$$\Gamma^\dagger := \{(y, x) | (x, y) \in \Gamma\}.$$

We have defined a map

$$\dagger : \mathrm{Morph}(X, Y) \to \mathrm{Morph}(Y, X) \tag{3.5}$$

for all objects X and Y which clearly satisfies

$$\dagger^2 = \mathrm{id} \tag{3.6}$$

and

$$(\Gamma_2 \circ \Gamma_1)^\dagger = \Gamma_1^\dagger \circ \Gamma_2^\dagger. \tag{3.7}$$

So $\dagger$ is a contravariant functor and satisfies our conditions for an involution. This makes our category **FinRel** of finite sets and relations into an involutive category.

3.3.5 Some notation.

In the category **FinRel** a morphism is a relation. So $\mathrm{Morph}(X, Y)$ is a subset of $X \times Y$. As we have seen, we can think of a relation as a generalization of the graph of a map which is a special kind of relation. The following definitions (some of which are borrowed from Alan Weinstein) will prove useful in other categorical settings: Let $\Gamma \in \mathrm{Morph}(X, Y)$

- X is called the **source** of Γ,
- Y is called the **target** of Γ,
- If T is a subset of X, then $\Gamma(T) := \{y | \exists x \in T \text{ such that } (x, y) \in \Gamma\}$ is called the **image** of T and is denoted by $\Gamma(T)$.
- $\Gamma(X)$ is called the **range** of Γ,
- The range of $\Gamma^\dagger$ is called the **domain** of Γ.
- Γ is **surjective** if if its range equals its target.

- Γ is **cosurjective** if its domain equals its source, i.e it is "defined everywhere".
- Γ is **injective** if for any $y \in Y$ there is at most one $x \in X$ with $(x, y) \in \Gamma$.
- Γ is **co-injective** if for any $x \in X$ there is at most one $y \in Y$ with $(x, y) \in \Gamma$, i.e. Γ is "single valued".
- Γ is called a **reduction** if it is surjective and co-injective,
- Γ is called a **coreduction** if it is injective and co-surjective, so it takes all the points of the source X into disjoint subsets of Y.

3.4 The linear symplectic category.

Let V_1 and V_2 be symplectic vector spaces with symplectic forms ω_1 and ω_2. We will let V_1^- denote the vector space V_1 equipped with the symplectic form $-\omega_1$. So $V_1^- \oplus V_2$ denotes the vector space $V_1 \oplus V_2$ equipped with the symplectic form $-\omega_1 \oplus \omega_2$.

A Lagrangian subspace Γ of $V_1^- \oplus V_2$ is called a **linear canonical relation**. The purpose of this section is to show that if we take the collection of symplectic vector spaces as objects, and the linear canonical relations as morphisms we get a category, cf. [GSIG].

Here composition is in the sense of composition of relations as in the category **FinRel**. In more detail: Let V_3 be a third symplectic vector space, let

$$\Gamma_1 \text{ be a Lagrangian subspace of } \quad V_1^- \oplus V_2$$

and let

$$\Gamma_2 \text{ be a Lagrangian subspace of } \quad V_2^- \oplus V_3.$$

Recall that as a *set* (see (3.1)) the composition

$$\Gamma_2 \circ \Gamma_1 \subset V_1 \times V_3$$

is defined by

$$(x, z) \in \Gamma_2 \circ \Gamma_1 \ \Leftrightarrow \ \exists\, y \in V_2 \text{ such that } (x, y) \in \Gamma_1 \text{ and } (y, z) \in \Gamma_2.$$

We must show that this is a Lagrangian subspace of $V_1^- \oplus V_3$. It will be important for us to break up the definition of $\Gamma_2 \circ \Gamma_1$ into two steps:

3.4.1 The space $\Gamma_2 \star \Gamma_1$.

Define

$$\Gamma_2 \star \Gamma_1 \subset \Gamma_1 \times \Gamma_2$$

to consist of all pairs $((x, y), (y', z))$ such that $y = y'$. We will restate this definition in two convenient ways. Let

$$\pi : \Gamma_1 \to V_2, \qquad \pi(v_1, v_2) = v_2$$

and

$$\rho : \Gamma_2 \to V_2, \qquad \rho(v_2, v_3) = v_2.$$

Let

$$\tau : \ \Gamma_1 \times \Gamma_2 \to V_2$$

be defined by

$$\tau(\gamma_1, \gamma_2) := \pi(\gamma_1) - \rho(\gamma_2). \tag{3.8}$$

Then $\Gamma_2 \star \Gamma_1$ is determined by the exact sequence

$$0 \to \Gamma_2 \star \Gamma_1 \to \Gamma_1 \times \Gamma_2 \xrightarrow{\tau} V_2 \to \text{Coker}\, \tau \to 0. \tag{3.9}$$

Another way of saying the same thing is to use the language of "fiber products" or "exact squares": Let $f : A \to C$ and $g : B \to C$ be maps, say between sets. Then we express the fact that $F \subset A \times B$ consists of those pairs (a, b) such that $f(a) = g(b)$ by saying that

$$\begin{array}{ccc} F & \longrightarrow & A \\ \big\downarrow & & \big\downarrow f \\ B & \underset{g}{\longrightarrow} & C \end{array}$$

is an **exact square** or a **fiber product diagram**.

Thus another way of expressing the definition of $\Gamma_2 \star \Gamma_1$ is to say that

$$\begin{array}{ccc} \Gamma_2 \star \Gamma_1 & \longrightarrow & \Gamma_1 \\ \big\downarrow & & \big\downarrow \pi \\ \Gamma_2 & \underset{\rho}{\longrightarrow} & V_2 \end{array} \tag{3.10}$$

is an exact square.

3.4.2 The transpose.

If $\Gamma \subset V_1^- \oplus V_2$ is a linear canonical relation, we define its transpose $\Gamma^\dagger$ just as in **FinRel**:

$$\Gamma^\dagger := \{(y, x) | (x, y) \in \Gamma\}. \tag{3.11}$$

Here $x \in V_1$ and $y \in V_2$ so $\Gamma^\dagger$ as defined is a linear Lagrangian subspace of $V_2 \oplus V_1^-$. But replacing the symplectic form by its negative does not change the set of Lagrangian subspaces, so $\Gamma^\dagger$ is also a Lagrangian subspace of $V_2^- \oplus V_1$, i.e. a linear canonical relation between V_2 and V_1. It is also obvious that just as in **FinRel** we have

$$\left(\Gamma^\dagger\right)^\dagger = \Gamma.$$

3.4.3 The projection $\alpha : \Gamma_2 \star \Gamma_1 \to \Gamma_2 \circ \Gamma_1$.

Consider the map

$$\alpha : (x, y, y, z) \mapsto (x, z). \tag{3.12}$$

By definition

$$\alpha : \Gamma_2 \star \Gamma_1 \to \Gamma_2 \circ \Gamma_1.$$

3.4.4 The kernel and image of a linear canonical relation.

Let V_1 and V_2 be symplectic vector spaces and let $\Gamma \subset V_1^- \times V_2$ be a linear canonical relation. Let

$$\pi : \Gamma \to V_2$$

be the projection onto the second factor. Define

- $\operatorname{Ker}\Gamma \subset V_1$ by $\operatorname{Ker}\Gamma = \{v \in V_1 | (v, 0) \in \Gamma\}$.
- $\operatorname{Im}\Gamma \subset V_2$ by $\operatorname{Im}\Gamma := \pi(\Gamma) = \{v_2 \in V_2 | \exists v_1 \in V_1 \text{ with } (v_1, v_2) \in \Gamma\}$.

Now $\Gamma^\dagger \subset V_2^- \oplus V_1$ and hence both $\ker \Gamma^\dagger$ and $\operatorname{Im}\Gamma$ are linear subspaces of the symplectic vector space V_2. We claim that

$$(\ker \Gamma^\dagger)^\perp = \operatorname{Im}\Gamma. \tag{3.13}$$

Here $\perp$ means perpendicular relative to the symplectic structure on V_2.

Proof. Let ω_1 and ω_2 be the symplectic bilinear forms on V_1 and V_2 so that $\tilde{\omega} = -\omega_1 \oplus \omega_2$ is the symplectic form on $V_1^- \oplus V_2$. So $v \in V_2$ is in $\operatorname{Ker}\Gamma^\dagger$ if and only if $(0, v) \in \Gamma$. Since Γ is Lagrangian, $(0, v) \in \Gamma \Leftrightarrow (0, v) \in \Gamma^\perp$ and

$$(0, v) \in \Gamma^\perp \Leftrightarrow 0 = -\omega_1(0, v_1) + \omega_2(v, v_2) = \omega_2(v, v_2) \ \forall \ (v_1, v_2) \in \Gamma.$$

But this is precisely the condition that $v \in (\operatorname{Im}\Gamma)^\perp$. □

The kernel of α consists of those $(0, v, v, 0) \in \Gamma_2 \star \Gamma_1$. We may thus identify

$$\ker \alpha = \ker \Gamma_1^\dagger \cap \ker \Gamma_2 \tag{3.14}$$

as a subspace of V_2.

If we go back to the definition of the map τ, we see that the image of τ is given by

$$\operatorname{Im}\tau = \operatorname{Im}\Gamma_1 + \operatorname{Im}\Gamma_2^\dagger, \tag{3.15}$$

a subspace of V_2. If we compare (3.14) with (3.15) we see that

$$\ker \alpha = (\operatorname{Im}\tau)^\perp \tag{3.16}$$

as subspaces of V_2 where $\perp$ denotes orthocomplement relative to the symplectic form ω_2 of V_2.

3.4.5 Proof that $\Gamma_2 \circ \Gamma_1$ is Lagrangian.

Since $\Gamma_2 \circ \Gamma_1 = \alpha(\Gamma_2 \star \Gamma_1)$ and $\Gamma_2 \star \Gamma_1 = \ker \tau$ it follows that $\Gamma_2 \circ \Gamma_1$ is a linear subspace of $V_1^- \oplus V_3$.

It is equally easy to see that $\Gamma_2 \circ \Gamma_1$ is an isotropic subspace of $V_1^- \oplus V_2$. Indeed, if (x, z) and (x', z') are elements of $\Gamma_2 \circ \Gamma_1$, then there are elements y and y' of V_2 such that

$$(x, y) \in \Gamma_1,\ (y, z) \in \Gamma_2,\ (x', y') \in \Gamma_1,\ (y', z') \in \Gamma_2.$$

Then

$$\omega_3(z, z') - \omega_1(x, x') = \omega_3(z, z') - \omega_2(y, y') + \omega_2(y, y') - \omega_1(x, x') = 0.$$

So we must show that $\dim \Gamma_2 \circ \Gamma_1 = \frac{1}{2} \dim V_1 + \frac{1}{2} \dim V_3$. It follows from (3.16) that

$$\dim\ \ker\ \alpha = \dim\ V_2 - \dim \operatorname{Im}\ \tau$$

and from the fact that $\Gamma_2 \circ \Gamma_1 = \alpha(\Gamma_2 \star \Gamma_1)$ that

$$\dim\ \Gamma_2 \circ \Gamma_1 = \dim\ \Gamma_2 \star \Gamma_1 - \dim\ \ker \alpha =$$

$$= \dim\ \Gamma_2 \star \Gamma_1 - \dim\ V_2 + \dim \operatorname{Im}\ \tau.$$

Since $\Gamma_2 \star \Gamma_1$ is the kernel of the map $\tau : \Gamma_1 \times \Gamma_2 \to V_2$ it follows that

$$\dim\ \Gamma_2 \star \Gamma_1 = \dim \Gamma_1 \times \Gamma_2 - \dim\ \operatorname{Im}\ \tau =$$

$$\frac{1}{2} \dim\ V_1 + \frac{1}{2} \dim\ V_2 + \frac{1}{2} \dim\ V_2 + \frac{1}{2} \dim\ V_3 - \dim\ \operatorname{Im}\ \tau.$$

Putting these two equations together we see that

$$\dim \Gamma_2 \circ \Gamma_1 = \frac{1}{2} \dim V_1 + \frac{1}{2} \dim V_3$$

as desired. We have thus proved

Theorem 3.4.1. *The composite* $\Gamma_2 \circ \Gamma_1$ *of two linear canonical relations is a linear canonical relation.*

The associative law can be proved exactly as for **FinRel**: given four symplectic vector spaces X, Y, Z, W we can form

$$\tilde{\Delta}^4_{XYZW} \subset [(X^- \times Y) \times (Y^- \times Z) \times (Z^- \times W)]^- \times (X^- \times W)$$

$$\tilde{\Delta}^4_{XYZW} = \{(xyyzzwxw)\}.$$

It is immediate to check that$\tilde{\Delta}^4_{XYZW}$ is a Lagrangian subspace, so

$$\tilde{\Delta}^4_{XYZW} \in \operatorname{Morph}((X^- \times Y) \times (Y^- \times Z) \times (Z^- \times W), X^- \times W).$$

If $\Gamma_1 \in \text{Morph}(X, Y)$, $\Gamma_2 \in \text{Morph}(Y, Z)$, and $\Gamma_3 \in \text{Morph}(Z, W)$ then

$$\Gamma_3 \circ (\Gamma_2 \circ \Gamma_1) = (\Gamma_3 \circ \Gamma_2) \circ \Gamma_1 = \tilde{\Delta}^4_{XYZW}(\Gamma_1 \times \Gamma_2 \times \Gamma_3),$$

as before. From this point of view the associative law is again a reflection of the fact that

$$(\Gamma_1 \times \Gamma_2) \times \Gamma_3 = \Gamma_1 \times (\Gamma_2 \times \Gamma_3) = \Gamma_1 \times \Gamma_2 \times \Gamma_3.$$

The diagonal Δ_V gives the identity morphism and so we have verified that

Theorem 3.4.2. LinSym *is a category whose objects are symplectic vector spaces and whose morphisms are linear canonical relations.*

3.4.6 Details concerning the identity $\tilde{\Delta}_{XYZ} \circ (\Gamma_1 \times \Gamma_2) = \Gamma_2 \circ \Gamma_1$.

Let X, Y, Z be symplectic vector spaces and $\Gamma_1 \in \text{Morph}(X, Y)$, $\Gamma_2 \in \text{Morph}(Y, Z)$. Since $\Gamma_1 \subset X^- \times Y$, $\Gamma_2 \subset Y^- \times Z$ so $\Gamma_1 \times \Gamma_2$ is a Lagrangian subspace of $X^- \times Y \times Y^- \times Z$ thought of as an element of $\text{Morph}(\text{pt.}, X^- \times Y \times Y^- \times Z)$.

Also

$$\tilde{\Delta}_{XYZ} \subset X^- \times Y \times Y^- \times Z \times X^- \times Z, \quad \tilde{\Delta}_{XYZ} = \{(x, y, y, z, x, z)\}.$$

So $\tilde{\Delta}_{XYZ} \star (\Gamma_1 \times \Gamma_2)$ consists of all $(x, y)(y', z), \overline{x}, \overline{y}, \overline{y}, \overline{z}$ such that $(x, y) \in \Gamma_1$, $(y', z) \in \Gamma_2$ and $\overline{x} = x, \overline{y} = y = y'$, $\overline{z} = z$. In other words,

$$\tilde{\Delta}_{XYZ} \star (\Gamma_1 \times \Gamma_2) = \{((x, y, y, z, x, z)|(x, y) \in \Gamma_1, \ (y, z) \in \Gamma_2\}.$$

Thus $\tilde{\Delta}_{XYZ} \star (\Gamma_1 \times \Gamma_2)$ is the kernel of the map

$$\tilde{\tau} : \tilde{\Delta}_{XYZ} \oplus (\Gamma_1 \times \Gamma_2) \to X \oplus Y \oplus Y \oplus Z$$

given by

$$\tilde{\tau}((x, y, y, z, x, z)(x_1, y_1)(y_2, z_2)) = (x - x_1, y - y_1, y - y_2, z - z_2).$$

The image of $\tilde{\tau}$ is

$$X \oplus (\Delta_Y + (\pi(\Gamma_1) \oplus \rho(\Gamma_2)) \oplus Z.$$

Here the middle expression is the subspace of $Y^- \oplus Y$ consisting of all $(y - y_1, y - y_2)$ with $y_1 \in \pi(\Gamma_1)$, $y_2 \in \rho(\Gamma_2)$. The symplectic orthogonal complement of the image of $\tilde{\tau}$ in $X^- \oplus Y \oplus Y^- \oplus Z$ is $0 \oplus Q \oplus 0$ where Q is the orthogonal complement of $\Delta_Y + (\pi(\Gamma_1) \oplus \rho(\Gamma_2)$ in $Y^- \oplus Y$.

From the general theory we know that this orthogonal complement is isomorphic to ker $\tilde{\alpha}$ where

$$\tilde{\alpha} : \tilde{\Delta}_{XYZ} \star (\Gamma_1 \times \Gamma_2) \to \tilde{\Delta}_{XYZ} \circ (\Gamma_1 \times \Gamma_2).$$

Since Δ_Y is a Lagrangian subspace of $Y^- \oplus Y$ we know that Q must be a subspace of Δ_Y and so consists of all (w, w) such that w is in the orthocomplement in Y of both $\pi(\Gamma_1)$ and $\rho(\Gamma_2)$. In other words w is such that $(0, w)$

is in the orthocomplement of Γ_1 in $X^- \times Y$ and so $(0,w) \in \Gamma_1$ and similarly $(w,0) \in \Gamma_2$. So $w \in \ker \alpha$ where

$$\alpha : \Gamma_2 \star \Gamma_1 \to \Gamma_1 \circ \Gamma_1.$$

In short,

Proposition 3.4.1. *We have an isomorphism from* $(\operatorname{Im} \tau)^\perp \cong \ker \alpha \to \ker \tilde{\alpha} \cong (\operatorname{Im} \tilde{\tau})^\perp$ *given by*

$$w \mapsto 0 \oplus (w,w) \oplus 0.$$

3.4.7 The category LinSym and the symplectic group.

The category **LinSym** is a vast generalization of the symplectic group because of the following observation: Let X and Y be symplectic vector spaces. Suppose that the Lagrangian subspace $\Gamma \subset X^- \oplus Y$ projects bijectively onto X under the projection of $X \oplus Y$ onto the first factor. This means that Γ is the graph of a linear transformation T from X to Y:

$$\Gamma = \{(x, Tx)\}.$$

T must be injective. Indeed, if $Tx = 0$ the fact that Γ is isotropic implies that $x \perp X$ so $x = 0$. Also T is surjective since if $y \perp \operatorname{im}(T)$, then $(0,y) \perp \Gamma$. This implies that $(0,y) \in \Gamma$ since Γ is maximal isotropic. By the bijectivity of the projection of Γ onto X, this implies that $y = 0$. In other words T is a bijection. The fact that Γ is isotropic then says that

$$\omega_Y(Tx_1, Tx_2) = \omega_X(x_1, x_2),$$

i.e. T is a symplectic isomorphism. If $\Gamma_1 = \operatorname{graph} T$ and $\Gamma_2 = \operatorname{graph} S$ then

$$\Gamma_2 \circ \Gamma_1 = \operatorname{graph} S \circ T$$

so composition of Lagrangian relations reduces to composition of symplectic isomorphisms in the case of graphs. In particular, if we take $Y = X$ we see that $\operatorname{Symp}(X)$ is a subgroup of Morph (X, X) in our category.

3.4.8 Reductions in the linear symplectic category.

Let X be an object in our category, i.e a symplectic vector space and let $Z \subset X$ be a coisotropic subspace of X. Since $Z^\perp \subset Z$, we can form the quotient space $B = Z/Z^\perp$ which is a symplectic vector space. Let $\pi : Z \to B$ denote the projection, $\iota : Z \to X$ the injection of Z as a subspace of X, and let ω_X and ω_B denote the symplectic forms on X and B. By definition,

$$\iota^* \omega_X = \pi^* \omega_B$$

so that the subset

$$\Gamma := \{(z, \pi(z)), z \in Z\} \subset X^- \times B$$

is isotropic. Let $k = \dim(Z^\perp)$. Since $\dim(Z) + \dim(Z^\perp) = \dim X$, we see that $\dim(Z) = \dim(X) - k$. On the other hand, $\dim(B) = \dim(Z) - k$. So $\dim(B) = \dim(X) - 2k$. So

$$\dim(\Gamma) = \dim(Z) = \dim(X) - k = \frac{1}{2}(\dim(X) + \dim(B)).$$

In other words, Γ is a Lagrangian subspace of $X^- \times B$, i.e. an element of $\mathrm{Morph}(X, B)$ which is clearly single valued and surjective, i.e. is a reduction.

Conversely, suppose that $\Gamma \in \mathrm{Morph}(X, B)$ is a reduction. Let $Z \subset X$ be the domain of Γ, so that Γ consists of all $(z, \pi(z))$ where $\pi : Z \to B$ is a surjective map. Let $V = \ker(\pi)$. Then since Γ is isotropic we see that $V^\perp \subset X$ contains Z. The dimension of Γ equals $\frac{1}{2}(\dim(X) + \dim(B)$. Let $k = \dim(Z) - \dim(B) = \dim(V)$. So

$$\dim(Z) = \dim(\Gamma) = \frac{1}{2}(\dim(X) + \dim(Z) - k)$$

implying that

$$\dim(Z) = \dim(X) - k = \dim(V^\perp).$$

So $V^\perp = Z$, i.e Z is co-isotropic. We have proved

Proposition 3.4.2. [Benenti and Tulszyjew [?], section 3.] *A reduction Γ in the linear symplectic category consists of a coisotropic subspace Z of a symplectic vector space X with quotient $B = Z/Z^\perp$ where $\Gamma \in \mathrm{Morph}(X, B)$ being the graph of the projection $\pi : Z \to B$.*

In fact, suppose that $\Gamma \in \mathrm{Morph}(X, B)$ is such that $\pi(\Gamma) = B$, where, recall, π is the projection of $\Gamma \subset X^- \times B$ onto the second factor. Then the projection ρ of Γ onto the first factor must be injective. Indeed, suppose that $(0, v) \in \Gamma$. Since Γ is isotropic, we must have $v \in B^\perp$ so $v = 0$. Thus

Proposition 3.4.3. *$\Gamma \in \mathrm{Morph}(X, Y)$ is a reduction if $\pi : \Gamma \to Y$ is surjective and hence (by applying †), $\Gamma \in \mathrm{Morph}(X, Y)$ is a co-reduction if $\rho : \Gamma \to X$ is surjective.*

We have the following result (a special case of a proposition due to Weinstein, [Wein11]):

Proposition 3.4.4. *Every morphism in the linear symplectic category can be written as the composition of a co-reduction with a reduction.*

Proof. Let Γ be a morphism from X to Y. Since Γ is a Lagrangian subspace of $X^- \times Y$, we can think of Γ as a morphism, call it γ, from pt.. to $X^- \times Y$. This is a coreduction. Hence so is $\mathbf{id}. \times \gamma$ which is a morphism from $X \times \mathrm{pt.}$ to $X \times X^- \times Y$. As a Lagrangian submanifold of $(X \times \mathrm{pt.})^- \times (X \times X^- \times Y) = X^- \times (X \times X^- \times Y)$ it consists of all points of the form

$$(x, x, x', y) \text{ with } (x', y) \in \Gamma. \tag{3.17}$$

Δ_X is a Lagrangian subspace of $X^- \times X$ which we can think of as a morphism ϵ_X from $X^- \times X$ to pt.. It is a reduction, hence so is $\epsilon_X \times \mathbf{id}_Y$. As a subset of $(X^- \times X \times Y)^- \times Y$ it consists of all points of the form

$$(x, x, y, y). \tag{3.18}$$

The composite of these two morphisms consists of the subset of $X \times Y = X \times \text{pt.} \times \text{pt.} \times Y$ given by those (x,y) such that there exists a $w = (x, x', y)$ with (w, y) of the form (3.18) so $x = x'$ with (x, w) of the form (3.17) so $(x', y) \in \Gamma$. So the composite is Γ.

□

3.4.9 Composition with reductions or co-reductions.

Suppose that $\Gamma \in \text{Morph}(X, B)$ is a reduction and so corresponds to a co-isotropic subspace $Z \subset X$, and let $V = Z^\perp$ be the kernel of the projection $\pi_\Gamma : Z \to B$. Let $\Lambda \in \text{Morph}(B, W)$. Since π is surjective, for any $(b, w) \in \Lambda$ there exists a $z \in Z$ with $(z, w) \in \Lambda \circ \Gamma$ with $\pi_\Gamma(z) = b$ and this z is determined up to an element of V. So

Proposition 3.4.5. *If $\Gamma \in \text{Morph}(X, B)$ is a reduction with $V = \ker(\pi_\Gamma)$ and $\Lambda \in \text{Morph}(B, W)$ then*

$$\Lambda \circ \Gamma = V \times \Lambda.$$

Hence, if $\Gamma \in \text{Morph}(B, X)$ is a co-reduction with $V = \ker \rho_\Gamma$ and $\Lambda \in \text{Morph}(W, B)$ then

$$\Gamma \circ \Lambda = \Lambda \times V.$$

3.5 The category of oriented linear canonical relations.

Recall that on an n-dimensional vector space V, its n-th exterior power $\wedge^n V$ is one dimensional. Hence $\wedge^n V \setminus \{0]$ has two components, and a choice of one of them is called an orientation of V. Put another way, any basis e of $\wedge^n V$ differs from any other basis by multiplication by a non-zero real number. This divides the set of bases into two equivalence classes, the elements in each equivalence class differ from one another by a positive multiple.

If

$$0 \to V' \to V \to V'' \to 0$$

is an exact sequence of vector spaces a basis of V' extends to a basis of V which then determines a basis of V''. So an orientation on any two of the three vector spaces determines an orientation on the third. An orientation on a vector space determines an orientation on its dual space.

A symplectic vector space carries a canonical orientation; indeed if ω is the symplectic form on a $2n$ dimensional vector space then ω^n is a non-zero element of $\wedge^{2n} V^*$, hence determines an orientation on V^* and hence on V.

Suppose that V_1, V_2, V_3 be three symplectic vector spaces with

$$\Gamma_1 \subset V_1^- \oplus V_2, \quad \Gamma_2 \subset V_2^- \oplus V_3.$$

Claim: *An orientation on* Γ_1 *and* Γ_2 *determines an orientation on* $\Gamma_2 \circ \Gamma_1$.

Proof. Let us first consider the case where the composition is transverse. We then have the exact sequence

$$0 \to \Gamma_2 \circ \Gamma_1 \to \Gamma_1 \oplus \Gamma_2 \xrightarrow{\tau} V_2 \to 0$$

so the orientations on Γ_1 and Γ_2 determine an orientation on $\Gamma_1 \oplus \Gamma_2$, which together with the canonical orientation on V_2 determine an orientation on $\Gamma_2 \circ \Gamma_1$.

The general case is only slightly more complicated: we have the exact sequences

$$\begin{array}{rcl} 0 \to \Gamma_2 \star \Gamma_1 \to & \Gamma_1 \oplus \Gamma_2 & \xrightarrow{\tau} \mathrm{Im}\tau \to 0 \\ 0 \to \ker\alpha \to & \Gamma_2 \star \Gamma_1 & \xrightarrow{\alpha} \Gamma_2 \circ \Gamma_1 \to 0 \\ 0 \to \ker\alpha \to & \mathrm{Im}\tau & \to \mathrm{Im}\tau/\ker\alpha \to 0. \end{array} \tag{3.19}$$

In the last sequence we know that by definition, ker α (considered as a subspace of V_2) is a subspace of Im τ and we proved that $\ker\alpha = \mathrm{Im}\tau^\perp$. So $\mathrm{Im}\tau/\ker\alpha$ is a symplectic vector space and hence has a canonical orientation. Thus a choice of orientation on, say, $\ker\alpha$ determines an orientation on $\mathrm{Im}\,\tau$. Such a choice then (together with the orientation on $\Gamma_1 \oplus \Gamma_2$) determines an orientation on $\Gamma_2 \star \Gamma_1$ by the first sequence and then an orientation on $\Gamma_2 \circ \Gamma_1$ by the second sequence. Had we made the opposite choice of orientation on $\ker\alpha$ this would have made the opposite choices of orientation on $\mathrm{Im}\,\tau$ and hence on $\Gamma_2 \star \Gamma_1$ from the first exact sequence, but then we would end up with the same orientation on $\Gamma_2 \circ \Gamma_1$ from the second exact sequence.

□

Proposition 3.5.1. *The set whose objects are symplectic vector spaces and whose morphisms are oriented linear canonical relations form a category.*

Proof. We must prove the associative law. For this we use the identity

$$\tilde{\Delta}_{XYZ} \circ (\Gamma_1 \times \Gamma_2) = \Gamma_2 \circ \Gamma_1 \tag{$*$}$$

(with $X = V_1, Y = V_2, Z = V_3$), together with the exact sequences (3.19) applied to $\tilde{\Delta}_{XYZ}$ and $\Gamma_1 \times \Gamma_2$. The space $\tilde{\Delta}_{XYZ}$ has a canonical orientation as it is isomorphic to the symplectic vector space $X \oplus Y \oplus Z$. From Proposition 3.4.1 we know that $\ker\alpha$ is isomorphic to $\ker\tilde{\alpha}$. So we conclude that the orientation induced on $\Gamma_2 \circ \Gamma_1$ is obtained from applying the construction above to $(*)$. Thus the associativity follows from our "universal" associative law in that the orientation on $\Gamma_3 \circ (\Gamma_2 \circ \Gamma_1)$ and on $(\Gamma_3 \circ \Gamma_2) \circ \Gamma_1$ both coincide with the orientation induced on

$$\tilde{\Delta}_{XYZW} \circ (\Gamma_1 \times \Gamma_2 \times \Gamma_3).$$

□

Chapter 4

The Symplectic "Category".

Let M be a symplectic manifold with symplectic form ω. Then $-\omega$ is also a symplectic form on M. We will frequently write M instead of (M,ω) and by abuse of notation we will let M^- denote the manifold M with the symplectic form $-\omega$.

Let (M_i,ω_i) $i=1,2$ be symplectic manifolds. A Lagrangian submanifold Γ of

$$\Gamma \subset M_1^- \times M_2$$

is called a **canonical relation**. So Γ is a subset of $M_1 \times M_2$ which is a Lagrangian submanifold relative to the symplectic form $\omega_2 - \omega_1$ in the obvious notation. So a canonical relation is a relation which is a Lagrangian submanifold.

For example, if $f : M_1 \to M_2$ is a symplectomorphism, then $\Gamma_f = \text{graph } f$ is a canonical relation.

If $\Gamma_1 \subset M_1 \times M_2$ and $\Gamma_2 \subset M_2 \times M_3$ we can form their composite

$$\Gamma_2 \circ \Gamma_1 \subset M_1 \times M_3$$

in the sense of the composition of relations. So $\Gamma_2 \circ \Gamma_1$ consists of all points (x,z) such that there exists a $y \in M_2$ with $(x,y) \in \Gamma_1$ and $(y,z) \in \Gamma_2$

Let us put this in the language of fiber products: Let

$$\pi : \Gamma_1 \to M_2$$

denote the restriction to Γ_1 of the projection of $M_1 \times M_2$ onto the second factor. Let

$$\rho : \Gamma_2 \to M_2$$

denote the restriction to Γ_2 of the projection of $M_2 \times M_3$ onto the first factor. Let

$$F \subset M_1 \times M_2 \times M_2 \times M_3$$

be defined by

$$F = (\pi \times \rho)^{-1}\Delta_{M_2}.$$

In other words, F is defined as the fiber product (or exact square)

$$\begin{array}{ccc} F & \xrightarrow{\iota_1} & \Gamma_1 \\ {\scriptstyle \iota_2}\downarrow & & \downarrow{\scriptstyle \pi}. \\ \Gamma_2 & \xrightarrow[\rho]{} & M_2 \end{array} \tag{4.1}$$

so

$$F \subset \Gamma_1 \times \Gamma_2 \subset M_1 \times M_2 \times M_2 \times M_3.$$

Let pr_{13} denote the projection of $M_1 \times M_2 \times M_2 \times M_3$ onto $M_1 \times M_3$ (projection onto the first and last components). Let π_{13} denote the restriction of pr_{13} to F. Then, as a *set*,

$$\Gamma_2 \circ \Gamma_1 = \pi_{13}(F). \tag{4.2}$$

The map pr_{13} is smooth, and hence its restriction to any submanifold is smooth. The problems are that

1. F defined as
$$F = (\pi \times \rho)^{-1}\Delta_{M_2},$$
i.e. by (4.1), need not be a submanifold, and

2. that the restriction π_{13} of pr_{13} to F need not be an embedding.

So we need some additional hypotheses to ensure that $\Gamma_2 \circ \Gamma_1$ is a submanifold of $M_1 \times M_3$. Once we impose these hypotheses we will find it easy to check that $\Gamma_2 \circ \Gamma_1$ is a Lagrangian submanifold of $M_1^- \times M_3$ and hence a canonical relation.

4.1 Clean intersection.

Assume that the maps

$$\pi : \Gamma_1 \to M_2 \quad \text{and} \quad \rho : \Gamma_2 \to M_2$$

defined above intersect cleanly.

Notice that $(m_1, m_2, m_2', m_3) \in F$ if and only if

- $m_2 = m_2'$,
- $(m_1, m_2) \in \Gamma_1$, and
- $(m_2', m_3) \in \Gamma_2$.

So we can think of F as the subset of $M_1 \times M_2 \times M_3$ consisting of all points (m_1, m_2, m_3) with $(m_1, m_2) \in \Gamma_1$ and $(m_2, m_3) \in \Gamma_2$. The clean intersection hypothesis involves two conditions. The first is that F be a manifold. The second is that the derived square be exact at all points. Let us state this second condition more explicitly: Let $m = (m_1, m_2, m_3) \in F$. We have the following vector spaces:

$$\begin{aligned} V_1 &:= T_{m_1} M_1, \\ V_2 &:= T_{m_2} M_2, \\ V_3 &:= T_{m_3} M_3, \\ \Gamma_1^m &:= T_{(m_1,m_2)}\Gamma_1, \quad \text{and} \\ \Gamma_2^m &:= T_{(m_2,m_3)}\Gamma_2. \end{aligned}$$

So

$$\Gamma_1^m \subset T_{(m_1,m_2)}(M_1 \times M_2) = V_1 \oplus V_2$$

is a linear Lagrangian subspace of $V_1^- \oplus V_2$. Similarly, Γ_2^m is a linear Lagrangian subspace of $V_2^- \oplus V_3$. The *clean intersection hypothesis* asserts that $T_m F$ is given by the exact square

$$\begin{array}{ccc} T_m F & \xrightarrow{d(\iota_1)_m} & \Gamma_1^m \\ {\scriptstyle d(\iota_2)_m}\downarrow & & \downarrow{\scriptstyle d\pi_{(m_1,m_2)}} \\ \Gamma_2^m & \xrightarrow[d\rho_{(m_2,m_3)}]{} & T_{m_2}M_2 \end{array} \tag{4.3}$$

In other words, $T_m F$ consists of all $(v_1, v_2, v_3) \in V_1 \oplus V_2 \oplus V_3$ such that

$$(v_1, v_2) \in \Gamma_1^m \quad \text{and} \quad (v_2, v_3) \in \Gamma_2^m.$$

The exact square (4.3) is of the form (3.10) that we considered in Section 3.4. We know from Section 3.4 that $\Gamma_2^m \circ \Gamma_1^m$ is a linear Lagrangian subspace of $V_1^- \oplus V_3$. In particular its dimension is $\frac{1}{2}(\dim M_1 + \dim M_3)$ which does not depend on the choice of $m \in F$. This implies the following: Let

$$\iota : F \to M_1 \times M_2 \times M_3$$

denote the inclusion map, and let

$$\kappa_{13} : M_1 \times M_2 \times M_3 \to M_1 \times M_3$$

denote the projection onto the first and third components. So

$$\kappa_{13} \circ \iota : F \to M_1 \times M_3$$

is a smooth map whose differential at any point $m \in F$ maps $T_m F$ onto $\Gamma_2^m \circ \Gamma_1^m$ and so has locally constant rank. Furthermore, the image of $T_m F$ is a Lagrangian subspace of $T_{(m_1,m3)}(M_1^- \times M_3)$. We have proved:

Theorem 4.1.1. *If the canonical relations* $\Gamma_1 \subset M_1^- \times M_2$ *and* $\Gamma_2 \subset M_2^- \times M_3$ *intersect cleanly, then their composition* $\Gamma_2 \circ \Gamma_1$ *is an immersed Lagrangian submanifold of* $M_1^- \times M_3$.

We must still impose conditions that will ensure that $\Gamma_2 \circ \Gamma_1$ is an honest submanifold of $M_1 \times M_3$. We will do this in the next section.

We will need a name for the manifold F we created out of Γ_1 and Γ_2 above. As in the linear case, we will call it $\Gamma_2 \star \Gamma_1$.

4.2 Composable canonical relations.

We recall a theorem from differential topology:

Theorem 4.2.1. *Let* X *and* Y *be smooth manifolds and* $f : X \to Y$ *is a smooth map of constant rank. Let* $W = f(X)$. *Suppose that* f *is proper and that for every* $w \in W$, $f^{-1}(w)$ *is connected and simply connected. Then* W *is a smooth submanifold of* Y.

We apply this theorem to the map $\kappa_{13} \circ \iota : F \to M_1 \times M_3$. To shorten the notation, let us define

$$\kappa := \kappa_{13} \circ \iota. \tag{4.4}$$

Theorem 4.2.2. *Suppose that the canonical relations* Γ_1 *and* Γ_2 *intersect cleanly. Suppose in addition that the map* κ *is proper and that the inverse image of every* $\gamma \in \Gamma_2 \circ \Gamma_1 = \kappa(\Gamma_2 \star \Gamma_1)$ *is connected and simply connected. Then* $\Gamma_2 \circ \Gamma_1$ *is a canonical relation. Furthermore*

$$\kappa : \Gamma_2 \star \Gamma_1 \to \Gamma_2 \circ \Gamma_1 \tag{4.5}$$

is a smooth fibration with compact connected fibers.

So we are in the following situation: We can not always compose the canonical relations $\Gamma_2 \subset M_2^- \times M_3$ and $\Gamma_1 \subset M_1^- \times M_2$ to obtain a canonical relation $\Gamma_2 \circ \Gamma_1 \subset M_1^- \times M_3$. We must impose some additional conditions, for example those of the theorem. So, following Weinstein, [Wein81] we put quotation marks around the word category to indicate this fact.

We will let $\mathcal{S}$ denote the "category" whose objects are symplectic manifolds and whose morphisms are canonical relations. We will call $\Gamma_1 \subset M_1^- \times M_2$ and $\Gamma_2 \subset M_2^- \times M_3$ **cleanly composable** if they satisfy the hypotheses of Theorem 4.2.2.

If $\Gamma \subset M_1^- \times M_2$ is a canonical relation, we will sometimes use the notation

$$\Gamma \in \mathrm{Morph}(M_1, M_2)$$

and sometimes use the notation

$$\Gamma : \; M_1 \twoheadrightarrow M_2$$

to denote this fact.

4.3 Transverse composition.

A special case of clean intersection is transverse intersection. In fact, in applications, this is a convenient hypothesis, and it has some special properties:

Suppose that the maps π and ρ are transverse. This means that

$$\pi \times \rho : \Gamma_1 \times \Gamma_2 \to M_2 \times M_2$$

intersects Δ_{M_2} transversally, which implies that the codimension of

$$\Gamma_2 \star \Gamma_1 = (\pi \times \rho)^{-1}(\Delta_{M_2})$$

in $\Gamma_1 \times \Gamma_2$ is $\dim M_2$. So with $F = \Gamma_2 \star \Gamma_1$ we have

$$\begin{aligned} \dim F &= \dim \Gamma_1 + \dim \Gamma_2 - \dim M_2 \\ &= \frac{1}{2}\dim M_1 + \frac{1}{2}\dim M_2 + \frac{1}{2}\dim M_2 + \frac{1}{2}\dim M_3 - \dim M_2 \\ &= \frac{1}{2}\dim M_1 + \frac{1}{2}\dim M_3 \\ &= \dim \Gamma_2 \circ \Gamma_1. \end{aligned}$$

So under the hypothesis of transversality, the map $\kappa = \kappa_{13} \circ \iota$ is an immersion. If we add the hypotheses of Theorem 4.2.2, we see that κ is a diffeomorphism.

For example, if Γ_2 is the graph of a symplectomorphism of M_2 with M_3 then $d\rho_{(m_2,m_3)} : T_{(m_2,m_3)}(\Gamma) \to T_{m_2}M_2$ is surjective at all points $(m_2, m_3) \in \Gamma_2$. So if $m = (m_1, m_2, m_2, m_3) \in \Gamma_1 \times \Gamma_2$ the image of $d(\pi \times \rho)_m$ contains all vectors of the form $(0, w)$ in $T_{m_2}M_2 \oplus T_{m_2}M_2$ and so is transverse to the diagonal. The manifold $\Gamma_2 \star \Gamma_1$ consists of all points of the form $(m_1, m_2, g(m_2))$ with $(m_1, m_2) \in \Gamma_1$, and

$$\kappa : (m_1, m_2, g(m_2)) \mapsto (m_1, g(m_2)).$$

Since g is one to one, so is κ. So the graph of a symplectomorphism is transversally composible with *any* canonical relation.

We will need the more general concept of "clean composability" described in the preceding section for certain applications.

4.4 Lagrangian submanifolds as canonical relations.

We can consider the "zero dimensional symplectic manifold" consisting of the distinguished point that we call "pt.". Then a canonical relation between pt. and a symplectic manifold M is a Lagrangian submanifold of pt. $\times M$ which may be identified with a Lagrangian submanifold of M. These are the "points" in our "category" $\mathcal{S}$.

Suppose that Λ is a Lagrangian submanifold of M_1 and $\Gamma \in \text{Morph}(M_1, M_2)$ is a canonical relation. If we think of Λ as an element of Morph(pt., M_1), then if Γ and Λ are composible, we can form $\Gamma \circ \Lambda \in \text{Morph}(\text{pt.}, M_2)$ which may be identified with a Lagrangian submanifold of M_2. If we want to think of it this way, we may sometimes write $\Gamma(\Lambda)$ instead of $\Gamma \circ \Lambda$.

We can mimic the construction of composition given in Section 3.3.2 for the category of finite sets and relations. Let M_1, M_2 and M_3 be symplectic manifolds and let $\Gamma_1 \in \text{Morph}(M_1, M_2)$ and $\Gamma_2 \in \text{Morph}(M_2, M_3)$ be canonical relations. So

$$\Gamma_1 \times \Gamma_2 \subset M_1^- \times M_2 \times M_2^- \times M_3$$

is a Lagrangian submanifold. Let

$$\tilde{\Delta}_{M_1,M_2,M_3} = \{(x,y,y,z,x,z)\} \subset M_1 \times M_2 \times M_2 \times M_3 \times M_1 \times M_3. \tag{4.6}$$

We endow the right hand side with the symplectic structure

$$M_1 \times M_2^- \times M_2 \times M_3^- \times M_1^- \times M_3 = (M_1^- \times M_2 \times M_2^- \times M_3)^- \times (M_1^- \times M_3).$$

Then $\tilde{\Delta}_{M_1,M_2,M_3}$ is a Lagrangian submanifold, i.e. an element of

$$\text{Morph}(M_1^- \times M_2 \times M_2^- \times M_3, M_1^- \times M_3).$$

Just as in Section 3.3.2,

$$\tilde{\Delta}_{M_1,M_2,M_3}(\Gamma_1 \times \Gamma_2) = \Gamma_2 \circ \Gamma_1.$$

It is easy to check that Γ_2 and Γ_1 are composible if and only if $\tilde{\Delta}_{M_1,M_2,M_3}$ and $\Gamma_1 \times \Gamma_2$ are composible.

4.5 The involutive structure on $\mathcal{S}$.

Let $\Gamma \in \text{Morph}(M_1, M_2)$ be a canonical relation. Just as in the category of finite sets and relations, define

$$\Gamma^\dagger = \{(m_2, m_1) | (m_1, m_2) \in \Gamma\}.$$

As a set it is a subset of $M_2 \times M_1$ and it is a Lagrangian submanifold of $M_2 \times M_1^-$. But then it is also a Lagrangian submanifold of

$$(M_2 \times M_1^-)^- = M_2^- \times M_1.$$

So

$$\Gamma^\dagger \in \text{Morph}(M_2, M_1).$$

Therefore $M \mapsto M, \Gamma \mapsto \Gamma^\dagger$ is a involutive functor on $\mathcal{S}$.

4.6 Reductions in the symplectic “category”.

In this section we recast the results of Sections 3.4.8 and 3.4.9 in the manifold setting.

4.6.1 Reductions in the symplectic "category" are reductions by coisotropics.

Let $Z \subset X$ be a coisotropic submanifold. The null distribution of $\iota_Z \omega_X$ is a foliation by Frobenius. Suppose that it is fibrating with base Y so we have $\pi : Z \to Y$ where the fiber dimension of π equals the codimension of $Z = k$, say. We have an induced symplectic form ω_Y on Y such that $\pi^* \omega_Y = \iota^* \omega_X$ so the subset

$$\{(z, \pi(z)) | z \in Z\} \subset X^- \times Y$$

is isotropic for the form $\omega_Y - \omega_X$. Its dimension is $\dim Z = \dim X - k = \frac{1}{2}(\dim X + \dim Y)$ since $\dim Y = \dim X - 2k$, so is Lagrangian. As a morphism it is surjective and single valued so is a reduction in the sense of Section 3.3.5.

Conversely, suppose that a morphism in our "category" is surjective with image Y and let Z be the pre-image of Y. So we are assuming that $Z \subset X$ is a submanifold with surjection $\pi : Z \to Y$. The Lagrangian submanifold Λ of $X^- \times Y$ consists of all $(z, \pi(z)), z \in Z$. Its dimension equals $\dim Z$ so we must have

$$\dim Z = \frac{1}{2} \dim X + \frac{1}{2} \dim Y.$$

Let $k := \dim Z - \dim Y$. Then we must have $\dim Z = \dim X - k$. Let V be the vertical bundle for the fibration π. Since Λ is isotropic, so that $\pi^* \omega_Y = \iota_Z^* \omega_X$ we see that the orthogonal complement $TV^\perp$ relative to ω_X to the tangent space TV contains TZ. But the dimension of this complement is $\dim X - k = \dim Z$ so Z is co-isotropic.

Thus we obtain symplectic "category" version of the Proposizione 3.10 of Benenti and Tulszyjew [Ben], namely that a reduction $\Gamma \in \mathrm{Morph}(X, Y)$ consists of a co-isotropic submanifold $Z \subset X$ with $\pi : Z \to Y$ the fibration associated to the null foliation $\iota_Z^* \omega_X$. Then Γ consists of all $(z, \pi(z))$.

4.6.2 The decomposition of any morphism into a reduction and a coreduction.

We next prove Weinstein's theorem, [Wein11] that any $f \in \mathrm{Morph}(X, Y)$ can be written as the transverse composition of a reduction and a coreduction. This is the manifold version of Proposition 3.4.4, but the proof is essentially identical:

Let f be a morphism from X to Y. Since f is a Lagrangian submanifold of $X^- \times Y$, we can think of f as a morphism $\gamma(f)$ from pt.. to $X^- \times X^- \times Y$. This is a coreduction. Hence so is $\mathbf{id}. \times \gamma(f)$ which is a morphism from $X \times \mathrm{pt.}$ to $X \times X^- \times Y$. As a Lagrangian submanifold of $(X \times \mathrm{pt.})^- \times (X \times X^- \times Y) = X^- \times (X \times X^- \times Y)$ it consists of all points of the form

$$(x, x, x', y) \text{ with } (x', y) \in f. \tag{4.7}$$

Δ_X is a Lagrangian subspace of $X^- \times X$ which we can think of as a morphism ϵ_X from $X^- \times X$ to pt.. It is a reduction, hence so is $\epsilon_X \times \mathbf{id}_Y$. As a subset of

$(X^- \times X \times Y)^- \times Y$ it consists of all points of the form

$$(x, x, y, y). \tag{4.8}$$

The composite of these two morphisms consists of the subset of $X \times Y = X \times \text{pt.} \times \text{pt.} \times Y$ given by those (x,y) such that there exists a $w = (x, x', y)$ with (w, y) of the form (4.8) so $x = x'$ with (x, w) of the form (3.17) so $(x', y) \in \Gamma$. So the composite is Γ.

4.6.3 Composition with reductions or co-reductions.

We now give the manifold version of Prop. 3.4.9. Suppose that $\Gamma \in \text{Morph}(X, B)$ is a reduction and so corresponds to a co-isotropic submanifold $Z \subset X$, and let V be a typical fiber the projection $\pi_\Gamma : Z \to B$. Let $\Lambda \in \text{Morph}(B, W)$. Since π is surjective, for any $(b, w) \in \Lambda$ there exists a $z \in Z$ with $(z, w) \in \Lambda \circ \Gamma$ with $\pi_\Gamma(z) = b$ and this z is determined up to an element of V. So

Proposition 4.6.1. *If* $\Gamma \in \text{Morph}(X, B)$ *is a reduction with* $V = \ker(\pi_\Gamma)$ *and* $\Lambda \in \text{Morph}(B, W)$ *then*

$$\Lambda \circ \Gamma \sim V \times \Lambda.$$

Hence, if $\Gamma \in \text{Morph}(B, X)$ *is a co-reduction with* $V sim \ker \rho_\Gamma$ *and* $\Lambda \in \text{Morph}(W, B)$ *then*

$$\Gamma \circ \Lambda = \Lambda \times V.$$

4.7 Canonical relations between cotangent bundles.

In this section we want to discuss some special properties of our "category" $\mathcal{S}$ when we restrict the objects to be cotangent bundles (which are, after all, special kinds of symplectic manifolds). One consequence of our discussion will be that $\mathcal{S}$ contains the category $\mathcal{C}^\infty$ whose objects are smooth manifolds and whose morphisms are smooth maps as a (tiny) subcategory. Another consequence will be a local description of Lagrangian submanifolds of the cotangent bundle which generalizes the description of horizontal Lagrangian submanifolds of the cotangent bundle that we gave in Chapter 1. We will use this local description to deal with the problem of passage through caustics that we encountered in Chapter 1.

We recall the following definitions from Chapter 1: Let X be a smooth manifold and T^*X its cotangent bundle, so that we have the projection $\pi : T^*X \to X$. The canonical one form α_X is defined by (1.8). We repeat the definition: If $\xi \in T^*X, x = \pi(\xi)$, and $v \in T_\xi(T^*X)$ then the value of α_X at v is given by

$$\langle \alpha_X, v \rangle := \langle \xi, d\pi_\xi v \rangle. \tag{1.8}$$

The symplectic form ω_X is given by

$$\omega_X = -d\alpha_X. \tag{1.10}$$

So if Λ is a submanifold of T^*X on which α_X vanishes and whose dimension is $\dim X$ then Λ is (a special kind of) Lagrangian submanifold of T^*X.

The conormal bundle.

An instance of this is the **conormal bundle** of a submanifold: Let $Y \subset X$ be a submanifold. Its conormal bundle

$$N^*Y \subset T^*X$$

consists of all $z = (x, \xi) \in T^*X$ such that $x \in Y$ and ξ vanishes on T_xY. If $v \in T_z(N^*Y)$ then $d\pi_z(v) \in T_xY$ so by (1.8) $\langle \alpha_X, v \rangle = 0$.

4.8 The canonical relation associated to a map.

Let X_1 and X_2 be manifolds and $f : X_1 \to X_2$ be a smooth map. We set

$$M_1 := T^*X_1 \quad \text{and} \quad M_2 := T^*X_2$$

with their canonical symplectic structures. We have the identification

$$M_1 \times M_2 = T^*X_1 \times T^*X_2 = T^*(X_1 \times X_2).$$

The graph of f is a submanifold of $X_1 \times X_2$:

$$X_1 \times X_2 \supset \text{graph}(f) = \{(x_1, f(x_1))\}.$$

So the conormal bundle of the graph of f is a Lagrangian submanifold of $M_1 \times M_2$. Explicitly,

$$N^*(\text{graph}(f)) = \{(x_1, \xi_1, x_2, \xi_2) | x_2 = f(x_1), \quad \xi_1 = -df^*_{x_1}\xi_2\}. \tag{4.9}$$

Let

$$\varsigma_1 : T^*X_1 \to T^*X_1$$

be defined by

$$\varsigma_1(x, \xi) = (x, -\xi).$$

Then $\varsigma_1^*(\alpha_{X_1}) = -\alpha_{X_1}$ and hence

$$\varsigma_1^*(\omega_{X_1}) = -\omega_{X_1}.$$

We can think of this as saying that ς_1 is a symplectomorphism of M_1 with M_1^- and hence

$$\varsigma_1 \times \text{id}$$

is a symplectomorphism of $M_1 \times M_2$ with $M_1^- \times M_2$. Let

$$\Gamma_f := (\varsigma_1 \times \text{id})(N^*(\text{graph}(f))). \tag{4.10}$$

Then Γ_f is a Lagrangian submanifold of $M_1^- \times M_2$. In other words,

$$\Gamma_f \in \text{Morph}(M_1, M_2).$$

Explicitly,

$$\Gamma_f = \{(x_1, \xi_1, x_2, \xi_2)|x_2 = f(x_1), \quad \xi_1 = df_{x_1}^* \xi_2\}. \tag{4.11}$$

Suppose that $g : X_2 \to X_3$ is a smooth map so that $\Gamma_g \in \text{Morph}(M_2, M_3)$. So

$$\Gamma_g = \{(x_2, \xi_2, x_3, \xi_3)|x_3 = g(x_2), \xi_2 = dg_{x_2}^* \xi_3.\}.$$

The maps

$$\pi : \Gamma_f \to M_2, \quad (x_1, \xi_1, x_2, \xi_2) \mapsto (x_2, \xi_2)$$

and

$$\rho : \Gamma_g \to M_2, \quad (x_2, \xi_2, x_3, \xi_3) \mapsto (x_2, \xi_2)$$

are transverse. Indeed at any point $(x_1, \xi_1, x_2, \xi_2, x_2, \xi_2, x_3, \xi_3)$ the image of $d\pi$ contains all vectors of the form $(0, w)$ in $T_{x_2,\xi_2}(T^*M_2)$, and the image of $d\rho$ contains all vectors of the form $(v, 0)$. So Γ_g and Γ_f are transversely composible. Their composite $\Gamma_g \circ \Gamma_f$ consists of all (x_1, ξ_1, x_3, ξ_3) such that there exists an x_2 such that $x_2 = f(x_1)$ and $x_3 = g(x_2)$ and a ξ_2 such that $\xi_1 = df_{x_1}^* \xi_2$ and $\xi_2 = dg_{x_2}^* \xi_3$. But this is precisely the condition that $(x_1, \xi_1, x_3, \xi_3) \in \Gamma_{g \circ f}$! We have proved:

Theorem 4.8.1. *The assignments*

$$X \mapsto T^*X$$

and

$$f \mapsto \Gamma_f$$

define a covariant functor from the category $\mathcal{C}^\infty$ of manifolds and smooth maps to the symplectic "category" $\mathcal{S}$. As a consequence the assignments $X \mapsto T^*X$ and

$$f \mapsto (\Gamma_f)^\dagger$$

define a contravariant functor from the category $\mathcal{C}^\infty$ of manifolds and smooth maps to the symplectic "category" $\mathcal{S}$.

We now study special cases of these functors in a little more detail:

4.9 Pushforward of Lagrangian submanifolds of the cotangent bundle.

Let $f : X_1 \to X_2$ be a smooth map, and $M_1 := T^*X_1$, $M_2 := T^*X_2$ as before. The Lagrangian submanifold $\Gamma_f \subset M_1^- \times M_2$ is defined by (4.11). In particular, it is a subset of $T^*X_1 \times T^*X_2$ and hence a particular kind of relation (in the sense of Chapter 3). So if A is any subset of T^*X_1 then $\Gamma_f(A)$ is a subset of T^*X_2 which we shall also denote by $df_*(A)$. So

$$df_*(A) := \Gamma_f(A), \qquad A \subset T^*X_1.$$

Explicitly,

$$df_*A = \{(y,\eta) \in T^*X_2 | \exists (x,\xi) \in A \text{ with } y = f(x) \text{ and } \xi = df_x^*\eta\}.$$

Now suppose that $A = \Lambda$ is a Lagrangian submanifold of T^*X_1. Considering Λ as an element of Morph(pt.$, T^*X_1$) we may apply Theorem 4.1.1. Let

$$\pi_1 : \; N^*(\text{graph}(f)) \to T^*X_1$$

denote the restriction to $N^*(\text{graph}(f))$ of the projection of $T^*X_1 \times T^*X_2$ onto the first component. Notice that $N^*(\text{graph}(f))$ is stable under the map $(x,\xi,y,\eta) \mapsto (x,-\xi,y,-\eta)$ and hence π_1 intersects Λ cleanly if and only if $\pi_1 \circ (\varsigma \times \text{id}) : \Gamma_f \to T^*X_1$ intersects Λ cleanly where, by abuse of notation, we have also denoted by π_1 restriction of the projection to Γ_f. So

Theorem 4.9.1. *If Λ is a Lagrangian submanifold and $\pi_1 : N^*(\text{graph}(f)) \to T^*X_1$ intersects Λ cleanly then $df_*(\Lambda)$ is an immersed Lagrangian submanifold of T^*X_2.*

If f has constant rank, then the dimension of $df_x^*T^*(X_2)_{f(x)}$ does not vary, so that $df^*(T^*X_2)$ is a sub-bundle of T^*X_1. If Λ intersects this subbundle transversally, then our conditions are certainly satisified. So

Theorem 4.9.2. *Suppose that $f : X_1 \to X_2$ has constant rank. If Λ is a Lagrangian submanifold of T^*X_1 which intersects $df^*T^*X_2$ transversaly then $df_*(\Lambda)$ is a Lagrangian submanifold of T^*X_2.*

For example, if f is an immersion, then $df^*T^*X_2 = T^*X_1$ so all Lagrangian submanifolds are transverse to $df^*T^*X_2$.

Corollary 4.9.1. *If f is an immersion, then $df_*(\Lambda)$ is a Lagrangian submanifold of T^*X_2.*

At the other extreme, suppose that $f : X_1 \to X_2$ is a fibration. Then $H^*(X_1) := df^*T^*N$ consists of the "horizontal sub-bundle", i.e those covectors which vanish when restricted to the tangent space to the fiber. So

Corollary 4.9.2. *Let $f : X_1 \to X_2$ be a fibration, and let $H^*(X_1)$ be the bundle of the horizontal covectors in T^*X_1. If Λ is a Lagrangian submanifold of T^*X_1 which intersects $H^*(X_1)$ transversally, then $df_*(\Lambda)$ is a Lagrangian submanifold of T^*X_2.*

An important special case of this corollary for us will be when $\Lambda = \text{graph } d\phi$. Then $\Lambda \cap H^*(X_1)$ consists of those points where the "vertical derivative", i.e. the derivative in the fiber direction vanishes. At such points $d\phi$ descends to give a covector at $x_2 = f(x_1)$. If the intersection is transverse, the set of such covectors is then a Lagrangian submanifold of T^*N. All of the next chapter will be devoted to the study of this special case of Corollary 4.9.2.

4.9.1 Envelopes.

Another important special case of Corollary 4.9.2 is the theory of envelopes, a classical subject which has more or less disappeared from the standard curriculum:

Let

$$X_1 = X \times S, \qquad X_2 = X$$

where X and S are manifolds and let $f = \pi : X \times S \to X$ be projection onto the first component.

Let

$$\phi : X \times S \to \mathbb{R}$$

be a smooth function having 0 as a regular value so that

$$Z := \phi^{-1}(0)$$

is a submanifold of $X \times S$. In fact, we will make a stronger assumption: Let $\phi_s : X \to \mathbb{R}$ be the map obtained by holding s fixed:

$$\phi_s(x) := \phi(x, s).$$

We make the stronger assumption that each ϕ_s has 0 as a regular value, so that

$$Z_s := \phi_s^{-1}(0) = Z \cap (X \times \{s\})$$

is a submanifold and

$$Z = \bigcup_s Z_s$$

as a set. The Lagrangian submanifold $N^*(Z) \subset T^*(X \times S)$ consists of all points of the form

$$(x, s, td\phi_X(x, s), td_S\phi(x, s)) \text{ such that } \phi(x, s) = 0.$$

Here t is an arbitrary real number. The sub-bundle $H^*(X \times S)$ consists of all points of the form

$$(x, s, \xi, 0).$$

So the transversality condition of Corollary 4.9.2 asserts that the map

$$z \mapsto d\left(\frac{\partial \phi}{\partial s}\right)$$

has rank equal to dim S on Z. The image Lagrangian submanifold $df_* N^*(Z)$ then consists of all covectors $td_X\phi$ where

$$\phi(x, s) = 0 \quad \text{and} \quad \frac{\partial \phi}{\partial s}(x, s) = 0,$$

a system of $p+1$ equations in $n+p$ variables, where $p = \dim S$ and $n = \dim X$

Our transversality assumptions say that these equations define a submanifold of $X \times S$. If we make the stronger hypothesis that the last p equations can be solved for s as a function of x, then the first equation becomes

$$\phi(x, s(x)) = 0$$

which defines a hypersurface $\mathcal{E}$ called the **envelope** of the surfaces Z_s. Furthermore, by the chain rule,

$$d\phi(\cdot, s(\cdot)) = d_X\phi(\cdot, s(\cdot)) + d_S\phi(\cdot, s(\cdot))d_X s(\cdot) = d_X\phi(\cdot, s(\cdot))$$

since $d_S\phi = 0$ at the points being considered. So if we set

$$\psi := \phi(\cdot, s(\cdot))$$

we see that under these restrictive hypotheses $df_* N^*(Z)$ consists of all multiples of $d\psi$, i.e.

$$df_*(N^*(Z)) = N^*(\mathcal{E})$$

is the normal bundle to the envelope.

In the classical theory, the envelope "develops singularities". But from our point of view it is natural to consider the Lagrangian submanifold $df_* N^*(Z)$. This will not be globally a normal bundle to a hypersurface because its projection on X (from T^*X) may have singularities. But as a submanifold of T^*X it is fine:

Examples:

- Suppose that S is an oriented curve in the plane, and at each point $s \in S$ we draw the normal ray to S at s. We might think of this line as a light ray propagating down the normal. The initial curve is called an "initial wave front" and the curve along which the light tends to focus is called the "caustic". Focusing takes place where "nearby normals intersect" i.e. at the envelope of the family of rays. These are the points which are the loci of the centers of curvature of the curve, and the corresponding curve is called the evolute.

- We can let S be a hypersurface in n-dimensions, say a surface in three dimensions. We can consider a family of lines emanating from a point source (possible at infinity), and reflected by S. The corresponding envelope is called the "caustic by reflection". In Descartes' famous theory of the rainbow he considered a family of parallel lines (light rays from the sun) which were refracted on entering a spherical raindrop, internally reflected by the opposite side and refracted again when exiting the raindrop. The corresponding "caustic" is the Descartes cone of 42 degrees.

- If S is a submanifold of $\mathbb{R}^n$ we can consider the set of spheres of radius r centered at points of S. The corresponding envelope consist of "all points at distance r from S". But this develops singularities past the radii of curvature. Again, from the Lagrangian or "upstairs" point of view there is no problem.

4.10 Pullback of Lagrangian submanifolds of the cotangent bundle.

We now investigate the contravariant functor which assigns to the smooth map $f : X_1 \to X_2$ the canonical relation

$$\Gamma_f^\dagger : \quad T^*X_2 \twoheadrightarrow T^*X_1.$$

As a subset of $T^*(X_2) \times T^*(X_1)$, $\Gamma_f^\dagger$ consists of all

$$(y, \eta, x, \xi)|\ y = f(x), \quad \text{and} \quad \xi = df_x^*(\eta). \tag{4.12}$$

If B is a subset of T^*X_2 we can form $\Gamma_f^\dagger(B) \subset T^*X_1$ which we shall denote by $df^*(B)$. So

$$df^*(B) := \Gamma_f^\dagger(B) = \{(x, \xi)|\exists\ b = (y, \eta) \in B \text{ with } f(x) = y,\ df_x^*\eta = \xi\}. \tag{4.13}$$

If $B = \Lambda$ is a Lagrangian submanifold, once again we may apply Theorem 4.1.1 to obtain a sufficient condition for $df^*(\Lambda)$ to be a Lagrangian submanifold of T^*X_1. Notice that in the description of $\Gamma_f^\dagger$ given in (4.12), the η can vary freely in $T^*(X_2)_{f(x)}$. So the issue of clean or transverse intersection comes down to the behavior of the first component. So, for example, we have the following theorem:

Theorem 4.10.1. *Let $f : X_1 \to X_2$ be a smooth map and Λ a Lagrangian submanifold of T^*X_2. If the maps f, and the restriction of the projection $\pi : T^*X_2 \to X_2$ to Λ are transverse, then $df^*\Lambda$ is a Lagrangian submanifold of T^*X_1.*

Here are two examples of the theorem:

- Suppose that Λ is a horizontal Lagrangian submanifold of T^*X_2. This means that restriction of the projection $\pi : T^*X_2 \to X_2$ to Λ is a diffeomorphism and so the transversality condition is satisfied for any f. Indeed, if $\Lambda = \Lambda_\phi$ for a smooth function ϕ on X_2 then

$$f^*(\Lambda_\phi) = \Lambda_{f^*\phi}.$$

- Suppose that $\Lambda = N^*(Y)$ is the normal bundle to a submanifold Y of X_2. The transversality condition becomes the condition that the map f is transversal to Y. Then $f^{-1}(Y)$ is a submanifold of X_1. If $x \in f^{-1}(Y)$ and $\xi = df_x^*\eta$ with $(f(x), \eta) \in N^*(Y)$ then ξ vanishes when restricted to $T(f^{-1}(Y))$, i.e. $(x, \xi) \in \mathcal{N}(f^{-1}(S))$. More precisely, the transversality asserts that at each $x \in f^{-1}(Y)$ we have $df_x(T(X_1)_x) + TY_{f(x)} = T(X_2)_{f(x)}$ so

$$T(X_1)_x / T(f^{-1}(Y))_x \cong T(X_2)_{f(x))} / TY_{f(x)}$$

and so we have an isomorphism of the dual spaces

$$N_x^*(f^{-1}(Y)) \cong N^*f(x)(Y).$$

In short, the pullback of $N^*(Y)$ is $N^*(f^{-1}(Y))$.

4.11 The moment map.

In this section we show how to give a categorical generalization of the classical moment map for a Hamiltonian group action. We begin with a review of the classical theory.

4.11.1 The classical moment map.

In this section we recall the classical moment map, especially from Weinstein's point of view.

Let (M, ω) be a symplectic manifold, K a connected Lie group and τ an action of K on M preserving the symplectic form. From τ one gets an infinitesimal action

$$\delta\tau : \mathfrak{k} \to \mathrm{Vect}(M) \tag{4.14}$$

of the Lie algebra, $\mathfrak{k}$, of K, mapping $\xi \in \mathfrak{k}$ to the vector field, $\delta\tau(\xi) =: \xi_M$. Here ξ_M is the infinitesimal generator of the one parameter group

$$t \mapsto \tau_{\exp -t\xi}.$$

The minus sign is to guarantee that $\delta\tau$ is a Lie algebra homomorphism.

In particular, for $p \in M$, one gets from (4.14) a linear map,

$$d\tau_p : \mathfrak{k} \to T_pM\,, \quad \xi \to \xi_M(p)\,; \tag{4.15}$$

and from ω_p a linear isomorphism,

$$T_p \to T_p^* \quad v \to i(v)\omega_p\,; \tag{4.16}$$

which can be composed with (4.15) to get a linear map

$$\tilde{d\tau_p} : \mathfrak{k} \to T_p^*M\,. \tag{4.17}$$

Definition 4.11.1. *A K-equivariant map*

$$\phi : M \to \mathfrak{k}^* \tag{4.18}$$

is a **moment map***, if for every $p \in M$:*

$$d\phi_p : T_pM \to \mathfrak{k}^* \tag{4.19}$$

is the transpose of the map (4.17).

The property (4.19) determines $d\phi_p$ at all points p and hence determines ϕ up to an additive constant, $c \in (\mathfrak{k}^*)^K$ if M is connected. Thus, in particular, if K is semi-simple, the moment map, if it exists, is unique. As for the existence of ϕ, the duality of (4.17) and (4.19) can be written in the form

$$i(\xi_M)\omega = d\langle \phi, \xi \rangle \tag{4.20}$$

for all $\xi \in \mathfrak{k}$; and this shows that the vector field, ξ_M, has to be Hamiltonian. If K is compact the converse is true. A sufficient condition for the existence of ϕ is that each of the vector fields, ξ_M, be Hamiltonian. (See for instance, [**?**], § 26.) An equivalent formulation of this condition will be useful below:

Definition 4.11.2. *A symplectomorphism, $f : M \to M$ is* **Hamiltonian** *if there exists a family of symplectomorphisms, $f_t : M \to M$, $0 \le t \le 1$, depending smoothly on t with $f_0 = id_M$ and $f_1 = f$, such that the vector field*

$$v_t = f_t^{-1} \frac{df_t}{dt}$$

is Hamiltonian for all t.

It is easy to see that ξ_M is Hamiltonian for all $\xi \in \mathfrak{k}$ if and only if the symplectomorphism, τ_g, is exact for all $g \in K$.

Our goal in this section is to describe a generalized notion of moment mapping in which there are no group actions involved. First, however, we recall a very suggestive way of thinking about moment mappings and the “moment geometry” associated with moment mappings, due to Alan Weinstein, [Wein81]. From the left action of K on T^*K one gets a trivialization

$$T^*K = K \times \mathfrak{k}^*$$

and via this trivialization a Lagrangian submanifold

$$\Gamma_\tau = \{(m, \tau_g m, g, \phi(m))\,;\, m \in M\,,\, g \in K\}\,,$$

of $M \times M^- \times T^*K$, which Weinstein calls *the moment Lagrangian.* He views this as a canonical relation between $M^- \times M$ and T^*K, i.e. as a morphism

$$\Gamma_\tau : M^- \times M \twoheadrightarrow T^*K\,.$$

4.11.2 Families of symplectomorphisms.

We now turn to the first stage of our generalization of the moment map, where the group action is replaced by a family of symplectomorphisms:

Let (M, ω) be a symplectic manifold, S an arbitrary manifold and f_s, $s \in S$, a family of symplectomorphisms of M depending smoothly on s . For $p \in M$ and $s_0 \in S$ let $g_{s_0,p} : S \to M$ be the map, $g_{s_0,p}(s) = f_s \circ f_{s_0}^{-1}(p)$. Composing the derivative of $g_{s_0,p}$ at s_0

$$(dg_{s_0,p})_{s_0} : T_{s_0}S \to T_pM \tag{4.21}$$

with the map (4.16) one gets a linear map

$$(d\tilde{g}_{s_0,p})_{s_0} : T_{s_0}S \to T_p^*M\,. \tag{4.22}$$

Now let Φ be a map of $M \times S$ into T^*S which is compatible with the projection, $M \times S \to S$ in the sense

$$\begin{array}{ccc} M \times S & \xrightarrow{\Phi} & T^*S \\ & \searrow & \downarrow \\ & & S \end{array}$$

commutes; and for $s_0 \in S$ let

$$\Phi_{s_0} : M \to T^*_{s_0} S$$

be the restriction of Φ to $M \times \{s_0\}$.

Definition 4.11.3. *Φ is a* **moment map** *if, for all s_0 and p,*

$$(d\Phi_{s_0})_p : T_p M \to T^*_{s_0} S \tag{4.23}$$

is the transpose of the map (4.22).

We will prove below that a sufficient condition for the existence of Φ is that the f_s's be Hamiltonian; and, assuming that Φ exists, we will consider the analogue for Φ of Weinstein's moment Lagrangian,

$$\Gamma_\Phi = \{(m, f_s(m), \Phi(m,s))\,;\, m \in M\,,\, s \in S\}\,, \tag{4.24}$$

and ask if the analogue of Weinstein's theorem is true: Is (4.24) a Lagrangian submanifold of $M \times M^- \times T^*S$?

Equivalently consider the imbedding of $M \times S$ into $M \times M^- \times T^*S$ given by the map

$$G : M \times S \to M \times M^- \times T^*S\,,$$

where $G(m,s) = (m, f_s(m), \Phi(m,s))$. Is this a Lagrangian imbedding? The answer is "no" in general, but we will prove:

Theorem 4.11.1. *The pull-back by G of the symplectic form on $M \times M^- \times T^*S$ is the pull-back by the projection, $M \times S \to S$ of a closed two-form, μ, on S.*

If μ is exact, i.e., if $\mu = d\nu$, we can modify Φ by setting

$$\Phi_{\text{new}}(m,s) = \Phi_{\text{old}}(m,s) - \nu_s\,,$$

and for this modified Φ the pull-back by G of the symplectic form on $M \times M^- \times T^*S$ will be zero; so we conclude:

Theorem 4.11.2. *If μ is exact, there exists a moment map, $\Phi : M \times S \to T^*S$, for which Γ_Φ is Lagrangian.*

The following converse result is also true.

Theorem 4.11.3. *Let Φ be a map of $M \times S$ into T^*S which is compatible with the projection of $M \times S$ onto S. Then if Γ_Φ is Lagrangian, Φ is a moment map.*

Remarks:

1. A moment map with this property is still far from being unique; however, the ambiguity in the definition of Φ is now a *closed* one-form, $\nu \in \Omega^1(S)$.

2. if $[\mu] \neq 0$ there is a simple expedient available for making Γ_Φ Lagrangian. One can modify the symplectic structure of T^*S by adding to the standard symplectic form the pull-back of $-\mu$ to T^*S.

3. Let $\mathcal{G}_e$ be the group of Hamiltonian symplectomorphisms of M. Then for every manifold, S and smooth map

$$F : S \to \mathcal{G}_e$$

one obtains by the construction above a cohomology class $[\mu]$ which is a homotopy invariant of the mapping F.

4. For a smooth map $F : S \to \mathcal{G}_e$, there exists an analogue of the character Lagrangian. Think of Γ_Φ as a canonical relation or “map”

$$\Gamma_\Phi : M^- \times M \twoheadrightarrow T^*S$$

and define the character Lagrangian of F to be the image with respect to Γ_Φ of the diagonal in $M^- \times M$.

Our proof of the results above will be an illustration of the principle: the more general the statement of a theorem the easier it is to prove. We will first generalize these results by assuming that the f_s's are canonical relations rather than canonical transformations, i.e., are morphisms in our category. Next we will get rid of morphisms altogether and replace $M \times M^-$ by a symplectic manifold M and canonical relations by Lagrangian submanifolds of M.

4.11.3 The moment map in general.

Let (M, ω) be a symplectic manifold. Let Z, X and S be manifolds and suppose that

$$\pi : Z \to S$$

is a fibration with fibers diffeomorphic to X. Let

$$G : Z \to M$$

be a smooth map and let

$$g_s : Z_s \to M, \quad Z_s := \pi^{-1}(s)$$

denote the restriction of G to Z_s. We assume that

$$g_s \text{ is a Lagrangian embedding} \tag{4.25}$$

and let

$$\Lambda_s := g_s(Z_s) \tag{4.26}$$

denote the image of g_s. Thus for each $s \in S$, the restriction of G imbeds the fiber, $Z_s = \pi^{-1}(s)$, into M as the Lagrangian submanifold, Λ_s. Let $s \in S$ and $\xi \in T_sS$. For $z \in Z_s$ and $w \in T_zZ_s$ tangent to the fiber Z_s

$$dG_z w = (dg_s)_z w \in T_{G(z)}\Lambda_s$$

so dG_z induces a map, which by abuse of language we will continue to denote by dG_z

$$dG_z : T_zZ/T_zZ_s \to T_mM/T_m\Lambda, \quad m = G(z). \tag{4.27}$$

But $d\pi_z$ induces an identification

$$T_zZ/T_z(Z_s) = T_sS. \tag{4.28}$$

Furthermore, we have an identification

$$T_mM/T_m(\Lambda_s) = T^*_m\Lambda_s \tag{4.29}$$

given by

$$T_mM \ni u \mapsto i(u)\omega_m(\cdot) = \omega_m(u, \cdot).$$

Finally, the diffeomorphism $g_s : Z_s \to \Lambda_s$ allows us to identify

$$T^*_m\Lambda_s \sim T^*_zZ_s, \quad m = G(z).$$

Via all these identifications we can convert (4.27) into a map

$$T_sS \to T^*_zZ_s\,. \tag{4.30}$$

Now let $\Phi : Z \to T^*S$ be a lifting of $\pi : Z \to S$, so that

$$\begin{array}{ccc} Z & \xrightarrow{\Phi} & T^*S \\ & \searrow_{\pi} & \downarrow \\ & & S \end{array}$$

commutes; and for $s \in S$ let

$$\Phi_s : Z_s \to T^*_sS$$

be the restriction of Φ to Z_s.

Definition 4.11.4. *Φ is a* **moment map** *if, for all s and all $z \in Z_s$,*

$$(d\Phi_s)_z : T_zZ_s \to T^*_sS \tag{4.31}$$

is the transpose of (4.30).

Note that this condition determines Φ_s up to an additive constant $\nu_s \in T^*_sS$ and hence, as in § 4.11.2, determines Φ up to a section, $s \to \nu_s$, of T^*S.

When does a moment map exist? By (4.30) a vector, $v \in T_sS$, defines, for every point, $z \in Z_s$, an element of T^*Z_s and hence defines a one-form on Z_s which we will show to be closed. We will say that G is *exact* if for all s and all $v \in T_sS$ this one-form is exact, and we will prove below that the exactness of G is a necessary and sufficient condition for the existence of Φ.

Given a moment map, Φ, one gets from it an imbedding

$$(G, \Phi) : Z \to M \times T^*S \tag{4.32}$$

and as in the previous section we can ask how close this comes to being a Lagrangian imbedding. We will prove

Theorem 4.11.4. *The pull-back by (4.32) of the symplectic form on $M \times T^*S$ is the pull-back by π of a closed two-form μ on S.*

The cohomology class of this two-form is an intrinsic invariant of G (doesn't depend on the choice of Φ) and as in the last section one can show that this is the only obstruction to making (4.32) a Lagrangian imbedding.

Theorem 4.11.5. *If $[\mu] = 0$ there exists a moment map, Φ, for which the imbedding (4.32) is Lagrangian.*

Conversely we will prove

Theorem 4.11.6. *Let Φ be a map of Z into T^*S lifting the map, π, of Z into S. Then if the imbedding (4.32) is Lagrangian Φ is a moment map.*

4.11.4 Proofs.

Let us go back to the map (4.30). If we hold s fixed but let z vary over Z_s, we see that each $\xi \in T_sS$ gives rise to a one form on Z_s. To be explicit, let us choose a trivialization of our bundle around Z_s so we have an identification

$$H : Z_s \times U \to \pi^{-1}(U)$$

where U is a neighborhood of s in S. If $t \mapsto s(t)$ is any curve on S with $s(0) = s,\ s'(0) = \xi$ we get a curve of maps $h_{s(t)}$ of $Z_s \to M$ where

$$h_{s(t)} = g_{s(t)} \circ H.$$

We thus get a vector field v^ξ along the map h_s

$$v^\xi : Z_s \to TM, \qquad v^\xi(z) = \frac{d}{dt} h_{s(t)}(z)_{|t=0}.$$

Then the one form in question is

$$\tau^\xi = h_s^*(i(v^\xi)\omega).$$

A direct check shows that this one form is exactly the one form described above (and hence is independent of all the choices). We claim that

$$d\tau^\xi = 0. \tag{4.33}$$

Indeed, the general form of the Weil formula (14.8) and the fact that $d\omega = 0$ gives

$$\left(\frac{d}{dt} h_{s(t)}^* \omega\right)_{|t=0} = dh_s^* i(v^\xi)\omega$$

and the fact that Λ_s is Lagrangian for all s implies that the left hand side and hence the right hand side is zero. Let us now assume that G is *exact*, i.e. that for all s and ξ the one form τ^ξ is exact. So

$$\tau^\xi = d\phi^\xi$$

for some C^∞ function ϕ^ξ on Z_s. The function ϕ^ξ is uniquely determined up to an additive constant on each Z_s (if Z_s is connected) which we can fix (in various ways) so that it depends smoothly on s and linearly on ξ. For example, if we have a cross-section $c : S \to Z$ we can demand that $\phi(c(s))^\xi \equiv 0$ for all s and ξ. Alternatively, we can equip each fiber Z_s with a compactly supported density dz_s which depends smoothly on s and whose integral over Z_s is one for each s. We can then demand that that $\int_{Z_s} \phi^\xi dz_s = 0$ for all ξ and s.

Suppose that we have made such choice. Then for fixed $z \in Z_s$ the number $\phi^\xi(z)$ depends linearly on ξ. Hence we get a map

$$\Phi_0 : Z \to T^*S, \qquad \Phi_0(z) = \lambda \Leftrightarrow \lambda(\xi) = \phi^\xi(z). \tag{4.34}$$

We shall see below (Theorem 4.11.8) that Φ_0 is a moment map by computing its derivative at $z \in Z$ and checking that it is the transpose of (4.30).

If each Z_s is connected, our choice determines ϕ^ξ up to an additive constant $\nu(s,\xi)$ which we can assume to be smooth in s and linear in ξ. Replacing ϕ^ξ by $\phi^\xi + \nu(s,\xi)$ has the effect of making the replacement

$$\Phi_0 \mapsto \Phi_0 + \nu \circ \pi$$

where $\nu : S \to T^*S$ is the one form $\langle \nu_s, \xi \rangle = \nu(s,\xi)$

Let ω_S denote the canonical two form on T^*S.

Theorem 4.11.7. *There exists a closed two form ρ on S such that*

$$G^*\omega - \Phi^*\omega_S = \pi^*\rho. \tag{4.35}$$

If $[\rho] = 0$ then there is a one form ν on S such that if we set

$$\Phi = \Phi_0 + \nu \circ \pi$$

then

$$G^*\omega - \Phi^*\omega_S = 0. \tag{4.36}$$

As a consequence, the map

$$\tilde{G} : Z \to M^- \times T^*S, \qquad z \mapsto (G(z), \Phi(z)) \tag{4.37}$$

is a Lagrangian embedding.

Proof. We first prove a local version of the theorem. Locally, we may assume that $Z = X \times S$. This means that we have an identification of Z_s with X for all s. By the Weinstein tubular neighborhood theorem we may assume (locally) that $M = T^*X$ and that for a fixed $s_0 \in S$ the Lagrangian submanifold Λ_{s_0} is the zero section of T^*X and that the map

$$G : X \times S \to T^*X$$

is given by

$$G(x,s) = d_X\psi(x,s)$$

where $\psi \in C^\infty(X \times S)$. In local coordinates $x_1, \ldots, x_k$ on X, this reads as

$$G(x,s) = \frac{\partial \psi}{\partial x_1} dx_1 + \cdots + \frac{\partial \psi}{\partial x_k} dx_k.$$

In terms of these choices, the maps $h_{s(t)}$ used above are given by

$$h_{s(t)}(x) = d_X \psi(x, s(t))$$

and so (in local coordinates) on X and on S the vector field v^ξ is given by

$$v^\xi(z) = \frac{d}{dt} h_{s(t)}(z)_{|t=0} =$$

$$\frac{\partial^2 \psi}{\partial x_1 \partial s_1} \xi_1 \frac{\partial}{\partial p_1} + \cdots + \frac{\partial^2 \psi}{\partial x_1 \partial s_r} \xi_r \frac{\partial}{\partial p_1} + \cdots + \frac{\partial^2 \psi}{\partial x_k \partial s_r} \xi_r \frac{\partial}{\partial p_k}$$

where $r = \dim S$. We can write this more compactly as

$$\frac{\partial \langle d_S \psi, \xi \rangle}{\partial x_1} \frac{\partial}{\partial p_1} + \cdots + \frac{\partial \langle d_S \psi, \xi \rangle}{\partial x_k} \frac{\partial}{\partial p_k}.$$

Taking the interior product of this with $\sum dq_i \wedge dp_i$ gives

$$-\frac{\partial \langle d_S \psi, \xi \rangle}{\partial x_1} dq_1 - \cdots - \frac{\partial \langle d_S \psi, \xi \rangle}{\partial x_k} dq_k$$

and hence the one form τ^ξ is given by

$$-d_X \langle d_S \psi, \xi \rangle.$$

so we may choose

$$\Phi(x,s) = -d_S \psi(x,s).$$

Thus

$$G^* \alpha_X = d_X \psi, \qquad \Phi^* \alpha_S = -d_S \psi$$

and hence

$$G^* \omega_X - \Phi^* \omega_S = -dd\psi = 0.$$

This proves a local version of the theorem with $\rho = 0$.

We now pass from the local to the global: By uniqueness, our global Φ_0 must agree with our local Φ up to the replacement $\Phi \mapsto \Phi + \mu \circ \pi$. So we know that

$$G^* \omega - \Phi_0^* \omega_S = (\mu \circ \pi)^* \omega_S = \pi^* \mu^* \omega_S.$$

Here μ is a one form on S regarded as a map $S \to T^*S$. But

$$d\pi^* \mu^* \omega_S = \pi^* \mu^* d\omega_S = 0.$$

So we know that $G^* \omega - \Phi_0^* \omega_S$ is a closed two form which is locally and hence globally of the form $\pi^* \rho$ where $d\rho = 0$. This proves (4.35).

Now suppose that $[\rho] = 0$ so we can write $\rho = d\nu$ for some one form ν on S. Replacing Φ_0 by $\Phi_0 + \nu \circ \pi$ replaces ρ by $\rho + \nu^*\omega_S$. But

$$\nu^*\omega_S = -\nu^* d\alpha_S = -d\nu = -\rho. \qquad \square$$

Remark. If $[\rho] \neq 0$ then we can not succeed by modifying Φ. But we can modify the symplectic form on T^*S replacing ω_S by $\omega_S - \pi_S^*\rho$ where π_S denotes the projection $T^*S \to S$.

4.11.5 The derivative of Φ.

We continue the current notation. So we have the map

$$\Phi : Z \to T^*S.$$

Fix $s \in S$. The restriction of Φ to the fiber Z_s maps $Z_s \to T_s^*S$. Since T_s^*S is a vector space, we may identify its tangent space at any point with T_s^*S itself. Hence for $z \in Z_s$ we may regard $d\Phi_z$ as a linear map from T_zZ to T_s^*S. So we write

$$d\Phi_z : T_zZ_s \to T_s^*S. \tag{4.38}$$

On the other hand, recall that using the identifications (4.28) and (4.29) we got a map

$$dG_z : T_sS \to T_m^*\Lambda, \quad m = G(z)$$

and hence composing with $d(g_s)_z^* : T_m^*\Lambda \to T_z^*Z_s$ a linear map

$$\chi_z := d(g_s)_z^* \circ dG_z : T_sS \to T_z^*Z. \tag{4.39}$$

Theorem 4.11.8. *The maps $d\Phi_z$ given by (4.38) and χ_z given by (4.39) are transposes of one another.*

Proof. Each $\xi \in T_sS$ gives rise to a one form τ^ξ on Z_s and by definition, the value of this one form at $z \in Z_s$ is exactly $\chi_z(\xi)$. The function ϕ^ξ was defined on Z_s so as to satisfy $d\phi^\xi = \tau^\xi$. In other words, for $v \in T_zZ$

$$\langle \chi_z(\xi), v\rangle = \langle d\Phi_z(v), \xi\rangle. \qquad \square$$

Corollary 4.11.1. *The kernel of χ_z is the annihilator of the image of the map (4.38). In particular z is a regular point of the map $\Phi : Z_s \to T_s^*S$ if the map χ_z is injective.*

Corollary 4.11.2. *The kernel of the map (4.38) is the annihilator of the image of χ_z.*

4.11.6 A converse.

The following is a converse to Theorem 4.11.7:

Theorem 4.11.9. *If $\Phi : Z \to T^*S$ is a lifting of the map $\pi : Z \to S$ to T^*S and (G, Φ) is a Lagrangian imbedding of*

$$Z \to M^- \times T^*S$$

then Φ is a moment map.

Proof. It suffices to prove this in the local model described above where $Z = X \times S$, $M = T^*X$ and $G(x, s) = d_X\psi(x, s)$. If $\Phi : X \times S \to T^*S$ is a lifting of the projection $X \times S \to X$, then (G, Φ) can be viewed as a section of $T^*(X \times S)$ i.e. as a one form β on $X \times S$. If (G, Φ) is a Lagrangian imbedding then β is closed. Moreover, the (1,0) component of β is $d_X\psi$ so $\beta - d\psi$ is a closed one form of type (0,1), and hence is of the form $\mu \circ \pi$ for some closed one form on S. this shows that

$$\Phi = d_S\psi + \pi^*\mu$$

and hence, as verfied above, is a moment map. □

4.11.7 Back to families of symplectomorphisms.

Let us now specialize to the case of a parametrized family of symplectomorphisms. So let (M, ω) be a symplectic manifold, S a manifold and

$$F : M \times S \to M$$

a smooth map such that

$$f_s : M \to M$$

is a symplectomophism for each s, where $f_s(m) = F(m, s)$. We can apply the results of the preceding section where now $\Lambda_s \subset M \times M^-$ is the graph of f_s (and the M of the preceding section is replaced by$M \times M^-$) and so

$$G : M \times S \to M \times M^-, \qquad G(m, s) = (m, F(m, s)). \tag{4.40}$$

Theorem 4.11.7 says that get a map

$$\Phi : M \times S \to T^*S$$

and a moment Lagrangian

$$\Gamma_\Phi \subset M \times M^- \times T^*S.$$

The equivariant situation.

Suppose that a compact Lie group K acts as fiber bundle automorphisms of $\pi : Z \to S$ and acts as symplectomorphisms of M. Suppose further that the fibers of Z are compact and equipped with a density along the fiber which is invariant under the group action. (For example, we can put any density on Z_s varying smoothly on s and then replace this density by the one obtained by averaging over the group.) Finally suppose that the map G is equivariant for the group actions of K on Z and on M. Then the map $\tilde{G}$ can be chosen to be equivariant for the actions of K on Z and the induced action of K on $M \times T^*S$.

More generally we want to consider situations where a Lie group K acts on Z as fiber bundle automorphisms and on M and where we know by explicit construction that the map $\tilde{G}$ can be chosen to be equivariant .

Hamiltonian group actions.

Let us specialize further by assuming that S is a Lie group K and that $F : M \times K \to M$ is a Hamiltonian group action. So we have a map

$$G : M \times K \to M \times M^-, \qquad (m, a) \mapsto (m, am).$$

Let K act on $Z = M \times K$ via its left action on K so $a \in K$ acts on Z as

$$a : (m, b) \mapsto (m, ab).$$

We expect to be able to construct $\tilde{G} : M \times K \to T^*K$ so as to be equivariant for the action of K on $Z = M \times K$ and the induced action of K on T^*K.

To say that the action is Hamiltonian with moment map $\Psi : M \to \mathfrak{k}^*$ is to say that

$$i(\xi_M)\omega = -d\langle \Psi, \xi \rangle.$$

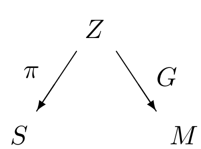
Thus under the left invariant identification of T^*K with $K \times \mathfrak{k}^*$ we see that Ψ determines a map

$$\Phi : M \times K \to T^*K, \qquad \Phi(m, a) = (a, \Psi(m)).$$

So our Φ of (4.34) is indeed a generalization of the moment map for Hamiltonian group actions.

4.12 Double fibrations.

The set-up described in § 4.11.2 has some legitimate applications of its own. For instance suppose that the diagram



is a double fibration: i.e., both π and G are fiber mappings and the map

$$(G, \pi) : Z \to M \times S$$

is an imbedding. In addition, suppose there exists a moment map $\Phi : Z \to T^*S$ such that

$$(G, \Phi) : Z \to M \times T^*S \tag{4.41}$$

is a Lagrangian imbedding. We will prove

Theorem 4.12.1. *The moment map $\Phi : Z \to T^*S$ is a co-isotropic immersion.*

Proof. We leave as an exercise the following linear algebra result:

Lemma 4.12.1. *Let V and W be symplectic vector spaces and Γ a Lagrangian subspace of $V \times W$. Suppose the projection of Λ into V is surjective. Then the projection of Γ into W is injective and its image is a co-isotropic subspace of W.*

To prove the theorem let Γ_Φ be the image of the imbedding (4.41). Then the projection, $\Gamma_\phi \to M$, is just the map, G; so by assumption it is a submersion. Hence by the lemma, the projection, $\Gamma_\Phi \to T^*S$, which is just the map, Φ, is a co-isotropic immersion.

The most interesting case of the theorem above is the case when Φ is an imbedding. Then its image, Σ, is a co-isotropic submanifold of T^*S and M is just the quotient of Σ by its null-foliation. This description of M gives one, in principle, a method for quantizing M as a Hilbert subspace of $L_2(S)$. (For examples of how this method works in practice, see [**?**].)

4.12.1 The moment image of a family of symplectomorphisms

As in §4.11.7 let M be a symplectic manifold and let $\{f_s\,,\ s \in S\}$ be an exact family of symplectomorphisms. Let

$$\Phi : M \times S \to T^*S$$

be the moment map associated with this family and let

$$\Gamma = \{(m, f_s(m))\,,\ \Phi(m,s)\,;\ (m,s) \in M \times S\} \tag{4.42}$$

be its moment Lagrangian. From the perspective of §4.4, Γ is a morphism or "map"

$$\Gamma : M^- \times M \Rightarrow T^*S$$

mapping the categorical "points" (Lagrangian submanifolds) of $M^- \times M$ into the categorical "points" (Lagrangian submanifolds) of T^*S. Let Λ_Φ be the image with respect to this "map" of the diagonal, Δ, in $M \times M$. In more prosaic terms this image is just the image with respect to Φ (in the usual sense) of the subset

$$X = \{(m,s) \in M \times S\,;\ f_s(m) = m\} \tag{4.43}$$

of $M \times S$. As we explained in §4.2 this image will be a Lagrangian submanifold of T^*S only if one imposes transversal or clean intersection hypotheses on Γ and Δ. More explicitly let

$$\rho : \Gamma \to M \times M \tag{4.44}$$

be the projection of Γ into $M \times M$. The the pre-image in Γ of Δ can be identified with the set (4.43), and if ρ intersects Δ cleanly, the set (4.43) is a submanifold of $M \times S$ and we know from Theorem 4.1.1 that:

Theorem 4.12.2. *The composition,*

$$\Phi \circ j : X \to T^*S\,, \tag{4.45}$$

of Φ *with the inclusion map,* j*, of* X *into* $M \times S$ *is a mapping of constant rank and its image,* Δ_Φ*, is an immersed Lagrangian submanifold of* T^*S*.*

Remarks.

1. If the projection (4.44) intersects Δ transversally one gets a stronger result, Namely in this case the map (4.45) is a Lagrangian immersion.

2. If the map (4.45) is proper and its level sets are simply connected, then Λ_Φ is an imbedded Lagrangian submanifold of T^*S, and (4.45) is a fiber bundle mapping with X as fiber and Λ_Φ as base.

Let's now describe what this "moment image", Λ_Φ, of the moment Lagrangian look like in some examples:

4.12.2 The character Lagrangian.

Let K be the standard n-dimensional torus and $\mathfrak{k}$ its Lie algebra. Given a Hamiltonian action, τ, of K on a compact symplectic manifold, M, one has its usual moment mapping, $\phi : M \to \mathfrak{k}^*$; and if K acts faithfully the image of ϕ is a convex n-dimensional polytope, $\mathbf{P}_\Phi$.

If we consider the moment map $\Phi : M \to T^*K = K \times \mathfrak{k}^*$ in the sense of §4.11.2, The image of Φ in the categorical sense can be viewed as a labeled polytope in which the open $(n-k)$-dimensional faces of $\mathbf{P}_\Phi$ are labeled by k-dimensional subgroups of K. More explicitly, since M is compact, there are a finite number of subgroups of K occurring as stabilizer groups of points. Let

$$K_\alpha\,, \quad \alpha = 1, \ldots, N \tag{4.46}$$

be a list of these subgroups and for each α let

$$M_{i,\alpha}\,, \quad i = 1, \ldots, k_\alpha \tag{4.47}$$

be the connected components of the set of points whose stabilizer group is K_α. Then the sets

$$\phi(M_{i,\alpha}) = \mathbf{P}_{i,\alpha} \tag{4.48}$$

in $\mathfrak{k}^*$ are the open faces of $\mathbf{P}$ and the categorical image, Λ_Φ, of the set of symplectomorphisms $\{\tau_a\,,\ a \in K\}$ is the disjoint union of the Lagrangian manifolds

$$\Lambda_{i,\alpha} = K_\alpha \times \mathbf{P}_{i,\alpha} \tag{4.49}$$

4.12.3 The period–energy relation.

If one replaces the group, $K = \mathbb{T}^n$ in this example by the non-compact group, $K = \mathbb{R}^n$ one can't expect Λ_Φ to have this kind of polyhedral structure; however, Λ_Φ does have some interesting properties from the dynamical systems perspective. If $H : M \to (\mathbb{R}^n)^*$ is the moment map associated with the action of $\mathbb{R}^n$ onto M, the coordinates, H_i, of H can be viewed as Poisson–commuting Hamiltonians, and the $\mathbb{R}^n$ action is generated by their Hamiltonian vector fields, ν_{H_i}, i.e., by the map

$$s \in \mathbb{R}^n \to f_s = (\exp s_1 \nu_{H_1}) \dots (\exp s_n \nu_{H_n})\,. \tag{4.50}$$

Suppose now that $H : M \to (\mathbb{R}^n)^*$ is a proper submersion. Then each connected component, Λ, of Λ_Φ in $T^*\mathbb{R}^n = \mathbb{R}^n \times (\mathbb{R}^n)^*$ is the graph of a map

$$H \to \left(\frac{\partial \psi}{\partial H_1}, \cdots, \frac{\partial \psi}{\partial H_n} \right)$$

over an open subset, U, of $(\mathbb{R}^n)^*$ with $\psi \in C^\infty(U)$, and, for $c \in U$, the element, $T = (T_1, \dots, T_n)$, $T_i = \frac{\partial \psi}{\partial H_i}(c)$, of $\mathbb{R}^n$ is the stabilizer of a connected component of periodic trajectories of the ν_{H_i}'s on the level set:

$$H_1 = c_1, \dots, H_n = c\,.$$

In particular all trajectories of ν_{H_i} have the same period, T_i, on this level set. This result is known in the theory of dynamical systems as the *period–energy relation*. In many examples of interest, the Legendre transform

$$\frac{\partial \psi}{\partial H} : U \to \mathbb{R}^n$$

is invertible, mapping U bijectively onto an open set, V, and in this case Λ is the graph of the "period mapping"

$$T \in V \to \frac{\partial \psi^*}{\partial T} \in (\mathbb{R}^n)^*$$

where ψ^* is the Legendre function dual to ψ.

4.12.4 The period–energy relation for families of symplectomorphisms.

We will show that something similar to this period–energy relation is true for families of symplectomorphisms providing we impose some rather strong assumptions on M and ω. Namely we will have to assume that ω is exact and that $H^1(M, \mathbb{R}) = 0$. Modulo these assumptions one can define, for a symplectomorphism, $f : M \to M$, and a fixed point, p of f, a natural notion of "the period of p".

The definition is the following. Choose a one-form, α, with $d\alpha = \omega$. Then

$$d(\alpha - f^*\alpha) = \omega - f^*\omega = 0$$

so

$$\alpha - f^*\alpha = d\psi \tag{4.51}$$

for some ψ in $C^\infty(M)$. (Unfortunately, ψ is only defined up to an additive constant, and one needs some "intrinsic" way of normalizing this constant. For instance, if ψ is bounded and M has finite volume one can require that the integral of ψ over M be zero, or if there is a natural base point, p_0, in M fixed by f, one can require that $\psi(p_0) = 0$.) Now, for every fixed point, p, set

$$T_p = \psi(p)\,. \tag{4.52}$$

This definition depends on the normalization we've made of the additive constant in the definition of ψ, but we claim that it's independent of the choice of α. In fact, if we replace α by $\alpha + dg$, $g \in C^\infty(M)$, ψ gets changed to $\psi + f^*g - g$ and at the fixed point, p,

$$\psi(p) + (f^*g - g)(p) = \psi(p)\,,$$

so the definition (4.42) doesn't depend on α.

There is also a dynamical systems method of defining these periods. By a variant of the mapping torus construction of Smale one can construct a contact manifold, W, which is topologically identical with the usual mapping torus of f, and on this manifold a contact flow having the following three properties.

1. M sits inside W and is a global cross-section of this flow.
2. f is the "first return" map.
3. If $f(p) = p$ the periodic trajectory of the flow through p has T_p as period.

Moreover, this contact manifold is unique up to contact isomorphism. (For details see [?] or [?].) Let's apply these remarks to the set-up we are considering in this paper. As above let $F : M \times S \to M$ be a smooth mapping such that for every s the map $f_s : M \to M$, mapping m to $F(m, s)$, is a symplectomorphism. Let us assume that

$$H^1(M \times S, \mathbb{R}) = 0.$$

Let π be the projection of $M \times S$ onto M. Then if α is a one-form on M satisfying $d\alpha = \omega$ and α_S is the canonical one-form on T^*S the moment map $\Phi : M \times S \to M$ associated with F has the defining property

$$\pi^*\alpha - F^*\alpha + \Phi^*\alpha_S = d\psi \tag{4.53}$$

for some ψ in $C^\infty(M \times S)$. Let's now restrict both sides of (4.53) to $M \times \{s\}$. Since Φ maps $M \times \{s\}$ into T_s^*, and the restriction of α_S to T_s^* is zero we get:

$$\alpha - f_s^*\alpha = d\psi_s \tag{4.54}$$

where $\psi_s = \psi_{|M\times\{s\}}$.

Next let X be the set, (4.43), i.e., the set:

$$\{(m,s) \in M \times S\,, \quad F(m,s) = m\}$$

and let's restrict (4.53) to X. If j is the inclusion map of X into $M \times S$, then $F \circ j = \pi$; so

$$j^*(\pi^*\alpha - F^*\alpha) = 0$$

and we get from (4.53)

$$j^*(\phi^*\alpha_S - d\psi) = 0\,. \tag{4.55}$$

The identities, (4.54) and (4.55) can be viewed as a generalization of the period–energy relation. For instance, suppose the map

$$\tilde{F} : M \times S \to M \times M$$

mapping (m,s) to $(m\,,\,F(m,s))$ is transversal to Δ. Then by Theorem 4.12.2 the map $\Phi \circ j : X \to T^*S$ is a Lagrangian immersion whose image is Λ_Φ. Since $\tilde{F}$ intersects Δ transversally, the map

$$\tilde{f}_s : M \to M \times M\,, \quad \tilde{f}_s(m) = (m, f_s(m))\,,$$

intersects Δ transversally for almost all s, and if M is compact, f_s is Lefschetz and has a finite number of fixed points, $p_i(s)$, $i = 1, \ldots, k$. The functions, $\psi_i(s) = \psi(p_i(s), s)$, are, by (4.54), the periods of these fixed points and by (4.55) the Lagrangian manifolds

$$\Lambda_{\psi_i} = \{(s,\xi) \in T^*S \quad \xi = d\psi_i(s)\}$$

are the connected components of Λ_Φ.

4.13 The category of exact symplectic manifolds and exact canonical relations.

4.13.1 Exact symplectic manifolds.

Let (M,ω) be a symplectic manifold. It is possible that the symplectic form ω is exact, that is, that $\omega = -d\alpha$ for some one form α. When this happens, we say that (M,α) is an **exact sympletic manifold**. In other words, an exact symplectic manifold is a pair consisting of a manifold M together with a one form α such that $\omega = -d\alpha$ is of maximal rank. The main examples for us, of course, are cotangent bundles with their canonical one forms. Observe that

Proposition 4.13.1. *No positive dimensional compact symplectic manifold can be exact.*

Indeed, if (M, ω) is a symplectic manifold with M compact, then

$$\int_M \omega^d > 0$$

where $2d = \dim M$ assuming that $d > 0$. But if $\omega = -d\alpha$ then

$$\omega^d = -d\left(\alpha \wedge \omega^{d-1}\right)$$

and so $\int_M \omega^d = 0$ by Stokes' theorem. $\square$

4.13.2 Exact Lagrangian submanifolds of an exact symplectic manifold.

Let (M, α) be an exact symplectic manifold and Λ a Lagrangian submanifold of (M, ω) where $\omega = -d\alpha$. Let

$$\beta_\Lambda := \iota_\Lambda^* \alpha \tag{4.56}$$

where

$$\iota_\Lambda : \Lambda \to M$$

is the embedding of Λ as a submanifold of M. So

$$d\beta_\Lambda = 0.$$

Suppose that β_Λ is exact, i.e. that $\beta_\Lambda = d\psi$ for some function ψ on Λ. (This will always be the case, for example, if Λ is simply connected.) We then call Λ an **exact** Lagrangian submanifold and ψ a choice of **phase function** for Λ.

Another important class of examples is where $\beta_\Lambda = 0$, in which case we can choose ψ to be locally constant. For instance, if $M = T^*X$ and $\Lambda = N^*(Y)$ is the conormal bundle to a submanifold $Y \subset X$ then we know that the restriction of α_X to $N^*(V)$ is 0.

4.13.3 The sub"category" of $\mathcal{S}$ whose objects are exact.

Consider the "category" whose objects are exact symplectic manifolds and whose morphisms are canonical relations between them. So let (M_1, α_1) and (M_2, α_2) be exact symplectic manifolds. Let

$$\mathrm{pr}_1 : M_1 \times M_2 \to M_1, \quad \mathrm{pr}_2 : M_1 \times M_2 \to M_2$$

be projections onto the first and second factors. Let

$$\alpha := -\,\mathrm{pr}_1^* \alpha_1 + \mathrm{pr}_2^* \alpha_2.$$

Then $-d\alpha$ gives the symplectic structure on $M_1^- \times M_2$.

To say that $\Gamma \in \mathrm{Morph}(M_1, M_2)$ is to say that Γ is a Lagrangian submanifold of $M_1^- \times M_2$. Let $\iota_\Gamma : \Gamma \to M_1^- \times M_2$ denote the inclusion map, and define, as above:

$$\beta_\Gamma := \iota_\Gamma^* \alpha.$$

We know that $d\beta_\Gamma = \iota^* d\alpha = 0$. So every canonical relation between cotangent bundles comes equipped with a closed one form.

Example: the canonical relation of a map.

Let $f : X_1 \to X_2$ be a smooth map and Γ_f the corresponding canonical relation from $M_1 = T^*X_1$ to $M_2 = T^*X_2$. By definition $\Gamma_f = (\varsigma_1 \times \mathrm{id})N^*(\mathrm{graph}(f))$ and we know that the canonical one form vanishes on any conormal bundle. Hence

$$\beta_{\Gamma_f} = 0.$$

So if Γ is a canonical relation coming from a smooth map, its associated one form vanishes. We want to consider an intermediate class of Γ's - those whose associated one forms are exact.

Before doing so, we must study the behavior of the β_Γ under composition.

4.13.4 Functorial behavior of β_Γ.

Let (M_i, α_i) $i = 1, 2, 3$ be exact symplectic manifolds and

$$\Gamma_1 \in \mathrm{Morph}(M_1, M_2), \quad \Gamma_2 \in \mathrm{Morph}(M_2, M_3)$$

be cleanly composible canonical relations. Recall that we defined

$$\Gamma_2 \star \Gamma_1 \subset \Gamma_1 \times \Gamma_2$$

to consist of all (m_1, m_2, m_2, m_3) and we have the fibration

$$\kappa : \Gamma_2 \star \Gamma_1 \to \Gamma_2 \circ \Gamma_2, \quad \kappa(m_1, m_2, m_2, m_3) = (m_1, m_3).$$

We also have the projections

$$\varrho_1 : \Gamma_2 \star \Gamma_1 \to \Gamma_1, \quad \varrho_1((m_1, m_2, m_2, m_3)) = (m_1, m_2)$$

and

$$\varrho_2 : \Gamma_2 \star \Gamma_1 \to \Gamma_2, \quad \varrho((m_1, m_2, m_2, m_3)) = (m_2, m_3).$$

We claim that

$$\kappa^* \beta_{\Gamma_2 \circ \Gamma_1} = \varrho_1^* \beta_{\Gamma_1} + \varrho_2^* \beta_{\Gamma_2}. \tag{4.57}$$

Proof. Let ρ_1 and π_1 denote the projections of Γ_1 onto M_1 and M_2, and let ρ_2 and π_2 denote the projections of Γ_2 onto M_2 and M_3, so that

$$\pi_1 \varrho_1 = \rho_2 \varrho_2$$

both maps sending (m_1, m_2, m_2, m_3) to m_2. So

$$\beta_{\Gamma_1} = -\rho_1^* \alpha_1 + \pi_1^* \alpha_2 \ \text{ and } \ \beta_{\Gamma_2} = -\rho_2^* \alpha_2 + \pi_2^* \alpha_3.$$

Thus

$$\varrho_1^* \beta_{\Gamma_1} + \varrho_2^* \beta_{\Gamma_2} = -\varrho_1^* \rho_1^* \alpha_1 + \varrho_2^* \pi_2^* \alpha_3 = \kappa^* \beta_{\Gamma_2 \circ \Gamma_1}. \quad \square$$

As a corollary we see that if $\beta_{\Gamma_i} = d\psi_i, \ \ i = 1, 2$ then

$$\kappa^* \beta_{\Gamma_2 \circ \Gamma_1} = d\left(\varrho_1^* \psi_1 + \varrho_2^* \psi_2\right).$$

So let us call a canonical relation **exact** if its associated (closed) one form is exact. We see that if we restrict ourselves to canonical relations which are exact, then we obtain a sub"category" of the "category" whose objects are exact symplectic manifolds and whose morphisms are exact canonical relations.

4.13.5 Defining the "category" of exact symplectic manifolds and canonical relations.

If Γ is an exact canonical relation so that $\beta_\Gamma = d\psi$, then ψ is only determined up to an additive constant (if Γ is connected). But we can *enhance* our sub"category" by specifying ψ. That is, we consider the "category" whose objects are exact symplectic manifolds and whose morphisms are pairs (Γ, ψ) where Γ is an exact canonical relation and $\beta_\Gamma = d\psi$. Then composition is defined as follows: If Γ_1 and Γ_2 are cleanly composible, then we define

$$(\Gamma_2, \psi_2) \circ (\Gamma_1, \psi_1) = (\Gamma_2 \circ \Gamma_1, \psi) \tag{4.58}$$

where the (local) additive constant in ψ is determined by

$$\kappa^*\psi = \varrho_1^*\psi_1 + \varrho_2^*\psi_2. \tag{4.59}$$

We shall call this enhanced sub"category" the "category" of **exact canonical relations**.

An important sub"category" of this "category" is where the objects are cotangent bundles with their canonical one forms.

The "category" of exact symplectic manifolds and conormal canonical relations.

As we saw above, the restriction of the canonical one form of a cotangent bundle to the conormal bundle of a submanifold of the base has the property that $\iota^*\alpha = 0$. So we can consider the subcategory of the "category" of exact symplectic manifolds and canonical relations by demanding that $\beta_\Gamma = 0$ and that $\psi = 0$. Of course, in this subcategory the ψ's occurring in (4.58) and (4.59) are all zero. We shall call this subcategory of the exact symplectic "category" the "category" of symplectic manifolds and **conormal** canonical relations. in honor of the conormal case.

The integral symplectic "category".

On the other hand in Chapter 12 we will make use of a slightly larger "category" than the "category" of exact symplectic manifolds and exact canonical relations. The objects in this larger "category" will still be exact symplectic manifolds (M, α). But a morphism between (M_1, α_1) and (M_2, α_2) will be a pair (Γ, f) where Γ is a Lagrangian submanifold of $M_1^- \times M_2$ and $f : \Gamma \to S^1$ is a C^∞ map satisfying

$$\iota_\Gamma^*\alpha = \frac{1}{2\pi i}\frac{df}{f}. \tag{4.60}$$

Here $\alpha = \pi_2^*\alpha_2 - \pi_1^*\alpha_1$ as before.

(Notice that if (Γ, ψ) is a morphism in the exact symplectic "category", then we get a morphism is this larger "category" by setting $f = e^{2\pi i\phi}$.) The condition (4.60) implies that $\iota_\Gamma^*\alpha$ defines an integral cohomology class which is the reason that we call this "category" the **integral symplectic "category"**. The composition law (generalizing the laws in (4.58) and (4.59)) is

$$(\Gamma_2, f_2) \circ (\Gamma_1, f_1) = (\Gamma, f)$$

where

$$\kappa^* f = (\rho_2^* f) \cdot (\rho_1^* f). \tag{4.61}$$

4.13.6 Pushforward via a map in the "category" of exact canonical relations between cotangent bundles.

As an illustration of the composition law (4.58) consider the case where Λ_Z is an exact Lagrangian submanifold of T^*Z so that the restriction of the one form of T^*Z to Λ is given by $d\psi_\Lambda$. We consider Λ as an element of $\mathrm{Morph}(\mathrm{pt.}, T^*Z)$ so we can take (Λ, ψ) as the (Γ_1, ψ_1) in (4.58). Let $f : Z \to X$ be a smooth map and take Γ_2 in (4.58) to be Γ_f. We know that the one form associated to Γ_f vanishes. In our enhanced category we must specify the function whose differential vanishes on Γ_f - that is we must pick a (local) constant c. So in (4.58) we have $(\Gamma_2, \psi_2) = (\Gamma_f, c)$. Assume that the Γ_f and Λ are composible. Recall that then $\Gamma_f \circ \Lambda_Z = df_*(\Lambda_Z)$ consists of all (x, ξ) where $x = f(z)$ and $(z, df^*(\xi)) \in \Lambda$. Then (4.58) says that

$$\psi(x, \xi) = \psi_\Lambda(z, \eta) + c. \tag{4.62}$$

In the next chapter and in Chapter 8 will be particularly interested in the case where f is a fibration. So we are given a fibration $\pi : Z \to X$ and we take $\Lambda_Z = \Lambda_\phi$ to be a horizontal Lagrangian submanifold of T^*Z. We will also assume that the composition in (4.58) is transversal. In this case the pushforward map $d\pi_*$ gives a diffeomorphism of Λ_ϕ with $\Lambda := df_*(\Lambda_\phi)$. In our applications, we will be given the pair (Λ, ψ) and we will regard (4.62) as *fixing the arbitrary constant* in ϕ rather than in Γ_f whose constant we take to be 0.

Chapter 5

Generating functions.

In this chapter we continue the study of canonical relations between cotangent bundles. We begin by studying the canonical relation associated to a map in the special case when this map is a fibration. This will allow us to generalize the local description of a Lagrangian submanifold of T^*X that we studied in Chapter 1. In Chapter 1 we showed that a *horizontal* Lagrangian submanifold of T^*X is locally described as the set of all $d\phi(x)$ where $\phi \in C^\infty(X)$ and we called such a function a "generating function". The purpose of this chapter is to generalize this concept by introducing the notion of a generating function relative to a fibration.

5.1 Fibrations.

In this section we will study in more detail the canonical relation associated to a fibration. So let X and Z be manifolds and

$$\pi : Z \to X$$

a smooth fibration. So (by equation (4.11))

$$\Gamma_\pi \in \mathrm{Morph}(T^*Z, T^*X)$$

consists of all $(z, \xi, x, \eta) \in T^*Z \times T^*X$ such that

$$x = \pi(z) \quad \text{and} \quad \xi = (d\pi_z)^*\eta.$$

Then

$$\mathrm{pr}_1 : \Gamma_\pi \to T^*Z, \quad (z, \xi, x, \eta) \mapsto (z, \xi)$$

maps Γ_π bijectively onto the sub-bundle of T^*Z consisting of those covectors which vanish on tangents to the fibers. We will call this sub-bundle the **horizontal sub-bundle** and denote it by H^*Z. So at each $z \in Z$, the fiber of the horizontal sub-bundle is

$$H^*(Z)_z = \{(d\pi_z)^*\eta,\ \eta \in T^*_{\pi(z)}X\}.$$

Let Λ_Z be a Lagrangian submanifold of T^*Z which we can also think of as an element of Morph(pt., T^*Z). We want to study the condition that Γ_π and Λ_Z be composable so that we be able to form

$$\Gamma_\pi(\Lambda_Z) = \Gamma_\pi \circ \Lambda_Z$$

which would then be a Lagrangian submanifold of T^*X. If $\iota : \Lambda_Z \to T^*Z$ denotes the inclusion map then the clean intersection part of the composibility condition requires that ι and pr_1 intersect cleanly. This is the same as saying that Λ_Z and H^*Z intersect cleanly in which case the intersection

$$F := \Lambda_Z \cap H^*Z$$

is a smooth manifold and we get a smooth map $\kappa : F \to T^*X$. The remaining hypotheses of Theorem 4.2.2 require that this map be proper and have connected and simply connected fibers.

A more restrictive condition is that intersection be transversal, i.e. that

$$\Lambda_Z \pitchfork H^*Z$$

in which case we always get a Lagrangian immersion

$$F \to T^*X, \quad (z, d\pi_z^*\eta) \mapsto (\pi(z), \eta).$$

The additional composibility condition is that this be an embedding.

Let us specialize further to the case where Λ_Z is a horizontal Lagrangian submanifold of T^*Z. That is, we assume that

$$\Lambda_Z = \Lambda_\phi = \gamma_\phi(Z) = \{(z, d\phi(z))\}$$

as in Chapter 1. When is

$$\Lambda_\phi \pitchfork H^*Z?$$

Now H^*Z is a sub-bundle of T^*Z so we have the exact sequence of vector bundles

$$0 \to H^*Z \to T^*Z \to V^*Z \to 0 \tag{5.1}$$

where

$$(V^*Z)_z = T_z^*Z/(H^*Z)_z = T_z^*(\pi^{-1}(x)), \quad x = \pi(z)$$

is the cotangent space to the fiber through z.

Any section $d\phi$ of T^*Z gives a section $d_{vert}\phi$ of V^*Z by the above exact sequence, and $\Lambda_\phi \pitchfork H^*Z$ if and only if this section intersects the zero section of V^*Z transversally. If this happens,

$$C_\phi := \{z \in Z | (d_{vert}\phi)_z = 0\}$$

is a submanifold of Z whose dimension is $\dim X$. Furthermore, at any $z \in C_\phi$

$$d\phi_z = (d\pi_z)^*\eta \quad \text{for a unique} \quad \eta \in T^*_{\pi(z)}X.$$

Thus Λ_ϕ and Γ_π are transversally composable if and only if

$$C_\phi \to T^*X, \quad z \mapsto (\pi(z), \eta)$$

is a Lagrangian embedding in which case its image is a Lagrangian submanifold

$$\Lambda = \Gamma_\pi(\Lambda_\phi) = \Gamma_\pi \circ \Lambda_\phi$$

of T^*X. When this happens we say that ϕ is a **transverse generating function of Λ with respect to the fibration** (Z, π).

If Λ_ϕ and Γ_π are merely cleanly composable, we say that ϕ is a **clean** generating function with respect to π.

If ϕ is a transverse generating function for Λ with respect to the fibration, π, and $\pi_1 : Z_1 \to Z$ is a fibration over Z, then it is easy to see that $\phi_1 = \pi_1^*\phi$ is a clean generating function for Λ with respect to the fibration, $\pi \circ \pi_1$; and we will show in the next section that there is a converse result: Locally, *every* clean generating function can be obtained in this way from a transverse generating function. For this reason it will suffice, for many of the things we'll be doing in this chapter, to work with transverse generating functions; and to simplify notation, we will henceforth,in this chapter, unless otherwise stated, use the terms "generating function" and "transverse generating function" interchangeably.

However, in the applications in Chapter 9, we will definitely need to use clean generating functions.

5.1.1 Transverse vs. clean generating functions.

Locally, we can assume that Z is the product, $X \times S$, of X with an open subset, S, of $\mathbb{R}^k$ with standard coordinates $s_1, \dots, s_k$. Then H^*Z is defined by the equations, $\eta_1 = \cdots = \eta_k = 0$, where the η_i's are the standard cotangent coordinates on T^*S; so $\Lambda_\phi \cap H^*Z$ is defined by the equations

$$\frac{\partial \phi}{\partial s_i} = 0\,, \quad i = 1, \dots, k\,.$$

Let C_ϕ be the subset of $X \times S$ defined by these equations. Then if Λ_ϕ intersects H^*Z cleanly, C_ϕ is a submanifold of $X \times S$ of codimension $r \leq k$; and, at every point $(x_0, s_0) \in C_\phi$, C_ϕ can be defined locally near (x_0, s_0) by r of these equations, i.e., modulo repagination, by the equations

$$\frac{\partial \phi}{\partial s_i} = 0\,, \quad i = 1, \dots, r\,.$$

Moreover these equations have to be independent: the tangent space at (x_0, s_0) to C_ϕ has to be defined by the equations

$$d\left(\frac{\partial \phi}{\partial s_i}\right)_{(x_0, \xi_0)} = 0\,, \quad i = 1, \dots, r\,.$$

Suppose $r < k$ (i.e., suppose this clean intersection is not transverse). Since $\partial\phi/\partial s_k$ vanishes on C_ϕ, there exist C^∞ functions, $g_i \in C^\infty(X \times S)$, $i = 1, \ldots, r$ such that

$$\frac{\partial \phi}{\partial s_k} = \sum_{i=1}^{r} g_i \frac{\partial \phi}{\partial s_i}\,.$$

In other words, if ν is the vertical vector field

$$\nu = \frac{\partial}{\partial s_k} - \sum_{i=1}^{r} g_i(x,s) \frac{\partial}{\partial s_i}$$

then $D_\nu \phi = 0$. Therefore if we make a change of vertical coordinates

$$(s_i)_{\text{new}} = (s_i)_{\text{new}}(x,s)$$

so that in these new coordinates

$$\nu = \frac{\partial}{\partial s_k}$$

this equation reduces to

$$\frac{\partial}{\partial s_k}\phi(x,s) = 0\,,$$

so, in these new coordinates,

$$\phi(x,s) = \phi(x, s_1, \ldots, s_{k-1})\,.$$

Iterating this argument we can reduce the number of vertical coordinates so that $k = r$, i.e., so that ϕ is a transverse generating function in these new coordinates. In other words, a clean generating function is just a transverse generating function to which a certain number of vertical "ghost variables" ("ghost" meaning that the function doesn't depend on these variables) have been added. The number of these ghost variables is called the **excess** of the generating function. (Thus for the generating function in the paragraph above, its excess is $k - r$.) More intrinsically the *excess is the difference between the dimension of the critical set C_ϕ of ϕ and the dimension of X.*

As mentioned above, unless specified otherwise, we assume in this Chapter that our generating function are transverse generating functions.

5.2 The generating function in local coordinates.

Suppose that X is an open subset of $\mathbb{R}^n$, that

$$Z = X \times \mathbb{R}^k$$

that π is projection onto the first factor, and that (x,s) are coordinates on Z so that $\phi = \phi(x,s)$. Then $C_\phi \subset Z$ is defined by the k equations

$$\frac{\partial \phi}{\partial s_i} = 0, \qquad i = 1, \ldots, k.$$

and the transversality condition is that these equations be functionally independent. This amounts to the hypothesis that their differentials

$$d\left(\frac{\partial\phi}{\partial s_i}\right) \qquad i = 1, \ldots k$$

be linearly independent. Then $\Lambda \subset T^*X$ is the image of the embedding

$$C_\phi \to T^*X, \quad (x, s) \mapsto \frac{\partial\phi}{\partial x} = d_X\phi(x, s).$$

5.3 Example - a generating function for a conormal bundle.

Suppose that

$$Y \subset X$$

is a submanifold defined by the k functionally independent equations

$$f_1(x) = \cdots = f_k(x) = 0.$$

Let $\phi : X \times \mathbb{R}^k \to \mathbb{R}$ be the function

$$\phi(x, s) := \sum_i f_i(x) s_i. \tag{5.2}$$

We claim that

$$\Lambda = \Gamma_\pi \circ \Lambda_\phi = N^*Y, \tag{5.3}$$

the conormal bundle of Y. Indeed,

$$\frac{\partial\phi}{\partial s_i} = f_i$$

so

$$C_\phi = Y \times \mathbb{R}^k$$

and the map

$$C_\phi \to T^*X$$

is given by

$$(x, s) \mapsto \sum s_i d_X f_i(x).$$

The differentials $d_X f_i$ span the conormal bundle to Y at each $x \in Y$ proving (5.3).

As a special case of this example, suppose that

$$X = \mathbb{R}^n \times \mathbb{R}^n$$

and that Y is the diagonal

$$\operatorname{diag}(X) = \{(x, x)\} \subset X$$

which may be described as the set of all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ satisfying

$$x_i - y_i = 0, \quad i = 1, \dots, n.$$

We may then choose

$$\phi(x, y, s) = \sum_i (x_i - y_i) s_i. \tag{5.4}$$

Now $\mathrm{diag}(X)$ is just the graph of the identity transformation so by Section 4.8 we know that $(\varsigma_1 \times \mathrm{id})(N^*(\mathrm{diag}(X))$ is the canonical relation giving the identity map on T^*X. By abuse of language we can speak of ϕ as the generating function of the identity canonical relation. (But we must remember the ς_1.)

5.4 Example. The generating function of a geodesic flow.

A special case of our generating functions with respect to a fibration is when the fibration is trivial, i.e. π is a diffeomorphism. Then the vertical bundle is trivial and we have no "auxiliary variables". Such a generating function is just a generating function in the sense of Chapter 1. For example, let X be a Riemannian manifold and let $\phi_t \in C^\infty(X \times X)$ be defined by

$$\phi_t(x, y) := \frac{1}{2t} d(x, y)^2, \tag{5.5}$$

where

$$t \neq 0.$$

Let us compute Λ_ϕ and $(\varsigma_1 \times \mathrm{id})(\Lambda_\phi)$. We first do this computation under the assumption that $X = \mathbb{R}^n$ and the metric occurring in (5.5) is the Euclidean metric so that

$$\begin{aligned}
\phi(x, y, t) &= \frac{1}{2t} \sum_i (x_i - y_i)^2 \\
\frac{\partial \phi}{\partial x_i} &= \frac{1}{t}(x_i - y_i) \\
\frac{\partial \phi}{\partial y_i} &= \frac{1}{t}(y_i - x_i) \quad \text{so} \\
\Lambda_\phi &= \{(x, \frac{1}{t}(x - y), y, \frac{1}{t}(y - x)\} \text{ and} \\
(\varsigma_1 \times \mathrm{id})(\Lambda_\phi) &= \{(x, \frac{1}{t}(y - x), y, \frac{1}{t}(y - x)\}.
\end{aligned}$$

In this last equation let us set $y - x = t\xi$, i.e.

$$\xi = \frac{1}{t}(y - x)$$

which is possible since $t \neq 0$. Then

$$(\varsigma_1 \times \mathrm{id})(\Lambda_\phi) = \{(x, \xi, x + t\xi, \xi)\}$$

which is the graph of the symplectic map

$$(x, \xi) \mapsto (x + t\xi, \xi).$$

If we identify cotangent vectors with tangent vectors (using the Eulidean metric) then $x + t\xi$ is the point along the line passing through x with tangent vector ξ a distance $t\|\xi\|$ out. The one parameter family of maps $(x, \xi) \mapsto (x + t\xi, \xi)$ is known as the geodesic flow. In the case of Euclidean space, the time t value of this flow is a diffeomorphism of T^*X with itself for every t. So long as $t \neq 0$ it has the generating function given by (5.5) with no need of auxiliary variables. When $t = 0$ the map is the identity and we need to introduce a fibration.

More generally, this same computation works on any "geodesically convex" Riemannian manifold, where:

A Riemannian manifold X is called **geodesically convex** if, given any two points x and y in X, there is a unique geodesic which joins them. We will show that the above computation of the generating function works for any geodesically convex Riemannian manifold. In fact, we will prove a more general result. Recall that geodesics on a Riemannian manifold can be described as follows: A Riemann metric on a manifold X is the same as a scalar product on each tangent space T_xX which varies smoothly with X. This induces an identification of TX with T^*X and hence a scalar product $\langle\ ,\ \rangle_x$ on each T^*X. This in turn induces the "kinetic energy" Hamiltonian

$$H(x, \xi) := \frac{1}{2}\langle \xi, \xi \rangle_x.$$

The principle of least action says that the solution curves of the corresponding vector field v_H project under $\pi : T^*X \to X$ to geodesics of X and every geodesic is the projection of such a trajectory.

An important property of the kinetic energy Hamiltonian is that it is quadratic of degree two in the fiber variables. We will prove a theorem (see Theorem 5.4.1 below) which generalizes the above computation and is valid for any Hamiltonian which is homogeneous of degree $k \neq 1$ in the fiber variables and which satisfies a condition analogous to the geodesic convexity theorem. We first recall some facts about homogeneous functions and Euler's theorem.

Consider the one parameter group of dilatations $t \mapsto \mathfrak{d}(t)$ on any cotangent bundle T^*X:

$$\mathfrak{d}(t) : T^*X \to T^*X : \qquad (x, \xi) \mapsto (x, e^t\xi).$$

A function f is homogenous of degree k in the fiber variables if and only if

$$\mathfrak{d}(t)^* f = e^{kt} f.$$

For example, the principal symbol of a k-th order linear partial differential operator on X is a function on T^*X with which is a polynomial in the fiber variables and is homogenous of degree k.

Let $\mathcal{E}$ denote the vector field which is the infinitesimal generator of the one parameter group of dilatations. It is called the **Euler vector field**. Euler's theorem (which is a direct computation from the preceding equation) says that f is homogenous of degree k if and only if

$$\mathcal{E}f = kf.$$

Let $\alpha = \alpha_X$ be the canonical one form on T^*X. From its very definition (1.8) it follows that

$$\mathfrak{d}(t)^*\alpha = e^t\alpha$$

and hence that

$$D_{\mathcal{E}}\alpha = \alpha.$$

Since $\mathcal{E}$ is everywhere tangent to the fiber, it also follows from (1.8) that

$$i(\mathcal{E})\alpha = 0$$

and hence that

$$\alpha = D_{\mathcal{E}}\alpha = i(\mathcal{E})d\alpha = -i(\mathcal{E})\omega$$

where $\omega = \omega_X = -d\alpha$.

Now let H be a function on T^*X which is homogeneous of degree k in the fiber variables. Then

$$\begin{aligned}
kH = \mathcal{E}H &= i(\mathcal{E})dH \\
&= i(\mathcal{E})i(v_H)\omega \\
&= -i(v_H)i(\mathcal{E})\omega \\
&= i(v_H)\alpha \quad \text{and} \\
(\exp v_H)^*\alpha - \alpha &= \int_0^1 \frac{d}{dt}(\exp tv_H)^*\alpha dt \quad \text{with} \\
\frac{d}{dt}(\exp tv_H)^*\alpha &= (\exp tv_H)^*\left(i(v_H)d\alpha + di(v_H)\alpha\right) \\
&= (\exp tv_H)^*\left(-i(v_H)\omega + di(v_H)\alpha\right) \\
&= (\exp tv_H)^*\left(-dH + kdH\right) \\
&= (k-1)(\exp tv_H)^*dH \\
&= (k-1)d(\exp tv_H)^*H \\
&= (k-1)dH
\end{aligned}$$

since H is constant along the trajectories of v_H. So

$$(\exp v_H)^*\alpha - \alpha = (k-1)dH. \tag{5.6}$$

Remark. In the above calculation we assumed that H was smooth on all of T^*X including the zero section, effectively implying that H is a polynomial in the fiber variables. But the same argument will go through (if $k > 0$) if all we assume is that H (and hence v_H) are defined on $T^*X\setminus$ the zero section, in which case H can be a more general homogeneous function on $T^*X\setminus$ the zero section.

Now $\exp v_H\ : T^*X \to T^*X$ is symplectic map. Let

$$\Gamma := \text{graph }(\exp v_H),$$

so $\Gamma \subset T^*X^- \times T^*X$ is a Lagrangian submanifold. Suppose that the projection $\pi_{X\times X}$ of Γ onto $X \times X$ is a diffeomorphism, i.e. suppose that Γ is horizontal. This says precisely that for every $(x,y) \in X \times X$ there is a unique $\xi \in T^*_xX$ such that

$$\pi \exp v_H(x,\xi) = y.$$

In the case of the geodesic flow, this is guaranteed by the condition of geodesic convexity.

Since Γ is horizontal, it has a generating function ϕ such that

$$d\phi = \text{pr}_2^*\,\alpha - \text{pr}_1^*\,\alpha$$

where $\text{pr}_i,\ i = 1,2$ are the projections of $T^*(X \times X) = T^*X \times T^*X$ onto the first and second factors. On the other hand pr_1 is a diffeomorphism of Γ onto T^*X. So

$$\text{pr}_1 \circ (\pi_{X\times X|\Gamma})^{-1}$$

is a diffeomorphism of $X \times X$ with T^*X.

Theorem 5.4.1. *Assume the above hypotheses. Then up to an additive constant we have*

$$\left(\text{pr}_1 \circ (\pi_{X\times X|\Gamma})^{-1}\right)^* [(k-1)H] = \phi.$$

In the case where $H = \frac{1}{2}\|\xi\|^2$ is the kinetic energy of a geodesically convex Riemann manifold, this says that

$$\phi = \frac{1}{2}d(x,y)^2.$$

Indeed, this follows immediately from (5.6). An immediate corollary (by rescaling) is that (5.5) is the generating function for the time t flow on a geodesically convex Riemannian manifold.

As mentioned in the above remark, the same theorem will hold if H is only defined on $T^*X \setminus \{0\}$ and the same hypotheses hold with $X \times X$ replaced by $X \times X \setminus \Delta$.

5.5 The generating function for the transpose.

Let

$$\Gamma \in \text{Morph}(T^*X, T^*Y)$$

be a canonical relation, let

$$\pi : Z \to X \times Y$$

be a fibration and ϕ a generating function for Γ relative to this fibration. In local coordinates this says that $Z = X \times Y \times S$, that

$$C_\phi = \{(x, y, s) | \frac{\partial \phi}{\partial s} = 0\},$$

and that Γ is the image of C_ϕ under the map

$$(x, y, s) \mapsto (-d_X\phi, d_Y\phi).$$

Recall that

$$\Gamma^\dagger \in \mathrm{Morph}(T^*Y, T^*X)$$

is given by the set of all (γ_2, γ_1) such that $(\gamma_1, \gamma_2) \in \Gamma$. So if

$$\kappa : X \times Y \to Y \times X$$

denotes the transposition

$$\kappa(x, y) = (y, x)$$

then

$$\kappa \circ \pi : Z \to Y \times X$$

is a fibration and $-\phi$ is a generating function for $\Gamma^\dagger$ relative to $\kappa \circ \pi$. Put more succinctly, if $\phi(x, y, s)$ is a generating function for Γ then

$$\psi(y, x, s) = -\phi(x, y, s) \text{ is a generating function for } \Gamma^\dagger. \tag{5.7}$$

For example, if Γ is the graph of a symplectomorphism, then $\Gamma^\dagger$ is the graph of the inverse diffeomorphism. So (5.7) says that $-\phi(y, x, s)$ generates the inverse of the symplectomorphism generated by $\phi(x, y, s)$.

This suggests that there should be a simple formula which gives a generating function for the composition of two canonical relations in terms of the generating function of each. This was one of Hamilton's great achievements - that, in a suitable sense to be described in the next section - the generating function for the composition is the sum of the individual generating functions.

5.6 The generating function for a transverse composition.

Let X_1, X_2 and X_3 be manifolds and

$$\Gamma_1 \in \mathrm{Morph}(T^*X_1, T^*X_2), \quad \Gamma_2 \in \mathrm{Morph}(T^*X_2, T^*X_3)$$

be canonical relations which are transversally composable. So we are assuming in particular that the maps

$$\Gamma_1 \to T^*X_2, \quad (p_1, p_2) \mapsto p_2 \quad \text{and} \quad \Gamma_2 \to T^*X_2, \quad (q_2, q_3) \mapsto q_2$$

are transverse.

Suppose that

$$\pi_1 : Z_1 \to X_1 \times X_2, \quad \pi_2 : Z_2 \to X_2 \times X_3$$

are fibrations and that $\phi_i \in C^\infty(Z_i),\ i = 1, 2$ are generating functions for Γ_i with respect to π_i.

From π_1 and π_2 we get a map

$$\pi_1 \times \pi_2 : Z_1 \times Z_2 \to X_1 \times X_2 \times X_2 \times X_3.$$

Let

$$\Delta_2 \subset X_2 \times X_2$$

be the diagonal and let

$$Z := (\pi_1 \times \pi_2)^{-1}(X_1 \times \Delta_2 \times X_3).$$

Finally, let

$$\pi : Z \to X_1 \times X_3$$

be the fibration

$$Z \to Z_1 \times Z_2 \to X_1 \times X_2 \times X_2 \times X_3 \to X_1 \times X_3$$

where the first map is the inclusion map and the last map is projection onto the first and last components. Let

$$\phi : Z \to \mathbb{R}$$

be the restriction to Z of the function

$$(z_1, z_2) \mapsto \phi_1(z_1) + \phi_2(z_2). \tag{5.8}$$

Theorem 5.6.1. *ϕ is a generating function for $\Gamma_2 \circ \Gamma_1$ with respect to the fibration $\pi : Z \to X_1 \times X_3$.*

Proof. We may check this in local coordinates where the fibrations are trivial to that

$$Z_1 = X_1 \times X_2 \times S, \qquad Z_2 = X_2 \times X_3 \times T$$

so

$$Z = X_1 \times X_3 \times (X_2 \times S \times T)$$

and π is the projection of Z onto $X_1 \times X_3$. Notice that X_2 has now become a factor in the *parameter space*. The function ϕ is given by

$$\phi(x_1, x_3, x_2, s, t) = \phi_1(x_1, x_2, s) + \phi_2(x_2, x_3, t).$$

For $z = (x_1, x_3, x_2, s, t)$ to belong to C_ϕ the following three conditions must be satisfied and be functionally independent:

- $\frac{\partial\phi_1}{\partial s}(x_1,x_2,s)=0$, i.e. $z_1=(x_1,x_2,s)\in C_{\phi_1}$.
- $\frac{\partial\phi_2}{\partial t}(x_2,x_3,t)=0$, i.e. $z_2=(x_2,x_3,t)\in C_{\phi_2}$ and
- $$\frac{\partial\phi_1}{\partial x_2}(x_1,x_2,s)+\frac{\partial\phi_2}{\partial x_2}(x_2,x_3,t)=0.$$

To show that these equations are functionally independent, we will rewrite them as the following system of equations on $X_1\times X_3\times X_2\times X_2\times S\times T$:

1. $\frac{\partial\phi_1}{\partial s}(x_1,x_2,s)=0$, i.e. $z_1=(x_1,x_2,s)\in C_{\phi_1}$,
2. $\frac{\partial\phi_2}{\partial t}(y_2,x_3,t)=0$, i.e. $z_2=(y_2,x_3,t)\in C_{\phi_2}$,
3. $x_2=y_2$ and
4. $$\frac{\partial\phi_1}{\partial x_2}(x_1,x_2,s)+\frac{\partial\phi_2}{\partial x_2}(y_2,x_3,t)=0.$$

It is clear that 1) and 2) are independent, and define the product $C_{\phi_1}\times C_{\phi_2}$ as a submanifold of $X_1\times X_3\times X_2\times X_2\times S\times T$. So to show that 1)-4) are independent, we must show that 3) and 4) are an independent system of equations on $C_{\phi_1}\times C_{\phi_2}$.

From the fact that ϕ_1 is a generating function for Γ_1, we know that the map

$$\gamma_1:\ C_{\phi_1}\to\Gamma_1,\qquad \gamma_1(p_1)=\left(x_1,-\frac{\partial\phi_1}{\partial x_1}(p_1),x_2,\frac{\partial\phi_1}{\partial x_2}(p_1)\right)$$

where

$$(x_1,x_2)=\pi_1(p_1)$$

is a diffeomorphism. Similarly, the map

$$\gamma_2:\ C_{\phi_1}\to\Gamma_2,\qquad \gamma_2(p_2)=\left(x_2,-\frac{\partial\phi_2}{\partial x_2}(p_2),x_3,\frac{\partial\phi_2}{\partial x_3}(p_2)\right)$$

where

$$(x_2,x_3)=\pi_2(p_2)$$

is a diffeomorphism.

So if we set $M_i:=T^*X_i,\ i=1,2,3$ we can write the preceding diffeomorphisms as

$$\gamma_i(p_i)=(m_i,m_{i+1}),i=1,2$$

where

$$m_i=(x_i,-\frac{\partial\phi_i}{\partial x_i}(p_i)),\qquad m_{i+1}=(x_{i+1},\frac{\partial\phi_i}{\partial x_{i+1}}(p_i)) \tag{5.9}$$

and the x_i are as above. We have the diffeomorphism

$$\gamma_1\times\gamma_2:\ C_{\phi_1}\times C_{\phi_2}\to\Gamma_1\times\Gamma_2$$

and the map

$$\kappa : \Gamma_1 \times \Gamma_2 \to M_2 \times M_2, \quad \kappa(m_1, m_2, n_2, m_3) = (m_2, n_2).$$

This map κ is assumed to be transverse to the diagonal Δ_{M_2}, and hence the map

$$\lambda : C_{\phi_1} \times C_{\phi_2} \to M_2 \times M_2, \quad \lambda := \kappa \circ (\gamma_1 \times \gamma_2)$$

is transverse to Δ_{M_2}. This transversality is precisely the functional independence of conditions 3) and 4) above.

The manifold $\Gamma_2 \star \Gamma_1$ was defined to be $\kappa^{-1}(\Delta_{M_2})$ and the second condition for transverse composibility was that the map

$$\rho : \Gamma_2 \star \Gamma_1 \to M_1^- \times M_3, \quad \rho(m_1.m_2, m_2, m_3) = (m_1, m_3)$$

be an embedding whose image is then defined to be $\Gamma_2 \circ \Gamma_1$. The diffeomorphism $\gamma_1 \times \gamma_2$ then shows that the critical set C_ϕ is mapped diffeomorphically onto $\Gamma_2 \star \Gamma_1$. Here ϕ is defined by (5.8). Call this diffeomorphism τ. So

$$\tau : C_\phi \cong \Gamma_2 \star \Gamma_1.$$

Thus

$$\rho \circ \tau : \ C_\phi \to \Gamma_2 \circ \Gamma_1$$

is a diffeomorphism, and (5.9) shows that this diffeomorphism is precisely the one that makes ϕ a generating function for $\Gamma_2 \circ \Gamma_1$. □

In the next section we will show that the arguments given above apply, essentially without change, to *clean* composition, yielding a clean generating function for the composite.

5.7 Generating functions for clean composition of canonical relations between cotangent bundles.

Suppose that the canonical relation, Γ_1 and Γ_2 are *cleanly* composable. Let $\phi_1 \in C^\infty(X_1 \times X_2 \times S)$ and $\phi_2 \in C^\infty(X_2 \times X_3 \times T)$ be *transverse* generating functions for Γ_1 and Γ_2 and as above let

$$\phi(x_1, x_3, x_2, s, t) = \phi_1(x_1, x_2, s) + \phi_2(x_2, x_3, t)\,.$$

We will prove below that ϕ is a clean generating function for $\Gamma_2 \circ \Gamma_1$ with respect to the fibration

$$X_1 \times X_3 \times (X_2 \times S \times T) \to X_1 \times X_3\,.$$

The argument is similar to that above: As above C_ϕ is defined by the three sets of equations:

1. $\frac{\partial \phi_1}{\partial s} = 0$

2. $\frac{\partial \phi_2}{\partial t} = 0$

3. $\frac{\partial \phi_1}{\partial x_2} + \frac{\partial \phi_2}{\partial x_2} = 0$.

Since ϕ_1 and ϕ_2 are transverse generating functions the equations 1 and 2 are an independent set of defining equations for $C_{\phi_1} \times C_{\phi_2}$. As for the equation 3, our assumption that Γ_1 and Γ_2 compose cleanly tells us that the mappings

$$\frac{\partial \phi_1}{\partial x_2} : C_{\phi_1} \to T^* X_2$$

and

$$\frac{\partial \phi_2}{\partial x_2} : C_{\phi_2} \to T^* X_2$$

intersect cleanly. In other words the subset, C_ϕ, of $C_{\phi_1} \times C_{\phi_2}$ defined by the equation $\frac{\partial \phi}{\partial x_2} = 0$, is a submanifold of $C_{\phi_1} \times C_{\phi_2}$, and its tangent space at each point is defined by the linear equation, $d\frac{\partial \phi}{\partial x_2} = 0$. Thus the set of equations, 1–3, are a *clean* set of defining equations for C_ϕ as a submanifold of $X_1 \times X_3 \times (X_2 \times S \times T)$. In other words ϕ is a clean generating function for $\Gamma_2 \circ \Gamma_1$.

The *excess*, ϵ, of this generating function is equal to the dimension of C_ϕ minus the dimension of $X_1 \times X_3$. One also gets a more intrinsic description of ϵ in terms of the projections of Γ_1 and Γ_2 onto $T^* X_2$. From these projections one gets a map

$$\Gamma_1 \times \Gamma_2 \to T^*(X_2 \times X_2)$$

which, by the cleanness assumption, intersects the conormal bundle of the diagonal cleanly; so its pre-image is a submanifold, $\Gamma_2 \star \Gamma_1$, of $\Gamma_1 \times \Gamma_2$. It's easy to see that

$$\epsilon = \dim \Gamma_2 \star \Gamma_1 - \dim \Gamma_2 \circ \Gamma_1 .$$

5.8 Reducing the number of fiber variables.

Let $\Lambda \subset T^* X$ be a Lagrangian manifold and let $\phi \in C^\infty(Z)$ be a generating function for Λ relative to a fibration $\pi : Z \to X$. Let

$$x_0 \in X,$$

let

$$Z_0 := \pi^{-1}(x_0),$$

and let

$$\iota_0 : Z_0 \to Z$$

be the inclusion of the fiber Z_0 into Z. By definition, a point $z_0 \in Z_0$ belongs to C_ϕ if and only if z_0 is a critical point of the restriction $\iota_0^* \phi$ of ϕ to Z_0.

Theorem 5.8.1. *If z_0 is a non-degenerate critical point of $\iota_0^*\phi$ then Λ is horizontal at*

$$p_0 = (x_0, \xi_0) = \frac{\partial \phi}{\partial x}(z_0).$$

Moreover, there exists an neighborhood U of x_0 in X and a function $\psi \in C^\infty(U)$ such that

$$\Lambda = \Lambda_\psi$$

on a neighborhood of p_0 and

$$\pi^*\psi = \phi$$

on a neighborhood U' of z_0 in C_ϕ.

Proof. (In local coordinates.) So $Z = X \times \mathbb{R}^k$, $\phi = \phi(x, s)$ and C_ϕ is defined by the k independent equations

$$\frac{\partial \phi}{\partial s_i} = 0, \qquad i = 1, \dots k. \tag{5.10}$$

Let $z_0 = (x_0, s_0)$ so that s_0 is a non-degenerate critical point of $\iota_0^*\phi$ which is the function

$$s \mapsto \phi(x_0, s)$$

if and only if the Hessian matrix

$$\left(\frac{\partial^2 \phi}{\partial s_i \partial s_j} \right)$$

is of rank k at s_0. By the implicit function theorem we can solve equations (5.10) for s in terms of x near (x_0, s_0). This says that we can find a neighborhood U of x_0 in X and a C^∞ map

$$g : U \to \mathbb{R}^k$$

such that

$$g(x) = s \Leftrightarrow \frac{\partial \phi}{\partial s_i} = 0,\ i = 1, \dots, k$$

if (x, s) is in a neighborhood of (x_0, s_0) in Z. So the map

$$\gamma : U \to U \times \mathbb{R}^k, \qquad \gamma(x) = (x, g(x))$$

maps U diffeomorphically onto a neighborhood of (x_0, s_0) in C_ϕ. Consider the commutative diagram

$$\begin{array}{ccc} U & \xrightarrow{\gamma} & C_\phi \\ \big\downarrow & & \big\downarrow{\scriptstyle d_X\phi} \\ X & \xleftarrow[\pi_X]{} & \Lambda \end{array}$$

where the left vertical arrow is inclusion and π_X is the restriction to Λ of the projection $T^*X \to X$. From this diagram it is clear that the restriction of π to the image of U in C_ϕ is a diffeomorphism and that Λ is horizontal at p_0. Also

$$\mu := d_X\phi \circ \gamma$$

is a section of Λ over U. Let

$$\psi := \gamma^*\phi.$$

Then

$$\mu = d_X\phi\circ\gamma = d_X\phi\circ\gamma + d_S\phi\circ\gamma = d\phi\circ\gamma$$

since $d_S\phi\circ\gamma\equiv 0$. Also, if $v\in T_xX$ for $x\in U$, then

$$d\psi_x(v) = d\phi_{\gamma(x)}(d\gamma_x(v)) = d\phi_{\gamma(x)}(v, dg_x(v))$$

$$= (d_X\phi)_{\gamma(x)}(v) = (d_X\phi\circ\gamma)(x)(v)$$

so

$$\langle\mu(x), v\rangle = \langle d\psi_x, v\rangle$$

so

$$\Lambda = \Lambda_\psi$$

over U and from $\pi : Z\to X$ and $\gamma\circ\pi = \mathrm{id}$ on $\gamma(U)\subset C_\phi$ we have

$$\pi^*\psi = \pi^*\gamma^*\phi = (\gamma\circ\pi)^*\phi = \phi$$

on $\gamma(U)$. □

We can apply the proof of this theorem to the following situation: Suppose that the fibration

$$\pi : Z\to X$$

can be factored as a succession of fibrations

$$\pi = \pi_1\circ\pi_0$$

where

$$\pi_0 : Z\to Z_1 \quad\text{and}\quad \pi_1 : Z_1\to X$$

are fibrations. Moreover, suppose that the restriction of ϕ to each fiber

$$\pi_0^{-1}(z_1)$$

has a unique non-degenerate critical point $\gamma(z_1)$. The map

$$z_1\mapsto\gamma(z_1)$$

defines a smooth section

$$\gamma : Z_1\to Z$$

of π_0. Let

$$\phi_1 := \gamma^*\phi.$$

Theorem 5.8.2. *ϕ_1 is a generating function for Λ with respect to π_1.*

Proof. (Again in local coordinates.) We may assume that

$$Z = X \times S \times T$$

and

$$\pi(x,s,t) = x, \quad \pi_0(x,s,t) = (x,s), \quad \pi_1(x,s) = x.$$

The condition for (x,s,t) to belong to C_ϕ is that

$$\frac{\partial \phi}{\partial s} = 0$$

and

$$\frac{\partial \phi}{\partial t} = 0.$$

This last condition has a unique solution giving t as a smooth function of (x,s) by our non-degeneracy condition, and from the definition of ϕ_1 it follows that $(x,s) \in C_{\phi_1}$ if and only if $\gamma(x,s) \in C_\phi$. Furthermore

$$d_X\phi_1(x,s) = d_X\phi(x,s,t)$$

along $\gamma(C_{\phi_1})$. □

For instance, suppose that $Z = X \times \mathbb{R}^k$ and $\phi = \phi(x,s)$ so that $z_0 = (x_0,s_0) \in C_\phi$ if and only if

$$\frac{\partial \phi}{\partial s_i}(x_0,s_0) = 0, \quad i = 1,\dots,k.$$

Suppose that the matrix

$$\left(\frac{\partial^2 \phi}{\partial s_i \partial s_j}\right)$$

is of rank r, for some $0 < r \le k$. By a linear change of coordinates we can arrange that the upper left hand corner

$$\left(\frac{\partial^2 \phi}{\partial s_i \partial s_j}\right), \quad 1 \le i,j, \le r$$

is non-degenerate. We can apply Theorem 5.8.2 to the fibration

$$X \times \mathbb{R}^k \to X \times \mathbb{R}^\ell, \quad \ell = k - r$$

$$(x, s_1, \dots s_k) \mapsto (x, t_1, \dots, t_\ell), \quad t_i = s_{i+r}$$

to obtain a generating function $\phi_1(x,t)$ for Λ relative to the fibration

$$X \times \mathbb{R}^\ell \to X.$$

Thus by reducing the number of variables we can assume that at $z_0 = (x_0,t_0)$

$$\frac{\partial^2 \phi}{\partial t_i \partial t_j}(x_0,t_0) = 0, \quad i,j = 1,\dots,\ell. \tag{5.11}$$

A generating function satisfying this condition will be said to be **reduced** at (x_0,t_0).

5.9 The existence of generating functions.

In this section we will show that every Lagrangian submanifold of T^*X can be described locally by a generating function ϕ relative to some fibration $Z \to X$.

So let $\Lambda \subset T^*X$ be a Lagrangian submanifold and let $p_0 = (x_0, \xi_0) \in \Lambda$. To simplify the discussion let us temporarily make the assumption that

$$\xi_0 \neq 0. \tag{5.12}$$

If Λ is horizontal at p_0 then we know from Chapter 1 that there is a generating function for Λ near p_0 with the trivial (i.e. no) fibration. If Λ is not horizontal at p_0, we can find a Lagrangian subspace

$$V_1 \subset T_{p_0}(T^*X)$$

which is horizontal and transverse to $T_{p_0}(\Lambda)$.

Indeed, to say that V_1 is horizontal, is to say that it is transverse to the Lagrangian subspace W_1 given by the vertical vectors at p_0 in the fibration $T^*X \to X$. By the Proposition in §2.2 we know that we can find a Lagrangian subspace which is transversal to both W_1 and $T_{p_0}(\Lambda)$.

Let Λ_1 be a Lagrangian submanifold passing through p_0 and whose tangent space at p_0 is V_1. So Λ_1 is a horizontal Lagrangian submanifold and

$$\Lambda_1 \pitchfork \Lambda = \{p_0\}.$$

In words, Λ_1 intersects Λ transversally at p_0. Since Λ_1 is horizontal, we can find a neighborhood U of x_0 and a function $\phi_1 \in C^\infty(U)$ such that $\Lambda_1 = \Lambda_{\phi_1}$. By our assumption (5.12)

$$(d\phi_1)_{x_0} = \xi_0 \neq 0.$$

So we can find a system of coordinates $x_1 \ldots, x_n$ on U (or on a smaller neighborhood) so that

$$\phi_1 = x_1.$$

Let $\xi_1 \ldots, \xi_n$ be the dual coordinates so that in the coordinate system

$$x_1 \ldots, x_n, \xi_1 \ldots, \xi_n$$

on T^*X the Lagrangian submanifold Λ_1 is described by the equations

$$\xi_1 = 1, \xi_2 = \cdots = \xi_n = 0.$$

Consider the canonical transformation generated by the function

$$\tau : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad \tau(x, y) = -x \cdot y.$$

The Lagrangian submanifold in $T^*\mathbb{R}^n \times T^*\mathbb{R}^n$ generated by τ is

$$\{(x, -y, y, -x)\}$$

so the canonical relation is

$$\{(x, \xi, \xi, -x)\}.$$

In other words, it is the graph of the linear symplectic transformation

$$\gamma : (x, \xi) \mapsto (\xi, -x).$$

So $\gamma(\Lambda_1)$ is (locally) the cotangent space at $y_0 = (1, 0, \ldots, 0)$. Since $\gamma(\Lambda)$ is transverse to this cotangent fiber, it follows that $\gamma(\Lambda)$ is horizontal. So in some neighborhood W of y_0 there is a function ψ such that

$$\gamma(\Lambda) = \Lambda_{-\psi}$$

over W. By equation (5.7) we know that

$$\tau^*(x, y) = -\tau(y, x) = y \cdot x$$

is the generating function for γ^{-1}. Furthermore, near p_0,

$$\Lambda = \gamma^{-1}(\Lambda_\psi).$$

Hence, by Theorem 5.6.1 the function

$$\psi_1(x, y) := y \cdot x - \psi(y) \tag{5.13}$$

is a generating function for Λ relative to the fibration

$$(x, y) \mapsto x.$$

Notice that this is a generalization of the construction of a generating function for a linear Lagrangian subspace transverse to the horizontal in Section 2.9.1.

We have proved the existence of a generating function under the auxiliary hypothesis (5.12). However it is easy to deal with the case $\xi_0 = 0$ as well. Namely, suppose that $\xi_0 = 0$. Let $f \in C^\infty(X)$ be such that $df(x_0) \neq 0$. Then

$$\gamma_f : T^*X \to T^*X, \quad (x, \xi) \mapsto (x, \xi + df)$$

is a symplectomorphism and $\gamma_f(p_0)$ satisfies (5.12). We can then form

$$\gamma \circ \gamma_f(\Lambda)$$

which is horizontal. Notice that $\gamma \circ \gamma_f$ is given by

$$(x, \xi) \mapsto (x, \xi + df) \mapsto (\xi + df, -x).$$

If we consider the generating function on $\mathbb{R}^n \times \mathbb{R}^n$ given by

$$g(x, z) = -x \cdot z + f(x)$$

then the corresponding Lagrangian submanifold is

$$\{(x, -z + df, z, -x)\}$$

so the canonical relation is

$$\{(x, z - df, z, -x)\}$$

or, setting $\xi = z + df$ so $z = \xi - df$ we get

$$\{(x, \xi, \xi + df, -x)\}$$

which is the graph of $\gamma \circ \gamma_f$. We can now repeat the previous argument to conclude that

$$y \cdot x - f(x) - \psi(y)$$

is a generating function for Λ. So we have proved:

Theorem 5.9.1. *Every Lagrangian submanifold of T^*X can be locally represented by a generating function relative to a fibration.*

Let us now discuss generating functions for canonical relations: So let X and Y be manifolds and

$$\Gamma \subset T^*X \times T^*Y$$

a canonical relation. Let $(p_0, q_0) = (x_0, \xi_0, y_0, \eta_0) \in \Gamma$ and assume now that

$$\xi_0 \neq 0, \quad \eta_0 \neq 0. \tag{5.14}$$

We claim that the following theorem holds

Theorem 5.9.2. *There exist coordinate systems $(U, x_1, \dots, x_n)$ about x_0 and $(V, y_1 \dots, y_k)$ about y_0 such that if*

$$\gamma_U : T^*U \to T^*\mathbb{R}^n$$

is the transform

$$\gamma_U(x, \xi) = (-\xi, x)$$

and

$$\gamma_V : T^*V \to T^*\mathbb{R}^k$$

is the transform

$$\gamma_V(y, \eta) = (-\eta, y)$$

then locally, near

$$p_0' := \gamma_U^{-1}(p_0) \quad \text{and} \quad q_0' := \gamma_V(q_0),$$

the canonical relation

$$\gamma_V^{-1} \circ \Gamma \circ \gamma_U \tag{5.15}$$

is of the form

$$\Gamma_\phi, \quad \phi = \phi(x, y) \in C^\infty(\mathbb{R}^n \times \mathbb{R}^k).$$

Proof. Let

$$M_1 := T^*X, \quad M_2 = T^*Y$$

and

$$V_1 := T_{p_0}M_1, \quad V_2 := T_{q_0}M_2, \quad \Sigma := T_{(p_0,q_0)}\Gamma$$

so that Σ is a Lagrangian subspace of

$$V_1^- \times V_2.$$

Let W_1 be a Lagrangian subspace of V_1 so that (in the linear symplectic category)

$$\Sigma(W_1) = \Sigma \circ W_1$$

is a Lagrangian subspace of V_2. Let W_2 be another Lagrangian subspace of V_2 which is transverse to $\Sigma(W_1)$. We may choose W_1 and W_2 to be horizontal subspaces of $T_{p_0}M_1$ and $T_{q_0}M_2$. Then $W_1 \times W_2$ is transverse to Σ in $V_1 \times V_2$ and we may choose a Lagrangian submanifold passing through p_0 and tangent to W_1 and similarly a Lagrangian submanifold passing through q_0 and tangent to W_2. As in the proof of Theorem 5.9.1 we can arrange local coordinates $(x_1 \ldots, x_n)$ on X and hence dual coordinates $(x_1, \ldots x_n, \xi_1, \ldots, \xi_n)$ around p_0 such that the Lagrangian manifold tangent to W_1 is given by

$$\xi_1 = 1, \quad \xi_2 = \cdots \xi_n = 0$$

and similarly dual coordinates on $M_2 = T^*Y$ such that the second Lagrangian submanifold (the one tangent to W_2) is given by

$$\eta_1 = 1, \quad \eta_2 = \cdots = \eta_k = 0.$$

It follows that the Lagrangian submanifold corresponding to the canonical relation (5.15) is horizontal and hence is locally of the form Γ_ϕ. $\square$

5.10 The Legendre transformation.

Coming back to our proof of the existence of a generating function for Lagrangian manifolds, let's look a little more carefully at the details of this proof. Let $X = \mathbb{R}^n$ and let $\Lambda \subset T^*X$ be the Lagrangian manifold defined by the fibration, $Z = X \times \mathbb{R}^n \xrightarrow{\pi} X$ and the generating function

$$\phi(x,y) = x \cdot y - \psi(y) \tag{5.16}$$

where $\psi \in C^\infty(\mathbb{R}^n)$. Then

$$(x,y) \in C_\phi \Leftrightarrow x = \frac{\partial \psi}{\partial y}(y)\,.$$

Recall also that $(x_0, y_0) \in C_\phi \Leftrightarrow$ the function $\phi(x_0, y)$ has a critical point at y_0. Let us suppose this is a *non-degenerate* critical point, i.e., that the matrix

$$\left(\frac{\partial^2 \phi}{\partial y_i \partial y_j}(x_0, y_0)\right) = \left(\frac{\partial \psi}{\partial y_i \partial y_j}(y_0)\right) \tag{5.17}$$

is of rank n. By Theorem 5.8.1 we know that there exists a neighborhood $U \ni x_0$ and a function $\psi^* \in C^\infty(U)$ such that

$$\begin{aligned}\psi^*(x) &= \phi(x,y) \text{ at } (x,y) \in C_\phi && (5.18)\\ \Lambda &= \Lambda_{\psi^*} && (5.19)\end{aligned}$$

locally, near the image $p_0 = (x_0, \xi_0)$ of the map $\frac{\partial \phi}{\partial x} : C_\phi \to \Lambda$. What do these three assertions say? Assertion (5.17) simply says that the map

$$y \to \frac{\partial \psi}{\partial y} \tag{5.20}$$

is a diffeomorphism at y_0. Assertion (5.18) says that

$$\psi^*(x) = xy - \psi(x) \tag{5.21}$$

at $x = \frac{\partial \psi}{\partial y}$, and assertion(5.19) says that

$$x = \frac{\partial \psi}{\partial y} \Leftrightarrow y = -\frac{\partial \psi^*}{\partial x} \tag{5.22}$$

i.e., the map

$$x \to -\frac{\partial \psi^*}{\partial x} \tag{5.23}$$

is the inverse of the mapping (5.20). The mapping (5.20) is known as the Legendre transform associated with ψ and the formulas (5.21)– (5.23) are the famous inversion formula for the Legendre transform. Notice also that in the course of our proof that (5.21) is a generating function for Λ we proved that ψ is a generating function for $\gamma(\Lambda)$, i.e., locally near $\gamma(p_0)$

$$\gamma(\Lambda) = \Lambda_{-\psi}\,.$$

Thus we've proved that locally near p_0

$$\Lambda_{\psi^*} = \gamma^{-1}(\Lambda_\psi)$$

where

$$\gamma^{-1} : T^*\mathbb{R}^n \to T^*\mathbb{R}^n$$

is the transform $(y, \eta) \to (x, \xi)$ where

$$y = \xi \quad \text{and} \quad x = -\eta\,.$$

This identity will come up later when we try to compute the semi-classical Fourier transform of the rapidly oscillating function

$$a(y) e^{i\frac{\psi(y)}{\hbar}}\,,\; a(y) \in C_0^\infty(\mathbb{R}^n)\,.$$

5.11 The Hörmander-Morse lemma.

In this section we will describe some relations between different generating functions for the same Lagrangian submanifold. Our basic goal is to show that if we have two generating functions for the same Lagrangian manifold they can be obtained (locally) from one another by applying a series of "moves", each of a very simple type.

Let Λ be a Lagrangian submanifold of T^*X, and let

$$Z_0 \stackrel{\pi_0}{\to} X, \quad Z_1 \stackrel{\pi_1}{\to} X$$

be two fibrations over X. Let ϕ_1 be a generating function for Λ with respect to $\pi_1 : Z_1 \to X$.

Proposition 5.11.1. *If*

$$f : Z_0 \to Z_1$$

is a diffeomorphism satisfying

$$\pi_1 \circ f = \pi_0$$

then

$$\phi_0 = f^*\phi_1$$

is a generating function for Λ with respect to π_0.

Proof. We have $d(\phi_1 \circ f) = d\phi_0$. Since f is fiber preserving, f maps C_{ϕ_0} diffeomorphically onto C_{ϕ_1}. Furthermore, on C_{ϕ_0} we have

$$d\phi_1 \circ f = (d\phi_1 \circ f)_{hor} = (d\phi_0)_{hor}$$

so f conjugates the maps $d_X\phi_i : C_{\phi_i} \to \Lambda, \quad i = 0, 1$. Since $d_X\phi_1$ is a diffeomorphism of C_{ϕ_1} with Λ we conclude that $d_X\phi_0$ is a diffeomorphism of C_{ϕ_0} with Λ, i.e. ϕ_0 is a generating function for Λ. □

Our goal is to prove a result in the opposite direction. So as above let $\pi_i : Z_i \to X,\ i = 0, 1$ be fibrations and suppose that ϕ_0 and ϕ_1 are generating functions for Λ with respect to π_i. Let

$$p_0 \in \Lambda$$

and $z_i \in C_{\phi_i},\ i = 0, 1$ be the pre-images of p_0 under the diffeomorphism $d\phi_i$ of C_{ϕ_i} with Λ. So

$$d_X\phi_i(z_i) = p_0, \qquad i = 0, 1.$$

Finally let $x_0 \in X$ be given by

$$x_0 = \pi_0(z_0) = \pi_1(z_1)$$

and let $\psi_i,\ i = 0, 1$ be the restriction of ϕ_i to the fiber $\pi_i^{-1}(x_0)$. Since $z_i \in C_{\phi_i}$ we know that z_i is a critical point for ψ_i. Let

$$d^2\psi_i(z_i)$$

be the Hessian of ψ_i at z_i.

Theorem 5.11.1. The Hörmander Morse lemma. *If $d^2\psi_0(z_0)$ and $d^2\psi_1(z_1)$ have the same rank and signature, then there exists neighborhood U_0 of z_0 in Z_0 and U_1 of z_1 in Z_1 and a diffeomorphism*

$$f : U_0 \to U_1$$

such that

$$\pi_1 \circ f = \pi_0$$

and

$$\phi_1 \circ f = f^*\phi_1 = \phi_0 + \text{ const. } .$$

Proof. We will prove this theorem in a number of steps. We will first prove the theorem under the additional assumption that Λ is horizontal at p_0. Then we will reduce the general case to this special case.

Assume that Λ is horizontal at $p_0 = (x_0, \xi_0)$. This implies that Λ is horizontal over some neighborhood of x_0. Let S be an open subset of $\mathbb{R}^k$ and

$$\pi : X \times S \to X$$

projection onto the first factor. Suppose that $\phi \in C^\infty(X \times S)$ is a generating function for Λ with respect to π so that

$$d_X\phi : \quad C_\phi \to \Lambda$$

is a diffeomorphism, and let $z_0 \in C_\phi$ be the pre-image of p_0 under this diffeomorphism, i.e.

$$z_0 = (d_X\phi)^{-1}(p_0).$$

We begin by proving that the vertical Hessian of ϕ at z_0 is non-degenerate.

Since Λ is horizontal at p_0 there is a neighborhood U of x_0 $\psi \in C^\infty(U)$ such that

$$d\psi : U \to T^*X$$

maps U diffeomorphically onto a neighborhood of p_0 in Λ. So

$$(d\psi)^{-1} \circ d_X\phi : \ C_\phi \to U$$

is a diffeomorphism. But $(d\psi)^{-1}$ is just the restriction to a neighborhood of p_0 in Λ of the projection $\pi_X : T^*X \to X$. So $\pi_X \circ d_X\phi : C_\phi \to X$ is a diffeomorphism (when restricted to $\pi^{-1}(U)$). But

$$\pi_X \circ d_X\phi = \pi_{|C_\phi}$$

so the restriction of π to C_ϕ is a diffeomorphism. So C_ϕ is horizontal at z_0, in the sense that

$$T_{z_0}C_\phi \cap T_{z_0}S = \{0\}.$$

So we have a smooth map

$$\mathbf{s} : U \to S$$

such that $x \mapsto (x, \mathbf{s}(x))$ is a smooth section of C_ϕ over U. We have

$$d_X\phi = d\phi \quad \text{at all points } \ (x, \mathbf{s}(x))$$

by the definition of C_ϕ and $d\psi(x) = d_X\phi(x, \mathbf{s}(x)) = d\phi(x, \mathbf{s}(x))$ so

$$\psi(x) = \phi(x, \mathbf{s}(x)) + \text{const.} \ . \tag{5.24}$$

The submanifold $C_\phi \subset Z = X \times S$ is defined by the k equations

$$\frac{\partial \phi}{\partial s_i} = 0, \quad i = 1, \ldots, k$$

and hence $T_{z_0}C_\phi$ is defined by the k independent linear equations

$$d\left(\frac{\partial \phi}{\partial s_i}\right) = 0, \quad i = 1, \ldots, k.$$

A tangent vector to S at z_0, i.e. a tangent vector of the form

$$(0, v), \quad v = (v^1, \ldots v^k)$$

will satisfy these equations if and only if

$$\sum_j \frac{\partial^2 \phi}{\partial s_i \partial s_j} v^j = 0, \quad i = 1, \ldots, k.$$

But we know that these equations have only the zero solution as no non-zero tangent vector to S lies in the tangent space to C_ϕ at z_0. We conclude that the vertical Hessian matrix

$$d_S^2\phi = \left(\frac{\partial^2 \phi}{\partial s_i \partial s_j}\right)$$

is non-degenerate.

We return to the proof of the theorem under the assumption that that Λ is horizontal at $p_0 = (x_0, \xi_0)$. We know that the vertical Hessians occurring in the statement of the theorem are both non-degenerate, and we are assuming that they are of the same rank. So the fiber dimensions of π_0 and π_1 are the same. So we may assume that $Z_0 = X \times S$ and $Z_1 = X \times S$ where S is an open subset of $\mathbb{R}^k$ and that coordinates have been chosen so that the coordinates of z_0 are $(0, 0)$ as are the coordinates of z_1. We write

$$\mathfrak{s}_0(x) = (x, \mathbf{s}_0(x)), \quad \mathfrak{s}_1(x) = (x, \mathbf{s}_1(x)),$$

where $\mathbf{s}_0$ and $\mathbf{s}_1$ are smooth maps $X \to \mathbb{R}^k$ with

$$\mathbf{s}_0(0) = \mathbf{s}_1(0) = 0.$$

Let us now take into account that the signatures of the vertical Hessians are the same at z_0. By continuity they must be the same at the points $(x, \mathbf{s}_0(x))$ and $(x, \mathbf{s}_1(x))$ for each $x \in U$. So for each fixed $x \in U$ we can make an affine change of coordinates in S and add a constant to ϕ_1 so as to arrange that

1. $\mathbf{s}_0(x) = \mathbf{s}_1(x) = 0$.
2. $\frac{\partial \phi_0}{\partial s_i}(x,0) = \frac{\partial \phi_1}{\partial s_i}(x,0),\ i = 1\dots,k$.
3. $\phi_0(x,0) = \phi_1(x,0)$.
4. $d_S^2\phi_0(x,0) = d_S^2\phi_1(x,0)$.

We can now apply Morse's lemma with parameters (see §14.14.3 for a proof) to conclude that there exists a fiber preserving diffeomorphism $f : U \times S \to U \times S$ with

$$f^*\phi_1 = \phi_0.$$

This completes the proof of Theorem 5.11.1 under the additional hypothesis that Lagrangian manifold Λ is horizontal.

Reduction of the number of fiber variables. Our next step in the proof of Theorem 5.11.1 will be an application of Theorem 5.8.2. Let $\pi : Z \to X$ be a fibration and ϕ a generating function for Λ with respect to π. Suppose we are in the setup of Theorem 5.8.2 which we recall with some minor changes in notation: We suppose that the fibration

$$\pi : Z \to X$$

can be factored as a succession of fibrations

$$\pi = \rho \circ \varrho$$

where

$$\rho : Z \to W \quad \text{and} \quad \varrho : W \to X$$

are fibrations. Moreover, suppose that the restriction of ϕ to each fiber

$$\rho^{-1}(w)$$

has a unique non-degenerate critical point $\gamma(w)$. The map

$$w \mapsto \gamma(w)$$

defines a smooth section

$$\gamma : W \to Z$$

of ρ. Let

$$\chi := \gamma^*\phi.$$

Theorem 5.8.2 asserts that χ is a generating function of Λ with respect to ϱ. Consider the Lagrangian submanifold

$$\Lambda_\chi \subset T^*W.$$

This is horizontal as a Lagrangian submanifold of T^*W and ϕ is a generating function for Λ_χ relative to the fibration $\rho : Z \to W$.

Now suppose that we had two fibrations and generating functions as in the hypotheses of Theorem 5.11.1 and suppose that they both factored as above with *the same* $\varrho : W \to X$ and the same χ. So we get fibrations $\varrho_0 : Z_0 \to W$ and $\varrho_1 : Z_1 \to W$. We could then apply the above (horizontal) version of Theorem 5.11.1 to conclude the truth of the theorem.

Since the ranks of $d^2\psi_0$ and $d^2\psi_1$ at z_0 and z_1 are the same, we can apply the reduction leading to equation (5.11) to each. So by the above argument Theorem 5.11.1 will be proved once we prove it for the reduced case.

Some normalizations in the reduced case. We now examine a fibration $Z = X \times S \to S$ and generating function ϕ and assume that ϕ is reduced at $z_0 = (x_0, s_0)$ so all the second partial derivatives of ϕ in the S direction vanish, i. e.

$$\frac{\partial^2 \phi}{\partial s_i \partial s_j}(x_0, s_0) = 0 \quad \forall i, j.$$

This implies that

$$T_{s_0} S \cap T_{(x_0,s_0)} C_\phi = T_{s_0} S.$$

i.e. that

$$T_{s_0} S \subset T_{(x_0,s_0)} C_\phi. \tag{5.25}$$

Consider the map

$$d_X \phi : X \times S \to T^* X, \qquad (x, s) \mapsto d_X \phi(x, s).$$

The restriction of this map to C_ϕ is just our diffeomorphism of C_ϕ with Λ. So the restriction of the differential of this map to any subspace of any tangent space to C_ϕ is injective. By (5.25) the restriction of the differential of this map to $T_{s_0} S$ at (x_0, s_0) is injective. In other words, by passing to a smaller neighborhood of (x_0, s_0) if necessary, we have an embedding

$$\begin{array}{ccc} X \times S & \xrightarrow{d_X \phi} & W \subset T^* X \\ \pi \downarrow & & \downarrow \pi_X \\ X & \xrightarrow[\text{id}]{} & X \end{array}$$

of $X \times S$ onto a subbundle W of $T^* X$.

Now let us return to the proof of our theorem. Suppose that we have two generating functions $\phi_i,\ i = 0, 1\ X \times S_i \to X$ and both are reduced at the points z_i of C_{ϕ_1} corresponding to $p_0 \in \Lambda$. So we have two embeddings

$$\begin{array}{ccc} X \times S_i & \xrightarrow{d_X \phi_i} & W_i \subset T^* X \\ \pi \downarrow & & \downarrow \pi_X \\ X & \xrightarrow[\text{id}]{} & X \end{array}$$

of $X \times S_i$ onto subbundles W_i of T^*X for $i = 0, 1$. Each of these maps the corresponding C_{ϕ_i} diffeomorphically onto Λ.

Let V be a tubular neighborhood of W_1 in T^*X and $\tau : V \to W_1$ a projection of V onto W_1 so we have the commutative diagram

$$\begin{array}{ccc} V & \xrightarrow{\tau} & W_1 \\ {\scriptstyle \pi_X}\downarrow & & \downarrow{\scriptstyle \pi_X} \\ X & \xrightarrow[\text{id}]{} & X \end{array} .$$

Let

$$\gamma := (d_X\phi_1)^{-1} \circ \tau.$$

So we have the diagram

$$\begin{array}{ccc} V & \xrightarrow{\gamma} & X \times S_1 \\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \pi_X} \\ X & \xrightarrow[\text{id}]{} & X \end{array}$$

and

$$\gamma \circ d_X\phi_1 = \text{id}\,.$$

We may assume that $W_0 \subset V$ so we get a fiber map

$$g := \gamma \circ d_X\phi_0 \quad g : X \times S_0 \to X \times S_1.$$

When we restrict g to C_{ϕ_0} we get a diffeomorphism of C_{ϕ_0} onto C_{ϕ_1}. By (5.25) we know that

$$T_{s_i}S_i \subset T_{z_i}C_{\phi_i}$$

and so dg_{z_0} maps $T_{s_0}S_0$ bijectively onto $T_{s_1}S_1$. Hence g is locally a diffeomorphism at z_0. So by shrinking X and S_i we may assume that

$$g : X \times S_0 \to X \times S_1$$

is a fiber preserving diffeomorphism. We now apply Proposition 5.11.1. So we replace ϕ_1 by $g^*\phi_1$. Then the two fibrations Z_0 and Z_1 are the same and $C_{\phi_0} = C_{\phi_1}$. Call this common submanifold C. Also $d_X\phi_0 = d_X\phi_1$ when restricted to C, and by definition the vertical derivatives vanish. So $d\phi_0 = d\phi_1$ on C, and so by adjusting an additive constant we can arrange that $\phi_0 = \phi_1$ on C.

Completion of the proof. We need to prove the theorem in the following situation:

- $Z_0 = Z_1 = X \times S$ and $\pi_0 = \pi_1$ is projection onto the first factor.
- The two generating functions ϕ_0 and ϕ_1 have the same critical set:

$$C_{\phi_0} = C_{\phi_1} = C.$$

- $\phi_0 = \phi_1$ on C.
- $d_S\phi_i = 0,\ i = 0,1$ on C and $d_X\phi_0 = d_X\phi_1$ on C.
- $$d\left(\frac{\partial\phi_0}{\partial s_i}\right) = d\left(\frac{\partial\phi_1}{\partial s_i}\right) \text{ at } z_0.$$

We will apply the Moser trick: Let

$$\phi_t := (1-t)\phi_0 + t\phi_1.$$

From the above we know that

- $\phi_t = \phi_0 = \phi_1$ on C.
- $d_S\phi_t = 0$ on C and $d_X\phi_t = d_X\phi_0 = d_X\phi_1$ on C.
- $$d\left(\frac{\partial\phi_t}{\partial s_i}\right) = d\left(\frac{\partial\phi_0}{\partial s_i}\right) = d\left(\frac{\partial\phi_1}{\partial s_i}\right) \text{ at } z_0.$$

So in a sufficiently small neighborhood of Z_0 the submanifold C is defined by the k independent equations

$$\frac{\partial\phi_t}{\partial s_i} = 0, \quad i = 1,\dots k.$$

We look for a vertical (time dependent) vector field

$$v_1 = \sum_i v_i(x,s,t)\frac{\partial}{\partial s_i}$$

on $X \times S$ such that

1. $D_{v_t}\phi_t = -\dot{\phi}_t = \phi_0 - \phi_1$ and
2. $v = 0$ on C.

Suppose we find such a v_t. Then solving the differential equations

$$\frac{d}{dt}f_t(m) = v_t(f_t(m)), \quad f_0(m) = m$$

will give a family of fiber preserving diffeomorphsms (since v_t is vertical) and

$$f_1^*\phi_1 - \phi_0 = \int_0^1 \frac{d}{dt}(f_t^*\phi_t)dt = \int_0^1 f_t^*[D_{v_t}\phi_t + \dot{\phi}_t]dt = 0.$$

So finding a vector field v_t satisfying 1) and 2) will complete the proof of the theorem. Now $\phi_0 - \phi_1$ vanishes to second order on C which is defined by the independent equations $\partial\phi_t/\partial s_i = 0$. So we can find functions

$$w_{ij}(x,s,t)$$

defined and smooth in some neighborhood of C such that

$$\phi_0 - \phi_1 = \sum_{ij} w_{ij}(x,s,t) \frac{\partial \phi_t}{\partial s_i} \frac{\partial \phi_t}{\partial s_j}$$

in this neighborhood. Set

$$v_i(x,s,t) = \sum_i w_{ij}(x,s,t) \frac{\partial \phi_t}{\partial s_j}.$$

Then condition 2) is clearly satisfied and

$$D_{v_t}\phi_t = \sum_{ij} w_{ij}(x,s,t) \frac{\partial \phi_t}{\partial si} \frac{\partial \phi_t}{\partial s_j} = \phi_0 - \phi_1 = -\dot{\phi}$$

as required. □

5.12 Changing the generating function.

We summarize the results of the preceding section as follows: Suppose that $(\pi_1 : Z_1 \to X, \phi_1)$ and $(\pi_2 : Z_2 \to X, \phi_2)$ are two descriptions of the same Lagrangian submanifold Λ of T^*X. Then locally one description can be obtained from the other by applying sequentially "moves" of the following three types:

1. **Adding a constant.** We replace ϕ_1 by $\phi_2 = \phi_1 + c$ where c is a constant.

2. **Equivalence.** There exists a diffeomorphism $g : Z_1 \to Z_2$ with

$$\pi_2 \circ g = \pi_1 \qquad \text{and} \qquad \phi_2 \circ g = \phi_1.$$

3. **Increasing (or decreasing) the number of fiber variables.** Here $Z_2 = Z_1 \times \mathbb{R}^d$ and

$$\phi_2(z,s) = \phi_1(z) + \frac{1}{2}\langle As, s\rangle$$

where A is a non-degenerate $d \times d$ matrix (or vice versa).

5.13 The Maslov bundle.

We wish to associate to each Lagrangian submanifold of a cotangent bundle a certain flat line bundle which will be of importance to us when we get to the symbol calculus in Chapter 8. We begin with a review of the Čech-theoretic description of flat line bundles.

5.13.1 The Čech description of locally flat line bundles.

Let Y be a manifold and $\mathbb{U} = \{U_i\}$ be an open cover of Y. Let

$$\mathbb{N}^1 = \{(i,j)|U_i \cap U_j \neq \emptyset\}.$$

A collection of of non-zero complex numbers $\{c_{ij}\}_{(i,j)\in\mathbb{N}^1}$ is called a (multiplicative) cocycle (relative to the cover $\mathbb{U}$) if

$$c_{ij} \cdot c_{jk} = c_{ik} \quad \text{whenever} \qquad U_i \cap U_j \cap U_k \neq \emptyset. \tag{5.26}$$

From this data one constructs a line bundle as follows: One considers the set

$$\amalg_i(U_i \times \mathbb{C})$$

and puts an an equivalence relation on it by declaring that

$$(p_i, a_i) \sim (p_j, a_j) \qquad \Leftrightarrow \qquad p_i = p_j \in U_i \cap U_j \text{ and } \; a_i = c_{ij}a_j.$$

Then

$$\mathbb{L} := \amalg_i(U_i \times \mathbb{C})/\sim$$

is a line bundle over Y. The constant functions

$$U_i \to 1 \in \mathbb{C}$$

form flat local sections of $\mathbb{L}$

$$s_i : U \to \mathbb{L}, \quad p \mapsto [(p, 1)]$$

and thus make $\mathbb{L}$ into a line bundle with flat connection over Y.

Any section s of $\mathbb{L}$ can be written over U_i as $s = f_i s_i$. If v is a vector field on Y, we may define $D_v s$ by

$$D_v s := (D_v f_i)s_i \quad \text{ on } \; U_i.$$

The fact that the transitions between s_i and s_j are constant shows that this is well defined.

5.13.2 The local description of the Maslov cocycle.

We first define the Maslov line bundle $\mathbb{L}_{\text{Maslov}} \to \Lambda$ in terms of a global generating function, and then show that the definition is invariant under change of generating function. We then use the local existence of generating functions to patch the line bundle together globally. Here are the details:

Suppose that ϕ is a generating function for Λ relative to a fibration $\pi : Z \to X$. Let z be a point of the critical set C_ϕ, let $x = \pi(z)$ and let $F = \pi^{-1}(x)$ be the fiber containing z. The restriction of ϕ to the fiber F has a critical point at z. Let $\text{sgn}^\#(z)$ be the signature of the Hessian at z of ϕ restricted to F. This gives an integer valued function on C_ϕ:

$$\text{sgn}^\# : C_\phi \to \mathbb{Z}, \qquad z \mapsto \text{sgn}^\#(z).$$

Notice that since the Hessian can be singular at points of C_ϕ this function can be quite discontinuous.

From the diffeomorphism $\lambda_\phi = d_X\phi$

$$\lambda_\phi : C_\phi \to \Lambda$$

we get a $\mathbb{Z}$ valued function sgn_ϕ on Λ given by

$$\mathrm{sgn}_\phi := \mathrm{sgn}^\sharp \circ \lambda_\phi^{-1}.$$

Let

$$s_\phi := e^{\frac{\pi i}{4}\mathrm{sgn}_\phi}.$$

So

$$s_\phi : \Lambda \to \mathbb{C}^*$$

taking values in the eighth roots of unity.

We define the Maslov bundle $\mathbb{L}_{\mathrm{Maslov}} \to \Lambda$ to be the trivial flat bundle having s_ϕ as its defining flat section.

Suppose that (Z_i, π_i, ϕ_i), $\quad i = 1, 2$ are two descriptions of Λ by generating functions which differ from one another by one of the three Hörmander moves of Section 5.12. We claim that

$$s_{\phi_1} = c_{1,2} s_{\phi_2} \tag{5.27}$$

for some constant $c_{1,2} \in \mathbb{C}^*$. So we need to check this for the three types of move of Section 5.12. For moves of type 1) and 2), i.e. adding a constant or equivalences this is obvious. For each of these moves there is no change in sgn_ϕ.

For a move of type 3) the $\mathrm{sgn}_1^\#$ and $\mathrm{sgn}_2^\#$ are related by

$$\mathrm{sgn}_1^\# = \mathrm{sgn}_2^\# + \text{signature of } A.$$

This proves (5.27), and defines the Maslov bundle when a global generating function exists.

In this discussion we have been tacitly assuming that ϕ is a transverse generating function of Λ. However, the definition of s_ϕ above makes sense as well for clean generating functions. Namely if $\phi \in \mathcal{C}^\infty(Z)$ is a clean generating function for Λ with respect to the fibration $\pi : Z \to X$ then as we showed in §5.11, π factors (locally) into fibrations with connected fibers

$$Z \xrightarrow{\pi_1} Z_1 \xrightarrow{\pi_2} X$$

and φ can be written as a pull-back $\varphi = \pi_1^*\varphi_1$ where $\varphi_1 \in \mathcal{C}^\infty(Z_1)$ is a transverse generating function for Λ with respect to π_2. Thus $C_\varphi = \pi_1^{-1}(C_{\varphi_1})$ and the signature map, $(\mathrm{sgn})^\# : C_\varphi \to \mathbb{Z}$ is just the pull-back of the signature map $C_{\varphi_1} \to \mathbb{Z}$ associated with φ_1. Moreover, the diffeomorphism

$$\lambda_{\varphi_1} : C_{\varphi_1} \to \Lambda$$

lifts to a fiber preserving map

$$\lambda_\varphi : C_\varphi \to \Lambda$$

and we can define, as above, a function

$$\mathrm{sgn}_\phi : \Lambda \to \mathbb{Z}$$

by requiring that $\mathrm{sgn}_\phi \circ \lambda = (\mathrm{sgn})^\#$ and then define s_ϕ as above to be the function $e^{\frac{i\pi}{4}\,\mathrm{sgn}_\phi}$.

5.13.3 The global definition of the Maslov bundle.

Now consider a general Lagrangian submanifold $\Lambda \subset T^*X$. Cover Λ by open sets U_i such that each U_i *is* defined by a generating function and that generating functions ϕ_i and ϕ_j are obtained from one another by one of the Hörmander moves. We get functions $s_{\phi_i} : U_i \to \mathbb{C}$ such that on every overlap $U_i \cap U_j$

$$s_{\phi_i} = c_{ij} s_{\phi_j}$$

with constants c_{ij} with $|c_{ij}| = 1$. Although the functions s_ϕ might be quite discontinuous, the c_{ij} in (5.27) are constant on $U_i \cap U_j$. On the other hand, the fact that $s_{\phi_i} = c_{ij} s_{\phi_j}$ shows that the cocycle condition (5.26) is satisfied. In other words we get a Čech cocycle on the one skeleton of the nerve of this cover and hence a flat line bundle.

5.13.4 The Maslov bundle of a canonical relation between cotangent bundles.

We have defined the Maslov bundle for any Lagrangian submanifold of any cotangent bundle. If

$$\Gamma \in \mathrm{Morph}(T^*X_1, T^*X_2)$$

is a canonical relation between cotangent bundles, so that Γ is a Lagrangian submanifold of

$$(T^*X_1)^- \times T^*X_2$$

then

$$(\varsigma_1 \times \mathrm{id})(\Gamma)$$

is a Lagrangian submanifold of

$$T^*X_1 \times T^*X_2 = T^*(X_1 \times X_2)$$

and hence has an associated Maslov line bundle. We then use the identification $\varsigma_1 \times \mathrm{id}$ to pull this line bundle back to Γ. In other words, we define

$$\mathbb{L}_{\mathrm{Maslov}}(\Gamma) := (\varsigma_1 \times \mathrm{id})^* \mathbb{L}_{\mathrm{Maslov}}((\varsigma_1 \times \mathrm{id})(\Gamma)). \tag{5.28}$$

5.13.5 Functoriality of the Maslov bundle.

Let X_1, X_2, and X_3 be differentiable manifolds, and let

$$\Gamma_1 \in \mathrm{Morph}(T^*X_1, T^*X_2) \quad \text{and} \quad \Gamma_2 \in \mathrm{Morph}(T^*X_2, T^*X_3)$$

be cleanly composable canonical relations. Recall that this implies that we have a submanifold

$$\Gamma_2 \star \Gamma_1 \subset T^*X_1 \times T^*X_2 \times T^*X_3$$

and a fibration (4.5)

$$\kappa : \Gamma_2 \star \Gamma_1 \to \Gamma_2 \circ \Gamma_1$$

with compact connected fibers. So we can form the line bundle

$$\kappa^*(\mathbb{L}_{\mathrm{Maslov}}(\Gamma_2 \circ \Gamma_1)) \to \Gamma_2 \star \Gamma_1.$$

On the other hand, $\Gamma_2 \star \Gamma_1$ consists of all (m_1, m_2, m_3) with

$$(m_1, m_2) \in \Gamma_1 \quad \text{and} \quad (m_2.m_3) \in \Gamma_2.$$

So we have projections

$$\mathrm{pr}_1 : \ \Gamma_2 \star \Gamma_1 \to \Gamma_1, \quad (m_1, m_2, m_3) \mapsto (m_1, m_2)$$

and

$$\mathrm{pr}_2 : \ \Gamma_2 \star \Gamma_1 \to \Gamma_2, \quad (m_1, m_2, m_3) \mapsto (m_2, m_3).$$

So we can also pull the Maslov bundles of Γ_1 and Γ_2 back to $\Gamma_2 \star \Gamma_1$. We claim that

$$\kappa^*\mathbb{L}_{\mathrm{Maslov}}(\Gamma_2 \circ \Gamma_1) \cong \mathrm{pr}_1^* \mathbb{L}_{\mathrm{Maslov}}(\Gamma_1) \otimes \mathrm{pr}_2^* \mathbb{L}_{\mathrm{Maslov}}(\Gamma_2) \tag{5.29}$$

as line bundles over $\Gamma_2 \star \Gamma_1$.

Proof. We know from Section 5.7 that we can locally choose generating functions ϕ_1 for Γ_1 relative to a fibration

$$X_1 \times X_2 \times S_1 \to X_1 \times S_2$$

and ϕ_2 for Γ_2 relative to a fibration

$$X_2 \times X_3 \times S_2 \to X_2 \times X_3$$

so that

$$\phi = \phi(x_1, x_2, x_3, s_1, s_2) = \phi_1(x_1, x_2, s_1) + \phi_2(x_2, x_3, s_2)$$

is a generating function for $\Gamma_2 \circ \Gamma_1$ relative to the fibration

$$X_1 \times X_3 \times X_2 \times S_1 \times S_2 \to X_1 \times X_3$$

(locally). We can consider the preceding equation as taking place over a neighborhood in $\Gamma_2 \star \Gamma_1$. Over such a neighborhood, the restrictions of the bundles on both sides of (5.29) are trivial, and we define the isomorphism in (5.29) to be given by

$$\mathrm{pr}_1^* s_{\phi_1} \otimes \mathrm{pr}_2^* s_{\Phi_2} \mapsto \kappa^* s_\phi. \tag{5.30}$$

We must check that this is well defined.

We may further restrict our choices of generating functions and neighborhoods for Γ_1 so that the passage from one to the other is given by one of the Hörmander moves, and similarly for Γ_2. A Hörmander move of type 1 on each factor just adds a constant to ϕ_1 and to ϕ_2 and hence adds the sum of these constants to ϕ, i.e. is a Hörmander move of type 1 on $\Gamma_2 \circ \Gamma_2$. Similarly for a Hörmander move of type 2. Also for Hörmander moves of type 3, we are adding a quadratic form in (additional) s variables to ϕ_1, and a quadratic form in t variables to ϕ_2 yielding a Hörmander move of type 3 to ϕ. This proves that (5.29) is well defined. □

5.14 Identifying the two definitions of the Maslov bundle.

We will use the functoriality above to show that the line bundle $\mathbb{L}_{\text{Maslov}}$ that we defined in §5.13.2 coincides with the line bundle that we defined in §2.8. Let $p_0 = (x_0, \xi_0)$ be a point of Λ. Without loss of generality we can assume $\xi_0 \neq 0$. Hence by §5.3 there exists a coordinate patch centered at x_0 and a generating function for Λ near p_0

$$\tilde{\psi} : U \times \mathbb{R}^n \to \mathbb{R}$$

having the form

$$\tilde{\psi}(x, y) = x \cdot y + \psi(y). \tag{5.31}$$

Then $C_{\tilde{\psi}}$ is the set

$$C_{\tilde{\psi}} : \quad x = -\frac{\partial \psi}{\partial y}$$

and

$$\lambda_\psi : C_{\tilde{\psi}} \to \Lambda$$

is the map

$$y \mapsto \left(-\frac{\partial \psi}{\partial y}, y \right).$$

Let Λ_1 be a Lagrangian submanifold which is horizontal and intersects Λ transversally at p_0. From Chapter 1 we know that $\Lambda_1 = \Lambda_\phi$ for some $\phi \in C^\infty(X)$, ie.e is the image of the map

$$U \ni x \mapsto \frac{\partial \phi}{\partial x}.$$

Since ϕ is a function on X and so does not involve any fiber variables, the section s_ϕ of $\mathbb{L}(\Lambda_1)$ associated with ϕ is the function $s_\phi = 1$. On the other hand, at every point $= \lambda_\psi(y) \in \Lambda$, the section of $\mathbb{L}_{\text{Maslov}}(\Lambda)$ associated with $\tilde{\psi}$ is the function

$$s_\psi = e^{\frac{\pi i}{4}\operatorname{sgn} d^2\psi}.$$

Let us now consider Λ and Λ_1 as canonical relations

$$\Lambda \in \operatorname{Morph}(\text{pt.}, T^*X), \quad \Lambda_1 \in \operatorname{Morph}(\text{pt.}, T^*X)$$

and consider the composition

$$\Lambda_1^\dagger \circ \Lambda \in \operatorname{Morph}(\text{pt.}, \text{pt.}). \tag{5.32}$$

Since composition of canonical relations corresponds to addition of their generating functions, we get a generating function

$$x \cdot y + \phi(x) + \psi(y)$$

for (5.32) with respect to the fibration

$$\mathbb{R}^{2n} \to \text{pt.}.$$

This has a critical point at $(x, y) = (x_0, \xi_0) = p_0$ and the composition for the sections $1 = s_\phi$ and s_ψ of the Maslov bundles $\mathbb{L}(\Lambda_1^\dagger)$ and $\mathbb{L}_{\text{Maslov}}(\Lambda)$ that we described in the preceding section gives us, for the composite section the element

$$e^{\frac{\pi i}{4}\operatorname{sgn} D} \in \mathbb{L}_{\text{pt.}.} = \mathbb{C} \tag{5.33}$$

where

$$D = \begin{pmatrix} A & I \\ I & B \end{pmatrix} \tag{5.34}$$

where

$$A = \left(\frac{\partial^2\psi}{\partial y_i \partial y_j}(\xi_0)\right)$$

and

$$B = \left(\frac{\partial^2\phi}{\partial x_i \partial x_j}(x_0)\right).$$

In particular, let us fix ϕ to be of the form

$$\phi(x) = \sum_i b_i x_i + \sum_{ij} b_{ij} x_i x_j$$

where the b_i are the coordinates of ξ_0 at $x = 0 = x_0$. Let us vary $B = (b_{ij})$ so that D stays non-degenerate which is the same as saying that Λ_ϕ stays transversal to Λ at p_0.

Let V be the tangent space at p_0 to the cotangent bundle of X, let M_1 be the tangent space to Λ at p_0 and let M_2 be the tangent space to the cotangent fiber $T_{x_0}X$ at p_0.

As we vary A, we get by (2.8) and (5.33) and (5.34) a map

$$f : \mathcal{L}(V, M_1, M_2) \to \mathbb{C}$$

satisfying the transformation law (2.17). Thus this function is an element of the Maslov line $\mathbb{L}_{\text{Maslov}}(p_0)$ that we defined in Section 2.8. Thus our composition formula (5.32) for $s_\phi \circ s_\psi$ gives us an identification of this line with the fiber of $\mathbb{L}(p_0)$ as defined in Section 5.13.3.

5.15 More examples of generating functions.

5.15.1 The image of a Lagrangian submanifold under geodesic flow.

Let X be a geodesically convex Riemannian manifold, for example $X = \mathbb{R}^n$. Let f_t denote geodesic flow on X. We know that for $t \neq 0$ a generating function for the symplectomorphism f_t is

$$\psi_t(x, y) = \frac{1}{2t} d(x, y)^2.$$

Let Λ be a Lagrangian submanifold of T^*X. Even if Λ is horizontal, there is no reason to expect that $f_t(\Lambda)$ be horizontal - caustics can develop. But our theorem about the generating function of the composition of two canonical relations will give a generating function for $f_t(\Lambda)$. Indeed, suppose that ϕ is a generating function for Λ relative to a fibration

$$\pi : X \times S \to X.$$

Then

$$\frac{1}{2} d(x, y)^2 + \psi(y, s)$$

is a generating function for $f_t(\Lambda)$ relative to the fibration

$$X \times X \times S \to X, \quad (x, y, s) \mapsto x.$$

5.15.2 The billiard map and its iterates.

Definition of the billiard map.

Let Ω be a bounded open convex domain in $\mathbb{R}^n$ with smooth boundary X. We may identify the tangent space to any point of $\mathbb{R}^n$ with $\mathbb{R}^n$ using the vector space structure, and identify $\mathbb{R}^n$ with $(\mathbb{R}^n)^*$ using the standard inner product. Then at any $x \in X$ we have the identifications

$$T_x X \cong T_x X^*$$

using the Euclidean scalar product on $T_x X$ and

$$T_x X = \{v \in \mathbb{R}^n |\ v \cdot n(x) = 0\} \tag{5.35}$$

where $n(x)$ denotes the inward pointing unit normal to X at x. Let $U \subset TX$ denote the open subset consisting of all tangent vectors (under the above identification) satisfying

$$\|v\| < 1.$$

For each $x \in X$ and $v \in T_xX$ satisfying $\|v\| < 1$ let

$$u := v + an(x) \quad \text{where } a := (1 - \|v\|^2)^{\frac{1}{2}}.$$

So u is the unique inward pointing unit vector at x whose orthogonal projection onto T_xX is v.

Consider the ray through x in the direction of u, i.e. the ray

$$x + tu, \quad t > 0.$$

Since Ω is convex and bounded, this ray will intersect X at a unique point y. Let w be the orthogonal projection of u on T_yX. So we have defined a map

$$\mathcal{B} : U \to U, \qquad (x, v) \mapsto (y, w)$$

which is known as the **billiard map**.

The generating function of the billiard map.

We shall show that the billiard map is a symplectomorphism by writing down a function ϕ which is its generating function.

Consider the function

$$\psi : \ \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \ \psi(x, y) = \|x - y\|.$$

This is smooth at all points (x, y), $x \neq y$. Let us compute $d_x\psi(v)$ at such a point (x, y) where $v \in T_xX$.

$$\frac{d}{dt}\psi(x + tv, y)_{|t=0} = \left(\frac{x - y}{\|y - x\|}, v\right)$$

where $(\ , \ )$ denotes the scalar product on $\mathbb{R}^n$. Identifying $T\mathbb{R}^n$ with $T^*\mathbb{R}^n$ using this scalar product, we can write that for all $x \neq y$

$$d_x\psi(x, y) = -\frac{y - x}{\|x - y\|}, \qquad d_y\psi(x, y) = \frac{y - x}{\|x - y\|}.$$

If we set

$$u = \frac{y - x}{\|x - y\|}, \quad t = \|x - y\|$$

we have

$$\|u\| = 1$$

and

$$y = x + tu.$$

Let ϕ be the restriction of ψ to $X \times X \subset \mathbb{R}^n \times \mathbb{R}^n$. Let

$$\iota : X \to \mathbb{R}^n$$

denote the embedding of X into $\mathbb{R}^n$. Under the identifications

$$T_x\mathbb{R}^n \cong T_x^*\mathbb{R}^n, \quad T_xX \cong T_x^*X$$

the orthogonal projection

$$T_x^*\mathbb{R}^n \cong T_x\mathbb{R}^n \ni u \mapsto v \in T_xX \cong T_x^*X$$

is just the map

$$d\iota_x^* : T_x^*\mathbb{R}^n \to T_x^*X, \quad u \mapsto v.$$

So

$$v = d\iota_x^* u = d\iota_x^* d_x\psi(x,y) = d_x\phi(x,y).$$

So we have verified the conditions

$$v = -d_x\phi(x,y), \qquad w = d_y\phi(x,y)$$

which say that ϕ is a generating function for the billiard map $\mathcal{B}$.

Iteration of the billiard map.

Our general prescription for the composite of two canonical relations says that a generating function for the composite is given by the sum of generating functions for each (where the intermediate variable is regarded as a fiber variable over the initial and final variables). Therefore a generating function for $\mathcal{B}^n$ is given by the function

$$\phi(x_0, x_1, \ldots, x_n) = \|x_1 - x_0\| + \|x_2 - x_1\| + \cdots + \|x_n - x_{n-1}\|.$$

5.15.3 The classical analogue of the Fourier transform.

We repeat a previous computation: Let $X = \mathbb{R}^n$ and consider the map

$$\mathfrak{F} : T^*X \to T^*X, \quad (x,\xi) \mapsto (-\xi, x).$$

The generating function for this symplectomorphism is

$$x \cdot y.$$

Since the transpose of the graph of a symplectomorphism is the graph of the inverse, the generating function for the inverse is

$$-y \cdot x.$$

So a generating function for the identity is

$$\phi \in C^\infty(X \times X, \times \mathbb{R}^n)$$

$$\phi(x,z,y) = (x - z) \cdot y.$$

5.15.4 Quadratic generating functions.

Reduced quadratic generating functions.

Let X and Y be vector spaces, $\pi : Y \to X$ a linear fibration and ϕ a homogenous quadratic generating function. The condition that ϕ be reduced says that the restriction of ϕ to the kernel of π vanishes. So let K be this kernel, i.e. we have the exact sequence

$$0 \to K \stackrel{\iota}{\to} Y \stackrel{\pi}{\to} X \to 0. \tag{5.36}$$

If $k \in K$ and $x \in X$, then $\phi(k, y)$ does not depend on the choice of y with $\pi y = x$, so we get a bilinear map

$$B: \; K \times X \to \mathbb{R}, \quad B(k, x) = \phi(k, y) \;\text{ where } \pi y = x.$$

We can consider B as a linear map

$$B: \; K \to X^*.$$

So $\operatorname{Im} B \subset X^*$ is a subspace of the (linear) Lagrangian subspace of $T^*X = X \oplus X^*$ determined by the generating function ϕ. The kernel of ϕ consists of "excess variables" so must vanish for the case that ϕ is transverse.

Let $W \subset X$ be the annihilator space of $\operatorname{Im} B$, i.e

$$W := (\operatorname{Im} B)^0.$$

Then the restriction on ϕ to $\pi^{-1}(W)$ depends only on the image of π, i.e. there is a quadratic form Q on W such that

$$Q(x_1, x_2) = \phi(y_1, y_2)$$

is independent of the choice of y_1, y_2 with $\pi Y_i = x_1, \; i = 1, 2$ when $x_1, x_2 \in W$.

Then

$$\Lambda = \Lambda_{W,Q} \oplus \operatorname{Im} B \tag{5.37}$$

where

$$\Lambda_{W,Q} = \{(x, dQ(x)), \quad x \in W\}$$

. In terms of coordinates, if $x_1, \ldots, x_k$ is a system of coordinates on W extended to a system of coordinates on X then Λ consists of all points of the form

$$\left(x_1, \ldots, x_k, 0, \ldots, 0; \frac{\partial Q}{\partial x_1}, \ldots, \frac{\partial Q}{\partial x_k}, \xi_{k+1} \ldots, \xi_n \right).$$

Reducing a homogeneous quadratic generating function.

More generally, consider the case where we have the exact sequence (5.36) and a homogeneous function quadratic function ϕ on Y, and hence a linear map

$$L_\phi : Y \to Y^*$$

such that
$$\phi(y) = \frac{1}{2}\langle L_\phi y, y\rangle.$$
Our general definition of generating function restricted to the case of homogeneous quadratic functions says that we first pass to the critical set which in this case corresponds to the subspace $C_\phi \subset Y$
$$C_\phi = \ker(\iota^* \circ L_\phi).$$
Taking the transpose of (5.36) we see that π^* is injective and $\ker \iota^* = \operatorname{Im} \pi^*$. Since $L_\phi(C_\phi) \subset \ker \iota^*$ we see that L_ϕ maps $C_\phi \to X^*$.

The general definition of a generating function then specializes in this case to the assertion that
$$\Lambda = \rho_\phi(C_\phi)$$
where $\rho_\phi : C_\phi \to TX^* = X \oplus X^*$ is given by
$$\rho_\phi(u) = (\pi(u), L_\phi(u)).$$
Let
$$K_0 := K \cap C_\phi,$$
so K_0 is the null space of the restriction on ϕ to K, i.e. $K_0 = K^\perp$ relative to the quadratic form $\phi \circ \iota$ on K.

In terms of the preceding paragraph, we know that ϕ is reduced if and only if $K_0 = K$.

Example: When Λ is transverse to X. Recall from Chapter 2 that in this case we can take $Y = X \oplus X^*$ so that $K = X^*$ and ϕ to have the form
$$\phi(x,\xi) = \langle \xi, x\rangle - P(\xi)$$
where P is a quadratic function on X^*. Let $L_P : X^* \to (X^*)^* = X$ be the linear map associated to P. We have $Y^* = X^* \oplus X$ and L_ϕ is given by
$$L_\phi(x,\xi) = (\xi, x - L_P(\xi)).$$
Hence
$$(\iota^* \circ L_\phi)(x,\xi) = x - L_P(\xi)$$
so that
$$C_\phi = \{(x,\xi)|x = L_\xi(\xi)\}.$$
The generating function ϕ in this case will be reduced if and only if $P \equiv 0$ in which case $\Lambda = X^*$ and $C_\phi = Y$.

If $P \not\equiv 0$ we may "reduce" the number of fiber variables by replacing Y by $Y_0 = C_\phi$. We then get the exact sequence
$$0 \to \ker L_P \to Y_0 \to X \to 0$$
which has the form (5.36) and (5.37) becomes
$$\Lambda = \{L_P(\xi), \xi).$$

Reduction.

In general, the quadratic form induced by $\phi \circ \iota$ on K/K_0 is non-degenerate. In particular, the restriction of $L_\phi \circ \iota$ to any complement K_1 of K_0 in K maps this complement surjectively onto $(K_0)^0 \subset K^*$, the null space of K_0, and from linear algebra, $(K_0)^0 = (K/K_0)^*$.

Let ι_1 denote the restriction of ι to K_1 and let $Y_0 = \ker(\iota_1^* \circ L_\phi)$. Clearly $\iota(K_0) \subset Y_0$.

Lemma 5.15.1. *$\pi_{|Y_0}$ maps Y_0 surjectively onto X.*

Proof. Let $x \in X$. Let $y \in Y$ be such that $\pi y = x$. Let $k^* = (\iota^* \circ L_\phi)(y)$. We can find a $k \in K_1$ such that $(L_\phi \circ \iota)(k) = k^*$. Then $y - \iota(k) \in Y_0$ and $\pi(y - \iota(k)) = x$. □

Let
$$\phi_0 := \phi_{|Y_0} \quad \iota_0 := \iota_{K_0}, \quad \text{and} \quad \pi_0 := \pi_{Y_0}.$$
So we have the exact sequence
$$0 \to K_0 \stackrel{\iota_0}{\to} Y_0 \stackrel{\pi_0}{\to} X \to 0. \tag{5.38}$$
If $y \in C_\phi$ then by definition, $\iota^* L_\phi(y) = 0$, so in the proof of the above lemma, we do not need to modify y. Hence

Proposition 5.15.1. *The sequence (5.38) is exact and and the function ϕ_0 is a reduced generating function for Λ.*

Chapter 6

The calculus of $\frac{1}{2}$-densities.

An essential ingredient in our symbol calculus will be the notion of a $\frac{1}{2}$- density on a canonical relation. We begin this chapter with a description of densities of arbitrary order on a vector space, then on a manifold, and then specialize to the study of $\frac{1}{2}$-densities. We study $\frac{1}{2}$-densities on canonical relations in the next chapter.

6.1 The linear algebra of densities.

6.1.1 The definition of a density on a vector space.

Let V be an n-dimensional vector space over the real numbers. A basis $\mathbf{e} = e_1, \dots, e_n$ of V is the same as an isomorphism $\ell_{\mathbf{e}}$ of $\mathbb{R}^n$ with V according to the rule

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto x_1 e_1 + \cdots + x_n e_n.$$

We can write this as

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto (e_1, \dots e_n) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

or even more succinctly as

$$\ell_{\mathbf{e}} : \quad \mathbf{x} \mapsto \mathbf{e} \cdot \mathbf{x}$$

where

$$\mathbf{x} := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad \mathbf{e} := (e_1, \dots, e_n).$$

The group $Gl(n) = Gl(n, \mathbb{R})$ acts on the set $\mathcal{F}(V)$ of all bases of V according to the rule

$$\ell_{\mathbf{e}} \mapsto \ell_{\mathbf{e}} \circ A^{-1}, \quad A \in Gl(n)$$

which is the same as the "matrix multiplication"

$$\mathbf{e} \mapsto \mathbf{e} \cdot A^{-1}.$$

This action is effective and transitive:

- If $\mathbf{e} = \mathbf{e} \cdot A^{-1}$ for some basis $\mathbf{e}$ then $A = I$, the identity matrix, and
- Given any two bases $\mathbf{e}$ and $\mathbf{f}$ these exists a (unique) A such that $\mathbf{e} = \mathbf{f} \cdot A$.

We shall use the word **frame** as being synonymous with the word "basis", especially when we want to talk of a basis with a particular property.

Let $\alpha \in \mathbb{C}$ be any complex number. A **density of order** α on V is a function

$$\rho : \mathcal{F}(V) \to \mathbb{C}$$

satisfying

$$\rho(\mathbf{e} \cdot A) = \rho(\mathbf{e}) |\det A|^{\alpha} \quad \forall A \in Gl(n), \ \mathbf{e} \in \mathcal{F}(V). \tag{6.1}$$

We will denote the space of all densities of order α on V by

$$|V|^{\alpha}.$$

This is a one dimensional vector space over the complex numbers. Indeed, if we fix one $\mathbf{f} \in \mathcal{F}(V)$, then every $\mathbf{e} \in \mathcal{F}(V)$ can be written uniquely as $\mathbf{e} = \mathbf{f} \cdot B, \ B \in Gl(n)$. So we may specify $\rho(f)$ to be any complex value and then define $\rho(e)$ to be $\rho(f) \cdot |\det B|^{\alpha}$. It is then easy to check that (6.1) holds. This shows that densities of order α exist, and since we had no choice once we specified $\rho(\mathbf{f})$ we see that the space of densities of order α on V form a one dimensional vector space over the complex numbers.

Let $L : V \to V$ be a linear map. If L is invertible and $\mathbf{e} \in \mathcal{F}(V)$ then $L\mathbf{e} = (Le_1, \ldots, Le_n)$ is (again) a basis of V. If we write

$$Le_j = \sum_i L_{ij} e_i$$

then

$$L\mathbf{e} = \mathbf{e}\mathrm{L}$$

where L is the matrix

$$\mathrm{L} := (L_{ij})$$

so if $\rho \in |V|^{\alpha}$ then

$$\rho(L\mathbf{e}) = |\det L|^{\alpha} \rho(\mathbf{e}).$$

We can extend this to all L, non necessarily invertible, where the right hand side is 0. So here is an equivalent definition of a density of order α on an n-dimensional real vector space:

A density ρ of order α is a rule which assigns a number $\rho(v_1,\ldots,v_n)$ to every n-tuplet of vectors and which satisfies

$$\rho(Lv_1,\ldots,Lv_n) = |\det L|^\alpha \rho(v_1,\ldots,v_n) \tag{6.2}$$

for any linear transformation $L : V \to V$. Of course, if the $v_1,\ldots,v_n$ are not linearly independent then

$$\rho(v_1,\ldots,v_n) = 0.$$

6.1.2 Multiplication.

If $\rho \in |V|^\alpha$ and $\tau \in |V|^\beta$ then we get a density $\rho\cdot\tau$ of order $\alpha+\beta$ given by

$$(\rho\cdot\tau)(\mathbf{e}) = \rho(\mathbf{e})\tau(\mathbf{e}).$$

In other words we have an isomorphism:

$$|V|^\alpha \otimes |V|^\beta \cong |V|^{\alpha+\beta}, \quad \rho\otimes\tau \mapsto \rho\cdot\tau. \tag{6.3}$$

6.1.3 Complex conjugation.

If $\rho \in |V|^\alpha$ then $\overline{\rho}$ defined by

$$\overline{\rho}(\mathbf{e}) = \overline{\rho(\mathbf{e})}$$

is a density of order $\overline{\alpha}$ on V. In other words we have an anti-linear map

$$|V|^\alpha \to |V|^{\overline{\alpha}}, \qquad \rho \mapsto \overline{\rho}.$$

This map is clearly an anti-linear isomorphism. Combined with (6.3) we get a sesquilinear map

$$|V|^\alpha \otimes |V|^\beta \to |V|^{\alpha+\overline{\beta}}, \rho\otimes\tau \mapsto \rho\cdot\overline{\tau}.$$

We will especially want to use this for the case $\alpha = \beta = \frac{1}{2} + is$ where s is a real number. In this case we get a sesquilinear map

$$|V|^{\frac{1}{2}+is} \otimes |V|^{\frac{1}{2}+is} \to |V|^1. \tag{6.4}$$

6.1.4 Elementary consequences of the definition.

There are two obvious but very useful facts that we will use repeatedly:

1. An element of $|V|^\alpha$ is completely determined by its value on a single basis $\mathbf{e}$.

2. More generally, suppose we are given a subset S of the set of bases on which a subgroup $H \subset Gl(n)$ acts transitively and a function $\rho : S \to \mathbb{C}$ such that (6.1) holds for all $A \in H$. Then ρ extends uniquely to a density of order α on V.

Here are some typical ways that we will use these facts:

Orthonormal frames: Suppose that V is equipped with a scalar product. This picks out a subset $\mathcal{O}(V) \subset \mathcal{F}(V)$ consisting of the orthonormal frames. The corresponding subgroup of $Gl(n)$ is $O(n)$ and every element of $O(n)$ has determinant ± 1. So any density of any order must take on a constant value on orthonormal frames, and item 2 above implies that any constant then determines a density of any order. We have trivialized the space $|V|^\alpha$ for all α. Another way of saying the same thing is that V has a preferred density of order α, namely the density which assigns the value one to any orthonormal frame. The same applies if V has any non-degenerate quadratic form, not necessarily positive definite.

Symplectic frames: Suppose that V is a symplectic vector space, so $n = \dim V = 2d$ is even. This picks out a collection of preferred bases, namely those of the form $e_1, \dots, e_d,\ f_1, \dots f_d$ where

$$\omega(e_i, e_j) = 0, \quad \omega(f_i, f_j) = 0. \quad \omega(e_i, f_j) = \delta_{ij}$$

where ω denotes the symplectic form. These are known as the symplectic frames. In this case $H = Sp(n)$ and every element of $Sp(n)$ has determinant one. So again $|V|^\alpha$ is trivialized. Again, another way of saying this is that a symplectic vector space has a preferred density of any order - the density which assigns the value one to any symplectic frame.

Transverse Lagrangian subspaces: Suppose that V is a symplectic vector space and that M and N are Lagrangian subspaces of V with $M \cap N = \{0\}$. *Any* basis $e_1, \dots e_d$ of M determines a dual basis $f_1, \dots f_d$ of N according to the requirement that

$$\omega(e_i, f_j) = \delta_{ij}$$

and then $e_1, \dots e_d, f_1 \dots f_d$ is a symplectic basis of V. If $C \in Gl(d)$ and we make the replacement

$$\mathbf{e} \mapsto \mathbf{e} \cdot C$$

then we must make the replacement

$$\mathbf{f} \mapsto \mathbf{f} \cdot (C^t)^{-1}.$$

So if ρ is a density of order α on M and τ is a density of order α on N they fit together to get a density of order zero (i.e. a constant) on V according to the rule

$$(\mathbf{e}, \mathbf{f}) = (e_1, \dots, e_d,\ f_1, \dots, f_d) \mapsto \rho(\mathbf{e})\tau(\mathbf{f})$$

on frames of the above dual type. The corresponding subgroup of $Gl(n)$ is a subgroup of $Sp(n)$ isomorphic to $Gl(d)$. So we have a canonical isomorphism

$$|M|^\alpha \otimes |N|^\alpha \cong \mathbb{C}. \tag{6.5}$$

Using (6.3) we can rewrite this as

$$|M|^\alpha \cong |N|^{-\alpha}.$$

Dual spaces: If we start with a vector space M we can make $M \oplus M^*$ into a symplectic vector space with M and M^* transverse Lagrangian subspaces and the pairing B between M and M^* just the standard pairing of a vector space with its dual space. So making a change in notation we have

$$|V|^\alpha \cong |V^*|^{-\alpha}. \tag{6.6}$$

Short exact sequences: Let

$$0 \to V' \to V \to V'' \to 0$$

be an exact sequence of linear maps of vector spaces. We can choose a preferred set of bases of V as follows : Let $(e_1, \dots, e_k)$ be a basis of V' and extend it to a basis $(e_1, \dots, e_k, e_{k+1}, \dots e_n)$ of V. Then the images of $e_i,\ i = k+1, \dots n$ under the map $V \to V''$ form a basis of V''. Any two bases of this type differ by the action of an $A \in Gl(n)$ of the form

$$A = \begin{pmatrix} A' & * \\ 0 & A'' \end{pmatrix}$$

so

$$\det A = \det A' \cdot \det A''.$$

This shows that we have an isomorphism

$$|V|^\alpha \cong |V'|^\alpha \otimes |V''|^\alpha \tag{6.7}$$

for any α.

Long exact sequences Let

$$0 \to V_1 \to V_2 \to \cdots V_k \to 0$$

be an exact sequence of vector spaces. Then using (6.7) inductively we get

$$\bigotimes_{j \text{ even}} |V_j|^\alpha \cong \bigotimes_{j \text{ odd}} |V_j|^\alpha \tag{6.8}$$

for any α.

6.1.5 Pullback and pushforward under isomorphism.

Let

$$L : V \to W$$

be an isomorphism of n-dimensional vector spaces. If

$$\mathbf{e} = (e_1, \dots, e_n)$$

is a basis of V then

$$L\mathbf{e} := (Le_1, \dots, Le_n)$$

is a basis of W and

$$L(\mathbf{e}\cdot A) = (L\mathbf{e})\cdot A \quad \forall A \in Gl(n).$$

So if $\rho \in |W|^\alpha$ then $L^*\rho$ defined by

$$(L^*\rho)(\mathbf{e}) := \rho(L\mathbf{e})$$

is an element of $|V|^\alpha$. In other words we have a **pullback isomorphism**

$$L^* : |W|^\alpha \to |V|^\alpha, \quad \rho \mapsto L^*\rho.$$

Applied to L^{-1} this gives a **pushforward isomorphism**

$$L_* : |V|^\alpha \to |W|^\alpha, \quad L_* = (L^{-1})^*.$$

6.1.6 Pairs of Lagrangian subspaces.

Here is another useful fact:

Let ℓ_1, ℓ_2 be Lagrangian subspaces of a symplectic vector space. We have the following two exact sequences:

$$0 \to \ell_1 \cap \ell_2 \to \ell_1 + \ell_2 \to (\ell_1 + \ell_2)/(\ell_1 \cap \ell_2) \to 0$$

and

$$0 \to \ell_1 \cap \ell_2 \to \ell_1 \oplus \ell_2 \to \ell_1 + \ell_2 \to 0.$$

Since $(\ell_1 + \ell_2)/(\ell_1 \cap \ell_2)$ is a symplectic vector space, the first exact sequence tells us that

$$|\ell_1 + \ell_2|^\alpha \sim |\ell_1 \cap \ell_2|^\alpha$$

and so the second exact sequence tells us that

$$|\ell_1|^\alpha \otimes |\ell_2|^\alpha \sim |\ell_1 \cap \ell_2|^{2\alpha}. \tag{6.9}$$

6.1.7 Spanning pairs of subspaces of a symplectic vector space.

Let M_1 and M_2 be (arbitrary) subspaces of a symplectic vector space V with the property that

$$M_1 + M_2 = V.$$

We then have the exact sequence

$$0 \to M_1 \cap M_2 \to M_1 \oplus M_2 \to V \to 0.$$

Since we have the trivialization $|V|^\alpha \cong \mathbb{C}$ determined by the symplectic structure, we get an isomorphism

$$|M_1|^\alpha \otimes |M_2|^\alpha \cong |M_1 \cap M_2|^\alpha. \tag{6.10}$$

6.1.8 Lefschetz symplectic linear transformations.

There is a special case of (6.5) which we will use a lot in our applications, so we will work out the details here. A linear map $L : V \to V$ on a vector space is called **Lefschetz** if it has no eigenvalue equal to 1. Another way of saying this is that $I - L$ is invertible. Yet another way of saying this is the following: Let

$$\operatorname{graph} L \subset V \oplus V$$

be the graph of L so

$$\operatorname{graph} L = \{(v, Lv) \quad v \in V\}.$$

Let

$$\Delta \subset V \oplus V$$

be the diagonal, i.e. the graph of the identity transformation. Then L is Lefschetz if and only if

$$\operatorname{graph} L \cap \Delta = \{0\}. \tag{6.11}$$

Now suppose that V is a symplectic vector space and we consider $V^- \oplus V$ as a symplectic vector space. Suppose also that L is a (linear) symplectic transformation so that graph L is a Lagrangian subspace of $V^- \oplus V$ as is Δ. Suppose that L is also Lefschetz so that (6.11) holds.

The isomorphism

$$V \to \operatorname{graph} L : \qquad v \mapsto (v, Lv)$$

pushes the canonical α-density on V to an α-density on graph L, namely, if $v_1, \dots, v_n$ is a symplectic basis of V, then this pushforward α density assigns the value one to the basis

$$((v_1, Lv_1), \dots, (v_n, Lv_n)) \quad \text{of } \operatorname{graph} L.$$

Let us call this α-density ρ_L. Similarly, we can use the map

$$\operatorname{diag} : V \to \Delta, \qquad v \mapsto (v, v)$$

to push the canonical α density to an α-density ρ_Δ on Δ. So ρ_Δ assigns the value one to the basis

$$((v_1, v_1), \dots, (v_n, v_n)) \qquad \text{of } \Delta.$$

According to (6.5)

$$|\operatorname{graph} L|^\alpha \otimes |\Delta|^\alpha \cong \mathbb{C}.$$

So we get a number $\langle \rho_L, \rho_\Delta \rangle$ attached to these two α-densities. We claim that

$$\langle \rho_L, \rho_\Delta \rangle = |\det(I - L)|^{-\alpha}. \tag{6.12}$$

Before proving this formula, let us give another derivation of (6.5). Let M and N be subspaces of a symplectic vector space W. (The letter V is currently overworked.) Suppose that $M\cap N=\{0\}$ so that $W=M\oplus N$ as a vector space and so by (6.7) we have

$$|W|^\alpha = |M|^\alpha\otimes|N|^\alpha.$$

We have an identification of $|W|^\alpha$ with $\mathbb{C}$ given by sending

$$|W|^\alpha \ni \rho_W \mapsto \rho_W(\mathbf{w})$$

where $\mathbf{w}$ is any symplectic basis of W. Combining the last two equations gives an identification of $|M|^\alpha\otimes|N|^\alpha$ with $\mathbb{C}$ which coincides with (6.5) in case M and N are Lagrangian subspaces. Put another way, let $\mathbf{w}$ be a symplectic basis of W and suppose that $A\in Gl(\dim W)$ is such that

$$\mathbf{w}\cdot A = (\mathbf{m},\mathbf{n})$$

where $\mathbf{m}$ is a basis of M and $\mathbf{n}$ is a basis of N. Then the pairing of of $\rho_M\in|M|^\alpha$ with $\rho_N\in|N|^\alpha$ is given by

$$\langle\rho_M,\rho_N\rangle = |\det A|^{-\alpha}\rho_M(\mathbf{m})\rho_N(\mathbf{n}). \tag{6.13}$$

Now let us go back to the proof of (6.12). If $\mathbf{e},\mathbf{f}=e_1,\dots,e_d,f_1\dots,f_d$ is a symplectic basis of V then

$$((\mathbf{e},0),(0,\mathbf{e}),(-\mathbf{f},0),(0,\mathbf{f}))$$

is a symplectic basis of $V^-\oplus V$. We have

$$((\mathbf{e},0),(0,\mathbf{e}),(-\mathbf{f},0),(0,\mathbf{f}))\begin{pmatrix} I_d & 0 & 0 & 0\\ 0 & 0 & I_d & 0\\ 0 & -I_d & 0 & 0\\ 0 & 0 & 0 & I_d\end{pmatrix} = ((\mathbf{e},0),(\mathbf{f},0),(0,\mathbf{e}),(0,\mathbf{f}))$$

and

$$\det\begin{pmatrix} I_d & 0 & 0 & 0\\ 0 & 0 & I_d & 0\\ 0 & -I_d & 0 & 0\\ 0 & 0 & 0 & I_d\end{pmatrix} = 1.$$

Let $\mathbf{v}$ denote the symplectic basis $\mathbf{e},\mathbf{f}$ of V so that we may write

$$((\mathbf{e},0),(\mathbf{f},0),(0,\mathbf{e}),(0,\mathbf{f})) = ((\mathbf{v},0),(0,\mathbf{v}))\,.$$

Write

$$\mathrm{L}v_j = \sum_i L_{ij}v_i,\quad \mathrm{L}=(L_{ij}).$$

Then

$$((\mathbf{v},0),(0,\mathbf{v}))\begin{pmatrix} I_d & I_d\\ \mathrm{L} & I_d\end{pmatrix} = ((\mathbf{v},\mathrm{L}\mathbf{v}),(\mathbf{v},\mathbf{v}))\,.$$

So taking

$$A = \begin{pmatrix} I_d & 0 & 0 & 0 \\ 0 & 0 & I_d & 0 \\ 0 & -I_d & 0 & 0 \\ 0 & 0 & 0 & I_d \end{pmatrix} \begin{pmatrix} I_n & I_n \\ \mathrm{L} & I_n \end{pmatrix}$$

we have

$$((\mathbf{e},0),(0,\mathbf{e}),(-\mathbf{f},0),(0,\mathbf{f}))\,A = ((\mathbf{v},L\mathbf{v}),(\mathbf{v},\mathbf{v}))\,.$$

So using this A in (6.13) proves (6.12) since

$$\det A = \det \begin{pmatrix} I_d & 0 & 0 & 0 \\ 0 & 0 & I_d & 0 \\ 0 & -I_d & 0 & 0 \\ 0 & 0 & 0 & I_d \end{pmatrix} \det \begin{pmatrix} I_n & I_n \\ \mathrm{L} & I_n \end{pmatrix} = \det(I_n - L).$$

We will now generalize (6.12). Let $L : V \to V$ be a linear symplectic map, and suppose that its fixed point set

$$U = V^L := \{v \in V | Lv = v\}$$

is a symplectic subspace of V, and let $U^\perp$ be its symplectic orthocomplement. So $U^\perp$ is invariant under L and is a symplectic subspace of V.

The decomposition $V = U \oplus U^\perp$ gives rise to the deccompositions

$$\Delta = \Delta_U \oplus \Delta_{U^\perp} \quad \text{and} \tag{6.14}$$

$$\operatorname{graph} L = \Delta_U \oplus \operatorname{graph}(L|U^\perp) \tag{6.15}$$

as Lagrangian subspaces of $U^- \oplus U$ and $(U^\perp)^- \oplus U^\perp$.

Let ρ_Δ and ρ_L be the elements of $|\Delta|^\alpha$ and $|\operatorname{graph} L|^\alpha$ as determined above from the canonical α densities on V. Then (6.14) and (6.15) imply that we can write

$$\rho_\Delta = \sigma_\Delta \otimes \tau_\Delta \tag{6.16}$$

$$\rho_L = \sigma_L \otimes \tau_L \tag{6.17}$$

with σ_Δ and $\sigma_L \in |\Delta_U|^\alpha$, with $\tau_\Delta \in |\Delta_{U^\perp}|^\alpha$ and and $\tau_L \in |\operatorname{graph}(L_{|U^\perp})|^\alpha$. Furthermore, we may identify Δ_U with U, which, by hypothesis, is a symplectic vector space and so carries a canonical density of order α. We may take σ_Δ and σ_L to be this canonical density of order α which then fixes τ_Δ and τ_L.

Now Δ and $\operatorname{graph} L$ are Lagrangian subspaces of $V^- \oplus V$ and their intersection is Δ_U which we identify with U. The isomorphism (6.9) gives us a map sending $\rho_\Delta \otimes \rho_L$ into $|U|^{2\alpha}$. From (6.16) and (6.17) we see that the image of (6.9)
$\rho_\Delta \otimes \rho_L$ is

$$|du|^{2\alpha}\langle \tau_L, \tau_\Delta \rangle$$

where $|du|^{2\alpha}$ is the canonical 2α density on U. Since the restriction of L to $U^\perp$ is Lefschetz, we may apply (6.12) to conclude (6.11)

Theorem 6.1.1. *If the fixed point set U of L is a symplectic subspace, then the isomorphism (6.9) determines a pairing sending the α density ρ_Δ on Δ and the α density ρ_L on* graph L *into 2α densities on U given by*

$$\langle \rho_L, \rho_\Delta \rangle = \left|\det(I_{U^\perp} - L_{|U^\perp}\right|^{-\alpha} du^{2\alpha} \tag{6.18}$$

where $du^{2\alpha}$ is the canonical 2α density on U determined by its symplectic structure.

6.2 Densities on manifolds.

Let $E \to X$ be a real vector bundle. We can then consider the complex line bundle

$$|E|^\alpha \to X$$

whose fiber over $x \in X$ is $|E_x|^\alpha$. The formulas of the preceding section apply pointwise.

We will be primarily interested in the tangent bundle TX. So $|TX|^\alpha$ is a complex line bundle which we will call the α-density bundle and a smooth section of $|TX|^\alpha$ will be called a smooth α-**density** or a **density of order** α.

Examples.

- Let $X = \mathbb{R}^n$ with its standard coordinates and hence the standard vector fields
$$\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}.$$
This means that at each point $p \in \mathbb{R}^n$ we have a preferred basis
$$\left(\frac{\partial}{\partial x_1}\right)_p, \ldots, \left(\frac{\partial}{\partial x_n}\right)_p.$$
We let
$$dx^\alpha$$
denote the α-density which assigns, at each point p, the value 1 to the above basis. So the most general smooth α-density on $\mathbb{R}^n$ can be written as
$$u \cdot dx^\alpha$$
or simply as
$$udx^\alpha$$
where u is a smooth function.

- Let X be an n-dimensional Riemannian manifold. At each point p we have a preferred family of bases of the tangent space - the orthonormal bases. We thus get a preferred density of order α - the density which assigns the value one to each orthonormal basis at each point.

- Let X be an n-dimensional orientable manifold and Ω a nowhere vanishing n-form on X. Then we get an α-density according to the rule: At each $p \in X$ assign to each basis $e_1, \ldots, e_n$ of T_pX the value

$$|\Omega(e_1, \ldots, e_n)|^\alpha.$$

We will denote this density by

$$|\Omega|^\alpha.$$

- As a special case of the preceding example, if M is a symplectic manifold of dimension $2d$ with symplectic form ω, take

$$\Omega = \omega \wedge \cdots \wedge \omega \qquad d \text{ factors.}$$

So every symplectic manifold has a preferred α-density for any α.

6.2.1 Multiplication of densities.

If μ is an α density and ν is a β density the we can multiply them (pointwise) to obtain an $(\alpha+\beta)$-density $\mu \cdot \nu$. Similarly, we can take the complex conjugate of an α-density to obtain an $\overline{\alpha}$-density.

6.2.2 Support of a density.

Since a density is a section of a line bundle, it makes sense to say that a density *is* or *is not* zero at a point. The **support** of a density is defined to be the closure of the set of points where it is *not* zero.

6.3 Pull-back of a density under a diffeomorphism.

If

$$f : X \to Y$$

is a diffeomorphism, then we get, at each $x \in X$, a linear isomorphism

$$df_x : T_xX \to T_{f(x)}Y.$$

A density ν of order α on Y assigns a density of order α (in the sense of vector spaces) to each T_yY which we can then pull back using df_x to obtain a density of order α on X. We denote this pulled back density by $f^*\nu$. For example, suppose that

$$\nu = |\Omega|^\alpha$$

for an n-form Ω on Y (where $n = \dim Y$). Then

$$f^*|\Omega|^\alpha = |f^*\Omega|^\alpha \tag{6.19}$$

where the $f^*\Omega$ occurring on right hand side of this equation is the usual pull-back of forms.

As an example, suppose that X and Y are open subsets of $\mathbb{R}^n$, then

$$dx^\alpha = |dx_1 \wedge \cdots \wedge dx_n|^\alpha, \qquad |dy|^\alpha = |dy_1 \wedge \cdots \wedge dy_n|^\alpha$$

and

$$f^*(dy_1 \wedge \cdots \wedge dy_n) = \det J(f) dx_1 \wedge \cdots \wedge dx_n$$

where $J(f)$ is the Jacobian matrix of f. So

$$f^* dy^\alpha = |\det J(f)|^\alpha dx^\alpha. \tag{6.20}$$

Here is a second application of (6.19). Let $f_t : X \to X$ be a one-parameter group of diffeomorphisms generated by a vector field v, and let ν be a density of order α on X. As usual, we define the Lie derivative $D_v\nu$ by

$$D_v\nu := \frac{d}{dt} f_t^* \nu_{|t=0}.$$

If $\nu = |\Omega|^\alpha$ then

$$D_v\nu = \alpha D_v|\Omega| \cdot |\Omega|^{\alpha-1}$$

and if X is oriented, then we can identify $|\Omega|$ with Ω on oriented bases, so

$$D_v|\Omega| = D_v\Omega = di(v)\Omega$$

on oriented bases. For example,

$$D_v dx^{\frac{1}{2}} = \frac{1}{2}(\text{div } v) dx^{\frac{1}{2}} \tag{6.21}$$

where

$$\text{div } v = \frac{\partial v_1}{\partial x_1} + \cdots + \frac{\partial v_n}{\partial x_n} \quad \text{if} \quad v = v_1 \frac{\partial}{\partial x_1} + \cdots + v_n \frac{\partial}{\partial x_n}.$$

6.4 Densities of order 1.

If we set $\alpha = 1$ in (6.20) we get

$$f^* dy = |\det J(f)| dx$$

or, more generally,

$$f^*(u dy) = (u \circ f) |\det J(f)| dx$$

which is the change of variables formula for a multiple integral. So if ν is a density of order one of compact support which is supported on a coordinate patch $(U, x_1, \ldots, x_n)$, and we write

$$\nu = g dx$$

then

$$\int \nu := \int_U g dx$$

is independent of the choice of coordinates. If ν is a density of order one of compact support we can use a partition of unity to break it into a finite sum of densities of order one and of compact support contained in coordinate patches

$$\nu = \nu_1 + \cdots + \nu_r$$

and $\int_X \nu$ defined as

$$\int_X \nu := \int \nu_1 + \cdots + \int \nu_r$$

is independent of all choices. In other words densities of order one (usually just called densities) are objects which can be integrated (if of compact support). Furthermore, if

$$f : X \to Y$$

is a diffeomorphism, and ν is a density of order one of compact support on Y, we have the general "change of variables formula"

$$\int_X f^* \nu = \int_Y \nu. \tag{6.22}$$

Suppose that α and β are complex numbers with

$$\alpha + \overline{\beta} = 1.$$

Suppose that μ is a density of order α and ν is a density of order β on X and that one of them has compact support. Then $\mu \cdot \overline{\nu}$ is a density of order one of compact support. So we can form

$$\langle \mu, \nu \rangle := \int_X \mu \overline{\nu}.$$

So we get an intrinsic sesquilinear pairing between the densities of order α of compact support and the densities of order $1 - \overline{\alpha}$.

6.5 The principal series representations of Diff(X).

So if $s \in \mathbb{R}$, we get a pre-Hilbert space structure on the space of smooth densities of compact support of order $\frac{1}{2} + is$ given by

$$(\mu, \nu) := \int_X \mu \overline{\nu}.$$

If $f \in \text{Diff}(X)$, i.e. if $f : X \to X$ is a diffeomorphism, then

$$(f^* \mu, f^* \nu) = (\mu, \nu)$$

and

$$(f \circ g)^* = g^* \circ f^*.$$

Let $\mathfrak{H}_s$ denote the completion of the pre-Hilbert space of densities of order $\frac{1}{2}+is$. The Hilbert space $\mathfrak{H}_s$ is known as the **intrinsic Hilbert space of order** s. The map

$$f \mapsto (f^{-1})^*$$

is a representation of Diff (X) on the space of densities or order $\frac{1}{2} + is$ which extends by completion to a unitary representation of Diff (X) on $\mathfrak{H}_s$. This collection of representations (parametrized by s) is known as the principal series of representations.

If we take $S = S^1 = \mathbb{PR}^1$ and restrict the above representations of Diff(X) to $G = PL(2, \mathbb{R})$ we get the principal series of representations of G.

We will concentrate on the case $s = 0$, i.e. we will deal primarily with densities of order $\frac{1}{2}$.

6.6 The push-forward of a density of order one by a fibration.

There is an important generalization of the notion of the integral of a density of compact support: Let

$$\pi : Z \to X$$

be a *proper* fibration. Let μ be a density of order one on Z. We are going to define

$$\pi_* \mu$$

which will be a density of order one on X. We proceed as follows: for $x \in X$, let

$$F = F_x := \pi^{-1}(x)$$

be the fiber over x. Let $z \in F$. We have the exact sequence

$$0 \to T_z F \to T_z Z \stackrel{d\pi_z}{\to} T_x X \to 0$$

which gives rise to the isomorphism

$$|T_z F| \otimes |T_x X| \cong |T_z Z|.$$

The density μ thus assigns to each z in the manifold F an element of

$$|T_z F| \otimes |T_x X|.$$

In other words, on the manifold F it is a density of order one with values in the fixed one dimensional vector space $|T_x X|$. Since F is compact, we can integrate this density over F to obtain an element of $|T_x X|$. As we do this for all x, we have obtained a density of order one on X.

Let us see what the operation $\mu \mapsto \pi_*\mu$ looks like in local coordinates. Let us choose local coordinates $(U, x_1, \ldots, x_n, s_1 \ldots, s_d)$ on Z and coordinates $y_1, \ldots, y_n$ on X so that

$$\pi : (x_1, \ldots, x_n, s_1, \ldots, s_d) \mapsto (x_1, \ldots, x_n).$$

Suppose that μ is supported on U and we write

$$\mu = udxds = u(x_1, \ldots, x_n, s_1 \ldots, s_d)dx_1 \ldots dx_n ds_1 \ldots ds_d.$$

Then

$$\pi_*\mu = \left(\int u(x_1, \ldots, x_n, s_1, \ldots, s_d) ds_1 \ldots ds_d \right) dx_1 \ldots dx_n. \tag{6.23}$$

In the special case that X is a point, $\pi_*\mu = \int_Z \mu$. Also, Fubini's theorem says that if

$$W \xrightarrow{\rho} Z \xrightarrow{\pi} X$$

are fibrations with compact fibers then

$$(\pi \circ \rho)_* = \pi_* \circ \rho_*. \tag{6.24}$$

In particular, if μ is a density of compact support on Z with $\pi : Z \to X$ a fibration then $\pi_*\mu$ is defined and

$$\int_X \pi_*\mu = \int_Z \mu. \tag{6.25}$$

If f is a C^∞ function on X of compact support and $\pi : Z \to X$ is a proper fibration then $\pi^* f$ is constant along fibers and (6.25) says that

$$\int_Z \pi^* f \mu = \int_X f \pi_* \mu. \tag{6.26}$$

In other words, the operations

$$\pi^* : C_0^\infty(X) \to C_0^\infty(Z)$$

and

$$\pi_* : C^\infty(|TZ|) \to C^\infty(|TX|)$$

are transposes of one another.

Chapter 7

The Enhanced Symplectic "Category".

Suppose that M_1, M_2, and M_3 are symplectic manifolds, and that

$$\Gamma_2 \in \text{Morph}(M_2, M_3) \quad \text{and} \qquad \Gamma_1 \in \text{Morph}(M_1, M_2)$$

are canonical relations which can be composed in the sense of Chapter 4. Let ρ_1 be a $\frac{1}{2}$-density on Γ_1 and ρ_2 a $\frac{1}{2}$-density on Γ_2. The purpose of this chapter is to define a $\frac{1}{2}$-density $\rho_2 \circ \rho_1$ on $\Gamma_2 \circ \Gamma_1$ and to study the properties of this composition. In particular we will show that the composition

$$(\Gamma_2, \rho_2) \times (\Gamma_1, \rho_1) \mapsto (\Gamma_2 \circ \Gamma_1, \rho_2 \circ \rho_1)$$

is associative when defined, and that the axioms for a "category" are satisfied.

7.1 The underlying linear algebra.

We recall some definitions from Section 3.4: Let V_1, V_2 and V_3 be symplectic vector spaces and let $\Gamma_1 \subset V_1^- \times V_2$ and $\Gamma_2 \subset V_2^- \times V_3$ be linear canonical relations. We let

$$\Gamma_2 \star \Gamma_1 \subset \Gamma_1 \times \Gamma_2$$

consist of all pairs $((x, y), (y', z))$ such that $y = y'$, and let

$$\tau : \ \Gamma_1 \times \Gamma_2 \to V_2$$

be defined by

$$\tau(\gamma_1, \gamma_2) := \pi(\gamma_1) - \rho(\gamma_2)$$

so that $\Gamma_2 \star \Gamma_1$ is determined by the exact sequence (3.9)

$$0 \to \Gamma_2 \star \Gamma_1 \to \Gamma_1 \times \Gamma_2 \stackrel{\tau}{\to} V_2 \to \text{Coker } \tau \to 0.$$

We also defined

$$\alpha : \Gamma_2 \star \Gamma_1 \to \Gamma_2 \circ \Gamma_1$$

by (3.12):

$$\alpha : (x, y, y, z) \mapsto (x, z).$$

Then $\ker \alpha$ consists of those $(0, v, v, 0) \in \Gamma_2 \star \Gamma_1$ and we can identify $\ker \alpha$ as a subspace of V_2. We proved that relative to the symplectic structure on V_2 we have (3.16):

$$\ker \alpha = (\operatorname{Im} \tau)^\perp$$

as subspaces of V_2. We are going to use (3.16) to prove

Theorem 7.1.1. *There is a canonical isomorphism*

$$|\Gamma_1|^{\frac{1}{2}} \otimes |\Gamma_2|^{\frac{1}{2}} \cong |\ker \alpha| \otimes |\Gamma_2 \circ \Gamma_1|^{\frac{1}{2}}. \tag{7.1}$$

Proof. It follows from (3.16) that we have an identification

$$(V_2/ \ker \alpha) \sim (V_2/(\operatorname{Im} \tau)^\perp) \sim (\operatorname{Im} \tau)^*.$$

From the short exact sequence

$$0 \to \ker \alpha \to V_2 \to V_2/ \ker \alpha \to 0$$

we get an isomorphism

$$|V_2|^{\frac{1}{2}} \sim |\ker \alpha|^{\frac{1}{2}} \otimes |V_2/ \ker \alpha|^{\frac{1}{2}}$$

and from the fact that V_2 is a symplectic vector space we have a canonical trivialization $|V_2|^{\frac{1}{2}} \cong \mathbb{C}$. Therefore

$$|\ker \alpha|^{\frac{1}{2}} \cong |V_2/ \ker \alpha|^{-\frac{1}{2}}.$$

But since $(V_2/ \ker \alpha) \cong (\operatorname{Im} \tau)^*$ we obtain an identification

$$|\ker \alpha|^{\frac{1}{2}} \cong |\operatorname{Im} \tau|^{\frac{1}{2}}. \tag{7.2}$$

From the exact sequence (3.9) we obtain the short exact sequence

$$0 \to \Gamma_2 \star \Gamma_1 \to \Gamma_1 \times \Gamma_2 \xrightarrow{\tau} \operatorname{Im} \tau \to 0$$

which gives an isomorphism

$$|\Gamma_1|^{\frac{1}{2}} \otimes |\Gamma_2|^{\frac{1}{2}} \cong |\Gamma_2 \star \Gamma_1|^{\frac{1}{2}} \otimes |\operatorname{Im} \tau|^{\frac{1}{2}}.$$

From the short exact sequence

$$0 \to \ker \alpha \to \Gamma_2 \star \Gamma_1 \to \Gamma_2 \circ \Gamma_1 \to 0$$

we get the isomorphism

$$|\Gamma_2 \star \Gamma_1|^{\frac{1}{2}} \cong |\Gamma_2 \circ \Gamma_1|^{\frac{1}{2}} \otimes |\ker \alpha|^{\frac{1}{2}}.$$

Putting these two isomorphisms together and using (7.2) gives (7.1). □

7.1.1 Transverse composition of $\frac{1}{2}$ densities.

Let us consider the important special case of (7.1) where τ is surjective and so $\ker\alpha = 0$. Then we have a short exact sequence

$$0 \to \Gamma_2 \star \Gamma_1 \to \Gamma_1 \times \Gamma_2 \overset{\tau}{\to} V_2 \to 0$$

and an isomorphism

$$\alpha : \Gamma_2 \star \Gamma_1 \cong \Gamma_2 \circ \Gamma_1$$

and so (7.1) becomes

$$|\Gamma_2 \circ \Gamma_1|^{\frac{1}{2}} \cong |\Gamma_1 \times \Gamma_2|^{\frac{1}{2}}. \tag{7.3}$$

So if we are given $\frac{1}{2}$-densities σ_1 on Γ_1 and σ_2 on Γ_2 we obtain a $\frac{1}{2}$-density $\sigma_2 \circ \sigma_1$ on $\Gamma_2 \circ \Gamma_1$.

Let us work out this "composition" explicitly in the case that Γ_2 is the graph of an isomorphism

$$S : V_2 \to V_3.$$

Then $\rho : \Gamma_2 \to V_2$ is an isomorphism, and so we can identify $\frac{1}{2}$-densities on Γ_2 with $\frac{1}{2}$-densities on V_2. Let us choose σ_2 to be the $\frac{1}{2}$-density on Γ_2 which is identified with the canonical $\frac{1}{2}$-density on V_2. So if $2d_2 = \dim V_2 = \dim V_3$ and $u_1, \dots, u_{2d_2}$ is a symplectic basis of V_2, then σ_2 assigns the value one to the basis

$$(u_1, Su_1), \dots, (u_{2d_2}, Su_{2d_2})$$

of Γ_2.

Let $2d_1 = \dim V_1$ and let

$$(e_1, f_1), \dots (e_{d_1+d_2}, f_{d_1+d_2})$$

be a basis of Γ_1. Then

$$(e_1, Sf_1), \dots (e_{d_1+d_2}, Sf_{d_1+d_2})$$

is a basis of $\Gamma_2 \circ \Gamma_1$. Under our identification of $\Gamma_2 \circ \Gamma_1$ with $\Gamma_2 \star \Gamma_1$ (which is a subspace of $\Gamma_1 \times \Gamma_2$) this is identified with the basis

$$[(e_1, f_1), (f_1, Sf_1)], \dots, [(e_{d_1+d_2}, f_{d_1+d_2}), (f_{d_1+d_2}), Sf_{d_1+d_2})]$$

of $\Gamma_2 \star \Gamma_1$. The space $\{0\} \times \Gamma_2$ is complementary to $\Gamma_2 \star \Gamma_1$ in $\Gamma_1 \times \Gamma_2$ and the basis

$$[(e_1, f_1), (f_1, Sf_1)], \dots, [(e_{d_1+d_2}, f_{d_1+d_2}), (f_{d_1+d_2}), Sf_{d_1+d_2})],$$
$$[(0,0), (u_1, Su_1)], \dots, [(0,0), (u_{2d_2}, Su_{2d_2})]$$

differs from the basis

$$[(e_1, f_f), (0,0)], \dots, [((e_{d_1+d_2}, f_{d_1+d_2}), (0,0)],$$

$$[(0,0), (u_1, Su_1)], \dots, [(0,0), (u_{2d_2}, Su_{2d_2})]$$

by multiplication by a matrix of the form

$$\begin{pmatrix} I & * \\ 0 & I \end{pmatrix}.$$

We conclude that

Proposition 7.1.1. *If Γ_2 is the graph of a symplectomorphism $S : V_2 \to V_3$ and $\sigma_2 \in |\Gamma_2|^{\frac{1}{2}}$ is identified with the canonical $\frac{1}{2}$-density on V_2, then $\sigma_2 \circ \sigma_1$ is given by $(\mathrm{id} \times S)_* \sigma_1$ under the isomorphism $\mathrm{id} \times S$ of Γ_1 with $\Gamma_2 \circ \Gamma_1$. In particular, if $S = \mathrm{id}$ then $\sigma_2 \circ \sigma_1 = \sigma_1$.*

7.2 Half densities and clean canonical compositions.

Let M_1, M_2, M_3 be symplectic manifolds and let $\Gamma_1 \subset M_1^- \times M_2$ and $\Gamma_2 \subset M_2^- \times M_3$ be canonical relations. Let

$$\pi : \ \Gamma_1 \to M_2, \ \pi(m_1, m_2) = m_2, \qquad \rho : \ \Gamma_2 \to M_2, \ \rho(m_2, m_3) = m_2,$$

and $\Gamma_2 \star \Gamma_1 \subset \Gamma_1 \times \Gamma_2$ the fiber product:

$$\Gamma_2 \star \Gamma_1 = \{(m_1, m_2, m_3) | (m_1, m_2) \in \Gamma_1, \quad (m_2, m_3) \in \Gamma_2\}.$$

Let

$$\alpha : \Gamma_2 \star \Gamma_1 \to M_1 \times M_3, \quad \alpha(m_1, m_2, m_3) = (m_1, m_3).$$

The image of α is the composition $\Gamma_2 \circ \Gamma_1$.

Recall that we say that Γ_1 and Γ_2 intersect cleanly if the maps ρ and π intersect cleanly. If π and ρ intersect cleanly then their fiber product $\Gamma_2 \star \Gamma_1$ is a submanifold of $\Gamma_1 \times \Gamma_2$ and the arrows in the exact square

$$\begin{array}{ccc} \Gamma_2 \star \Gamma_1 & \longrightarrow & \Gamma_1 \\ \downarrow & & \downarrow \pi \\ \Gamma_2 & \xrightarrow[\rho]{} & M_2 \end{array}$$

are smooth maps. Furthermore the differentials of these maps at any point give an exact square of the corresponding linear canonical relations. In particular, α is of constant rank and $\Gamma_2 \circ \Gamma_1$ is an immersed canonical relation. If we further assume that

1. α is proper and

2. the level sets of α are connected and simply connected,

then $\Gamma_2 \circ \Gamma_1$ is an embedded Lagrangian submanifold of $M_1^- \times M_2$ and

$$\alpha : \Gamma_2 \star \Gamma_1 \to \Gamma_2 \circ \Gamma_1$$

is a fiber map with proper fibers. So our key identity (7.1) holds at the tangent space level: Let $m = (m_1, m_2, m_3) \in \Gamma_2 \star \Gamma_1$ and $q = \alpha(m) \in \Gamma_2 \circ \Gamma_1$ and let $F_q = \alpha^{-1}(q)$ be the fiber of α passing through m. We get an isomorphism

$$|T_m F_q| \otimes |T_q(\Gamma_2 \circ \Gamma_1)|^{\frac{1}{2}} \cong |T_{m_1,m_2}\Gamma_1|^{\frac{1}{2}} \otimes |T_{(m_2,m_3)}\Gamma_2|^{\frac{1}{2}}. \tag{7.4}$$

This means that if we are given half densities ρ_1 on Γ_1 and ρ_2 on Γ_2 we get a half density on $\Gamma_2 \circ \Gamma_1$ by integrating the expression obtained from the left hand side of the above isomorphism over the fiber. This gives us the composition law for half densities. Once we establish the associative law and the existence of the identity we will have have *enhanced* our symplectic category so that now the morphisms consist of pairs (Γ, ρ) where Γ is a canonical relation and where ρ is a half density on Γ.

Notice that if the composition $\Gamma_2 \circ \Gamma_1$ is transverse, then integration is just pointwise evaluation as in Section 7.1.1. In particular, we may apply Proposition 7.1.1 pointwise if Γ_2 is the graph of a symplectomorphism. In particular, if $\Gamma_2 = \Delta(X_2)$ is the diagonal in $X_2 \times X_2$ and we use the canonical $\frac{1}{2}$-density σ_Δ coming from the identification of $\Delta(X_2)$ with the symplectic manifold X_2 with its canonical $\frac{1}{2}$-density, then $(\Delta(X_2), \sigma_\Delta) \circ (\Gamma_1 \sigma_1) = (\Gamma_1, \sigma_1)$. This shows that $(\Delta(X_2), \sigma_\Delta)$ acts as the identity for composition on the left at X_2, and using the involutive structure (see below) implies that it is also an identity for composition on the right. This establishes the existence of the identity. For the associative law, we use the trick of of reducing the associative law for compostion to the associative law for direct product as in Section 3.3.2:

7.3 Rewriting the composition law.

We will rewrite the composition law in the spirit of Sections 3.3.2 and 4.4: If $\Gamma \subset M^- \times M$ is the graph of a symplectomorphism, then the projection of Γ onto the first factor is a diffeomorphism. The symplectic form on M determines a canonical $\frac{1}{2}$-density on M, and hence on Γ. In particular, we can apply this fact to the identity map, so $\Delta \subset M^- \times M$ carries a canonical $\frac{1}{2}$-density. Hence, the submanifold

$$\tilde{\Delta}_{M_1,M_2,M_3} = \{(x, y, y, z, x, z)\} \subset M_1 \times M_2 \times M_2 \times M_3 \times M_1 \times M_3$$

as in (4.6) carries a canonical $\frac{1}{2}$-density $\tau_{1,2,3}$. Then we know that

$$\Gamma_2 \circ \Gamma_1 = \tilde{\Delta}_{M_1,M_2,M_3} \circ (\Gamma_1 \times \Gamma_2)$$

and it is easy to check that

$$\rho_2 \circ \rho_1 = \tau_{123} \circ (\rho_1 \times \rho_2).$$

Similarly,

$$(\Gamma_3 \circ \Gamma_2) \circ \Gamma_1 = \Gamma_3 \circ (\Gamma_2 \circ \Gamma_1) = \tilde{\Delta}_{M_1,M_2,M_3,M_4} \circ (\Gamma_1 \times \Gamma_2 \times \Gamma_3)$$

and $\tilde{\Delta}_{M_1,M_2,M_3,M_4}$ carries a canonical $\frac{1}{2}$-density $\tau_{1,2,3,4}$ with

$$(\rho_3 \circ \rho_2) \circ \rho_1 = \rho_3 \circ (\rho_2 \circ \rho_1) = \tau_{1.2.3.4} \circ (\rho_1 \times \rho_2 \times \rho_3).$$

This establishes the associative law.

7.4 Enhancing the category of smooth manifolds and maps.

Let X and Y be smooth manifolds and $E \to X$ and $F \to Y$ be vector bundles. According to Atiyah and Bott, a morphism from $E \to X$ to $F \to Y$ consists of a smooth map

$$f : X \to Y$$

and a section

$$r \in C^\infty(\mathrm{Hom}(f^*F, E)).$$

We described the finite set analogue of this concept in Section **??**. If s is a smooth section of $F \to Y$ then we get a smooth section of $E \to X$ via

$$(f, r)^* s(x) := r(s(f(x)), \quad x \in X.$$

We want to specialize this construction of Atiyah-Bott to the case where E and F are the line bundles of $\frac{1}{2}$-densities on the tangent bundles. So we say that r is an enhancement of the smooth map $f : X \to Y$ or that (f, r) is an enhanced smooth map if r is a smooth section of the line bundle

$$\mathrm{Hom}(|f^*TY|^{\frac{1}{2}}, |TX|^{\frac{1}{2}}).$$

The composition of two enhanced maps

$$(f, r) : (E \to X) \to (F \to Y) \quad \text{and} \quad (g, r') : (F \to Y) \to (G \to Z)$$

is $(g \circ f, r \circ r')$ where, for $\tau \in |T_{g(f(x))}Z)|^{\frac{1}{2}}$

$$(r \circ r')(\tau) = r(r'(\tau)).$$

We thus obtain a category whose objects are the line bundles of $\frac{1}{2}$-densities on the tangent bundles of smooth manifolds and whose morphisms are enhanced maps.

If ρ is a $\frac{1}{2}$-density on Y and (f, r) is an enhanced map then we get a $\frac{1}{2}$-density on X by the Atiyah-Bott rule

$$(f, r)^* \rho(x) = r(\rho(f(x)) \in |T_x X|^{\frac{1}{2}}.$$

Then we know that the assignment $(f, r) \mapsto (f, r)^*$ is functorial. We now give some examples of enhancement of particular kinds of maps:

7.4.1 Enhancing an immersion.

Suppose $f : X \to Y$ is an immersion. We then get the conormal bundle $N_f^* X$ whose fiber at x consists of all covectors $\xi \in T^*_{f(x)}Y$ such that $df_x^* \xi = 0$. We have the exact sequence

$$0 \to T_x X \stackrel{df_x}{\to} T_{f(x)}Y \to N_x Y \to 0.$$

Here $N_x Y$ is *defined* as the quotient $T_{f(x)}Y/df_x(T_x X)$. The fact that f is an immersion is the statement that df_x is injective. The space $(N_f^* X_x)$ is the dual space of $N_x Y$. From the exact sequence above we get the isomorphism

$$|T_{f(x)}Y|^{\frac{1}{2}} \cong |N_x Y|^{\frac{1}{2}} \otimes |T_x X|^{\frac{1}{2}}.$$

So

$$\mathrm{Hom}(|T_{f(x)}Y|^{\frac{1}{2}}, |T_x X|^{\frac{1}{2}}) \cong |T_x X|^{\frac{1}{2}} \otimes |T_{f(x)}Y|^{-\frac{1}{2}} \cong |N_x Y|^{-\frac{1}{2}} \cong |(N_f^* X)_x|^{\frac{1}{2}}.$$

Conclusion. Enhancing an immersion is the same as giving a section of $|N_f^* X|^{\frac{1}{2}}$.

7.4.2 Enhancing a fibration.

Suppose that $\pi : Z \to X$ is a submersion. If $z \in Z$, let V_z denote the tangent space to the fiber $\pi^{-1}(x)$ at z where $x = \pi(z)$. Thus V_z is the kernel of $d\pi_z : T_z Z \to T_{\pi(z)}X$. So we have an exact sequence

$$0 \to V_z \to T_z Z \to T_{\pi(z)}X \to 0$$

and hence the isomorphism

$$|T_z Z|^{\frac{1}{2}} \cong |V_z|^{\frac{1}{2}} \otimes |T_{\pi(z)}X|^{\frac{1}{2}}.$$

So

$$\mathrm{Hom}(|T_{\pi(z)}X|^{\frac{1}{2}}, |T_z Z|^{\frac{1}{2}}) \cong |T_{\pi(z)}X|^{-\frac{1}{2}} \otimes |T_z Z|^{\frac{1}{2}} \cong |V_z|^{\frac{1}{2}}. \tag{7.5}$$

Conclusion. Enhancing a fibration is the same as giving a section of $|V|^{\frac{1}{2}}$ where V denotes the vertical sub-bundle of the tangent bundle, i.e. the sub-bundle tangent to the fibers of the fibration.

7.4.3 The pushforward via an enhanced fibration.

Suppose that $\pi : Z \to X$ is a fibration with compact fibers and r is an enhancement of π so that r is given by a section of the line-bundle $|V|^{\frac{1}{2}}$ as we have just seen. Let ρ be a $\frac{1}{2}$-density on Z. From the isomorphism

$$|T_z Z|^{\frac{1}{2}} \cong |V_z|^{\frac{1}{2}} \otimes |T_{\pi(z)}X|^{\frac{1}{2}}$$

we can regard ρ as section of $|V|^{\frac{1}{2}} \otimes \pi^*|TX|^{\frac{1}{2}}$ and hence

$$r \cdot \rho$$

is a section of $|V| \otimes \pi^*|TX|^{\frac{1}{2}}$. Put another way, for each $x \in X$, $r \cdot \rho$ gives a density (of order one) on $\pi^{-1}(x)$ with values in the fixed vector space $|T_xX|^{\frac{1}{2}}$. So we can integrate this density of order one over the fiber to obtain

$$\pi_*(r \cdot \rho)$$

which is a $\frac{1}{2}$-density on X. If the enhancement r of π is understood, we will denote the push-forward of the $\frac{1}{2}$-density ρ simply by

$$\pi_*\rho.$$

We have the obvious variants on this construction if π is not proper. We can construct $\pi_*(r\cdot\rho)$ if either r or ρ are compactly supported in the fiber direction.

An enhanced fibration $\pi = (\pi, r)$ gives a pull-back operation π^* from half densities on X to $\frac{1}{2}$-densities on Z. So if μ is a $\frac{1}{2}$-density on X and ν is a $\frac{1}{2}$-density on Z then

$$\nu \cdot \pi^*\mu$$

is a density on Z. If μ is of compact support and if ν is compactly supported in the fiber direction, then $\nu \cdot \pi^*\mu$ is a density (of order one) of compact support on Z which we can integrate over Z. We can also form

$$(\pi_*\nu) \cdot \mu.$$

which is a density (of order one) which is of compact support on X. It follows from Fubini's theorem that

$$\int_Z \nu \cdot \pi^*\mu = \int_X (\pi_*\nu) \cdot \mu.$$

7.5 Enhancing a map enhances the corresponding canonical relation.

Let $f : X \to Y$ be a smooth map. We can enhance this map by giving a section r of $\mathrm{Hom}(|TY|^{\frac{1}{2}}, |TX|^{\frac{1}{2}})$. On the other hand, we can construct the canonical relation

$$\Gamma_f \in \mathrm{Morph}(T^*X, T^*Y)$$

as described in Section 4.8. Enhancing this canonical relation amounts to giving a $\frac{1}{2}$-density ρ on Γ_f. In this section we show how the enhancement r of the map f gives rise to a $\frac{1}{2}$-density on Γ_f.

Recall (4.11) which says that

$$\Gamma_f = \{(x_1, \xi_1, x_2, \xi_2)|x_2 = f(x_1), \quad \xi_1 = df^*_{x_1}\xi_2\}.$$

From this description we see that Γ_f is a vector bundle over X whose fiber over $x \in X$ is $T^*_{f(x)}Y$. So at each point $z = (x, \xi_1, y, \eta) \in \Gamma_f$ we have the isomorphism

$$|T_z\Gamma_f|^{\frac{1}{2}} \cong |T_xX|^{\frac{1}{2}} \otimes |T_\eta(T^*_{f(x)}Y)|^{\frac{1}{2}}.$$

But $(T^*_{f(x)}Y)$ is a vector space, and at any point η in a vector space W we have a canonical identification of $T_\eta W$ with W. So at each $z \in \Gamma_f$ we have an isomorphism

$$|T_z\Gamma_f|^{\frac{1}{2}} \cong |T_xX|^{\frac{1}{2}} \otimes |T_\eta(T^*_{f(x)}Y)|^{\frac{1}{2}} = \mathrm{Hom}(|T_{f(x)}Y|^{\frac{1}{2}}, |T_xX|^{\frac{1}{2}})$$

and at each x, $r(x)$ is an element of $\mathrm{Hom}(|T_{f(x)}Y|^{\frac{1}{2}}, |T_xX|^{\frac{1}{2}})$. So r gives rise to a $\frac{1}{2}$-density on Γ_f.

7.6 The involutive structure of the enhanced symplectic "category".

Recall that if $\Gamma \in \mathrm{Morph}(M_1, M_2)$ then we defined $\Gamma^\dagger \in (M_2, M_1)$ be

$$\Gamma^\dagger = \{(y, x) | (x, y) \in \Gamma\}.$$

We have the switching diffeomorphism

$$s : \Gamma^\dagger \to \Gamma, \qquad (y, x) \mapsto (x, y),$$

and so if ρ is a $\frac{1}{2}$-density on Γ then $s^*\rho$ is a $\frac{1}{2}$-density on $\Gamma^\dagger$. We define

$$\rho^\dagger = \overline{s^*\rho}. \tag{7.6}$$

Starting with an enhanced morphism (Γ, ρ) we define

$$(\Gamma, \rho)^\dagger = (\Gamma^\dagger, \rho^\dagger).$$

We show that $\dagger : (\Gamma, \rho) \mapsto (\Gamma, \rho)^\dagger$ satisfies the conditions for a involutive structure. Since $s^2 = \mathrm{id}$ it is clear that $\dagger^2 = \mathrm{id}$. If $\Gamma_2 \in \mathrm{Morph}(M_2, M_1)$ and $\Gamma_1 \in \mathrm{Morph}(M_1, M_2)$ are composible morphsims, we know that the composition of (Γ_2, ρ_2) with (Γ_1, ρ_1) is given by

$$(\tilde{\Delta}_{M_1,M_2,M_3}, \tau_{123}) \circ (\Gamma_1 \times \Gamma_2, \rho_1 \times \rho_2).$$

where

$$\tilde{\Delta}_{M_1,M_2,M_3} = \{(x, y, y, z, x, z) | x \in M_1, y \in M_2, z \in M_3\}$$

and τ_{123} is the canonical (real) $\frac{1}{2}$-density arising from the symplectic structures on M_1, M_2 and M_3. So

$$s : (\Gamma_2 \circ \Gamma_1)^\dagger = \Gamma_1^\dagger \circ \Gamma_2^\dagger \to \Gamma_2 \circ \Gamma_1$$

is given by applying the operator S switching x and z

$$S : \tilde{\Delta}_{M_3,M_2,M_1} \to \tilde{\Delta}_{M_1,M_2,M_3},$$

applying the switching operators $s_1 : \Gamma_1^\dagger \to \Gamma_1$ and $s_2 : \Gamma_2^\dagger \to \Gamma_2$ and also switching the order of Γ_1 and Γ_2. Pull-back under switching the order of Γ_1 and Γ_2 sends $\rho_1 \times \rho_2$ to $\rho_2 \times \rho_1$, applying the individual s_1^* and s_2^* and taking complex conjugates sends $\rho_2 \times \rho_1$ to $\rho_2^\dagger \times \rho_1^\dagger$. Also

$$S^* \tau_{123} = \tau_{321}$$

and τ_{321} is real. Putting all these facts together shows that

$$((\Gamma_2, \rho_2) \circ (\Gamma_1, \rho_1))^\dagger = (\Gamma_1, \rho_1)^\dagger \circ (\Gamma_2, \rho_2)^\dagger$$

proving that † satisfies the conditions for a involutive structure.

Let M be an object in our "category", i.e. a symplectic manifold. A "point" of M in our enchanced "category" will consist of a Lagrangian submanifold $\Lambda \subset M$ thought of as an element of Morph(pt., M) (in $\mathcal{S}$) together with a $\frac{1}{2}$-density on Λ. If (Λ, ρ) is such a point, then $(\Lambda, \rho)^\dagger = (\Lambda^\dagger, \rho^\dagger)$ where we now think of the Lagrangian submanifold $\Lambda^\dagger$ as an element of Morph(M, pt.).

Suppose that (Λ_1, ρ_1) and (Λ_2, ρ_2) are "points" of M and that $\Lambda_2^\dagger$ and Λ_1 are composible. Then $\Lambda_2^\dagger \circ \Lambda_1$ in $\mathcal{S}$ is an element of Morph(pt., pt.) which consists of a (single) point. So in our enhanced "category" $\tilde{\mathcal{S}}$

$$(\Lambda_2, \rho_2)^\dagger (\Lambda_1, \rho_1)$$

is a $\frac{1}{2}$-density on a point, i.e. a complex number. We will denote this number by

$$\langle (\Lambda_1, \rho_1), (\Lambda_2, \rho_2) \rangle .$$

7.6.1 Computing the pairing $\langle (\Lambda_1, \rho_1), (\Lambda_2, \rho_2) \rangle$.

This is, of course, a special case of the computation of Section 7.2, where $\Gamma_2 \circ \Gamma_1$ is a point.

The first condition that $\Lambda_2^\dagger$ and Λ_1 be composible is that Λ_1 and Λ_2 intersect cleanly as submanifolds of M. Then the F of (7.4) is $F = \Lambda_1 \cap \Lambda_2$ so (7.4) becomes

$$|T_p F| = |T_p(\Lambda_1 \cap \Lambda_2)| \cong |T_p \Lambda_1|^{\frac{1}{2}} \otimes |T_p \Lambda_2|^{\frac{1}{2}} \tag{7.7}$$

and so ρ_1 and $\overline{\rho_2}$ multiply together to give a density $\rho_1 \overline{\rho_2}$ on $\Lambda_1 \cap \Lambda_2$. A second condition on composibility requires that $\Lambda_1 \cap \Lambda_2$ be compact. The pairing is thus

$$\langle (\Lambda_1, \rho_1), (\Lambda_2, \rho_2) \rangle = \int_{\Lambda_1 \cap \Lambda_2} \rho_1 \overline{\rho_2}. \tag{7.8}$$

7.6.2 † and the adjoint under the pairing.

In the category of whose objects are Hilbert spaces and whose morphisms are bounded operators, the adjoint $A^\dagger$ of a operator $A : H_1 \to H_2$ is defined by

$$\langle Av, w\rangle_2 = \langle v, A^\dagger w\rangle_1, \tag{7.9}$$

for all $v \in H_1, w \in H_2$ where $\langle\ ,\rangle_i$ denotes the scalar product on $H_i,\ i = 1, 2$. This can be given a more categorical interpretation as follows: A vector u in a Hilbert space H determines and is determined by a bounded linear map from $\mathbb{C}$ to H,

$$z \mapsto zu.$$

In other words, if we regard $\mathbb{C}$ as the pt. in the category of Hilbert spaces, then we can regard $u \in H$ as an element of of Morph(pt., H). So if $v \in H$ we can regard $v^\dagger$ as an element of Morph(H, pt.) where

$$v^\dagger(u) = \langle u, v\rangle.$$

So if we regard † as the primary operation, then the scalar product on each Hilbert space is determined by the preceding equation - the right hand side is *defined* as being equal to the left hand side. Then equation (7.9) is a consequence of the associative law and the laws $(A \circ B)^\dagger = B^\dagger \circ A^\dagger$ and $\dagger^2 = \mathrm{id}$.. Indeed

$$\langle Av, w\rangle_2 := w^\dagger \circ A \circ v = (A^\dagger \circ w)^\dagger \circ v =: \langle v, A^\dagger w\rangle_1.$$

So once we agree that a $\frac{1}{2}$-density on pt. is just a complex number, we can conclude that the analogue of (7.9) holds in our enhanced category $\tilde{\mathcal{S}}$: If (Λ_1, ρ_1) is a "point " of M_1 in our enhanced category, and if (Λ_2, ρ_2) is a "point " of M_2 and if $(\Gamma, \tau) \in \mathrm{Morph}(M_1, M_2)$ then (assuming that the various morphisms are composible) we have

$$\langle ((\Gamma, \tau) \circ (\Lambda_1, \rho_1), (\Lambda_2, \rho_2)\rangle_2 = \left\langle (\Lambda_1, \rho_1), ((\Gamma, \tau)^\dagger \circ (\Lambda_2, \rho_2)\right\rangle_1. \tag{7.10}$$

7.7 The symbolic distributional trace.

We consider a family of symplectomorphisms as in Section 4.11.7 and follow the notation there. In particular we have the family $\Phi : M \times S \to S$ of symplectomorphisms and the associated moment Lagrangian

$$\Gamma := \Gamma_\Phi \subset M \times M^- \times T^*S.$$

7.7.1 The $\frac{1}{2}$-density on Γ.

Since M is symplectic it has a canonical $\frac{1}{2}$ density. So if we equip S with a half density ρ_S we get a $\frac{1}{2}$ density on $M \times M^- \times S$ and hence a $\frac{1}{2}$ density ρ_Γ making Γ into a morphism

$$(\Gamma, \rho_\Gamma) \in \mathrm{Morph}(M^- \times M, T^*S)$$

in our enhanced symplectic category.

Let $\Delta \subset M^- \times M$ be the diagonal. The map

$$M \to M^- \times M \qquad m \mapsto (m, m)$$

carries the canonical $\frac{1}{2}$-density on M to a $\frac{1}{2}$-density, call it ρ_Δ on Δ enhancing Δ into a morphism

$$(\Delta, \rho_\Delta) \in \text{Morph}(\text{pt.}., M^- \times M).$$

The generalized trace in our enhanced symplectic "category".

Suppose that Γ and Δ are composable. Then we get a Lagrangian submanifold

$$\Lambda = \Gamma \circ \Delta$$

and a $\frac{1}{2}$-density

$$\rho_\Lambda := \rho_\Gamma \circ \rho_\Delta$$

on Λ. The operation of passing from $F : M \times S \to M$ to (Λ, ρ_Λ) can be regarded as the symbolic version of the distributional trace operation in operator theory.

7.7.2 Example: The symbolic trace.

Suppose that we have a single symplectomorphism $f : M \to M$ so that S is a point as is T^*S. Let

$$\Gamma = \Gamma_f = \text{graph } f = \{(m, f(m)),\ m \in M\}$$

considered as a morphism from $M \times M^-$ to a point. Suppose that Γ and Δ intersect transversally so that $\Gamma \cap \Delta$ is discrete. Suppose, in fact, that it is finite. We have the $\frac{1}{2}$-densities ρ_Δ on $T_m\Delta$ and $T_m\Gamma$ at each point m of of $\Gamma \cap \Delta$. Hence, by (6.12), the result is

$$\sum_{m \in \Delta \cap \Gamma} |\det(I - df_m)|^{-\frac{1}{2}}. \tag{7.11}$$

7.7.3 General transverse trace.

Let S be arbitrary. We examine the meaning of the hypothesis that that the inclusion $\iota : \Delta \to M \times M$ and the projection $\Gamma \to M \times M$ be transverse.

Since Γ is the image of $(G, \Phi) : M \times S \to M \times M \times T^*S$, the projection of Γ onto $M \times M$ is just the image of the map G given in (4.40). So the transverse composibility condition is

$$G \pitchfork \Delta. \tag{7.12}$$

The fiber product of Γ and Δ can thus be identified with the "fixed point submanifold" of $M \times S$:

$$\mathfrak{F} := \{(m, s) | f_s(m) = m\}.$$

The transversality assumption guarantees that this is a submanifold of $M \times S$ whose dimension is equal to $\dim S$. The transversal version of our composition law for morphisms in the category $\mathcal{S}$ asserts that

$$\Phi : \mathfrak{F} \to T * S$$

is a Lagrangian immersion whose image is

$$\Lambda = \Gamma \circ \Delta.$$

Let us assume that $\mathfrak{F}$ is connected and that Φ is a Lagrangian imbedding. (More generally we might want to assume that $\mathfrak{F}$ has a finite number of connected components and that Φ restricted to each of these components is an imbedding. Then the discussion below would apply separately to each component of $\mathfrak{F}$.)

Let us derive some consequences of the transversality hypothesis $G \pitchfork \Delta$. By the Thom transverslity theorem, there exists an open subset

$$S_O \subset S$$

such that for every $s \in S_O$, the map

$$g_s : M \to M \times M, \quad g_s(m) = G(m, s) = (mf_s(m))$$

is transverse to Δ. So for $s \in S_O$,

$$g_s^{-1}(\Delta) = \{m_i(s), i = 1, \dots, r\}$$

is a finite subset of M and the m_i depend smoothly on $s \in S_O$. For each i, $\Phi(m_i(s)) \in T_s^* S$ then depends smoothly on $s \in S_O$. So we get one forms

$$\mu_i := \Phi(m_i(s)) \tag{7.13}$$

parametrizing open subsets Λ_i of Λ. Since Λ is Lagrangian, these one forms are closed. So if we assume that $H^1(S_O) = \{0\}$, we can write

$$\mu_i = d\psi_i$$

for $\psi_i \in C^\infty(S_O)$ and

$$\Lambda_i = \Lambda_{\psi_i}.$$

The maps

$$S_O \to \Lambda_i, \quad s \mapsto (s, d\psi_i(s))$$

map S_O diffeomorphically onto Λ_i. The pull-backs of the $\frac{1}{2}$-density $\rho_\Lambda = \rho_\Gamma \circ \rho_\Delta$ under these maps can be written as

$$h_i \rho_S$$

where ρ_S is the $\frac{1}{2}$-density we started with on S and where the h_i are the smooth functions

$$h_i(s) = \left|\det(I - df_{m_i})\right|^{-\frac{1}{2}}. \tag{7.14}$$

In other words, on the generic set S_O where g_s is transverse to Δ, we can compute the symbolic trace $h(s)$ of g_s as in the preceding section. At points not in S_O, the "fixed points coalesce" so that g_s is no longer transverse to Δ and the individual g_s no longer have a trace as individual maps. Nevertheless, the parametrized family of maps have a trace as a $\frac{1}{2}$-density on Λ which need not be horizontal over points of S which are not in S_O.

7.7.4 Example: Periodic Hamiltonian trajectories.

Let (M, ω) be a symplectic manifold and

$$H : M \to \mathbb{R}$$

a proper smooth function with no critical points. Let $v = v_H$ be the corresponding Hamiltonian vector field, so that

$$i(v)\omega = -dH.$$

The fact that H is proper implies that v generates a global one parameter group of transformations, so we get a Hamiltonian action of $\mathbb{R}$ on M with Hamiltonian H, so we know that the function Φ of (4.34) (determined up to a constant) can be taken to be

$$\Phi : M \times \mathbb{R} \to T^*\mathbb{R} = \mathbb{R} \times \mathbb{R}, \qquad \Phi(m, t) = (t, H(m)).$$

The fact that $dH_m \neq 0$ for any m implies that the vector field v has no zeros.

Notice that in this case the transversality hypothesis of the previous example is never satisfied. For if it were, we could find a dense set of t for which $\exp tv : M \to M$ has isolated fixed points. But if m is fixed under $\exp tv$ then every point on the orbit $(\exp sv)m$ of m is also fixed under $\exp tv$ and we know that this orbit is a curve since v has no zeros.

So the best we can do is assume clean intersection: Our Γ in this case is

$$\Gamma = \{m, (\exp sv)m, s, H(m))\}.$$

If we set $f_s = \exp sv$ we write this as

$$\Gamma = \{(m, f_s(m), s, H(m))\}.$$

The assumption that the maps $\Gamma \to M \times M$ and

$$\iota : \Delta \to M \times M$$

intersect cleanly means that the fiber product

$$X = \{(m, s) \in M \times \mathbb{R} | f_s(m) = m\}$$

is a manifold and that its tangent space at (m, s) is

$$\{(v, c) \in T_m M \times \mathbb{R} | v = (df_s)_m(v) + cv(m) \tag{7.15}$$

since

$$dF_{(m,s)}\left(v, c\frac{\partial}{\partial t}\right) = (df_s)_m(v) + cv(m).$$

The map $\Phi : X \to T^*S$ is of constant rank, and its image is an immersed Lagrangian submanifold of T^*S. One important consequence is:

The energy-period relation.

The restriction of $dt \wedge dH = \Phi^*(dt \wedge d\tau)$ vanishes. Thus if c is a regular value of H, then on every connected component of $H^{-1}(C) \cap X$ all trajectories of v have the same period. For this reason Λ is called the period Lagrangian.

The linear Poincaré map.

At each $m \in M$, let

$$W_m^0 := \{w \in T_mM |\ dH(w) = 0.\}$$

Since $dH(v) \equiv 0$, we have $v(m) \in W_m^0$ and since f_s preserves H and v we see that $(f_s)_m : T_mM \to T_mM$ induces a map

$$P_{m,s} : W_m^0/\mathbb{R}v(m) \to W_{m,s}^0/\mathbb{R}v$$

called the **linear Poincaré map**.

Let us make the genericity assumption

$$\det(I - P_{m,s}) \neq 0. \tag{7.16}$$

This means the following: Let $t \mapsto \gamma(t) = f_t(m)$ be the trajectory of $f_t = \exp tv$ through m. We know that the flow f_t preserves the hypersurface $H = H(m)$. Let Y be a transverss slice to γ through m on this hypersurface. If m' is a point of Y near m, then the trajectory through m' will intersect Y again at some point $p(m')$ at some time s' near s, and this map $p : Y \to Y$ is known as the Poincaré map of the flow (restricted to the hypersurface and relative to the choice of slice). Then $P_{m,s}$ can be identified with the differential of this Poincarë map, and our genericity assumption (7.16) says that m is a non-degenerate fixed point of p.

By (7.15), the genericity assumption (7.16) implies that

1. $\dim X = 2$,
2. $H : X \to \mathbb{R}$ is a sumbersion, and
3. $X \cap H^{-1}(c)$ is a disjoint union of periodic trajectories of v. In other words, if $X_i,\ i = 1, 2, \dots$ are the connected components of X and

$$(m_i, s) \in H^{-1}(c) \cap X_i$$

then

$$H^{-1}(c) \cap X_i = \gamma_i^c$$

where γ_i^c is the periodic trajectory of $v = v_H$ through m_i or period $s = T_i(c)$.

Remarks.

- If $m' = f_t(m)$ is a second point on the trajectory through m, then the maps $P_{m,s}$ and $P_{m',s}$ are conjugate. Hence $det(I - P_{m,s}) = \det(I - P_{m',s}$ so condition (7.16) depends on the periodic trajectory, not on the choice of a specific point on this trajectory.

- If m lies on a periodic trajectory γ_i then it will have a first return time $T_i^\sharp > 0$, the smallest positive s for which $f_s(m) = m, \quad m \in \gamma_i$. All other return times will be integer multiples of $T_i^\sharp$.

- The moment map $\Phi : M \times \mathbb{R} \to T^*S$ maps X_i onto the period Lagrangian

$$\Lambda_i = \{(t, \tau), t = T_i(\tau)\}.$$

  This map is a fiber mapping with compact fibers and the fiber above (t, τ) can be identified with the periodic trajectory γ_i.

Let us equip $\mathbb{R}$ with its standard $\frac{1}{2}$-density $|dt|^{\frac{1}{2}}$. We will obtain a $\frac{1}{2}$-density σ_i on Λ_i which will involve fiber integration over the fibration by periodic trajectories described above. If we use τ as a coordinate on Λ_i via the map $\tau \mapsto (t, \tau), \quad t = T_i(\tau)$ then a computation similar to the one we gave above for a single symplectomorphism shows that the induced $\frac{1}{2}$-density on Λ_i is given by

$$T_i^\sharp(\tau) |\det(I - P_{\gamma_i(\tau)})|^{-\frac{1}{2}}. \tag{7.17}$$

7.8 The Maslov enhanced symplectic "category".

Let X be a manifold, $\Lambda \subset T^*X$ a Lagrangian submanifold, $\pi : Z \to X$ a fibration and $\phi \in C^\infty(Z)$ a generating function for λ with respect to π.

For each $z \in C_\phi$ let $\operatorname{sgn} \phi(z)$ denote the signature of the quadratic form

$$d^2 \left(\phi \left| \pi^{-1}(\pi(z)\right)_z .$$

Let $s_\phi : C_\phi \to \mathbb{C}$ be the function

$$s_\phi := \exp \frac{\pi i}{4} \operatorname{sgn} \phi. \tag{7.18}$$

Under the identification $\lambda_\phi : C_\phi \to \Lambda$ we will regard s_ϕ as a function on Λ.

In Section 5.13 we defined the Maslov bundle $\mathbb{L}_{\text{Maslov}} \to \Lambda$ to be the trivial flat line bundle whose flat sections are constant multiples of s_ϕ.

More generally, if Λ does not admit a global generating function, we can cover Λ by open sets U_i on each of which we have a generating function ϕ_i, and we showed in Section 5.13.3 that the s_{ϕ_i}'s patch together to give a globally defined flat line bundle $\mathbb{L}_{\text{Maslov}} \to \Lambda$.

We can define this bundle for canonical canonical relations

$$\Gamma : T^*X_1 \Rightarrow T^*X_2$$

by regarding Γ as a Lagrangian submanifold of $(T^*X_1)^- \times T^*X_2$. As we showed in Section 5.13.5 it has the same functorial behavior with respect to clean composition of canonical relations as does the bundle of $\frac{1}{2}$-densities, compare (5.29) with (7.4).

So we enhance our symplectic "category" even further by defining

$$\mathbb{L}_\Lambda := \mathbb{L}_{\mathrm{Maslov}}(\Lambda) \otimes |T\Lambda|^{\frac{1}{2}} \tag{7.19}$$

$$\mathbb{L}_\Gamma := \mathbb{L}_{\mathrm{Maslov}}(\Gamma) \otimes |T\Gamma|^{\frac{1}{2}}, \tag{7.20}$$

where the objects are now pairs (Λ, σ), where σ is a section of $\mathbb{L}_\Lambda$ and morphisms are pairs (Γ, τ) with τ is a section of $\mathbb{L}_\Gamma$ and the composition law (when defined, i.e. under the hypotheses for clean composition) is given by combining the composition laws (5.29) and (7.4).

As we will see in the next chapter, this enhanced "category" will play a fundamental role in the theory of semi-classical Fourier integral operators.

Chapter 8

Oscillatory $\frac{1}{2}$-densities.

Let (Λ, ψ) be an exact Lagrangian submanifold of T^*X. Let

$$k \in \mathbb{Z}.$$

The plan of this chapter is to associate to (Λ, ψ) and to k a space

$$I^k(X, \Lambda, \psi)$$

of rapidly oscillating $\frac{1}{2}$-densities on X and to study the properties of these spaces. If Λ is horizontal with

$$\Lambda = \Lambda_\phi, \quad \phi \in C^\infty(X),$$

and

$$\psi = \phi \circ (\pi_X)_{|\Lambda}$$

this space will consist of $\frac{1}{2}$-densities of the form

$$\hbar^k a(x, \hbar) e^{i\frac{\phi(x)}{\hbar}} \rho_0$$

where ρ_0 is a fixed non-vanishing $\frac{1}{2}$-density on X and where

$$a \in C^\infty(X \times \mathbb{R}).$$

In other words, so long as $\Lambda = \Lambda_\phi$ is horizontal and $\psi = \phi \circ (\pi_X)_{|\Lambda}$, our space will consist of the $\frac{1}{2}$-densities we studied in Chapter 1.

As we saw in Chapter 1, one must take into account, when solving hyperbolic partial differential equations, the fact that caustics develop as a result of the Hamiltonian flow applied to initial conditions. So we will need a more general definition. We will make a more general definition, locally, in terms of a general generating function relative to a fibration, and then show that the class $I^k(X, \Lambda, \psi)$ of oscillating $\frac{1}{2}$-densities on X that we obtain this way is independent of the choice of generating functions.

This will imply that we can associate to every exact canonical relation between cotangent bundles (and every integer k) a class of (oscillatory) integral operators which we will call the semi-classical Fourier integral operators associated to the canonical relation. We will find that if we have two transversally composible canonical relations, the composition of their semi-classical Fourier integral operators is a semi-classical Fourier integral operator associated to the composition of the relations. We will then develop a symbol calculus for these operators and their composition.

For expository reasons, we will begin by carrying out the discussion in terms of transverse generating functions, which limits our symbol calculus to the case of transverse composition. Since, in the applications, we will need to allow clean compositions of canonical relations, we will go back and give the local description of the class $I^k(X, \Lambda, \psi)$ in terms of clean generating functions which will then allow us to give a symbol calculus for the semi-classical operators associated to clean composition of canonical relations.

In order not to overburden the notation, we will frequently write Λ instead of (Λ, ψ). But a definite choice of ψ will always be assumed. So, for example, we will write $I^k(X, \Lambda)$ instead of $I^k(X, \Lambda, \psi)$ for the class of $\frac{1}{2}$-densities that we will introduce over the next few sections.

A key ingredient in the study of an element of $I^k(X, \Lambda)$ is its symbol. Initially, we will define the "symbol" in terms of a (transverse) generating function as a function on Λ. Although this definition definitely depends on the choice of presentation of Λ by generating functions, we will find that the assertion that the symbol of an element of $I^k(X, \Lambda)$ vanishes at $p \in \Lambda$ *does* have invariant significance. So if we let $I^k_p(X, \Lambda)$ denote the set of all elements of $I^k(X, \Lambda)$ whose (non-intrinsic) symbol vanishes at p, we obtain an intrinsically defined line bundle $\mathbb{L}$ over Λ where

$$\mathbb{L}_p = I^k(X, \Lambda)/I^k_p(X, \Lambda).$$

We will find that this definition is independent of k.

(For the experts, our line bundle $\mathbb{L}$ can be identified with the line bundle of half-densities on Λ tensored with the Maslov bundle. But our point is to deal with intrinsically defined objects from the start.)

We then will have a symbol map from $I^k(X, \Lambda)$ to sections of $\mathbb{L}$ and will find that $I^k(X, \Lambda)/I^{k+1}(X, \Lambda)$ is isomorphic to sections of $\mathbb{L}$. We will also find that the study of $I^k(X, \Lambda)/I^{k+\ell}(X, \Lambda)$ is associated with a sheaf $\mathcal{E}^\ell$ on Λ giving rise to the concept of microlocalization.

8.1 Definition of $I^k(X, \Lambda)$ in terms of a generating function.

Let $\pi : Z \to X$ be a fibration which is enhanced in the sense of Section 7.4.2. Recall that this means that we are given a smooth section r of $|V|^{\frac{1}{2}}$ where V is the vertical sub-bundle of the tangent bundle of Z. We will assume that r

vanishes nowhere. If ν is a $\frac{1}{2}$-density on Z which is of compact support in the vertical direction, then recall from Section 7.4.3 that we get from this data a push-forward $\frac{1}{2}$-density $\pi_*\nu$ on X.

Now suppose that ϕ is a global generating function for (Λ,ψ) with respect to π. Recall that this means that we have fixed the arbitrary constant in ϕ so that

$$\psi(x,\xi)=\phi(z)$$

if $d\phi_z=\pi_z^*\xi$ where $\pi(z)=x,\ z\in C_\phi$. See the discussion following equation (4.62). Let

$$d:=\dim Z-\dim X.$$

We define $I_0^k(X,\Lambda,\phi)$ to be the space of all compactly supported $\frac{1}{2}$-densities on X of the form

$$\mu=\hbar^{k-\frac{d}{2}}\pi_*\left(ae^{i\frac{\phi}{\hbar}}\tau\right) \tag{8.1}$$

where $a=a(z,\hbar)$

$$a\in C_0^\infty(Z\times\mathbb{R})$$

and where τ is a nowhere vanishing $\frac{1}{2}$-density on Z. Then define $I^k(X,\Lambda,\phi)$ to consist of those $\frac{1}{2}$-densities μ such that $\rho\mu\in I_0^k(X,\Lambda,\phi)$ for every $\rho\in C_0^\infty(X)$.

It is clear that $I^k(X,\Lambda,\phi)$ does not depend on the choice of the enhancement r of π or on the choice of τ.

8.1.1 Local description of $I^k(X,\Lambda,\phi)$.

Suppose that $Z=X\times S$ where S is an open subset of $\mathbb{R}^d$ and π is projection onto the first factor. We may choose our fiber $\frac{1}{2}$-density to be the Euclidean $\frac{1}{2}$-density $ds^{\frac{1}{2}}$ and τ to be $\tau_0\otimes ds^{\frac{1}{2}}$ where τ_0 is a nowhere vanishing $\frac{1}{2}$-density on X. Then $\phi=\phi(x,s)$ and the push forward in (8.1) becomes the oscillating integral

$$\left(\int_S a(x,s,\hbar)e^{i\frac{\phi}{\hbar}}ds\right)\tau_0. \tag{8.2}$$

8.1.2 Independence of the generating function.

Let $\pi_i:Z_i\to X,\ \phi_i$ be two fibrations and generating functions for the same Lagrangian submanifold $\Lambda\subset T^*X$. We wish to show that $I^k(X,\Lambda,\phi_1)=I^k(X,\Lambda,\phi_2)$. By a partition of unity, it is enough to prove this locally. According to Section 5.12, since the constant is fixed by (4.62), it is enough to check this for two types of change of generating functions, 1) equivalence and 2) increasing the number of fiber variables. Let us examine each of the two cases:

Equivalence.

There exists a diffeomorphism $g:\ Z_1\to Z_2$ with

$$\pi_2\circ g=\pi_1 \qquad \text{and} \qquad \phi_2\circ g=\phi_1.$$

Let us fix a non-vanishing section r of the vertical $\frac{1}{2}$-density bundle $|V_1|^{\frac{1}{2}}$ of Z_1 and a $\frac{1}{2}$-density τ_1 on Z_1. Since g is a fiber preserving map, these determine vertical $\frac{1}{2}$-densities and $\frac{1}{2}$-densities $g_* r$ and $g_* \tau_1$ on Z_2. If $a \in C_0^\infty(Z_2 \times \mathbb{R})$ then the change of variables formula for an integral implies that

$$\pi_{2,*} a e^{i\frac{\phi_2}{\hbar}} g_* \tau_1 = \pi_{1,*} g^* a e^{i\frac{\phi_1}{\hbar}} \tau_1$$

where the push-forward $\pi_{2,*}$ on the left is relative to $g_* r$ and the push-forward on the right is relative to r. □

Increasing the number of fiber variables.

We may assume that $Z_2 = Z_1 \times S$ where S is an open subset of $\mathbb{R}^m$ and

$$\phi_2(z,s) = \phi_1(z) + \frac{1}{2}\langle As, s\rangle$$

where A is a symmetric non-degenerate $m \times m$ matrix. We write Z for Z_1. If d is the fiber dimension of Z then $d+m$ is the fiber dimension of Z_2. Let r be a vertical $\frac{1}{2}$-density on Z so that $r \otimes ds^{\frac{1}{2}}$ is a vertical $\frac{1}{2}$-density on Z_2. Let τ be a $\frac{1}{2}$ density on Z so that $\tau \otimes ds^{\frac{1}{2}}$ is a $\frac{1}{2}$-density on Z_2. We want to consider the expression

$$\hbar^{k-\frac{d+m}{2}} \pi_{2*} a_2(z,s,\hbar) e^{i\frac{\phi_2(z,s)}{\hbar}} (\tau \otimes ds^{\frac{1}{2}}).$$

Let $\pi_{2,1} : Z \times S \to Z$ be projection onto the first factor so that

$$\pi_{2*} = \pi_{1*} \circ \pi_{2,1*}$$

and the operation $\pi_{2,1*}$ sends

$$a_2(z,s,\hbar) e^{i\frac{\phi_2}{\hbar}} \tau \otimes ds^{\frac{1}{2}} \mapsto b(z,\hbar) e^{i\frac{\phi_1}{\hbar}} \tau$$

where

$$b(z,\hbar) = \int a_2(z,s,\hbar) e^{i\frac{\langle As,s\rangle}{2\hbar}} ds.$$

We now apply the Lemma of Stationary Phase (see Chapter 15) to conclude that

$$b(z,\hbar) = \hbar^{m/2} a_1(z,\hbar)$$

and in fact

$$a_1(z,\hbar) = c_A a_2(z,0,\hbar) + O(\hbar), \tag{8.3}$$

where c_A is a non-zero constant depending only on A. □

8.1.3 The global definition of $I^k(X,\Lambda)$.

Let (Λ,ψ) be an exact Lagrangian submanifold of T^*X. We can find a locally finite open cover of Λ by open sets Λ_i such that each Λ_i is defined by a generating function ϕ_i relative to a fibration $\pi_i : Z_i \to U_i$ where the U_i are open subsets of X. We let $I_0^k(X,\Lambda)$ consist of those $\frac{1}{2}$-densities which can be written as a finite sum of the form

$$\mu = \sum_{j=1}^{N} \mu_{i_j}, \quad \mu_{i_j} \in I_0^k(X,\Lambda_{i_j}).$$

By the results of the preceding section we know that this definition is independent of the choice of open cover and of the local descriptions by generating functions.

We then define the space $I^k(X,\Lambda)$ to consist of those $\frac{1}{2}$-densities μ on X such that $\rho\mu \in I_0^k(X,\Lambda)$ for every C^∞ function ρ on X of compact support.

8.2 Semi-classical Fourier integral operators.

Let X_1 and X_2 be manifolds, let

$$X = X_1 \times X_2$$

and let

$$M_i = T^*X_i, \quad i = 1,2.$$

Finally, let (Γ,Ψ) be an exact canonical relation from M_1 to M_2 so

$$\Gamma \subset M_1^- \times M_2.$$

Let

$$\varsigma_1 : M_1^- \to M_1, \quad \varsigma_1(x_1,\xi_1) = (x_1,-\xi_1)$$

so that

$$\Lambda := (\varsigma_1 \times \mathrm{id})(\Gamma)$$

and

$$\psi = \Psi \circ (\varsigma_1 \times \mathrm{id})$$

gives an exact Lagrangian submanifold (Λ,ψ) of

$$T^*X = T^*X_1 \times T^*X_2.$$

Associated with (Λ,ψ) we have the space of compactly supported oscillatory $\frac{1}{2}$-densities $I_0^k(X,\Lambda)$. Choose a nowhere vanishing density on X_1 which we will denote (with some abuse of language) as dx_1 and similarly choose a nowhere vanishing density dx_2 on X_2. We can then write a typical element μ of $I_0^k(X,\Lambda)$ as

$$\mu = u(x_1,x_2,\hbar)dx_1^{\frac{1}{2}}dx_2^{\frac{1}{2}}$$

where u is a smooth function of compact support in all three “variables”.

Recall that $L^2(X_i)$ is the intrinsic Hilbert space of L^2 half densities on X_i. Since u is compactly supported, we can define the integral operator

$$F_\mu = F_{\mu,\hbar}: \quad L^2(X_1) \to L^2(X_2)$$

by

$$F_\mu(f dx_1^{\frac{1}{2}}) = \left(\int_{X_1} f(x_1) u(x_1, x_2, \hbar) dx_1 \right) dx_2^{\frac{1}{2}}. \tag{8.4}$$

We will denote the space of such operators by

$$\mathcal{F}_0^m(\Gamma)$$

where

$$m = k + \frac{n_2}{2}, \quad n_2 = \dim X_2,$$

and call them compactly supported **semi-classical Fourier integral operators**. In other words, $F_{\mu,\hbar}$ is a compactly supported semi-classical Fourier integral operator of degree m if and only if its Schwartz kernel belongs to $I_0^{m-\frac{n_2}{2}}$.

We could, more generally, demand merely that $u(x_1, x_2, \hbar) dx_1^{\frac{1}{2}}$ be an element of $L^2(X_1)$ in this definition, in which case we would drop the subscript 0.

8.2.1 Composition of semi-classical Fourier integral operators.

Let X_1, X_2 and X_3 be manifolds, let $M_i = T^*X_i,\ i = 1, 2, 3$ and let

$$(\Gamma_1, \Psi_1) \in \mathrm{Morph}_{exact}(M_1, M_2), \qquad (\Gamma_2, \Psi_2) \in \mathrm{Morph}_{exact}(M_2, M_3)$$

be exact canonical relations. Let

$$F_1 \in \mathcal{F}_0^{m_1}(\Gamma_1) \quad \text{and} \quad F_2 \in \mathcal{F}_0^{m_2}(\Gamma_2).$$

Theorem 8.2.1. *If Γ_2 and Γ_1 are transversally composible, then*

$$F_2 \circ F_1 \in \mathcal{F}_0^{m_1+m_2}((\Gamma_2, \psi_2) \circ (\Gamma_1, \psi_1)). \tag{8.5}$$

where the composition of exact canonical relations is given in (4.58) and (4.59).

Proof. By partition of unity we may assume that we have fibrations

$$\pi_1 : X_1 \times X_2 \times S_1 \to X_1 \times X_2, \qquad \pi_2 : X_2 \times X_3 \times S_2 \to X_2 \times X_3$$

where S_1 and S_2 are open subsets of $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$ and that ϕ_1 and ϕ_2 are generating functions for Γ_1 and Γ_2 with respect to these fibrations. We also fix nowhere vanishing $\frac{1}{2}$-densities $dx_i^{\frac{1}{2}}$ on $X_i,\ i = 1, 2, 3$. So F_1 is an integral operator with respect to a kernel of the form (8.4) where

$$u_1(x_1, x_2, \hbar) = \hbar^{k_1 - \frac{d_1}{2}} \int a_1(x_1, x_2, s_1, \hbar) e^{i \frac{\phi_1(x_1, x_2, s_1)}{\hbar}} ds_1$$

where
$$k_1 = m_1 - \frac{n_2}{2}, \quad n_2 = \dim X_2$$
and F_2 has a similar expression (under the change $1 \mapsto 2,\ 2 \mapsto 3$). So their composition is the integral operator
$$f dx_1^{\frac{1}{2}} \mapsto \left(\int_{X_1} f(x_1) u(x_1, x_3, \hbar) dx_1 \right) dx_3^{\frac{1}{2}}$$
where
$$u(x_1, x_3, \hbar) = \hbar^{k_1 + k_2 - \frac{d_1+d_2}{2}} \times$$
$$\int a_1(x_1, x_2, s_1, \hbar) a_2(x_2, x_3, s_2, \hbar) e^{i \frac{\phi_1 + \phi_2}{\hbar}} ds_1 ds_2 dx_2. \tag{8.6}$$
By Theorem 5.6.1 $\phi_1(x_1, x_2, s_1) + \phi_2(x_2, x_3, s_2)$ is a generating function for $\Gamma_2 \circ \Gamma_1$ with respect to the fibration
$$X_1 \times X_3 \times (X_2 \times S_1 \times S_2) \to X_1 \times X_3,$$
and by (4.59) this is a generating function for $(\Gamma_2, \Psi_2) \circ (\Gamma_1, \Psi_1)$. Since the fiber dimension is $d_1 + d_2 + n_2$ and the exponent of $\hbar$ in the above expression is $k_1 + k_2 - \frac{d_1+d_2}{2}$ we obtain (8.5). □

8.3 The symbol of an element of $I^k(X,\Lambda)$.

Let $\Lambda = (\Lambda, \psi)$ be an exact Lagrangian submanifold of T^*X. We have attached to Λ the space $I^k(X,\Lambda)$ of oscillating $\frac{1}{2}$-densities. The goal of this section is to give an intrinsic description of the quotient
$$I^k(X,\Lambda)/I^{k+1}(X,\Lambda)$$
as sections of a line bundle $\mathbb{L} \to \Lambda$.

8.3.1 A local description of $I^k(X,\Lambda)/I^{k+1}(X,\Lambda)$.

Let S be an open subset of $\mathbb{R}^d$ and suppose that we have a generating function $\phi = \phi(x,s)$ for Λ with respect to the fibration
$$X \times S \to X, \quad (x,s) \mapsto x.$$
Fix a C^∞ nowhere vanishing $\frac{1}{2}$-density ν on X so that any other smooth $\frac{1}{2}$-density μ on X can be written as
$$\mu = u\nu$$
where u is a C^∞ function on X.

The critical set C_ϕ is defined by the d independent equations
$$\frac{\partial \phi}{\partial s_i} = 0, \qquad i = 1, \ldots, d \tag{8.7}$$

The fact that ϕ is a generating function of Λ asserts that the map

$$\lambda_\phi : C_\phi \to T^*X, \qquad (x,s) \mapsto (x, d\phi_X(x,s)) \tag{8.8}$$

is a diffeomorphism of C_ϕ with Λ. To say that $\mu = u\nu$ belongs to $I_0^k(X,\Lambda)$ means that the function $u(x,\hbar)$ can be expressed as the oscillatory integral

$$u(x,\hbar) = \hbar^{k-\frac{d}{2}} \int a(x,s,\hbar) e^{i\frac{\phi(x,s)}{\hbar}} ds, \quad \text{where} \quad a \in C_0^\infty(X \times S \times \mathbb{R}). \tag{8.9}$$

Proposition 8.3.1. *If $a(x,s,0) \equiv 0$ on C_ϕ then $\mu \in I_0^{k+1}(X,\Lambda)$.*

Proof. If $a(x,s,0) \equiv 0$ on C_ϕ then by the description (8.7) of C_ϕ we see that we can write

$$a = \sum_{j=1}^d a_j(x,s,\hbar)\frac{\partial \phi}{\partial s_j} + a_0(x,s,\hbar)\hbar.$$

We can then write the integral (8.9) as $v + u_0$ where

$$u_0(x,\hbar) = \hbar^{k+1-\frac{d}{2}} \int a_0(x,s,\hbar) e^{i\frac{\phi(x,s)}{\hbar}} ds$$

so

$$\mu_0 = u_0\nu \in I_0^{k+1}(X,\Lambda)$$

and

$$\begin{aligned} v &= \hbar^{k-\frac{d}{2}} \sum_{j=1}^d \int a_j(x,s,\hbar) \frac{\partial \phi}{\partial s_j} e^{i\frac{\phi}{\hbar}} ds \\ &= -i\hbar^{k+1-\frac{d}{2}} \sum_{j=1}^d \int a_j(x,s,\hbar) \frac{\partial}{\partial s_j} e^{i\frac{\phi}{\hbar}} ds \\ &= i\hbar^{k+1-\frac{d}{2}} \sum_{j=1}^d \int \left(\frac{\partial}{\partial s_j} a_j(x,s,\hbar) \right) e^{i\frac{\phi}{\hbar}} ds \end{aligned}$$

so

$$v = i\hbar^{k+1-\frac{d}{2}} \int b(x,s,\hbar) e^{i\frac{\phi}{\hbar}} ds \qquad \text{where} \qquad b = \sum_{j=1}^d \frac{\partial a_j}{\partial s_j}. \tag{8.10}$$

This completes the proof of Proposition 8.3.1. □

This proof can be applied inductively to conclude the following sharper result:

Proposition 8.3.2. *Suppose that $\mu = u\nu \in I_0^k(X,\Lambda)$ where u is given by (8.9) and for $i = 0, \ldots, 2\ell - 1$*

$$\frac{\partial^i a}{\partial \hbar^i}(x,s,0)$$

vanishes to order $2(\ell - i)$ on C_ϕ. Then

$$\mu \in I_0^{k+2\ell+1}(X,\Lambda).$$

As a corollary we obtain:

Proposition 8.3.3. *If a vanishes to infinite order on C_ϕ then $\mu \in I^\infty(X,\Lambda)$, i.e.*

$$\mu \in \bigcap_k I^k(X,\Lambda).$$

8.3.2 The local definition of the symbol.

We now make a tentative definition of the symbol, one that depends on the presentation (Z,π,ϕ) of the Lagrangian manifold, and also on the choices of non-vanishing half densities: If $\mu \in I^k(X,\Lambda)$ we define the function $\sigma_\phi \in C^\infty(\Lambda)$ by

$$\sigma_\phi(\mu)(x,\xi) = a(x,s,0) \text{ where } (x,s) \in C_\phi \text{ and } \lambda_\phi(x,s) = (x,\xi). \tag{8.11}$$

Strictly speaking, we should also include the choice of non-vanishing half-densities in the notation for σ but this would clutter up the page too much.

The symbol as just defined depends on the presentation of Λ and on the choices of non-vanishing half-densities. However, we claim that

Proposition 8.3.4. *If $p \in \Lambda$, the assertion that $(\sigma_\phi(\mu))(p) = 0$ has an intrinsic significance, i.e. is independent of all the above choices.*

Proof. Changing the choice of non-vanishing half-densities clearly multiplies $\sigma_\phi(\mu)(p)$ by a non-zero factor. So we must investigate the dependence on the presentation. As in Section 8.1.2, we must check what happens for the two Hörmander moves: For the case of equivalence this is obvious. When increasing the number fiber variables as in Section 8.1.2 (and with the notation of that section) we have $C_{\phi_2} = C_{\phi_1} \times \{0\}$ and setting $\hbar = 0$ in (8.3) shows that $\sigma_{\phi_1}(\mu) = c_A\sigma_{\phi_2}(\mu)$ where $c_A \neq 0$. □

8.3.3 The intrinsic line bundle and the intrinsic symbol map.

With the above notation, define

$$I_p^k(X,\Lambda) := \left\{\mu \in I^k(X,\Lambda) | \sigma_\phi(\mu)(p) = 0\right\}.$$

According to Prop. 8.3.4, this is independent of all the choices that went into the definition of σ_ϕ. So we have defined a line bundle

$$\mathbb{L} \to \Lambda$$

where

$$\mathbb{L}_p := I^k(X,\Lambda)/I_p^k(X,\Lambda). \tag{8.12}$$

Multiplication by $\hbar^{\ell-k}$ is an isomorphism of $I^k(X,\Lambda)$ onto $I^\ell(X,\Lambda)$ and it is easy to check that this isomorphism maps $I_p^k(X,\Lambda)$ onto $I_p^\ell(X,\Lambda)$, so we see that the above definition is independent of k.

The choice of data that went into the definition of σ_ϕ gives a trivialization of $\mathbb{L}$ and shows that $\mathbb{L} \to \Lambda$ is indeed a smooth line bundle. It also shows the following: let us define the **intrinsic symbol map**

$$\sigma : \ I^k(X,\Lambda) \to \text{sections of } \ \mathbb{L}$$

by

$$\sigma(\mu)_p := [\mu]_p = \mu / I^k_p(X,\Lambda) \in \mathbb{L}_p$$

i.e. $\sigma(\mu)_p$ is the equivalence class of μ mod $I^k_p(X,\Lambda)$. Then $\sigma(\mu)$ is a smooth section of $\mathbb{L}$. In other words,

$$\sigma : I^k(X,\Lambda) \to C^\infty(\mathbb{L}).$$

The following proposition now follows from Prop. 8.3.1:

Proposition 8.3.5. *If $\mu \in I^k(X,\Lambda)$ and $\sigma(\mu) \equiv 0$ then $\mu \in I^{k+1}(X,\Lambda)$.*

We will soon prove the converse to this proposition and hence conclude that σ induces an isomorphism of $I^k(X,\Lambda)/I^{k+1}(X,\Lambda)$ with $C^\infty(\mathbb{L})$.

8.4 Symbols of semi-classical Fourier integral operators.

Let X_1 and X_2 be manifolds, with

$$n_2 = \dim X_2$$

and let

$$\Gamma \in \mathrm{Morph}(T^*X_1, T^*X_2)$$

be an exact canonical relation. Let

$$\Lambda = (\varsigma_1 \times \mathrm{id})(\Gamma)$$

where $\varsigma(x_1,\xi_1) = (x_1,-\xi_1)$ so that Λ is an exact Lagrangian submanifold of $T^*(X_1 \times X_2)$. We have associated to Γ the space of compactly supported semi-classical Fourier integral operators

$$\mathcal{F}^m_0(\Gamma)$$

where $F \in \mathcal{F}^m_0(\Gamma)$ is an integral operator with kernel

$$\mu \in I_0^{m-\frac{n_2}{2}}(X_1 \times X_2, \Lambda).$$

We have the line bundle $\mathbb{L}_\Lambda \to \Lambda$ and we define the line bundle $\mathbb{L}_\Gamma \to \Gamma$ to be the pull-back under $\varsigma \otimes \mathrm{id}$ of the line-bundle $\mathbb{L}_\Lambda$:

$$\mathbb{L}_\Gamma \to \Gamma \ := \ (\varsigma \otimes \mathrm{id})^*_{|\Gamma}(\mathbb{L}_\Lambda). \tag{8.13}$$

Similarly, if $F \in \mathcal{F}_0^m(\Gamma)$ corresponds to

$$\mu \in I_0^{m-\frac{n_2}{2}}(X_1 \times X_2, \Lambda)$$

we define the symbol of F to be

$$\sigma(F) = (\varsigma \otimes \mathrm{id})^*_{|\Gamma}\sigma(\mu). \tag{8.14}$$

8.4.1 The functoriality of the symbol.

We recall some results from Section 5.6: Let X_1, X_2 and X_3 be manifolds and

$$\Gamma_1 \in \mathrm{Morph}(T^*X_1, T^*X_2), \quad \Gamma_2 \in \mathrm{Morph}(T^*X_2, T^*X_3)$$

be canonical relations which are transversally composible. So we are assuming in particular that the maps

$$\Gamma_1 \to T^*X_2, \quad (p_1, p_2) \mapsto p_2 \quad \text{and} \quad \Gamma_2 \to T^*X_2, \quad (q_2, q_3) \mapsto q_2$$

are transverse.

Suppose that

$$\pi_1 : Z_1 \to X_1 \times X_2, \quad \pi_2 : Z_2 \to X_2 \times X_3$$

are fibrations and that $\phi_i \in C^\infty(Z_i),\ i = 1, 2$ are generating functions for Γ_i with respect to π_i.

From π_1 and π_2 we get a map

$$\pi_1 \times \pi_2 : Z_1 \times Z_2 \to X_1 \times X_2 \times X_2 \times X_3.$$

Let

$$\Delta_2 \subset X_2 \times X_2$$

be the diagonal and let

$$Z := (\pi_1 \times \pi_2)^{-1}(X_1 \times \Delta_2 \times X_3).$$

Finally, let

$$\pi : Z \to X_1 \times X_3$$

be the fibration

$$Z \to Z_1 \times Z_2 \to X_1 \times X_2 \times X_2 \times X_3 \to X_1 \times X_3$$

where the first map is the inclusion map and the last map is projection onto the first and last components. Let

$$\phi : Z \to \mathbb{R}$$

be the restriction to Z of the function (5.8)

$$(z_1, z_2) \mapsto \phi_1(z_1) + \phi_2(z_2).$$

Then (Theorem 5.6.1) ϕ is a generating function for

$$\Gamma := \Gamma_2 \circ \Gamma_1$$

with respect to the fibration $\pi : Z \to X_1 \times X_3$.

Suppose that we have chosen trivializing data for semi-classical Fourier integral operators as in Section 8.2, and, more particularly, as in the proof of Theorem 8.2.1. So $F = F_2 \circ F_1$ corresponds to $\mu \in I^k(X, \Lambda)$ given by (8.6). We have the diffeomorphism

$$\kappa : \Gamma_2 \star \Gamma_1 \to \Gamma_2 \circ \Gamma_1$$

where

$$\Gamma_2 \star \Gamma_1 = \{(m_1, m_2.m_3)|(m_1, m_2) \in \Gamma_1,\ (m_2, m_3) \in \Gamma_2\}.$$

We also have the projections

$$\mathrm{pr}_1 : \Gamma_2 \star \Gamma_1 \to \Gamma_1, \quad (m_1, m_2, m_3) \mapsto (m_1, m_2)$$

and

$$\mathrm{pr}_2 : \Gamma_2 \star \Gamma_1 \to \Gamma_2, \quad (m_1, m_2, m_3) \mapsto (m_2, m_3).$$

Our choices of trivializing data give trivializations of $\mathbb{L}_1 \to \Gamma_1$ and of $\mathbb{L}_2 \to \Gamma_2$ and hence of

$$\mathrm{pr}_1^* \mathbb{L}_1 \otimes \mathrm{pr}_2^* \mathbb{L}_2 \to \Gamma_2 \star \Gamma_1.$$

Also, our choice of dx_2 gives a choice of trivializing data for (Z, π, ϕ) representing Γ. Indeed, in terms of local product representations $Z_1 = X_1 \times X_2 \times S_1$ and $Z_2 = X_2 \times X_2 \times S_2$ we now have the half-density $dx_2^{\frac{1}{2}} \otimes ds_1^{\frac{1}{2}} \otimes ds_2^{\frac{1}{2}}$ on $X_2 \times S_1 \times S_2$.

We have the diffeomorphism $\gamma := \gamma_\phi : C_\phi \to \Gamma$ and the maps

$$\gamma_i : C_{\phi_i} \to \Gamma_i, \quad i = 1, 2$$

as in the proof of Theorem 5.6.1. We have the immersion

$$\iota : C_\phi \to C_{\phi_1} \times C_{\phi_2}$$

given by

$$\iota(x_1, x_3, x_2, s, t) = ((x_1, x_2, s), (x_2, x_3, t)).$$

The amplitude in (8.6) is

$$a(x_1, x_3, x_2, s, t) = a_1(x_1, x_2, s) a_2(x_2, x_3, t)$$

so

$$a_{|C_\phi} = \iota^* \left(a_{1|C_{\phi_1}} \cdot a_{2|C_{\phi_2}} \right) \tag{8.15}$$

We have

$$\sigma_\phi(F) = (\gamma^{-1})^* a_{|C_\phi, \hbar=0}$$

with similar expressions for $\sigma_{\phi_1}(F_1)$ and $\sigma_{\phi_2}(F_2)$. Also, if $j : \Gamma_2 \star \Gamma_1 \to \Gamma_1 \times \Gamma_2$ denotes the injection

$$j(m_1, m_2, m_3) = ((m_1, m_2), (m_2, m_3))$$

then

$$j \circ \kappa^{-1} \circ \gamma = (\gamma_1 \times \gamma_2) \circ \iota$$

as maps from C_ϕ to $\Gamma_1 \times \Gamma_2$. In other words,

$$\iota \circ \gamma^{-1} \circ \kappa = (\gamma_1^{-1} \times \gamma_2^{-1}) \circ j$$

as maps from $\Gamma_2 \star \Gamma_1$ to $C_{\phi_1} \times C_{\phi_2}$. Setting $\hbar = 0$ in (8.15) we see that

$$\kappa^* \sigma_\phi(F) = j^* \left(\sigma_{\phi_1}(F_1)\sigma_{\phi_2}(F_2)\right). \tag{8.16}$$

In this equation, the data entering into the definition of σ_ϕ must be chosen consistently with the data defining σ_{ϕ_1} and σ_{ϕ_2}. But we see from this equation that if

$$p = \kappa(p_1, p_2, p_3), \qquad (p_1, p_2, p_3) \in \Gamma_2 \star \Gamma_1$$

then

$$\sigma_\phi(F)(p) = 0 \Leftrightarrow \quad \text{either } \sigma_{\phi_1}(F_1)(p_1, p_2) = 0 \quad \text{or} \quad \sigma_{\phi_2}(F_2)(p_2, p_3) = 0. \tag{8.17}$$

The condition of vanishing or not vanishing of the symbol is intrinsic, as we have seen. Let

$$\mathbb{L} \to \Gamma_2 \circ \Gamma_1, \ \ \mathbb{L}^1 \to \Gamma_1 \quad \text{and} \quad \mathbb{L}^2 \to \Gamma_2$$

be the intrinsic line bundles so that

$$\mathbb{L}^1_{(p_1,p_2)} = \mathcal{F}^{m_1}(\Gamma_1)/\mathcal{F}^{m_1}_{(p_1,p_2)}(\Gamma_1)$$

where $\mathcal{F}^{m_1}_{(p_1,p_2)}$ denotes those elements of $\mathcal{F}^{m_1}(\Gamma_1)$ whose symbols vanish at (p_1, p_2) with similar notation for $\mathbb{L}^2$ and $\mathbb{L}$.

Then (8.17) says the following: If $F_1 \in \mathcal{F}^{m_1}(\Gamma_1)$ and $F_2 \in \mathcal{F}^{m_2}(\Gamma_2)$ then $\sigma(F_2 \circ F_1)(p) = 0$ if and only if either $\sigma(F_1)(p_1, p_2) = 0$ or $\sigma(F_2)(p_2, p_3) = 0$ (or both). Thus composition of operators induces an isomorphism

$$\mathbb{L}_p \cong \mathbb{L}^1_{(p_1,p_2)} \otimes \mathbb{L}^2_{(p_2,p_3)}. \tag{8.18}$$

We have proved the following theorem:

Theorem 8.4.1. *Composition of semi-classical Fourier integral operators induces multiplication of their symbols in the following sense: Let*

$$\Gamma_1 \in \mathrm{Morph}(T^*X_1, T^*X_2), \quad \Gamma_2 \in \mathrm{Morph}(T^*X_2, T^*X_3)$$

be exact canonical relations and

$$\mathbb{L}^1 \to \Gamma_1, \quad \mathbb{L}^2 \to \Gamma_2$$

their associated intrinsic line bundles. Suppose that Γ_2 *and* Γ_1 *are transversally composible and let*

$$\Gamma = \Gamma_2 \circ \Gamma_1$$

and $\mathbb{L} \to \Gamma$ *its line bundle. Let*

$$\kappa : \Gamma_2 \star \Gamma_1 \to \Gamma$$

be the diffeomorphism $\kappa(p_1, p_2, p_3) = (p_1, p_3)$ *and* $j : \Gamma_2 \star \Gamma_1 \to \Gamma_1 \times \Gamma_2$ *the immersion* $j(p_1, p_2, p_3) = ((p_1, p_2), (p_2, p_3))$*. Then we have a canonical isomorphism*

$$\kappa^* \mathbb{L} \cong j^* (\mathbb{L}_1 \otimes \mathbb{L}_2). \tag{8.19}$$

If $F_1 \in \mathcal{F}^{m_1}(\Gamma_1)$ *and* $F_2 \in \mathcal{F}^{m_2}(\Gamma_1)$ *(so that* $F_2 \circ F_1 \in \mathcal{F}^{m_1+m_2}(\Gamma)$*) then*

$$\kappa^*(\sigma(F_2 \circ F_1)) = j^*(\sigma(F_1)\sigma(F_2)) \tag{8.20}$$

under the isomorphism (8.19).

We can now prove the converse to Prop. 8.3.1:

Proposition 8.4.1. *Let* μ *be an element of* $I^k(X, \Lambda)$ *and* $\sigma(\mu) = \sigma_k(\mu)$ *denote its symbol (as an element of* $I^k(X, \Lambda)$*). If* $\mu \in I^{k+1}(X, \Lambda)$ *then*

$$\sigma(\mu) \equiv 0.$$

Proof. Let us first prove this for the case that Λ is horizontal. So (locally) we can assume that $\Lambda = \Lambda_\phi$. So the fibration is trivial, and hence the critical set C_ϕ is X itself and the diffeomorphism $\lambda_\phi : X \to \Lambda_\phi$ is just the map $x \mapsto d\phi_x$. Any $\mu \in I^k(X, \Lambda)$ is of the form

$$\mu = \hbar^k a(x, \hbar) e^{i\frac{\phi}{\hbar}} dx^{\frac{1}{2}}$$

(with no integration) and

$$\sigma_\phi(\mu) = (\lambda_\phi^{-1})^* a(x, 0).$$

To say that $\mu \in I^{k+1}(X, \Lambda_\phi)$ means that μ is of the form

$$\hbar^{k+1} b(x, \hbar) e^{i\frac{\phi}{\hbar}}.$$

This implies that $a(x, \hbar) = \hbar b(x, \hbar)$, so setting $\hbar = 0$ shows that $\sigma(\mu) \equiv 0$. So the Proposition is trivially true when Λ is horizontal.

Now to the general case. Given any Lagrangian submanifold $\Lambda \subset T^*X$ and any $p \in \Lambda$, we can find a horizontal Lagrangian submanifold Λ_ϕ such that

$$\Lambda_\phi \cap \Lambda = \{p\}$$

and such that this intersection is transverse. Let $\mu_1 \in I^0(X, \Lambda_\phi)$ so that

$$\mu_1 = a_1(x, \hbar) e^{i\frac{\phi}{\hbar}} dx^{\frac{1}{2}}$$

and we choose μ_1 so that $a_1(x,0)$ does not vanish. In other words, $\sigma(\mu_1)$ is nowhere vanishing. We think of Λ as an element of Morph(pt., T^*X) and of $\Lambda_\phi^\dagger$ as an element of Morph(T^*X, pt.). This is a transverse composition, so for $\mu \in I^k(X,\Lambda)$ we have

$$F_{\mu_1} \circ F_\mu = F_\nu, \quad \text{where} \quad \nu \in I^{k+\frac{n}{2}}(\text{pt.})$$

so

$$\nu = \hbar^k c(\hbar)$$

and

$$\sigma(\nu) = \sigma(\mu_1)(p)\sigma(\mu)(p) = c(0).$$

If μ were actually in $I^{k+1}(X,\Lambda)$ we would conclude that $\nu \in I^{k+1}(\text{pt.})$ so

$$\nu = \hbar^{k+1} c_1(\hbar)$$

implying that $c(\hbar) = \hbar c_1(\hbar)$ so $\sigma(\nu) = c(0) = 0$. Since $\sigma(\mu_1)(p) \neq 0$ we conclude that $\sigma(\mu)(p) = 0$. Since we can do this for every $p \in \Lambda$ we conclude that $\sigma(\mu) \equiv 0$. □

Putting together Propositions 8.3.1 and 8.4.1 we obtain:

Theorem 8.4.2. *The symbol map σ induces a bijection*

$$I^k(X,\Lambda)/I^{k+1}(X,\Lambda) \to C^\infty(\mathbb{L}).$$

8.5 The Keller-Maslov-Arnold description of the line bundle $\mathbb{L}$

Let X be an n-dimensional manifold and $\Lambda \subseteq T^*X$ an exact Lagrangian submanifold. In §8.3 we proved that there exists an intrinsically defined line bundle $\mathbb{L} \to \Lambda$ and symbol map

$$\sigma_{\mathbb{L}} : I^k(X,\Lambda) \to \mathcal{C}^\infty(\mathbb{L}) \tag{8.21}$$

which is surjective and has kernel $I^{k+1}(X,\mathbb{L})$. In this section we will show that $\mathbb{L} \cong \mathbb{L}_{\text{Maslov}} \otimes |T\Lambda|^{\frac{1}{2}}$ and give a much more concrete description of this map. We'll begin by reviewing some material in §7.4–7.5 on "enhancing" fibrations. Let $Z \xrightarrow{\pi} X$ be a fibration and let V be the vertical sub-bundle of TZ. An *enhancement* of π is the choice of a non-vanishing section, v_π, of the $\frac{1}{2}$-density bundle, $|V|^{\frac{1}{2}}$. This enhancement does two things for us: it gives us a non-vanishing $\frac{1}{2}$ density, ρ_π, on the canonical relation Γ_π, and it also enables us to define a fiber integration operation

$$\pi_* : \mathcal{C}_0^\infty(|TZ|^{\frac{1}{2}}) \to \mathcal{C}_0^\infty(|TX|^{\frac{1}{2}})\,.$$

Now let Λ be an exact Lagrangian submanifold of T^*X and $\phi : Z \to \mathbb{R}$ a generating function for Λ with respect to π. Then by definition

$$\Lambda = \Gamma_\pi \circ \Lambda_\phi$$

where Λ_ϕ is the Lagrangian submanifold, $\{(q, d\phi_q)\,, \quad q \in Z\}$, of T^*Z. So if we are given a $\frac{1}{2}$ density, v, on Λ_ϕ we can associate with it a $\frac{1}{2}$ density $\rho_\pi \circ v$ on Λ by the composition described in (7.1). In particular let $\mu \in I^k(X, \Lambda)$ be the oscillatory $\frac{1}{2}$-density (8.1), i.e.,

$$\mu = (2\pi h)^{k-\frac{d}{2}} \pi_* \nu \tag{8.22}$$

where $\nu \in I^0(Z, \Lambda_\phi)$ is the oscillatory half-density, $\nu := a(z, \hbar) e^{i\frac{\phi}{\hbar}} \tau$.

Let us denote by $\wp$ the projection of Λ_φ onto Z. We define the "Symbol" of ν to be the $\frac{1}{2}$ density, $\sigma(\nu) = \wp^*(a(z, 0)\tau)$ on Λ, and we define the Symbol of μ to be the product

$$\sigma(\mu) := s_\phi \rho_\pi \circ \sigma(\nu) \tag{8.23}$$

where s_ϕ is the section of $\mathbb{L}_{\text{Maslov}}$ associated with ϕ, (see Sec. 5.13.2.)

We will show below that this "Symbol" is intrinsically defined. Assuming this for the moment, we now show that the "Symbol" map we've just defined:

$$\sigma : I^k(X, \Lambda) \to \mathcal{C}^\infty(\mathbb{L}_{\text{Maslov}} \otimes |T\Lambda|^{\frac{1}{2}}) \tag{8.24}$$

coincides with the map (8.21). In particular, this will show that the line bundle $\mathbb{L}$ of (8.21) can be identified with $\mathbb{L}_{\text{Maslov}} \otimes |T\Lambda|^{\frac{1}{2}}$.

To prove this we show that this map is surjective and that its kernel is $I^{k+1}(X, \Lambda)$.

To see that this is the case let's go back to §5.1 and recall how the composition $\Gamma_\pi \circ \Lambda_\phi$ is defined. As in §5.1 let H^*Z be the *horizontal* sub-bundle of T^*Z. Then one has canonical identifications

$$\Gamma_\pi = H^*Z$$

and

$$\Gamma_\pi \circ \Lambda_\phi = \Lambda_\phi \cap H^*Z.$$

The assumption that Γ_π and Λ_ϕ are transversally composable simply says that this intersection is transversal. i.e., that every point we have

$$T_p\Lambda_\phi \cap T_pH^*Z = T_p(\Lambda_\phi \cap H^*Z)$$

and

$$T_p\Lambda_\phi + T_pH^*Z = T_p(T^*Z)\,.$$

So at every p we are in the situation of (6.10). In other words one has a short exact sequence

$$0 \to T_p(\Lambda_\phi \cap H^*Z) \to T_p\Lambda_\phi \oplus T_pH^*Z \to T_pT^*Z \to 0\,.$$

Moreover, T^*Z is a symplectic manifold, so $|T_pT^*Z|^{\frac{1}{2}} \cong \mathbb{C}$, so (taking $\alpha = \frac{1}{2}$ in (6.10—))from this short exact sequence we get an isomorphism

$$|T_p(\Lambda_\phi \cap H^*Z)|^{\frac{1}{2}} = |T_p\Lambda_\phi|^{\frac{1}{2}} \otimes |T_pH^*Z|^{\frac{1}{2}}$$

and from the $\frac{1}{2}$-densities $\sigma(\nu)(p)$ and $\rho_\pi(p)$, we get a $\frac{1}{2}$-density

$$\sigma(\nu)(p)\sigma(\pi)(p) \in |T_p(\Lambda_\phi \cap H^*Z)|^{\frac{1}{2}}.$$

From the diffeomorphism,

$$\Lambda_\phi \cap H^*Z \to \Gamma_\pi \circ \Lambda\,,$$

mapping $p = (q, d\varphi(q))$ to $\lambda_\varphi(q)$, this maps to the $\frac{1}{2}$-density, $(\rho_\pi \circ \sigma(\nu))(\lambda_\phi(q))$ in $|T_{\lambda_\varphi(p)}\Gamma_\pi \circ \Lambda_\phi|^{\frac{1}{2}}$.

Now recall that by (8.22) $\sigma(\nu) = \wp^*a(z,0)\wp^*\tau$ where $\wp^*\tau$ is a non-vanishing $\frac{1}{2}$-density on Λ_ϕ. Hence

$$\sigma_\pi \circ \sigma(\nu) = (\lambda_\phi^{-1})^*(a(z,0)|C_\phi)\sigma_\pi \circ \wp^*\tau \tag{8.25}$$

where $(\lambda_\phi^{-1})^*a(z,0)|C_\varphi$ is the "provisional symbol" of μ and $\rho_\pi \circ \wp^*\tau$ is a non-vanishing $\frac{1}{2}$-density on Λ. Thus it's clear that the symbol mapping (8.24) is surjective and that its kernel is $I^{k+1}(X,\Lambda)$.

This proves that we have the identification $\mathbb{L} \cong \mathbb{L}_{\text{Maslov}} \otimes |T\Lambda|^{\frac{1}{2}}$ and that under this identification, the map "Symbol" coincides with the intrinsic symbol map defined earlier, *assuming that "Symbol" is intrinsically defined.*

We will now show that the symbol (8.25) is intrinsically defined, i.e., doesn't depend on our choice of defining data z, π, ϕ, ν, σ. To check this it suffices to show that (8.23) is unchanged if we apply a sequence of Hörmander moves to these data:

1. Let us first consider what happens if we replace these data by diffeomorphic data: Z_1, π_1, ϕ_1, ν_1, σ_1 where $f : Z_1 \to Z$ is a diffeomorphism with the properties $\pi \cdot f = \pi_1$, $\phi \circ f = \phi_1$, $f^*\nu = \nu_1$ and $f^*\sigma = \sigma_1$. Since $f : Z_1 \to Z$ is a diffeomorphism it lifts to a symplectomorphism, $f^\# : T^*Z_1 \to T^*Z$ and $(f^\#)^*\sigma(\nu) = \sigma(\nu_1)$. Moreover since $\pi \circ f = \pi_1$ and $f^*\sigma = \sigma^1$, $f^\#$ maps $H^*Z_1 = \Gamma_\pi$ diffeomorphically onto $H^*Z = \Gamma_\pi$ and maps σ_{π_1} onto σ_π. Thus $\sigma_\pi \circ \sigma(\nu) = \sigma_{\pi_1} \circ \sigma(\nu_1)$. Also since $\phi \circ f = \phi_1$ the signature functions $(\text{sgn})^\# : C_\phi \to \mathbb{Z}$ and $(\text{sgn})^\# : C_{\phi_1} \to \mathbb{Z}$ (see 5.13.2) are intertwined by f and hence $s_\phi = s_{\phi_1}$. Thus

$$\sigma(\mu) = s_\phi \rho_\pi \circ \sigma(\nu) = s_{\phi_1}\rho_{\pi_1} \circ \sigma(\nu_1)\,. \tag{8.26}$$

2. The situation is a bit more complicated for the Hörmander move that increases the number of fiber variables. Let $Q = \mathbb{R}^\ell \to \mathbb{R}$ be a non-degenerate quadratic form, and let us replace Z by $Z_1 = Z \times \mathbb{R}^\ell$, π by $\pi_1 = \pi \circ \rho$, where ρ is the projection of $Z \times \mathbb{R}^\ell$ onto Z, replace $\varphi(z)$ by $\varphi_1(z,s) = \varphi(z) + Q(s)$, replace σ by $\sigma_1 = \sigma|ds|^{\frac{1}{2}}$ and ν by the expression

$$\nu_1 = (2\pi h)^{k-\frac{d+\ell}{2}}a_1(z,s,h)e^{\frac{i\varphi_1}{h}}\tau|ds|^{\frac{1}{2}}C_Q$$

where $a_1(z,0,h) = a(z,h)$ and

$$c_Q = e^{-\frac{i\pi}{4}\operatorname{sgn} Q} |\det Q|^{\frac{1}{2}} .$$

By stationary phase

$$\rho_* \nu_1 = \nu + O(h^{k+1-\frac{d}{2}})$$

and hence

$$\mu_1 = (\pi_1)_* \nu_1 = \mu + O(h^{k+1}) .$$

On the other hand we claim that

$$\sigma_\rho \circ \sigma(\nu_1) = e^{-\frac{\pi i}{4}\operatorname{sgn} Q} \sigma(\nu) .$$

Indeed to check this it suffices to check this for the fibration $\rho : \mathbb{R}^\ell \to pt$, for the generating function, $\varpi_1 = Q(s)$, for the fiber $\frac{1}{2}$-density, $\sigma_1 = ds^{\frac{1}{2}}$ and for $\nu_1 = e^{\frac{iQ(s)}{\hbar}} |ds|^{\frac{1}{2}}$, i.e., to show that, in this case, $\sigma_\rho \circ \sigma(\nu_1) = |\det Q|^{-\frac{1}{2}}$, and we'll leave this as an exercise. Thus

$$\sigma_{\pi_1} \circ \sigma(\nu_1) = e^{-\frac{i\pi}{4}\operatorname{sgn} Q} \sigma_\pi \circ \sigma(\nu) \tag{8.27}$$

On the other hand since $\varphi_1(z,s) = \varphi(z) + Q(s)$, $s_{\varphi_1} = e^{\frac{i\pi}{4}\operatorname{sgn} Q} s_\varphi$ so we again get

$$\sigma(\mu) = s_\varphi \sigma_\pi \circ \sigma(\nu) s_{\varphi_1} \sigma_\pi \circ \sigma(\nu) = s_{\varphi_1} \circ \sigma(\nu_1). \tag{8.28}$$

Since every Hörmander move is a succession of the two elementary Hörmander moves described above this proves that $\sigma(\mu)$ is intrinsically defined.

Remark. The definition of $\mathbb{L}$ that we've given in this section is due to Hörmander, but the presence of the phase factor, s_φ, in this definition has antecedents in earlier work of Joe Keller in geometric optics and of Maslov–Arnold on the fundamental group of Lagrangian manifolds, $\Lambda \subseteq T^*X$.

8.6 Microlocality.

We have identified $I^k(X,\Lambda)/I^{k+1}(X,\Lambda)$ as the space of smooth sections of a line bundle $\mathbb{L}$ over Λ. What about higher quotients of the form

$$I^k(X,\Lambda)/I^{k+\ell}(X,\Lambda),\ \ell > 1?$$

We will find in this section that $I^k(X,\Lambda)/I^{k+\ell}(X,\Lambda)$ can be identified with elements of a sheaf on Λ. As usual, we will first describe this identification via the choice of some local data, and then describe what happens when we change our choice.

So we start with a (local) presentation (Z,π,ϕ) of Λ where $Z = X \times \mathbb{R}^d$, where π is projection onto the first factor, and where we have chosen densities $ds = ds_1 \cdots ds_d$ on $\mathbb{R}^d$ and dx on X. Then $\mu \in I^k(X,\Lambda)$ means that

$$\mu = u dx^{\frac{1}{2}}$$

where

$$u = \hbar^{k-\frac{d}{2}} \int_{\mathbb{R}^k} a(z,\hbar) e^{i\frac{\phi}{\hbar}} ds$$

where $a \in C_0^\infty(Z)$.

Recall that $\mu \in I^{k+1}(X,\Lambda)$ if and only if $a(x,s,0)_{|C_\phi} \equiv 0$. Let us now examine what the condition $\mu \in I^{k+2}(X,\Lambda)$ says about a. The fact that $a(x,s,0)_{|C_\phi} = 0$ tells us that we can write

$$a(x,s,0) = \sum_k a_k(x,s) \frac{\partial \phi}{\partial s_k}$$

and hence that we can write

$$a(x,s,\hbar) = \sum_k a_k(x,s) \frac{\partial \phi}{\partial s_k} + \hbar b(x,s,\hbar).$$

Then

$$\begin{aligned} \int_{\mathbb{R}^k} a(z,\hbar) e^{i\frac{\phi}{\hbar}} ds &= \int_{\mathbb{R}^k} \sum_k a_k(x,s) \frac{\partial \phi}{\partial s_k} e^{i\frac{\phi}{\hbar}} ds + \hbar \int_{\mathbb{R}^k} b(z,\hbar) e^{i\frac{\phi}{\hbar}} ds \\ &= -i\hbar \int_{\mathbb{R}^k} \sum_k a_k(x,s) \frac{\partial}{\partial s_k} \left(e^{i\frac{\phi}{\hbar}} \right) ds + \hbar \int_{\mathbb{R}^k} b(z,\hbar) e^{i\frac{\phi}{\hbar}} ds \\ &= i\hbar \int_{\mathbb{R}^k} \sum_k \frac{\partial a_k(x,s)}{\partial s_k} e^{i\frac{\phi}{\hbar}} ds + \hbar \int_{\mathbb{R}^k} b(z,\hbar) e^{i\frac{\phi}{\hbar}} ds \end{aligned}$$

So define the operator r_ϕ by

$$r_\phi(a) := i \sum_k \frac{\partial a_k(x,s)}{\partial s_k} + b. \tag{8.29}$$

Then we can write $\mu \in I^{k+1}(X,\Lambda)$ as $\mu = u dx^{\frac{1}{2}}$ where

$$u = \hbar^{k+1-\frac{d}{2}} \int r_\phi a(x,s,\hbar) e^{i\frac{\phi}{\hbar}} ds,$$

and hence

$$\mu \in I^{k+2}(X,\Lambda) \Leftrightarrow (r_\phi a(x,s,0))_{|C_\phi} = 0.$$

Notice that the operator r_ϕ involves a and its first two partial derivatives.

Iterating this argument proves

Proposition 8.6.1. *If $\mu \in I^k(X,\Lambda)$ and $\ell \geq 0$ then*

$$\mu \in I^{k+\ell}(X,\Lambda) \Leftrightarrow (r_\phi^j a)(z,0)_{|C_\phi} = 0 \quad \textit{for } 0 \leq j \leq \ell. \tag{8.30}$$

We now examine what this proposition tells us about $I^k/I^{k+\ell}$. For this we make some further choices:

Let $\mathcal{O}$ be tubular neighborhood of C_ϕ in Z, so that we have a retraction map

$$\mathrm{r} : \mathcal{O} \to C_\phi$$

and let $\rho \in C_0^\infty$ be a function which is:

- identically one in a neighborhhood of C_ϕ,
- with supp $\rho \subset \mathcal{O}$ and such that
- r: supp $\rho \to C_\phi$ is proper.

If $\mu = \hbar^k(\pi_* a(z,\hbar)e^{i\frac{\phi}{\hbar}}dz)dx^{\frac{1}{2}}$ and $\nu = \hbar^k(\pi_*\rho(z)a(z,\hbar)e^{i\frac{\phi}{\hbar}}dz)dx^{\frac{1}{2}}$ then $\mu - \nu \in I^\infty(X,\Lambda)$ since $\rho a = a$ in a neighborhhod of C_ϕ.

Proposition 8.6.2. *Every $\mu \in I^k(X,\Lambda)$ has a unique expression modulo $I^{k+\ell}(X,\Lambda)$ of the form*

$$\mu = \hbar^{k-\frac{d}{2}}\pi_*\left(\rho(z)\left(\sum_{j=0}^{\ell-1}\mathrm{r}^*\mathrm{a}_j\hbar^j\right)e^{i\frac{\phi}{\hbar}}dz\right)dx^{\frac{1}{2}}$$

with

$$\mathrm{a}_j \in C_0^\infty(C_\phi).$$

Proof. Let $\mu \in I^k(X,\Lambda) = \hbar^k\pi_*(a(z,\hbar)e^{i\frac{\phi}{\hbar}}dz)dx^{\frac{1}{2}}$. Let

$$\mathrm{a}_0 := a(z,0)_{|C_\pi}$$

and

$$\mu_1 := \mu - \hbar^{k-\frac{d}{2}}\pi_*\left(\rho(z)\mathrm{r}^*(\mathrm{a}_0)(z)e^{i\frac{\phi}{\hbar}}dz\right)dx^{\frac{1}{2}}.$$

Then $\mu_1 \in I^{k+1}(X,\Lambda)$ so

$$\mu_1 = \hbar^{k-\frac{d}{2}}\pi_*(\mathrm{b}(z,\hbar)e^{i\frac{\phi}{\hbar}}dz)dx^{\frac{1}{2}}$$

for some $\mathrm{b} \in C_0^\infty(Z\times\mathbb{R})$. Set $\mathrm{a}_1 := \mathrm{b}(z,0)_{|C_\phi}$ and

$$\mu_2 := \mu - \hbar^{k+1-\frac{d}{2}}\pi_*\left(\rho(z)(\mathrm{r}^*\mathrm{a}_1)(z)e^{i\frac{\phi}{\hbar}}dz\right)dx^{\frac{1}{2}}.$$

Then $\mu_2 \in I^{k+2}(X,\Lambda)$. Continue. ☐

Let us define

$$\sigma_{\mathcal{O}}^\ell : I^k(X,\Lambda) \to \bigoplus_{j=0}^{\ell-1}\hbar^j C_0^\infty(\Lambda)$$

by

$$\mu \mapsto (\lambda_\phi^{-1})^*\left(\sum_{j=0}^{\ell-1}\hbar^j\mathrm{a}_j.\right)$$

This map is independent of the choice of cutoff function ρ. Indeed, if we had two cutoff functions, they would agree in some neighborhood of C_ϕ and hence give the same a_j.

We need to investigate how $\sigma^\ell_{\mathcal{O}}$ depends on the choice of the tubular neighborhood $\mathcal{O}$. So let $\mathcal{O}_1$ and $\mathcal{O}_2$ be two such tubular neighborhoods. Let us set

$$\sigma^\ell_1 := \sigma^\ell_{\mathcal{O}_1} \quad \text{and} \quad \sigma^\ell_2 := \sigma^\ell_{\mathcal{O}_2}.$$

Proposition 8.6.3. *There exists a differential operator*

$$P : \bigoplus_{j=0}^{\ell-1} \hbar^j C^\infty_0(\Lambda) \to \bigoplus_{j=0}^{\ell-1} \hbar^j C^\infty_0(\Lambda)$$

of degree $2\ell - 2$ *such that*

$$\sigma^\ell_2 = P \circ \sigma^\ell_1.$$

Proof. Since the maps σ^ℓ_i, $\;i = 1, 2,\;$ are independent of the choice of cutoff functions, we may choose a common cutoff function ρ supported in $\mathcal{O}_1 \cap \mathcal{O}_2$. Suppose that

$$\mathrm{g} = \sum_{j=0}^{\ell-1} \hbar^j \mathrm{a}_j$$

where $\mathrm{a}_0, \dots \mathrm{a}_{\ell-1}$ are elements of $C^\infty_0(C_\phi)$ and that

$$\mu = \hbar^{k-\frac{d}{2}} \pi_* \left(\rho \,(\mathrm{r}_1^* \mathrm{g})\, e^{i\frac{\phi}{\hbar}} dz \right) dx^{\frac{1}{2}}$$

so that

$$\sigma^\ell_1(\mu) = (\lambda_\phi^{-1})^* \mathrm{g}.$$

Let

$$\begin{aligned} \nu &= \mu - \hbar^{k-\frac{d}{2}} \pi_* \left(\rho \,(\mathrm{r}_2^* \mathrm{g})\, e^{i\frac{\phi}{\hbar}} dz \right) dx^{\frac{1}{2}} \\ &= \hbar^{k-\frac{d}{2}} \pi_* \left(\rho \,(\mathrm{r}_1^* \mathrm{g} - \mathrm{r}_2^* \mathrm{g})\, e^{i\frac{\phi}{\hbar}} dz \right) dx^{\frac{1}{2}}. \end{aligned}$$

If we set

$$\tilde{\mathrm{g}} := \rho(\mathrm{r}_1^* \mathrm{g} - \mathrm{r}_2^* \mathrm{g})$$

then since $\tilde{\mathrm{g}}$ vanishes on C_ϕ we know that

$$\nu = \hbar^{k+1-\frac{d}{2}} \pi_* \left(r_\phi \tilde{\mathrm{g}} e^{i\frac{\phi}{\hbar}} dz \right) dx^{\frac{1}{2}}.$$

So define the operator P_1 by

$$P_1 \mathrm{g} = (r_\phi \tilde{\mathrm{g}})_{|C_\phi}.$$

We know that P_1 is a second order differential operator. Set

$$\mathrm{g}_1 := P_1 \mathrm{g}.$$

We have shown that

$$\mu = \hbar^{k-\frac{d}{2}} \pi_* \left(\rho \, (\mathrm{r}_2^* \mathrm{g}) \, e^{i\frac{\phi}{\hbar}} dz \right) dx^{\frac{1}{2}} \mod I^{k+1}(X, \Lambda).$$

In fact,

$$\mu = \hbar^{k-\frac{d}{2}} \pi_* \left(\rho \, (\mathrm{r}_2^* \mathrm{g}) \, e^{i\frac{\phi}{\hbar}} dz \right) dx^{\frac{1}{2}} \mu_1$$

where

$$\mu_1 = \hbar^{k+1-\frac{d}{2}} \pi_* \left(\rho \, (\mathrm{r}_1^* \mathrm{g}_1) \, e^{i\frac{\phi}{\hbar}} dz \right) dx^{\frac{1}{2}}.$$

Continuing in this way proves the proposition. □

8.6.1 The microsheaf.

Let U be an open subset of Λ. We define the subset

$$I_U^{k+\ell}(X, \Lambda)$$

by saying that for $\mu \in I^k(X, \Lambda)$ that

$$\mu \in I^{k+\ell}(X, \Lambda) \quad \Leftrightarrow \quad \sigma^\ell(\mu) \equiv 0 \quad \text{on} \quad U. \tag{8.31}$$

In order for this to make sense, we need to know that the condition $\sigma^\ell(\mu) \equiv 0$ is independent of the presentation. (We already know that it is independent of the tubular neighborhood $\mathcal{O}$ of C_ϕ.)

So we need to check this for each of the two Hörmander moves:

- Equivalence: In this case we have $(Z_1, \pi_1 \phi_1, dz_1)$ together with $(Z_2, \pi_2, \phi_2, dz_2)$ and a diffeomorphism $\psi : Z_1 \to Z_2$ such that

  $$\pi_1 = \pi_2 \circ \psi, \quad \phi_1 = \phi_2 \circ \psi, \quad \text{and } dz_1 = \psi^* dz_2.$$

  In this case, we choose

  $$\mathcal{O}_1 = \psi^{-1}(\mathcal{O}_2), \ \mathrm{r}_1 = \mathrm{r}_2 \circ \psi, \quad \text{and} \quad \rho_1 = \rho_2 \circ \psi$$

  and the result is obvious.

- $Z_2 = Z_1 \times \mathbb{R}^m$, $\pi_2 = \pi_1 \circ \mathrm{r}$ where $\mathrm{r} : Z_1 \times \mathbb{R}^m \to Z_1$ is projection onto the first factor, and

  $$\phi_2 = \phi_1 + Q$$

  where Q is a non-degenerate quadratic form on $\mathbb{R}^m$. In this case

  $$C_{\phi_2} = C_{\phi_1} \times \{0\}.$$

  We may choose our densities so that

  $$dz_2 = dz_1 \otimes ds$$

where ds is Lebesgue measure on $\mathbb{R}^n$. If $\mathrm{r}_1 : \mathcal{O}_1 \to C_{\phi_1}$ is a tubular neighborhood of C_{ϕ_1} we choose

$$\mathcal{O}_2 = \mathcal{O}_1 \times \mathbb{R}^m$$

and $\mathrm{r}_2 = \iota \circ \mathrm{r}_1 \circ \mathrm{r}$ where $\iota : Z_1 \to Z_1 \times \mathbb{R}^m$ is the injection $\iota(z) = z \times \{0\}$. If ρ_1 is a cutoff function for r_1 we chose ρ_2 to be of the form

$$\rho_2(z,s) = \rho_1(z)\rho(s)$$

where $\rho \in C_0^\infty(\mathbb{R}^m)$ which is identically one near the origin. Then the result is also obvious.

We now define

$$\mathcal{E}^\ell(U) := I^k(X,\Lambda)/I_U^{k+\ell}(X,\Lambda). \tag{8.32}$$

If $V \subset U$ is an open set, then $I_U^{k+\ell}(X,\Lambda) \subset I_V^{k+\ell}(X,\Lambda)$ so we get a projection

$$\mathcal{E}^\ell(U) \to \mathcal{E}^\ell(V)$$

and it is routine to check that the axioms for a sheaf are satisfied.

Notice that

- Multiplication by a power of $\hbar$ shows that $\mathcal{E}^\ell(U)$ is independent of k.
- For $\ell = 1$ the sheaf $\mathcal{E}^1$ is the sheaf of sections of $\mathbb{L}$.
- There is an intrinsic symbol map $\sigma_U^\ell : I^k(X,\Lambda) \to \mathcal{E}^\ell(U)$.
- In the whole discussion, we can let $\ell = \infty$.
- In particular, if $\mu \in I^k(X,\Lambda)$, we will say that $\mu \equiv 0$ on U if $\sigma_U^\infty(\mu) = 0$.
- For semi-classical Fourier integral operators $\mathcal{F}^k(\Gamma)$ we similarly get a sheaf on Γ.

8.6.2 Functoriality of the sheaf $\mathcal{E}^\ell$.

We return to the situation and the notation of Section 8.4.1. Let U_1 be an open subset of Γ_1 and let U_2 be an open subset of Γ_2. Then

$$\mathrm{pr}_1^{-1}(U_1)$$

is an open subset of $\Gamma_2 \star \Gamma_1$ as is $\mathrm{pr}_2^{-1}(U_2)$.

Let $F_1 \in \mathcal{F}^{m_1}(\Gamma_1)$ and $F_2 \in \mathcal{F}^{m_2}(\Gamma_2)$ so that $F_2 \circ F_1 \in \mathcal{F}^{m_1+m_2}(\Gamma_2 \circ \Gamma_1)$.

Theorem 8.6.1. *If* $\sigma_{U_1}^{\ell_1}(F_1) = 0$ *and* $\sigma_{U_2}^{\ell_2}(F_2) = 0$ *then*

$$\sigma_U^{\ell_1+\ell_2}(F_2 \circ F_1) = 0$$

where

$$U = \kappa\left(\mathrm{pr}_1^{-1}(U_1) \cap \mathrm{pr}_2^{-1}(U_2)\right).$$

Proof. This is a local assertion. So let $(Z_1, \pi_1\phi_1)$ be a presentation of Γ_1 and (Z_2, π_2, ϕ_2) be presentations of Γ_1 and Γ_2 where

$$Z_1 = X_1 \times X_2 \times \mathbb{R}^{d_1}, \qquad Z_2 = X_2 \times X_3 \times \mathbb{R}^{d_2}$$

with the obvious projections. Let

$$\mathrm{r}_1 : \mathcal{O}_1 \to C_{\phi_1} \quad \text{and} \quad \mathrm{r}_2 \to \mathcal{O}_2$$

be tubular neighborhoods and let ρ_1 and ρ_2 be cutoff functions. Define

$$W_1 := \mathrm{r}_1^{-1}\gamma_{\phi_1}^{-1}(U_1) \qquad W_2 := \mathrm{r}_2^{-1}\gamma_{\phi_2}^{-1}(U_2).$$

Then F^1 is of the form (8.4) where

$$u_1 = \hbar^{m_1 - \frac{n_2}{2} - \frac{d_1}{2}} \int_{W_1} \rho_1 \mathrm{r}_1^* \gamma_{\phi_1}^* \tilde{a_1} e^{i\frac{\phi_1}{\hbar}} ds_1, \quad n_2 = \dim X_2$$

where

$$\tilde{a}_1 \in C_0^\infty(\Gamma_1 \times \mathbb{R})$$

with a similar expression for F^2.

If we set

$$a_1 := \rho_1 \mathrm{r}_1^* \gamma_{\phi_1}^* \tilde{a_1}, \qquad a_2 := \rho_2 \mathrm{r}_2^* \gamma_{\phi_2}^* \tilde{a_2}$$

then the composition $F = F^2 \circ F^1$ is of the form (8.4) where u is given by (8.6).

Now our assumptions about F^1 and F^2 say that $\tilde{a}_i = \hbar^{\ell_i}\tilde{b}_i$ which implies that

$$a_i = \hbar^{\ell_i} b_i \quad \text{on} \quad W_i, \quad i = 1.2.$$

So

$$a = \hbar^{\ell_1+\ell_2} b_1(x_1, x_2, s_1, \hbar) b_2(x_2, x_3, s_2, \hbar)$$

on the set

$$W := \{(x_1, x_3, s_1, s_2, x_2) |\ (x_1, x_2, s_1) \in W_1 \text{ and } (x_2, x_3, s_2) \in W_2\}.$$

But the set $\gamma_\phi(W \cap C_\phi)$ is precisely the set U of the theorem. □

Corollary 8.6.1. *Composition of semi-classical Fourier integral operators induces a a map*

$$\mathcal{E}^\ell_{\Gamma_1}(U_1) \otimes \mathcal{E}^\ell_{\Gamma_2}(U_2) \to \mathcal{E}^\ell_\Gamma(U).$$

Proof. By the theorem, $F = F_2 \circ F_1$ lies in $\mathcal{F}^{k+\ell}$ if either $F_1 \in \mathcal{F}^{k_1+\ell}$ or $F_2 \in \mathcal{F}^{k_2+\ell}$. □

In sheaf theoretical terms we can state this corollary as

Theorem 8.6.2. *Composition of semi-classical Fourier integral operators induces a a morphism of sheaves*

$$\mathrm{pr}_1^* \mathcal{E}^\ell_{\Gamma_1} \otimes \mathrm{pr}_2^* \mathcal{E}^\ell_{\Gamma_2} \to \kappa^* \mathcal{E}^\ell_\Gamma.$$

8.7 Semi-classical pseudo-differential operators.

We want to apply the results of the preceding few sections to the case $X_1 = X_2 = X_3 = X$ and $\Gamma_1 = \Gamma_2 = \Delta$ where $\Delta_M \subset M^- \times M$ is the diagonal where $M = T^*X$. Since

$$\Delta_M = \Gamma_f, \quad f = \mathrm{id} : X \to X$$

we know that the composition

$$\Delta_M \circ \Delta_M = \Delta_M$$

is transverse.

We define

$$\Psi^k(X) := \mathcal{F}^k(\Delta_M). \tag{8.33}$$

Theorem 8.5 allows us to conclude that

$$F_1 \in \Psi^k(X) \ \text{ and } \quad F_2 \in \Psi^\ell(X) \quad \Rightarrow F_2 \circ F_1 \in \Psi^{k+\ell}(X).$$

So we define

$$\Psi(X) = \bigcup \Psi^j(X)$$

and conclude that $\Psi(X)$ is a filtered algebra. It is called the **algebra of semi-classical pseudo-differential operators** on X.

8.7.1 The line bundle and the symbol.

We can identify M with Δ_M via the map

$$\mathrm{diag} : M \to \Delta_M, \quad m \mapsto (m, m)$$

and we can identify M with $\Delta_M \star \Delta_M$ under the map

$$m \mapsto (m, m, m).$$

Under these identifications, the maps κ, pr_1 and pr_2 all become the identity map. So if we define

$$\mathbb{L}_M := \mathrm{diag}^* \mathbb{L}_{\Delta_M}$$

then (8.19) says that we have a canonical isomorphism

$$\mathbb{L}_M \cong \mathbb{L}_M \otimes \mathbb{L}_M$$

which implies that we have a canonical trivialization of $\mathbb{L}_M$.

In other words, under these identifications, we have a symbol map

$$\sigma : \Psi^k(X) \to C^\infty(M)$$

with kernel $\Psi^{k+1}(X)$

If $P_1 \in \Psi^{k_1}(X)$ and $P_2 \in \Psi^{k_2}(X)$ equation (8.20) becomes

$$\sigma(P_2 \circ P_1) = \sigma(P_1)\sigma(P_2).$$

8.7.2 The commutator and the bracket.

If $P_1 \in \Psi^{k_1}(X)$ and $P_2 \in \Psi^{k_2}(X)$ then

$$\sigma(P_2 \circ P_1) = \sigma(P_1)\sigma(P_2) = \sigma(P_1 \circ P_2)$$

so

$$\sigma(P_1 \circ P_2 - P_2 \circ P_1) = 0$$

which implies that

$$P_1 \circ P_2 - P_2 \circ P_1 \in \Psi^{k_1+k_2-1}(X).$$

Consider the symbol of $(P_1 \circ P_2 - P_2 \circ P_1)$ thought of as an element of $\Psi^{k_1+k_2-1}(X)$. We claim that this expression depends only on $\sigma(P_1)$ and $\sigma(P_2)$. Indeed, if we replace P_1 by $P_1 + Q_1$ where $Q_1 \in \Psi^{k_1-1}$ then $(P_1 \circ P_2 - P_2 \circ P_1)$ is replaced by

$$(P_1 \circ P_2 - P_2 \circ P_1) + (Q_1 \circ P_2 - P_2 \circ Q_1)$$

and the second term in parentheses is in $\Psi^{k_1+k_2-2}(X)$. Similarly if we replace P_2 by $P_2 + Q_2$. Thus there is a well defined bracket operation $[\ ,\ ]$ on $C^\infty(M)$ where

$$[f_1, f_2] = \sigma(P_1 \circ P_2 - P_2 \circ P_1)$$

(thought of as an element of $\Psi^{k_1+k_2-1}(X)$ when $f_1 = \sigma(P_1)$ and $f_2 = \sigma(P_2)$).

(This is a general phenomenon: if R is a filtered ring whose associated graded ring is commutative, then the graded ring inherits bracket structure.)

We will find that, up to a scalar factor, this bracket is the same as the Poisson bracket coming from the symplectic structure on M, see (8.46) below.

8.7.3 $I(X, \Lambda)$ as a module over $\Psi(X)$.

Let Λ be an exact Lagrangian submanifold of $M = T^*X$ thought of as an element of Morph(pt., M). Then we have the transversal composition

$$\Delta_M \circ \Lambda = \Lambda.$$

Thus we have the composition

$$P\mu := P \circ \mu. \quad P \in \Psi^{k_1}(X),\ \mu \in I^{k_2}(X, \Lambda)$$

where, on the right, μ is thought of as a semi-classical Fourier integral operator from pt. to X. It follows from (8.5) that

$$P\mu \in I^{k_1+k_2}(X, \Lambda). \tag{8.34}$$

In other words, $I(X, \Lambda) = \bigcup_\ell I^\ell(X, \Lambda)$ is a filtered module over the filtered algebra $\Psi(X)$.

Let us examine the symbol maps and the sheaves associated to this module structure: We begin by examining the various maps that occur in Theorem 8.4.1: We have the identification

$$\Delta_M \star \Lambda \to \Lambda, \quad (\text{pt.}, \lambda, \lambda) \mapsto \lambda.$$

Under this identification the map

$$\kappa : \Delta_M \star \Lambda \to \Delta_M \circ \Lambda, \quad (\text{pt.}, \lambda, \lambda) \mapsto \lambda$$

becomes the identity map. The map

$$\text{pr}_1 : \Delta_M \star \Lambda \to \Lambda, \quad (\text{pt.}, \lambda, \lambda) \mapsto (\text{pt.}, \lambda)$$

becomes the identity map, and the map

$$\text{pr}_2 : \Delta_M \star \Lambda \to \Delta_M, \quad (\text{pt.}, \lambda, \lambda) \mapsto (\lambda, \lambda)$$

becomes the inclusion $\iota \to M$ when we identify Δ_M with M. The map

$$j : \Delta_M \star \Lambda \to \Lambda \times \Delta_M, \quad (\text{pt.}, \lambda, \lambda) \mapsto ((\text{pt.}, \lambda), (\lambda, \lambda))$$

becomes

$$j = \text{id} \times \iota.$$

Then the left side of (8.19) is just $\mathbb{L}_\Lambda$ and the right hand side of (8.19) is $\mathbb{L}_\Lambda \otimes \mathbb{C}$ since $\mathbb{L}_M$ is the trivial bundle.

Equation (8.20) then becomes

$$\sigma(P\mu) = \iota^* (\sigma(P)) \, \sigma(\mu) \tag{8.35}$$

where $\sigma(P)$ is a function on $M = T^*X$ in view of our identification of M with Δ_M.

8.7.4 Microlocality.

If U is an open subset of $M = T^*X$ we define

$$\Psi_U^{k+\ell} := \mathcal{F}_U^{k+\ell}(\Delta_M)$$

(Again we are identifying M with Δ_M.) So

$$\Psi_U^\infty = \mathcal{F}_U^\infty(\Delta_M).$$

In particular, if $P \in \Psi^k(X)$ we say that $P = 0$ on U if $P \in \Psi_U^\infty$.

It follows from Theorem 8.6.1 that

Proposition 8.7.1. *If P_1 and $P_2 \in \Psi(X)$ and either P_1 or P_2 are zero on U then $P_1 P_2$ is zero on U.*

We define the **microsupport** of $P \in \Psi(X)$ as follows:

Definition 8.7.1. *We say that $p \in T^*X$ is* **not** *in the microsupport of P if there is an open set U containing p such that $P = 0$ on U.*

Let Λ be an exact Lagrangian submanifold of T^*X and $U \subset T^*X$ an open subset.

It follows from Theorem 8.6.1 that

$$P \in \Psi_U^{k_1+\ell} \text{ and } \mu \in I^{k_2}(X, \Lambda) \Rightarrow P\mu \in I_{U\cap\Lambda}^{k_1+k_2+\ell}(X, \Lambda). \tag{8.36}$$

Taking $\ell = \infty$ in this equation says that

Proposition 8.7.2. *If $P = 0$ on U then*

$$P\mu \in I^{\infty}_{U\cap\Lambda}(X, \Lambda).$$

8.7.5 The semi-classical transport operator.

Let $\iota : \Lambda \to T^*X$ be an exact Lagrangian submanifold, let $\mu \in I^{k_2}(X, \Lambda)$ and $P \in \Psi^{k_1}(X)$. Suppose that

$$\iota^* P \equiv 0.$$

It then follows from (8.35) that

$$\sigma(P\mu) = 0,$$

so

$$P\mu \in I^{k_1+k_2+1}(X, \Lambda).$$

We can then consider the symbol of $P\mu$, thought of as an element of $I^{k_1+k_2+1}(X, \Lambda)$.

Suppose we start with a section $s \in C^\infty(\mathbb{L}_\Lambda)$ and choose a $\mu \in I^{k_2}(\Lambda)$ such that

$$\sigma(\mu) = s.$$

We can then compute the symbol of $P\mu$ thought of as an element of $I^{k_1+k_2+1}(X, \Lambda)$. This gives a section, $\sigma_{k_1+k_2+1}(P\mu)$ of $\mathbb{L}_\Lambda$. We claim that $\sigma_{k_1+k_2+1}(P\mu)$ is independent of the choice of μ. Indeed, choosing a different μ amounts to replacing μ by $\mu + \nu$ where $\nu \in I^{k_2+1}(X, \Lambda)$ and

$$P\nu \in I^{k_1+k_2+2}(X, \Lambda)$$

so

$$\sigma_{k_1+k_2+1}(P(\mu + \nu)) = \sigma_{k_1+k_2+1}(P\mu).$$

We have thus defined an operator

$$L_P : C^\infty(\mathbb{L}_\Lambda) \to C^\infty(\mathbb{L}_\Lambda)$$

where

$$L_P(s) := \sigma_{k_1+k_2+1}(P\mu) \text{ if } \mu \in I^{k_2}(X, \Lambda) \text{ with } \sigma_{k_2}(\mu) = s. \tag{8.37}$$

Once again, multiplication by a power of $\hbar$ shows that the definition of L_P is independent of the choice of k_2.

Let us examine what happens when we replace s by fs where $f \in C^\infty(\Lambda)$: Choose $Q \in \Psi^0(X)$ with $\sigma(Q) = f$. Then

$$\begin{aligned} L_P(fs) &= L_P(\sigma(Q)\sigma(\mu)) \\ &= L_P\sigma(Q\mu) \\ &= \sigma(P(Q\mu)) \\ &= \sigma(Q(P\mu)) + \sigma((P \circ Q - Q \circ P)\mu) \\ &= fL_Ps + [p, f]s. \end{aligned}$$

where

$$p := \sigma(P).$$

Let us now use equation (8.46) (to be proved below) which says that

$$[p, f] = \frac{1}{i}\{p, f\}.$$

We know that since p vanishes on Λ, the corresponding vector field X_p is tangent to Λ, so

$$[p, f] = D_Y f$$

where Y is the restriction of X_p to Λ. So

$$L_P(fs) = fL_Ps + \frac{1}{i}(D_Y f)s.$$

Suppose we choose a connection ∇ on $\mathbb{L}_\Lambda$ so

$$\nabla_Z(fs) = f\nabla_Z s + (D_Z f)s$$

for any vector field Z on Λ. Thus

$$\left(L_P - \frac{1}{i}\nabla_Y\right)(fs) = f\left(L_P - \frac{1}{i}\nabla_Y\right)s.$$

This says that the operator $\left(L_P - \frac{1}{i}\nabla_Y\right))$ commutes with multiplication by functions, and hence is itself multiplication by a function:

$$\left(L_P - \frac{1}{i}\nabla_Y\right)s = \sigma_{\text{sub}}(P, \nabla)s.$$

Fixing ∇ (and writing $\sigma_{\text{sub}}(P)$ instead of $\sigma_{\text{sub}}(P, \nabla)$) we have

$$L_Ps = \frac{1}{i}\nabla_Y s + \sigma_{\text{sub}}(P)s. \tag{8.38}$$

This now allows us to carry out the program of chapter I, with differential operators replaced by semi-classical pseudo-differential operators. Suppose we are interested in finding an oscillatory half density μ which satisfies the equation

$$P\mu = 0$$

(in the sense of oscillatory half-densities). The first step is to solve the eikonal equation, as in Chapter I. This involves some hyperbolicity condition, as in Chapter I. Suppose we have done this, and so have found an exact Lagrangian submanifold Λ on which $\sigma(P) = 0$, and furthermore Λ is the flow out under the vector field X_p of an initial isotropic submanifold S.

For any $\mu \in I^{k_2}(X, \Lambda)$ we know that $P\mu \in I^{k_1+k_2+1}(X, \Lambda)$. We want to do better. We want to find μ such that $P\mu \in I^{k_1+k_2+2}(X, \Lambda)$. This means that want to choose μ so that its symbol satisfies $L_p s = 0$. According to (8.38), this amounts to solving the equation

$$\nabla_Y s + i\sigma_{\text{sub}}(P)s = 0$$

which is an ordinary first order differential homogeneous linear differential equation along the trajectories of Y. If we choose an initial section s_S of $\mathbb{L}_\Lambda$ along S, then there is a unique solution of this differential equation. Call the corresponding oscillatory half density μ_1. So

$$\mu_1 \in I^{k_1}(X, \Lambda) \quad \text{and} \quad P\mu_1 \in I^{k_1+k_2+2}(X, \Lambda).$$

We would now like to find $\mu_2 \in I^{k_2+1}(X, \Lambda)$ such that

$$P(\mu_1 + \mu_2) \in I^{k_1+k_2+3}(X, \Lambda)$$

which is the same as requiring that

$$\sigma_{k_1+k_2+2}(P\mu_1) + \sigma_{k_1+k_2+2}(P\mu_2) = 0$$

which amounts to finding a section s_2 of $\mathbb{L}_\Lambda$ such that

$$L_P s_2 = -\sigma_{k_1+k_2+2}(P\mu_1).$$

This amounts to an inhomogeneous linear differential equation along the trajectories on Y which we can solve once we have prescribed initial conditions along S. Continuing in this way, we can find

$$\mu_1 + \cdots + \mu_N$$

with prescribed initial conditions such that

$$P(\mu_1 + \cdots + \mu_N) \in I^{k_1+k_2+N+1}(X, \Lambda).$$

If we now choose

$$\mu \sim \sum_j \mu_j$$

then

$$P\mu = 0 \mod O(\hbar^\infty),$$

where we can prescribe initial values along S.

Since everything was intrinsically defined, we have no problems with caustics. However we do have to explain the relation between the semi-classical pseudodifferential operators discussed in this chapter, and the differential operators and the semi-classical differential operators discussed in Chapter I. We shall do this in Section 8.10.

8.8 The local theory.

Let
$$X \subset \mathbb{R}^n$$
be an open convex subset,
$$M = T^*X$$
and
$$\Delta_M \subset M \times M$$
the diagonal,
$$Z = X \times X \times \mathbb{R}^n$$
with $\pi : Z \to X \times X$ given by
$$\pi(x, y, \xi) = (x, y)$$
and
$$\phi(x, y, \xi) = (y - x) \cdot \xi.$$
Then we know that (Z, π, ϕ) is a generating function for Δ_M with
$$C_\phi = \{(x, y, \xi) | x = y\}.$$
So we may identify C_ϕ with
$$X \times \mathbb{R}^n.$$
Also, we identify Δ_M with $M = T^*X$ which is identified with $X \times \mathbb{R}^n$. Under these identifications the map
$$\gamma_\phi : C_\phi \to \Delta_M$$
becomes the identity map.

We will also choose the standard Lebesgue densities dx on X and $d\xi$ on $\mathbb{R}^n$ with their corresponding half-densities.

To get a local symbol calculus for $\Psi(X)$ we must choose a tubular neighborhood $\mathcal{O}$ of C_ϕ and a projection $\mathrm{pr} : \mathcal{O} \to C_\phi$. Three standard choices are to take $\mathcal{O} = Z$ and the projections $\mathrm{pr} : Z \to C_\phi$ to be
$$\mathrm{pr}_R(x, y, \xi) := (x, \xi) \tag{8.39}$$
$$\mathrm{pr}_L(x, y, \xi) := (y, \xi) \tag{8.40}$$
$$\mathrm{pr}_W(x, y, \xi) := \left(\frac{x + y}{2}, \xi\right) \tag{8.41}$$
The first choice, pr_R, gives rise to the semi-classical analogue of the right symbol calculus in the theory of pseudo-differential operators. The second choice, pr_L, gives the analogue of the left symbol calculus while the third choice gives rise to the analogue of the Weyl calculus.

In this section we will focus on pr_R. Choose a cutoff function $\rho \in C_0^\infty(\mathbb{R}^n)$ with $\rho(x) \equiv 1$ for $\|x\| \leq 1$. We now apply Proposition 8.6.2 to conclude that every $P \in \Psi^k(X)$ can be written uniquely mod Ψ^∞ as an integral operator K where

$$K: \; C_0^\infty(X) \to C^\infty(\mathbb{R}^n), \quad (Kf)(x) = \int K(x,y,\hbar) f(y) dy$$

where

$$K(x,y,\hbar) = \hbar^{k-n} \int \rho(y-x) \, (\mathrm{pr}_R^* a) \, e^{i \frac{(y-x)\cdot \xi}{\hbar}} d\xi, \quad a = a(x,\xi,\hbar) \in C_0^\infty(C_\phi \times \mathbb{R})$$

in other words,

$$K(x,y,\hbar) = \hbar^{k-n} \int \rho(y-x) a(x,\xi,\hbar) e^{i \frac{(y-x)\dot{\xi}}{\hbar}} d\xi. \tag{8.42}$$

Definition 8.8.1. *The function $\hbar^{k-n} a(x,\xi,\hbar)$ is called the (right)* **total symbol** *of P.*

8.8.1 The composition law for symbols.

Given $P_1 \in \Psi^{k_1}(X)$ and $P_2 \in \Psi^{k_2}(X)$ we will work out the formula for the total symbol of their composition $P_2 \circ P_1$ in terms of the total symbols of P_1 and P_2 by an application of the formula of stationary phase. The final result will be formula (8.45) below. We will give an alternative derivation of the composition laws using the semi-classical Fourier transform in the next chapter.

So suppose that

$$\begin{aligned} K_1(z,y,\hbar) &= \int \rho(z-y) a_1(z,\xi,\hbar) e^{i \frac{(z-y)\cdot\xi}{\hbar}} d\xi \\ K_2(x,z,\hbar) &= \int \rho(x-z) a_2(x,\eta,\hbar) e^{i \frac{(x-z)\cdot\eta}{\hbar}} d\eta \end{aligned}$$

so

$$\int K_2(x,z,\hbar) K_1(z,y,\hbar) dz =$$

$$\int \rho(x-z)\rho(z-y) a_2(x,\eta,\hbar) a_1(z,\xi,\hbar) e^{i\frac{\phi}{\hbar}} d\eta d\xi dz \tag{8.43}$$

where

$$\phi(x,y,z,\eta,\xi) = (x-z)\cdot\eta + (z-y)\cdot\xi = x\cdot\eta - y\cdot\xi + (\xi-\eta)\cdot z.$$

Make the change of variables

$$\eta_1 := \eta - \xi, \qquad z_1 := z - x$$

so that in terms of these new variables

$$\begin{aligned} \phi(x,y,z_1,\xi,\eta_1) &= x\cdot(\eta_1+\xi) - y\cdot\xi + (z_1+x)\cdot(\xi-\eta_1-\xi) \\ &= x\cdot\eta_1 + x\cdot\xi - y\cdot\xi - z_1\cdot\eta_1 - x\cdot\eta_1 \\ &= (x-y)\cdot\xi - z_1\cdot\eta_1. \end{aligned}$$

So (8.43) becomes

$$\int \tilde{a}(x,y,\xi,\hbar)e^{i\frac{(x-y)\cdot\xi}{\hbar}}d\xi$$

where

$$\tilde{a}(x,y,\xi,\hbar) =$$

$$\int \rho(-z_1)\rho(z_1+x-y)a_2(x,\eta_1+\xi,\hbar)a_1(z_1+x,\xi,\hbar)e^{-i\frac{z_1\cdot\eta_1}{\hbar}}d\eta_1 dz_1. \tag{8.44}$$

If we set $w=(z_1,\eta_1)$, this integral has the form

$$\int f(w)e^{i\frac{Aw,w)}{2\hbar}}dw$$

where A is the non-singular symmetric matrix

$$A = \begin{pmatrix} 0 & -I \\ -I & 0 \end{pmatrix}$$

where I is the $n\times n$ identity matrix. The formula of stationary phase says that (in general) an integral of the form $I(\hbar) = \int_{\mathbb{R}^m} f(w)e^{i\frac{Aw,w)}{2\hbar}}dw$ has the asymptotic expansion

$$I(\hbar) \sim \left(\frac{\hbar}{2\pi}\right)^{\frac{m}{2}} \gamma_A a(\hbar)$$

where

$$\gamma_A = |\det A|^{-\frac{1}{2}}\, e^{\frac{pii}{4}\operatorname{sgn} A}$$

and

$$a(\hbar) \sim \left(\exp\left(-i\frac{\hbar}{2}b(D)\right)f\right)(0)$$

where

$$b(D) = \sum_{k\ell} b_{k\ell}D_{x_k}D_{x_\ell}$$

with $B=(b_{k\ell})=A^{-1}$.

In our case $m=2n$, $|\det A|=1$, $\operatorname{sgn} A=0$ so $\gamma_A=1$ and $B=A$ so

$$b(D) = -2D_{\eta_1}\cdot D_{z_1}$$

and so (8.44) has the asymptotic expansion

$$\left(\frac{\hbar}{2\pi}\right)^n (\exp(i\hbar D_{\eta_1}\cdot D_{z_1})\,\rho(z_1)\rho(z_1+x-y)a_2(x,\eta_1+\xi,\hbar)a_1(z_1+x,\xi,\hbar)$$

evaluated at $z_1=x_1=0$. Any (non-trivial) derivative of $\rho(z_1)$ vanishes near $z_1=0$ since ρ is identically one there. So $\tilde{a}$ has the asymptotic expansion

$$\left(\frac{\hbar}{2\pi}\right)^n \sum_\beta (i\hbar)^{|\beta|}\frac{1}{\beta!}D_\xi^\beta a_2(x,\xi,\hbar)D_x^\beta\left[\rho(x-y)a_1(x,\xi,\hbar)\right].$$

Once again, any non-trivial derivative of $\rho(x-y)$ vanishes if $|x-y|\leq 1$. So (in terms of the above notation) we have proved

Theorem 8.8.1. *The kernel K of the compostite $P_2 \circ P_1$ has the form*

$$K(x,y,\hbar) = \left(\frac{\hbar}{2\pi}\right)^n \int \rho(x-y)a(x,\xi,\hbar)e^{-\frac{(x-y)\cdot\xi}{\hbar}}d\xi$$

where a has the asymptotic expansion

$$\sum_\beta (i\hbar)^{|\beta|}\frac{1}{\beta!}D_\xi^\beta a_2(x,\xi,\hbar)D_x^\beta a_1(x,\xi,\hbar). \tag{8.45}$$

Let us examine the first two terms in this expansion. They are

$$a_2a_1 + i\hbar\sum \frac{\partial a_2}{\partial \xi_j}\frac{\partial a_1}{\partial x_j}.$$

Interchanging P_1 and P_2 and subtracting shows that the bracket introduced in Section 8.7.1 is related to the Poisson bracket by

$$[\ ,\ ] = -i\{\ ,\ \}. \tag{8.46}$$

8.9 The semi-classical Fourier transform.

Let $X = \mathbb{R}^n$ and consider the function

$$\rho :\ X \times X \to \mathbb{R}, \quad \rho(x,y) = -x\cdot y.$$

Let $\Gamma_\rho \in \text{Morph}(T^*X, T^*X)$ be the corresponding canonical relation, so Γ_ρ consists of all (x,ξ,y,η) with

$$\xi = -\frac{\partial\rho}{\partial x}, \qquad \eta = \frac{\partial\rho}{\partial y}.$$

In other words

$$\xi = y, \quad \eta = -x$$

so Γ_ρ is the graph of the symplectomorphism

$$\mathfrak{J} :\ (x,\xi) \mapsto (\xi,-x).$$

Define the **semi-classical Fourier transform** to be the integral operator $\mathfrak{F}_\hbar$, where, for $f \in C_0^\infty(X)$

$$(\mathfrak{F}_\hbar f)(y) := \frac{1}{(2\pi\hbar)^{n/2}}\int f(x)e^{-i\frac{x\cdot y}{\hbar}}dx.$$

So $\mathfrak{F}_\hbar$ is a semi-classical Fourier integral operator associated to Γ_ρ. In terms of the usual Fourier transform $f \mapsto \hat{f}$ where

$$\hat{f}(z) = \frac{1}{(2\pi)^{n/2}}\int f(x)e^{-ix\cdot z}dx$$

we have

$$(\mathfrak{F}_\hbar f)(y) = \hbar^{-n/2}\hat{f}\left(\frac{y}{\hbar}\right).$$

The Fourier inversion formula says that

$$f(w) = \frac{1}{(2\pi)^{n/2}}\int \hat{f}(z)e^{iw\cdot z}dz.$$

Setting $z = y/\hbar$ this gives

$$f(w) = \frac{1}{(2\pi\hbar)^{n/2}}\int (\mathfrak{F}_\hbar f)(y)e^{i\frac{w\cdot y}{\hbar}}dy.$$

In other words, the semi-classical Fourier integral operator

$$g \mapsto \frac{1}{(2\pi\hbar)^{n/2}}\int g(y)e^{i\frac{w\cdot y}{\hbar}}dy$$

associated to the canonical transformation

$$\mathfrak{J}^{-1}: \ (x,\xi) \mapsto (-\xi, x)$$

is the inverse of $\mathfrak{F}_\hbar$. So we will denote the semi-classical Fourier integral operator

$$g \mapsto \frac{1}{(2\pi\hbar)^{n/2}}\int g(y)e^{\frac{w\cdot y}{\hbar}}dy \ \text{ by } \ \mathfrak{F}_\hbar^{-1}.$$

For example, let $P \in \Psi(\mathbb{R}^n)$ so that $P(fdx^{\frac{1}{2}}) = (Kf)dx^{\frac{1}{2}}$ where

$$(Kf)(x) = \int K(x,y,\hbar)f(y)dy$$

where

$$k(x,y,\hbar) = \int \rho(x-y)a(x,\xi,\hbar)e^{i\frac{(x-y)\cdot\xi}{\hbar}}d\xi.$$

Ignoring the cutoff factor, this has the form

$$(2\pi\hbar)^{-n/2}\int a(x,\xi,\hbar)e^{i\frac{x\cdot\xi}{\hbar}}(\mathfrak{F}_\hbar f)(\xi)d\xi. \tag{8.47}$$

So

$$P = A \circ \mathfrak{F}_\hbar \tag{8.48}$$

where (absorbing the powers of 2π) A is the operator whose Schwartz kernel is the oscillatory function

$$a(x,y,\hbar)e^{i\frac{x\cdot y}{\hbar}}.$$

In particular, A is a semi-classical Fourier integral operator associated with the symplectomorphism $\mathfrak{J}^{-1}$.

8.9.1 The local structure of oscillatory $\frac{1}{2}$-densities.

Let X be a manifold and $\Lambda \subset T^*X$ be an exact Lagrangian submanifold, and let

$$p_0 = (x_0, \xi_0) \in \Lambda$$

with $\xi_0 \neq 0$. According to the argument in Section 5.9, there are canonical Darboux coordinates

$$x_1, \ldots, x_n,\ \xi_1, \ldots, \xi_n$$

in a neighborhood V of p_0 such that the horizontal Lagrangian foliation

$$\xi_1 = c_1, \ldots \xi_n = c_n$$

is transverse to Λ. Let $\nu \in I^\ell(X, \Lambda)$ be microlocally supported in V.

We will use these coordinates and (by restriction) we may assume that $\Lambda \subset T^*(\mathbb{R}^n)$. As above, let $\mathfrak{J}$ denote the symplectomorphism

$$\mathfrak{J}(x, \xi) = (\xi, -x).$$

So $\mathfrak{J}(\Lambda)$ is *horizontal* , i.e.

$$\Lambda = \Lambda_{-\phi}$$

for some $\phi \in C^\infty(\mathbb{R}^n)$.

Since

$$\mathfrak{J}\left(\frac{\partial \phi}{\partial \xi}, \xi\right) = \left(\xi, -\frac{\partial \phi}{\partial \xi}\right),$$

we see that $\mathfrak{J}(\Lambda)$ is the image of the set

$$\left\{(x, \xi) \in T^*(\mathbb{R}^n) \quad | \; x = \frac{\partial \phi}{\partial \xi}\right\}.$$

As the inverse semi-classical Fourier transform $\mathfrak{F}_\hbar^{-1}$ is a Fourier integral operator of degree zero associated to the graph of $\mathfrak{J}^{-1}$ we know that

$$\mu \in I^\ell(\mathbb{R}^n, \Lambda_{-\phi}) \Leftrightarrow \nu = \mathfrak{F}_\hbar^{-1}\mu \in I^\ell(\mathbb{R}^n, \Lambda).$$

If we write μ in the form

$$\hbar^\ell b(\xi, \hbar) e^{\frac{-i\phi(\xi)}{\hbar}}$$

then

$$\nu = \mathfrak{F}_\hbar^{-1}\mu = \frac{\hbar^{\ell - \frac{n}{2}}}{(2\pi)^{n/2}} \int b(\xi, \hbar) e^{i\frac{x\cdot\xi - \phi(\xi)}{\hbar}} d\xi \tag{8.49}$$

gives the local expression for an element of $I^\ell(X, \Lambda)$.

8.9.2 The local expression of the module structure of $I(X, \Lambda)$ over $\Psi(X)$.

Continuing with the notation of previous sections, let

$$P = A \circ \mathfrak{F}_\hbar \in \Psi(X)$$

and

$$\nu = \mathfrak{F}_\hbar^{-1} \circ \mu \in I^\ell(X, \Lambda).$$

Then

$$P \circ \nu = A \circ \mu.$$

More explicitly $P \circ \nu$ has the expression

$$\hbar^{k+\ell-\frac{n}{2}} \int a(x, \xi, \hbar) b(\xi, \hbar) e^{i\frac{x\cdot\xi - \phi(\xi)}{\hbar}} d\xi. \tag{8.50}$$

8.9.3 Egorov's theorem.

As an application of the theorems of this section, consider the following situation: Let

$$\gamma : T^*X_1 \to T^*X_2$$

be a symplectomorphism, and set

$$\Gamma_1 := \text{graph } \gamma, \quad \Gamma_2 := \text{graph } \gamma^{-1}.$$

Suppose that F_1 is a semi-classical Fourier operator associated to Γ_1 and that $F_2 = F_1^{-1}$ on some open subset $U \subset T^*X_1$, meaning that for every $B \in \Psi^0(X)$ with microsupport in U, we have

$$F_2 F_1 B = B.$$

Theorem 8.9.1. [Egorof.] *For any $A \in \Psi^k(X_2)$ with microsupport in $\gamma(U)$,*

$$F_2 \circ A \circ F_1 \in \Psi^k(X_1)$$

and

$$\sigma(F_2 A F_1) = \gamma^* (\sigma(A)). \tag{8.51}$$

Proof. The first assertion follows from the fact that $\Gamma_2 \circ \Delta_{T^*X_2} \circ \Gamma_1 = \Delta_{T^*X_1}$.

As to (8.51), let $(x, \xi, y, \eta) \in \Gamma_1$ so from $F_2 \circ F_1 = I$ on U we get

$$\sigma(F_2)(y, \eta, x, \xi)\sigma(F_1)(x, \xi, y, \eta) = 1$$

for $(x, \xi) \in U$.

Now

$$\sigma(F_2 A F_1)(x, \xi)$$

$$= \sigma(F_2)(y,\eta,x,\xi)\sigma(A)(y,\eta)\sigma(F_1)(x,\xi,y,\eta).$$

Since A is a semi-classical pseudo-differential operator, $\sigma(A)$ is just a scalar, so we can pull the middle term out of the product, and use to preceding equation to conclude that

$$\sigma(F_2AF_1)(x,\xi) = \sigma(A)(y,\eta)$$

where (x,ξ) is related to (y,η) by $(x,\xi,y,\eta) \in \Gamma_1$, i.e. $(y,\eta) = \gamma(x,\xi)$. This is precisely the assertion of (8.51). □

8.10 Semi-classical differential operators and semi-classical pseudo-differential operators.

Recall from Chapter I that a semi-classical differential operator on $\mathbb{R}^n$ (of degree 0) has the expression

$$P = P(x,D,\hbar) = \sum a_\alpha(x,\hbar)(\hbar D)^\alpha, \quad a_\alpha \in C^\infty(X \times \mathbb{R}).$$

The right symbol of P is defined as

$$p(x,\xi,\hbar) := \sum a_\alpha(x,\hbar)\xi^\alpha$$

so that

$$\begin{aligned} P\left(e^{i\frac{x\cdot\xi}{\hbar}}\right) &= \sum a_\alpha(x,\hbar)(\hbar D)^\alpha \left(e^{i\frac{x\cdot\xi}{\hbar}}\right) \\ &= e^{i\frac{x\cdot\xi}{\hbar}} \sum a_\alpha(x,\hbar)\xi^\alpha \\ &= p(x,\xi,\hbar)e^{i\frac{x\cdot\xi}{\hbar}}. \end{aligned}$$

Proposition 8.10.1. *If P is a semi-classical differential operator and $f \in C_0^\infty(\mathbb{R}^n)$ then*

$$(Pf)(x) = (2\pi\hbar)^{n/2} \int p(x,\xi,\hbar)e^{i\frac{x\cdot\xi}{\hbar}}(\mathfrak{F}_\hbar f)(\xi)d\xi.$$

Proof. This follows from the semi-classical Fourier inversion formula

$$f(x) = (2\pi\hbar)^{-n/2} \int e^{i\frac{x\cdot\xi}{\hbar}}(\mathfrak{F}_\hbar f)(\xi)d\xi$$

and the above formula $Pe^{i\frac{x\cdot\xi}{\hbar}} = pe^{i\frac{x\cdot\xi}{\hbar}}$ by passing P under the integral sign. □

If we compare this proposition with (8.47), we see that the (right) symbol of a semi-classical differential operator plays the same role as the (right) symbol of a semi-classical pseudo-differential operator.

The composition of a semi-classical differential operator with a semi-classical pseudo-differential operator.

Theorem 8.10.1. *Let P be a semi-classical differential operator on $\mathbb{R}^n$ with right symbol $p = p(x,\xi,\hbar)$ and let Q be a semiclassical pseudo-differential operator on $\mathbb{R}^n$ with right symbol $q = q(x,\xi,\hbar)$. Then $P \circ Q$ is a semi-classical pseudo-differential operator with right symbol*

$$r(x,\xi,\hbar) \sim \sum_\alpha \frac{1}{\alpha!}\left(\frac{\partial}{\partial\xi}\right)^\alpha p\ (\hbar D_x)^\alpha q. \tag{8.52}$$

Remark. Notice that (except for the placement of powers of $\hbar$ and i) this is the same as formula (8.45) for the composition of two semi-classical pseudo-differential operators.

Proof. Notice that for any $f \in C^\infty(\mathbb{R}^n)$, Leibnitz's rule gives

$$(\hbar D_{x_j})[e^{i\frac{x\cdot\xi}{\hbar}} f] = e^{i\frac{x\cdot\xi}{\hbar}}[\hbar D_{x_j} + \xi_j]f$$

and hence by induction

$$(\hbar D_x)^\alpha[e^{i\frac{x\cdot\xi}{\hbar}} f] = e^{i\frac{x\cdot\xi}{\hbar}}[\hbar D_x + \xi]^\alpha f.$$

Applied to the formula

$$(Qf)(x) = (2\pi\hbar)^{-n/2}\int q(x,\xi,\hbar)e^{i\frac{x\cdot\xi}{\hbar}}(\mathfrak{F}_\hbar f)(\xi)d\xi$$

gives

$$(P(Qf))(x) = \int r(x,\xi,\hbar)e^{i\frac{x\cdot\xi}{\hbar}}(\mathfrak{F}_\hbar f)(\xi)dx$$

where

$$b(x,\xi,\hbar) = p(x,\hbar D_x + \xi,\hbar)q = \sum_\alpha \frac{1}{\alpha!}\left(\frac{\partial}{\partial\xi}\right)^\alpha p\ (\hbar D_x)^\alpha q$$

by the multinomial theorem. □

The action of a semi-classical differential operator on oscillatory $\frac{1}{2}$ densities.

Let P be a semi-classical differential operator

$$P = \sum_\alpha a_\alpha(x,\hbar)(\hbar D)^\alpha$$

so P has right symbol p.

Let ν be a semi-classial Fourier integral operator as given by (8.49). Once again, differentiating under the integral sign shows that $P\nu$ is given by

$$\hbar^{\ell-\frac{n}{2}}\int p(x,\xi,\hbar)b(\xi,\hbar)e^{i\frac{x\cdot\xi-\phi(\xi)}{\hbar}}d\xi. \tag{8.53}$$

Notice that (with $k = 0$ and a replaced by $p0$ this is the same as (8.50). This shows that $I(X, \Lambda)$ is a module over the ring of semi-classical differential operators.

8.10.1 Semi-classical differential operators act microlocally as semi-classical pseudo-differential operators.

Let $K \subset \mathbb{R}^n$ be a compact subset. Let

$$\mathrm{pr}_{\mathbb{R}^n} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$$

denote projection onto the first factor. We want to consider the action of the semiclassical differential operator P on the set of $\nu \in I(X, \Lambda)$ of the form(8.49) where

$$\mathrm{pr}_{\mathbb{R}^n} \mathrm{Supp}(b) \subset K. \tag{8.54}$$

Let $\rho \in C_0^\infty(\mathbb{R}^n)$ have the property that $\rho(\xi) = 1$ on K. Define the operator $\rho(\hbar D)$ on the set of μ satisfying (8.54) by

$$\rho(\hbar D) = \mathfrak{F}_\hbar^{-1} \circ \rho(\xi) \circ \mathfrak{F}_\hbar. \tag{8.55}$$

More explicitly, (and dropping the half density factors)

$$(\rho(D)f)(x) = (2\pi\hbar)^{-n/2} \int e^{i\frac{x\cdot\xi}{\hbar}} (\mathfrak{F}_\hbar f)(\xi) d\xi.$$

Then

$$P\rho(D)\nu = P\mathfrak{F}_\hbar^{-1}\rho(\xi)\mathfrak{F}_\hbar\mu = P\mathfrak{F}_\hbar^{-1}\mathfrak{F}_\hbar\mu = P\nu.$$

In short, $P = P\rho(D)$ microlocally in a neighborhood of a point of Λ.

Applied to $\Psi(X)$ regarded as a module over itself, we see that microlocally, in a neighborhood of any point of T^*X we can write $P = P\rho(\hbar D)$. This answers the issue raised at the end of Section 8.7.5 and we may apply the method of that section to the solution of (semi-classical) hyperbolic differential equations.

Application: The semi-classical wave equation.

Let

$$P = \sum_{|\alpha| \le r} a_\alpha(x, \hbar)(\hbar D)^\alpha$$

be a zero-th order semi-classical partial differential operator on $X := \mathbb{R}^n$. In this section we show how to apply the methods we have developed to solve the following problem:

Construct semi-classical operators

$$U(t) \in \Psi^0(X), \qquad -\infty < t < \infty$$

with

$$U(0) = \rho(\hbar D), \qquad \rho \in C_0^\infty(\mathbb{R}^n)$$

satisfying

$$\frac{1}{i}\frac{\partial}{\partial t}U(t) = PU(t) \mod \hbar^\infty.$$

In other words, we want to construct a semi-classical version of the wave operator

$$e^{itP}\rho(\hbar D)$$

and show that this is indeed a semi-classical pseudo-differential operator.

If $\mu = \mu(x, y, t, \hbar)$ is to be the Schwartz kernel of our desired $U(t)$, then the initial condition says that

$$\mu(x, y, 0, \hbar) = \hbar^{-n} \int \rho(\xi) e^{i\frac{(x-y)\cdot\xi}{\hbar}} d\xi, \tag{8.56}$$

while the wave equation requires that

$$\left(\frac{1}{i}\hbar\frac{\partial}{\partial t} - \hbar P\right)\mu = 0. \tag{8.57}$$

Condition (8.56) implies that

$$\mu(0) \in I^{-n}(X \times X, \Delta_X).$$

The leading symbol of the operator

$$\left(\frac{1}{i}\hbar\frac{\partial}{\partial t} - \hbar P\right)$$

occurring in (8.57) is just τ, the dual variable to t, and so the corresponding Hamiltonian vector field is $\frac{\partial}{\partial t}$.

Hence, if we take $\Lambda_0 = \Delta_X \times (0, 0) \subset T^*(X \times X) \times T^*\mathbb{R}$, the flowout by $\frac{\partial}{\partial t}$ of $\Lambda_)$ is just the subset given by $\tau = 0$ of $T^*(X \times X) \times T^*\mathbb{R}$. We can now apply the method of the transport equation as developed above to get a solution of (8.57) with initial condition (8.56) with $\mu \in I^{-n}(X \times X \times \mathbb{R}, \Lambda)$.

If ι_a denotes the injection

$$\iota_a : X \times X \to X \times X \times \mathbb{R}, \quad \iota_a(x, y) = (x, y, a)$$

then

$$\Gamma^\dagger_{\iota_a} \circ \Lambda = \Delta_X$$

so

$$\iota_a^*\mu \in I^{-n}(X \times X, \Delta_X)$$

proving that the corresponding operator $U(a)$ is indeed an element of $\Psi^0(X)$.

The construction of U that we just gave shows the power of the symbolic method. In fact, we will need more explicit information about $U(t)$ which will follow from more explicit local methods that we will develop in the next chapter.

8.10.2 Pull-back acts microlocally as a semi-classical Fourier integral operator.

Let X and Y be smooth manifolds and

$$G : X \to Y$$

as smooth map. Associated to G is the canonical relation $\Gamma_G \in \mathrm{Morph}(T^*X, T^*Y)$ where

$$(x, \xi, y, \eta) \in \Gamma_G \ \Leftrightarrow y = G(x) \ \text{ and } \xi = dG_x^* \eta.$$

We have the pull-back operator

$$G^* : C^\infty(Y) \to C^\infty(X).$$

We would like to think of G^* as being associated to the transpose canonical relation $\Gamma_G^\dagger$. But G^* is not a semi-classical Fourier integral operator. The point of this section is to show that microlocally it is.

Since we are making micro-local assertions, we may assume that $Y = \mathbb{R}^n$. Let $\rho = \rho(\xi)$ a smooth function of compact support, and $\rho(\hbar D)$ the operator sending $f \in C_0^\infty(Y)$ into $\rho(\hbar D) f$ where

$$(\rho(\hbar D) f)(x) = \hbar^{-n} \int \rho(\xi) e^{i\frac{(x-y)\cdot\zeta}{\hbar}} f(y) dy d\zeta.$$

Then $G^* \circ \rho(\hbar D)$ sends f into the function

$$x \mapsto \hbar^{-n} \int \rho(\zeta) e^{i\frac{(G(x)-y)\cdot\zeta}{\hbar}} f(y) dy d\zeta.$$

Let $g_i(x)$ denote the i-th coordinate of $G(x)$. The function

$$\psi(y, x, \zeta) := (G(x) - y) \cdot \zeta = \sum_i (g_i(x) - y_i)\zeta_i$$

is a generating function for $\Gamma_G^\dagger$. Indeed the condition $d_\zeta \psi = 0$ gives $y = G(x)$ and then the horizontal derivatives $D_{Y\times X}$ give $(\eta, G^*\eta)$ for $\eta = \sum_i \zeta_i dy_i$. In other words, $G^* \circ \rho(\hbar D)$ is a semi-classical Fourier integral operator of order $\frac{n_1 - n_2}{2}$ associated to $\Gamma_G^\dagger$. $\square$

8.11 Description of the space $I^k(X, \Lambda)$ in terms of a clean generating function.

In this section we give a local description of the space $I^k(X, \Lambda)$ in terms of a clean generating function. We refer back to Section 5.1.1 for notation and results concerning clean generating functions, and, in particular, for the concept of the excess, e, of a generating function.

So let (π, ϕ) be a clean presentation of Λ of excess e where $\pi : X \times \mathbb{R}^d \to X$ is projection onto the first factor . Recall that C_ϕ denotes the set where $\frac{\partial \phi}{\partial s_i} = 0$ where $s_1, \dots, s_d$ are the coordinates on $\mathbb{R}^d$. In Section 5.1.1 we proved

Proposition 8.11.1. *There exists a neighborhood U of C_ϕ and an embedding*

$$f : U \to X \times \mathbb{R}^d$$

such that

$$\pi \circ f = \pi$$

and

$$\phi = f^* \pi_1^* \phi_1 \tag{8.58}$$

where

$$\pi_1 : X \times \mathbb{R}^d \to X \times \mathbb{R}^{d_1}, \qquad d_1 = d - e$$

is the projection

$$\pi_1(x, s_1, \ldots, s_d) = (x, s_1, \ldots, s_{d_1})$$

and where ϕ_1 is a transverse generating function for Λ with respect to the projection $\pi_2 : X \times \mathbb{R}^{d_1} \to X$ onto the first factor. In particular, we have $f(C_\phi) = \pi_1^{-1}(C_{\phi_1})$ and the map

$$\wp_\phi : C_\phi \to \Lambda, \qquad (x,s) \mapsto \left(x, \frac{\partial \phi}{\partial x}\right)$$

factors as

$$\wp_\phi = \wp_{\phi_1} \circ \pi_1 \circ f.$$

Now let $a = a(x, s, \hbar) \in C_0^\infty(U \times \mathbb{R})$ and let

$$\mu = F_{a,\phi} := \hbar^{k - \frac{d}{2} + \frac{e}{2}} \int a(x, s, \hbar) e^{\frac{i\phi}{\hbar}} ds.$$

Notice that the class of such μ when $e = 0$ (i.e. for transverse generating functions) is precisely the space we denoted by $I_0^k(X, \Lambda, \phi)$ in Section 8.1. We can use the Proposition to show that we haven't enlarged the space $I_0^k(X, \Lambda)$ by allowing e to be unequal to zero.

Indeed, letting

$$J = \left(\frac{\partial f_i}{\partial s_j}\right)$$

where $f(x,s) = (x, f_1(x,s), \ldots, f_d(x,s))$, we can, by the change of variables formula, rewrite the above expression for μ as

$$\mu = \hbar^{k - \frac{d}{2} + \frac{e}{2}} \int \tilde{a}(x, s, \hbar) e^{\frac{i\pi_1^* \phi_1}{\hbar}} ds$$

where

$$\tilde{a} := (f^{-1})^* \left(a \left|\det J\right|^{-1}\right).$$

So if we set

$$a_1(x, s_1, \ldots, s_{d_1}, \hbar) := \int \tilde{a}(x, s_1, \ldots, s_d, \hbar) ds_{d_1+1} \cdots s_d, \tag{8.59}$$

then

$$\mu = \hbar^{k-\frac{d_1}{2}} \int a_1(x, s_1, \dots, s_{d_1}, \hbar) e^{\frac{i\phi_1}{\hbar}} ds_1 \cdots ds_{d_1}. \tag{8.60}$$

Since ϕ_1 is a transverse generating function for Λ we see that we have not enlarged the space $I_0^k(X, \Lambda)$.

Notice that it follows from the above definitions of $\tilde{a}$ and a_1 that if $a(x, s, 0) \equiv 0$ then $\mu \in I_0^{k+1}$.

If we now go back to the local definition of the symbol as given in Section 8.3.2, i.e.

$$\sigma_{\phi_1}(\mu) = \wp_{\phi_1}^{-1} a_1(x, s_1, \dots, s_{d_1}, 0)_{|C_{\phi_1}},$$

see equation (8.11), we see that

$$\sigma_{\phi_1}(\mu) = (\wp_\phi)_* \, a(x, s, 0)_{|C_\phi} \tag{8.61}$$

where $\wp_\phi = \pi_1 \circ f$ and $(\wp_\phi)_*$ is fiber integration with respect to the fiber density along the fiber $f^* ds$.

8.12 The clean version of the symbol formula.

We will now say all this more intrinsically. Let $\pi : Z \to X$ be a fibration and $\phi : Z \to \mathbb{R}$ a generating function for Λ with respect to π. Then ϕ is a *clean* generating function if and only if the canonical relations,

$$\Lambda_\phi : \text{pt.} \Rightarrow T^*Z \quad \text{and} \quad \Gamma_\pi : T^*Z \Rightarrow T^*X$$

intersect cleanly, in which case $\Lambda = \Gamma_\phi \circ \Lambda_\pi$. If in addition we are given a fiber $\frac{1}{2}$-density, σ, on Z this gives us a push-forward operation:

$$\pi_* : \mathcal{C}_0^\infty(|TZ|^{\frac{1}{2}}) \to \mathcal{C}_0^\infty(|TX|^{\frac{1}{2}})$$

and a $\frac{1}{2}$-density, σ_π, on Γ_π. Now let

$$\nu = (2\pi h)^{k-\frac{d+e}{2}} a(z, h) e^{\frac{i\phi}{h}} \tau$$

be an element of $I^{k-\frac{d+e}{2}}(Z, \Lambda_\phi)$, where a is in $\mathcal{C}_0^\infty(Z \times \mathbb{R})$ and τ is a non-vanishing $\frac{1}{2}$-density on Z. Then, by what we proved above, $\pi_* \nu = \mu$ is in $I^k(X, \Lambda)$. We will prove that, just as in the transverse case the symbol of μ is given by the formula

$$\sigma(\mu) = s_\phi \; \sigma_\pi \circ \sigma(\nu) \tag{8.62}$$

where s_ϕ is the section of $\mathbb{L}_{\text{Maslov}}(\Lambda)$ associated with ϕ. (For the definition of s_ϕ when ϕ is a *clean* generating function see §5.13.2 .)To prove this we will first suppose that $\Lambda \subseteq T^*X$ is horizontal, i.e., $\Lambda = \Lambda_\psi$ for some $\psi \in \mathcal{C}^\infty(X)$ and that $\phi = \psi \circ \pi$. Then Λ_ϕ sits inside H^*Z, so

$$\Gamma_\pi \star \Lambda_\phi = H^*Z \cap \Lambda_\phi = \Lambda_\phi$$

and the fibration

$$\Gamma_\pi \star \Lambda_\phi \to \Gamma_\pi \circ \Lambda_\phi = \Lambda \tag{8.63}$$

is just the restriction to Λ_ϕ of the fibration

$$H^* Z = \pi^* T^* X \to X\,.$$

In other words if we denote by

$$\wp_\phi : \Lambda_\phi \to Z$$

and

$$\wp : \Lambda \to X$$

the projection of Λ_ϕ onto Z and Λ onto X, the map (8.63) is just the map

$$\pi_\Lambda : \Lambda_\phi \to \Lambda\,, \quad \pi_\Lambda = \wp^{-1} \circ \pi \circ \wp\,. \tag{8.64}$$

In particular the fibers of this map coincide with the fibers of π, so our enhancement of π gives us an enhancement of π_Λ, and hence a push-forward operation

$$(\pi_\Lambda)_* : \mathcal{C}_0^\infty(|T\Lambda_\phi|^{\frac{1}{2}}) \to \mathcal{C}_0^\infty(|T\Lambda|^{\frac{1}{2}})$$

and it is easily checked that, for $\sigma \in \mathcal{C}_0^\infty(|T\Lambda_\phi|^{\frac{1}{2}})$,

$$\sigma_\pi \circ \sigma = (\pi_\Lambda)_* \sigma\,. \tag{8.65}$$

Thus given $\nu = (2\pi h)^k a(z,h) e^{i\frac{\phi}{h}} \tau$ in $I^k(Z, \lambda_\phi)$

$$\begin{aligned}\sigma_\pi \circ \sigma(\nu) &= \sigma_\pi \circ \wp_\phi^*(a(z,0)\tau) \\ &= \wp_\psi^* \pi_* a(z,0)\end{aligned}$$

so if

$$\mu = \pi_* \nu = (2\pi\hbar)^k (\pi_* a(z,\hbar)\tau) e^{i\frac{\psi}{\hbar}}$$

we get for the symbols of μ and ν

$$\sigma(\mu) = \sigma_\pi \circ \sigma(\nu) = s_\phi \sigma_\pi \circ \sigma(\nu)\,. \tag{8.66}$$

(We can insert the factor, s_ϕ, into the last term because, $\phi = \pi^*\psi$ involves no fiber variables and hence $s_\phi \equiv 1$.)

Let us now turn to the general case. As we observed above, the fibration, $\pi : Z \to X$ can be factored (locally) into a pair of fibrations

$$Z \xrightarrow{\pi_2} Z_1 \xrightarrow{\pi_1} Z$$

such that $\phi = \phi_1 \circ \pi_2$ and $\phi_1 : Z_1 \to \mathbb{R}$ is a transverse generating function for Λ with respect to π_1. Moreover, if we enhance these two fibrations by equipping them with fiber $\frac{1}{2}$-densities this gives us an enhancement of π having the properties

$$\sigma_\pi = \sigma_{\pi_2} \circ \sigma_{\pi_1}$$

and

$$\pi_* = (\pi_1)_* (\pi_2)_*$$

and the assertion (8.62) follows from the transversal version of this result and the result we've just proved.

8.13 Clean composition of Fourier integral operators.

Let X_1, X_2, X_3 be manifolds and $M_i = T^*X_i,\ i = 1, 2, 3$. Let

$$\Gamma_i \subset M_i^- \times M_{i+1}, \quad i = 1, 2$$

be exact canonical relations with phase functions ψ_i. Suppose that Γ_2 and Γ_1 are cleanly composable, so that $\Gamma_2 \star \Gamma_1$ is a C^∞ manifold and

$$\kappa : \Gamma_2 \star \Gamma_1 \to \Gamma_2 \circ \Gamma_1$$

is a smooth fibration with connected fibers. Let e be the fiber dimension of this fibration.

Suppose that
$$Z_i = X_i \times X_{i+1} \times \mathbb{R}^{d_i}, \quad i = 1, 2$$
that
$$\pi_i : Z_i \to X_i \times X_{i+1}, \quad i = 1, 2$$
and $\phi_i \in C^\infty(Z_i)$ are such that $(\pi_i, \phi_i),\ i = 1, 2$ are *transverse* presentations of (Γ_i, ψ_i).

Let
$$Z = X_1 \times X_3 \times (X_2 \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}), \quad \pi : Z \to X_1 \times X_3.$$
We know that the function ϕ on Z given by

$$\phi(x_1, x_3; x_2, s_1, s_2) = \phi_1(x_1, x_2, s_1) + \phi_2(x_2, x_3, s_2)$$

is a clean generating function for $\Gamma_2 \circ \Gamma_1$ with respect to π.

The diffeomorphisms $\wp_{\phi_i} : C_{\phi_i} \to \Gamma_i,\ i = 1, 2$ give us a diffeomorphism

$$\gamma_\phi : C_\phi \to \Gamma_2 \star \Gamma_1$$

where γ_ϕ is the composition

$$C_\phi \to C_{\phi_1} \times C_{\phi_2} \to \Gamma_1 \times \Gamma_2$$

and also gives us an identification

$$\Gamma_2 \star \Gamma_1 = \Gamma_\pi \star \Lambda_\phi. \tag{8.67}$$

We have the factorization

$$\wp_\phi = \kappa \circ \gamma_\phi.$$

Now suppose that $\mu_j,\ j = 1, 2$ are the Schwartz kernels of Fourier integral operators F_j of order k_j associated with Γ_j and that they have the local description

$$\mu_j = \hbar^{k_j - \frac{n_{j+1}}{2} - \frac{d_j}{2}} \int a(x_j, x_{j+1}, s_j, \hbar) e^{\frac{i\phi_j}{\hbar}} d\mathbf{s}_j, \quad j = 1, 2.$$

Then the operator $F_2 \circ F_1$ has Schwartz kernel

$$\mu = \int \mu_1(x_1, x_2, \hbar)\mu_2(x_2, x_3, \hbar) dx_2 = \hbar^{k_1 + k_2 - \frac{n_3}{2} - \frac{d_1 + d_2 + n_2}{2}} \int a_1 a_2 e^{\frac{i\phi}{\hbar}} d\mathbf{s}_1 d\mathbf{s_2} dx_2.$$

By the results of the preceding section, we know that

$$\mu \in I^{k_1 + k_2 - \frac{n_3}{2} - \frac{e}{2}}(X_1 \times X_3, \Gamma, \psi)$$

where $n_3 = \dim X_3$. Hence we conclude

Theorem 8.13.1. *The operator $F_2 \circ F_1$ is a Fourier integral operator of order $k_1 + k_2 - \frac{e}{2}$ associated with the canonical relation $\Gamma_2 \circ \Gamma_1$*

8.13.1 A more intrinsic description.

We can describe the construction above more intrinsically as follows. If π_i is the fibration of Z_i over $X_i \times X_{i+1}$ then $\pi_1 \times \pi_2$ is a fibration of $Z_1 \times Z_2$ over the product $X_1 \times X_2 \times X_2 \times X_3$ and Z is the preimage in $Z_1 \times Z_2$ of the set $X_1 \times \Delta_2 \times X_3$ where Δ_2 is the diagonal in $X_2 \times X_2$. Therefore $\pi : Z \to X_1 \times X_3$ is the composite map

$$\pi = \gamma \circ (\pi_1 \times \pi_2) \circ \iota \tag{8.68}$$

where ι is the inclusion of Z in $Z_1 \times Z_2$ and γ is the projection,

$$\gamma : X_1 \times \Delta_2 \times X_3 \to X_1 \times X_3 \,.$$

We will now show how to "enhance" the fibration, π, to make it into a morphism of $\frac{1}{2}$-densities. By the definition above the conormal bundle of Z in $Z_1 \times Z_2$ can be identified with the pull-back to Z of the cotangent bundle, T^*X_2, via the map

$$Z \to X_1 \times \Delta_2 \times X_3 \to \Delta_2 = X_2$$

the first arrow being the map, $(\pi_1 \circ \pi_2) \circ \iota$. Therefore, by Section 7.4.1, enhancing ι amounts to fixing a non-vanishing section of $|T^*X_2|^{\frac{1}{2}}$. On the other hand the fiber of γ is X_2 so enhancing γ amounts to fixing a section of $|TX_2|^{\frac{1}{2}}$. Thus the constant section, 1, of $|T^*X_2|^{\frac{1}{2}} \otimes |TX_2|^{\frac{1}{2}}$ gives one a simultaneous enhancement of γ and ι. Therefore from (8.68) we conclude

Proposition 8.13.1. *Enhancements of the fibrations, π_1 and π_2, automatically give one an enhancement for π.*

Fixing such enhancements the Schwartz kernel of F_i has a global description as a push-forward

$$h^{k_i-\frac{n_{i+1}}{2}-\frac{d_i}{2}}(\pi_i)_*\nu_i e^{\frac{i\phi_i}{h}} \tag{8.69}$$

where $\nu_i(z,h)$ is a globally defined $\frac{1}{2}$-density on Z_i depending smoothly on h. As for the Schwartz kernel of $F_2\circ F_1$ the formula for it that we described above can be written more intrinsically as

$$h^k\pi_*\iota^*((\nu_1\otimes\nu_2)e^{i\frac{\phi_1+\phi_2}{h}})\,. \tag{8.70}$$

where

$$k=k_1+k_2-\frac{d_1+d_2+n_2}{2}-\frac{n_3}{2}.$$

(Note that since we've enhanced ι the pull-back operation, ι^*, is well-defined as an operation on $\frac{1}{2}$-densities and since we've enhanced π the same is true of the operation, π_*.) We'll make use of (8.70) in the next section to compute the intrinsic symbol of $F_2\circ F_1$.

8.13.2 The composition formula for symbols of Fourier integral operators when the underlying canonical relations are cleanly composable.

From the intrinsic description of the Schwartz kernel of $F_2\circ F_1$ given by (8.70) and the results of Section **??** we get a simple description of the intrinsic symbol of $F_2\circ F_1$. The enhancing of π gives us a $\frac{1}{2}$-density, σ_π, on Γ_π and from the symbol of $\nu=\iota^*(\nu_1\otimes\nu_2)$ we get a $\frac{1}{2}$-density, $\sigma(\nu)$, on Λ_φ, and from these data we get by Theorem 36 of §7.1 an object, $\sigma_\pi*\nu$, on $\Gamma_\pi\star\Lambda_\phi$ of the form $\kappa^*\alpha\otimes\beta$ where α is a $\frac{1}{2}$-density on $\Gamma_\pi\circ\Lambda_\phi$ and β is a density on the fibers of the fibration, $\kappa:\Gamma_\pi\star\Lambda_\varphi\to\Gamma_\pi\circ\Lambda_\varphi$. Hence we can integrate β over fibers to get a complex-valued function, $\pi_*\beta$, on $\Gamma_\pi\circ\Lambda_\varphi$, and Theorem **??** of § **??** tells us that the composite symbol

$$\sigma_\pi\circ\sigma(\nu)=\alpha\pi_*\beta$$

is, modulo Maslov factors, the intrinsic symbol of $F_2\circ F_1$. On the other hand the symbol, σ_i, of F_i is a $\frac{1}{2}$-density on Γ_i, and from the $\frac{1}{2}$-densities, σ_1 and σ_2 we again get, by §7.1 , an object $\sigma_2\star\sigma_1$ on $\Gamma_2\star\Gamma_1$ which is the pull-back of a $\frac{1}{2}$-density on $\Gamma_2\circ\Gamma_1$ times a density on the fibers of the fibration, $\Gamma_2\star\Gamma_1\to\Gamma_2\circ\Gamma_1$, and the fiberwise integral of this object is the composite $\frac{1}{2}$-density $\sigma_2\circ\sigma_1$ on $\Gamma_2\circ\Gamma_1$. However as we observed above $\Gamma_2\star\Gamma_1=\Gamma_\pi\star\Lambda_\phi$, $\Gamma_2\circ\Gamma_1=\Gamma_\pi\circ\Lambda_\phi$, and the fibrations of $\Gamma_2\star\Gamma_1$ over $\Gamma_2\circ\Gamma_1$ and of $\Gamma_\pi\star\Lambda_\phi$ over $\Gamma_\pi\circ\Lambda_\phi$ are the same. Finally, a simple computation in linear algebra (which we'll omit) also shows that the objects $\sigma_2\star\sigma_1$ and $\sigma_\pi\star\sigma(\nu)$ are the same. As for Maslov factors, let Z be the preimage of $X_1\times\Delta_{X_2}\times X_3$ in $Z_1\times Z_2$ and let s_{ϕ_i}, $i=1,2$ be the section of $\mathbb{L}_{\text{Maslov}}(\Gamma_i)$ associated with ϕ_1. By the composition formula for sections of Maslov bundles described in Section 5.13.5

$$s_{\phi_2} \circ s_{\phi_1} = s_\phi$$

where ϕ is the restriction of $\phi_1 + \phi_2$ to Z and s_ϕ is the section of $\mathbb{L}_{\text{Maslov}}(\Gamma)$ associated to ϕ. Hence we have proved

Theorem 8.13.2. *The intrinsic symbol*

$$\sigma(F) = s_\phi \sigma_\pi \star \sigma(\nu)$$

of the Fourier integral operator $F = F_2 \circ F_1$ is the composition

$$\sigma(F_2) \circ \sigma(F_1)$$

of the M-enhanced symbols $\sigma(F_i) = s_{\phi_i}\sigma_{\pi_i} \circ \sigma(\nu_i),\ i = 1, 2$.

8.14 An abstract version of stationary phase.

As an application of the clean intersection ideas above, we'll discuss in this section an abstract version of the lemma of stationary phase. We'll begin by quickly reviewing the results of the previous two sections. Let X_i, $i = 1, 2, 3$, be manifolds and let $M_i = T^*X_i$. Assume we are given exact canonical relations

$$\Gamma_i : M_i \twoheadrightarrow M_{i+1}\,, \quad i = 1, 2$$

and assume that Γ_1 and Γ_2 are cleanly composable. Then we have a fibration

$$\kappa : \Gamma_2 \star \Gamma_1 \to \Gamma_2 \circ \Gamma_1 =: \Gamma$$

and the fiber dimension, e, of this fibration is the *excess* of this clean composition. If $F_i \in \mathcal{F}^{k_i}(\Gamma_i)$, $i = 1, 2$ is a Fourier integral operator with microsupport on Γ_i, then as we showed above $F_2 \circ F_1$ is in the space $\mathcal{F}^k(\Gamma_2 \circ \Gamma_1)$ where $k = k_1 + k_2 - \frac{e}{2}$. Moreover if $\varphi_{\Gamma_i} \in \mathcal{C}^\infty(\Gamma_i)$, $i = 1, 2$ are phase functions on Γ_1 and Γ_2, the associated phase function $\varphi_\Gamma \in \mathcal{C}^\infty(\Gamma)$ is defined by

$$\kappa^*\varphi_\Gamma = \gamma_1^*\varphi_{\Gamma_1} + \gamma_2^*\varphi_{\Gamma_2}\,. \tag{8.71}$$

Recall that

$$\Gamma_2 \star \Gamma_1 = \{(P_1, P_2, P_3), (P_i, P_{i+1}) \in \Gamma_i\}$$

and that $\gamma_i : \Gamma_2 * \Gamma_1 \to \Gamma_i$ is the projection

$$(P_1, P_2, P_3) \to (P_i, P_{i+1}), \quad i = 1, 2.$$

We now apply these facts to the following special case: Let X and Y be differentiable manifolds of dimensions m and n and let

$$f : X \to Y$$

be a C^∞ map and let

$$\Gamma_f = \{(x,\xi,y,\eta)\,;\, y = f(x)\,,\, \xi = df_x^*\eta\}\,.$$

Then $f^* : C^\infty(Y) \to C^\infty(X)$ can be viewed as a semi-classical F.I.O.

$$f^* \in \mathcal{F}^r(\Gamma_f^\sharp), \quad r = \frac{m-n}{2}$$

in the sense that for every $P \in \Psi^0(Y)$

$$f^*P \in \mathcal{F}^r(\Gamma_f^*)\,.$$

Moreover, suppose f is a fiber mapping with compact fibers. Then if we fix volume densities dx and dy on X and Y we get a fiber integration map

$$f_* : C^\infty(X) \to C^\infty(Y)$$

with the defining property that

$$\int f^*\varphi\psi\, dx = \int \varphi f_* \cdot \psi\, dy$$

for all $\varphi \in C_0^\infty(Y)$ and $\psi \in \mathcal{C}^\infty(X)$. In other words, f_* is just the transpose of f^*. Since transposes of semi-classical F.I.O.'s are also semi-classical F.I.O.'s we conclude that

$$f_* \in \mathcal{F}^0(\Gamma_f)$$

in the sense that $Pf_* \in \mathcal{F}^0(\Gamma_f)$ for all $P \subset \Psi^0(Y)$.

We want to apply these remarks to the following simple setup. Let X be a manifold and $Y \subset X$ a compact manifold of codimension n. Then we have an inclusion map $\iota : Y \to X$ and a projection map $\pi : Y \to \text{pt.}$. Equipping Y with a volume density, dy, we get from these maps Fourier integral operators

$$\iota^* : C^\infty(X) \to C^\infty(Y)$$

and

$$\pi_* : C^\infty(Y) \to C^\infty(\text{pt.}) = \mathbb{C}$$

associated with the canonical relations

$$\Gamma_\iota^\dagger = \{(x,\xi,y,\eta)\,,\, y = x\,,\, \eta = (d\iota)_y^*\xi\}$$

and

$$\Gamma_\pi = \{(y,\eta)\,,\, y \in Y\,,\, \eta = 0\}\,,$$

i.e., $\eta \in (d\pi_y)^*T^*\,\text{pt.} \Leftrightarrow \eta = 0$. Then

$$\Gamma_\pi \circ \Gamma_\iota^\dagger = \{(y,\xi)\,,\, y \in Y\,,\, \xi \in T_y^*X\,,\, (d\iota_y)^*\xi = 0\}$$

is just the conormal bundle $\Gamma = N^*Y$ in T^*X. Moreover its easy to see that this set coincides with $\Gamma_\pi \star \Gamma_\iota^\dagger$, so Γ_π and $\Gamma_\iota^\dagger$ are transversally composable. Therefore $\pi_*\iota^*$ is a semi-classical Fourier integral operator. Moreover since

$$\iota^* \in \mathcal{F}^{-\frac{n}{2}}(\Gamma_\iota^\dagger) \qquad n = \dim X$$

and

$$\begin{aligned} \pi_* &\in \mathcal{F}^o(\Gamma_\pi) \qquad 0 = \dim \mathrm{pt.}\,, \\ \pi_*\iota^* &\in \mathcal{F}^{-\frac{n}{2}}(\Gamma) \end{aligned}$$

where

$$\frac{n}{2} = -\frac{1}{2}(\dim X) + \frac{\dim Y}{2}\,.$$

Remark.

Since Γ is a conormal bundle $\iota_\Gamma^*\alpha_X = 0$ so Γ is exact with phase function $\varphi_\Gamma \equiv 0$. We'll make use of this fact below.

Now let $\Lambda \subseteq T^*X$ be an exact Lagrangian manifold with phase function φ_Λ. As is our wont, we'll regard Λ as a canonical relation

$$\Lambda : \mathrm{pt.} \Rightarrow T^*X$$

and Γ as a canonical relation

$$\Gamma : T^*X \Rightarrow \mathrm{pt.}$$

and composing these canonical relations we get the relation

$$\mathrm{pt.} \Rightarrow \mathrm{pt.}$$

and sitting over it the relation

$$\Gamma \star \Lambda$$

which is just the set of triples

$$(\mathrm{pt.}, p, \mathrm{pt.})$$

with $(\mathrm{pt.}, p) \in \Lambda$ and $(p, \mathrm{pt.}) \in \Gamma$, i.e., if we go back to thinking of Λ and Γ as Lagrangian manifolds in T^*X:

$$\Gamma \star \Lambda = \Gamma \cap \Lambda\,.$$

Therefore in this example Γ and Λ are cleanly composable iff Γ and Λ intersect cleanly in T^*X. Let's assume this is the case. Then taking

$$\mu \in I^k(X, \Lambda, \varphi)$$

and viewing μ as the Schwartz kernel of the operator

$$F_\mu : \mathcal{C}^\infty(\text{pt.}) \to \mathcal{C}^\infty(X)\,, \quad c \to c\mu$$

we get by composition of F.I.O.'s

$$\pi_*\iota^*\mu \in I^{k+\ell+\frac{n}{2}-\frac{e}{2}}(pt, \varphi_{\text{pt.}})$$

where $e = \dim \Gamma \cap \Lambda$ and

$$\begin{aligned} k+\ell+\frac{n}{2}-\frac{e}{2} &= k - \frac{\dim X}{2} + \frac{\dim Y}{2} + \frac{\dim X}{2} - \frac{e}{2} \\ &= k + \frac{m}{2} - \frac{e}{2} \qquad m = \frac{\dim Y}{2} \end{aligned}$$

and $\varphi_{\text{pt.}}$ satisfies $\kappa^*\varphi_{\text{pt.}} = \gamma_1^*\varphi_1 + \gamma_2^*\varphi_2$ where φ_1 and φ_2 are the phase functions on Γ and Λ and γ_1 and γ_2 are the inclusion maps,

$$\Gamma \cap \Lambda \to \Gamma$$

and

$$\Gamma \cap \Lambda \to \Lambda\,.$$

Thus since $\varphi_1 = 0$ and $\varphi_2 = \varphi$ our formula for composition of phase functions tells us

Lemma 8.14.1. *The restriction of φ to $\Lambda \cap \Gamma$ is constant and $\varphi_{\text{pt.}} = \varphi(p)$ where p is any point on $\Lambda \cap \Gamma$.*

Thus summarizing, we've proved

Theorem 8.14.1. *The integral*

$$\pi_*\iota^*\mu = \int_Y (\iota^*\mu)\, dy$$

has an asymptotic expansion

$$e^{i\frac{\varphi_{\text{pt.}}}{h}} h^{k+\frac{m}{2}-\frac{e}{2}} \sum_{i=0}^{\infty} a_i h^i\,. \tag{8.72}$$

This is, in semi-classical analysis, the *abstract lemma of stationary phase.*

Remark.

If Γ and Λ intersect cleanly in N connected components

$$(\Gamma \cap \Lambda)_r\,, \qquad r = 1, \ldots$$

one gets a slightly generalized version of (8.72)

$$\pi_*\iota^*\mu \sim \sum_{r=1}^{m} e^{\frac{i\varphi_r(\text{pt.})}{h}} h^{k+\frac{m}{2}-e_r} \sum_{i=0}^{\infty} a_{i,r} h^r \tag{8.73}$$

where $\varphi_r(\text{pt.}) = \varphi(p_r)$, $p_r \in (\Lambda \cap \Gamma)_r$ and $e_r = \dim \Lambda \cap \Gamma_r$.

Chapter 9

Pseudodifferential Operators.

In this chapter we will give a brief account of the "classical" theory of semi-classical pseudo-differential operators: pseudo-differential operators whose symbols satisfy appropriate growth conditions at infinity. We will show that most of the main properties of these operators can be deduced, via microlocalization, from properties of the semi-classical pseudo-differential operators with compact support that we introduced in Chapter 8.

9.1 Semi-classical pseudo-differential operators with compact microsupport.

In §8.6 we defined a class of operators which we called "semi-classical pseudo-differential operators". A more appropriate description of these operators is "semi-classical pseudo-differential operators with compact microsupport". On open subsets of $\mathbb{R}^n$ they are integral operators of the form

$$A : C^\infty(U) \to C^\infty(U), \quad \phi \mapsto \int K_A(x, y, \hbar)\phi(y)dy$$

where $K_A(x, y, \hbar)$ is an oscillatory integral

$$K_A(x, y, \hbar) = \int a(x, y, \xi, \hbar)e^{i\frac{(x-y)\cdot\xi}{\hbar}} d\xi \tag{9.1}$$

with amplitude

$$a \in C_0^\infty(U \times U \times \mathbb{R}^n \times \mathbb{R}).$$

By the general theory of oscillatory integrals, these are "semi-classical Fourier integral operators associated to the identity map of T^*U to itself". We know from the general theory that their definition is coordinate invariant. However, since these operators will play a fundamental role in this chapter, here is a short proof of this fact:

Let $f : V \to U \subset \mathbb{R}^n$ be a diffeomorphism, and let

$$B = f^* A (f^{-1})^*$$

so that B is an integral operator with kernel

$$K_B(x, y, \xi, \hbar) = \int a(f(x), f(y), \xi, \hbar) e^{i\frac{(f(x)-f(y))\cdot\xi}{\hbar}} \left|\det Df(y)\right| d\xi.$$

Define f_{ij} by

$$f_i(x) - f_i(y) = \sum_j f_{ij}(x, y)(x_j - y_j)$$

and let F be the matrix $F = (f_{ij})$. So

$$(f(x) - f(y)) \cdot \xi = (F(x, y)(x - y)) \cdot \xi = (x - y) \cdot F^\dagger(x, y)\xi$$

and the above expression for K_b can be written as

$$K_B = \int b(x, y, \xi, \hbar) e^{i\frac{(x-y)\cdot\xi}{\hbar}} d\xi$$

where

$$b(x, y, x, \xi, \hbar) = a(f(x), f(y), (F^\dagger)^{-1}(x, y)\xi, \hbar) \left|\det \left(F(x, y)^{-1} Df(y)\right)\right|. \quad \Box \tag{9.2}$$

Equation (9.2) shows how this changes under a diffeomorphism, and, in particular that it is intrinsically defined.

Moreover, since

$$f_i(x) - f_i(y) = Df_i(x - y) + O(\|x - y\|^2),$$

equation (9.2) also shows that

$$b(y, y, \xi, 0) = a\left(f(y), f(y), Df(y)^\dagger \xi, 0\right).$$

In other words, it shows that the leading symbol of $f^* A(f^{-1})^*$ is $g^*\sigma(A)(x, \xi)$ where $g : T^*V \to T^*U$ is the diffeomorphism of cotangent bundles corrresponding to the diffeomorphism f.

So this gives us an elementary proof of the a property of pseudo-differential operators that we proved in Chapter 8 - that their leading symbols are intrinsically defined as functions on the cotangent bundle.

Let us define the **microsupport** of A to be the closure in T^*U of the set of points, (x, ξ), at which $D_y^\alpha D_\hbar^N a(x, x, \xi, 0) \neq 0$ for some α and N.

We will let $\Psi_0(U)$ denote the set of semi-classical pseudo-differential operators with compact microsupport in U, and by $\Psi_{00}(U)$ the subset of $\Psi_0(U)$ consisting of semi-classical pseudo-differential operators with microsupport in the set $\xi \neq 0$.

More generally, if X is an n-dimensional manifold, we denote the analogous objects on X by $\Psi_0(X)$ and by $\Psi_{00}(X)$. Our proof above that the definition of of semi-classical pseudo-differential operators with compact microsupport is coordinate invariant justifies this definition.

9.2 Classical ΨDO's with polyhomogeneous symbols.

Our goal in this chapter is to get rid of the "compact microsupport" condition and show that $\Psi_0(X)$ is a subalgebra of a much larger class of of semi-classical pseudo-differential operators.

As a first step in this direction, we will give in this section a somewhat unorthodox description of the class of classical pseudo-differential operators with polyhomogeneous symbols, the standard house and garden variety of pseudo-differential operators of Kohn-Nirenberg, Hörmander, et al. (See for instance, [HorIII].) Our description is based on an observation that we made in §8.10: Let X be a manifold and let $A : C^\infty(X) \to C^\infty(X)$ be a *differential* operator. We saw that if $P \in \Psi_0(X)$ then $AP \in \Psi_0(X)$.

Now let $A : C_0^\infty(X) \to C^{-\infty}(X)$ be a continuous operator in the distributional sense, i.e. admitting as Schwartz kernel a generalized function

$$K_A \in C^{-\infty}(X \times X)$$

(relative to some choice of smooth density).

We will convert the observation we made above about differential operators into a definition:

Definition 9.2.1. *A is a* **classical pseudo-differential operator with polyhomogeneous symbol** *if,*

$$AP \in \Psi_{00}(X)$$

for every $P \in \Psi_{00}(X)$.

Remarks.

- We will explain at the beginning of the next section why we cannot replace $\Psi_{00}(X)$ by $\Psi_0(X)$ in this definition.
- From the results of §8.10 we know that differential operators belong to this class.

Here are some other examples: Assume for the moment that

$$K_A \in C^\ell(X \times X)$$

for some $\ell \geq 0$. Pre- and post-multiplying K_A by compactly supported smooth cut-off functions, we may assume that $X = \mathbb{R}^n$. We may write

$$K_A(x, y) = K(x, x - y)$$

where $K(x, w) = K_A(x, x - w)$.

Let P be the zero-th order semi-classical pseudo-differential operator

$$P = \psi(x)\rho(\hbar D)$$

where $\psi(w) \equiv 1$ on the set where $K(x,w) \neq 0$ and $\rho = \rho(\xi) \in C_0^\infty(\mathbb{R}^n)$.

The Schwartz kernel of P is

$$\psi(x)\left(\hbar^{-n}\int \rho(\xi)e^{i\frac{(x-y)\cdot\xi}{\hbar}}d\xi\right)$$

and hence the Schwartz kernel of AP is

$$\hbar^{-n}\int K(x,x-z)e^{\frac{i(z-y)\cdot\xi}{\hbar}}\rho(\xi)dzd\xi.$$

For fixed x, let us make the change of variables $w = z - x$. The above integral then becomes

$$\int K(x,-w)e^{i\frac{w\cdot\xi}{\hbar}}e^{i\frac{(x-y)\cdot\xi}{\hbar}}\rho(\xi)dwd\xi.$$

This equals

$$(2\pi)^{n/2}\int \hat{K}\left(x,\frac{\xi}{\hbar}\right)\rho(\xi)e^{i\frac{(x-y)\cdot\xi}{\hbar}}d\xi \tag{9.3}$$

where $\hat{K}$ is the Fourier transform of K with respect to w:

$$\hat{K}(x,\zeta) = \frac{1}{(2\pi)^{n/2}}\int K(x,w)e^{-iw\cdot\zeta}dw.$$

Suppose that ρ is supported on the set

$$\epsilon \leq \|\xi\| \leq \frac{1}{\epsilon}$$

and is identically one on the set

$$2\epsilon < \|\xi\| < \frac{1}{2\epsilon}.$$

Then $P \in \Psi_{00}(\mathbb{R}^n)$, so in order for AP to be a semi-classical pseudo-differential operator with compact microsupport, $\hat{K}$ has to have a semi-classical expansion

$$\hat{K}\left(x,\frac{\xi}{\hbar}\right) \sim \hbar^{-k}\sum_{i=0}^{\infty}F_i(x,\xi)\hbar^i$$

on the set $2\epsilon < \|\xi\| \leq \frac{1}{2\epsilon}$, for some k.

Letting $\hbar = \frac{1}{\|\xi\|}$ and writing

$$\xi = \|\xi\|\cdot\frac{\xi}{\|\xi\|}$$

this becomes the more conventional expression

$$\hat{K}(x,\xi) \sim \sum a_i(x,\xi) \tag{9.4}$$

for $\|\xi\| >> 0$ where

$$a_i(x,\xi) = \|\xi\|^k F_i\left(x, \frac{\xi}{\|\xi\|}\right)\|\xi\|^{-i} \tag{9.5}$$

is a homogeneous symbol of degree $-i+k$. In other words, A is a classical pseudo-differential operator with polyhomogeneous symbol $a(x,\xi) = \hat{K}(x,\xi)$. (For the standard definition of these objects, see [HorIII] p. 67.)

Notice that since $K(x,\cdot) \in C_0^\ell$, k has to be less than $-\frac{n}{2} - \ell$.

We now prove a converse result - that if A is a classical pseudo-differential operator with polyhomogeneous symbol

$$a(x,\xi) \sim \sum_{i=0}^{\infty} a_i(x,\xi) \tag{9.6}$$

which is compactly supported in x and of degree $k < -n$ then A is a polyhomogeneous pseudo-differential operator in our sense.

Let

$$K(x,w) = \frac{1}{(2\pi)^{n/2}}\int a(x,\xi)e^{iw\cdot\xi}d\xi,$$

be the inverse Fourier transform of a with respect to ξ. We recall the following facts about the Fourier transform:

Lemma 9.2.1. *If $-k > n+\ell$ then $K(x,\cdot) \in C^\ell$.*

Proof. For $|\alpha| \leq \ell$,

$$|(D_w)^\alpha K(x,w)| \leq \frac{1}{(2\pi)^{n/2}}\int |a(x,\xi)\xi^\alpha|\, d\xi$$

is bounded. Indeed, the integrand on the right is bounded by $\langle\xi\rangle^{k+\ell}$ and $k+\ell < -n$. □

Lemma 9.2.2. *On the set $w_j \neq 0$,*

$$K(x,w) = w_j^{-N}\frac{1}{(2\pi)^{n/2}}\int\left(i\frac{\partial}{\partial\xi_j}\right)^N a(x,\xi)e^{iw\cdot\xi}d\xi$$

for all N.

Proof. Use the identity

$$\left(-i\frac{\partial}{\partial\xi_j}\right)^n e^{iw\cdot\xi} = w_j^N e^{iw\cdot\xi}$$

and integration by parts. □

Lemma 9.2.3. *If a is a polyhomogeneous symbol of degree k, then*

$$\left(i\frac{\partial}{\partial \xi_j}\right)^N a(x,\xi)$$

is a polyhomogeneous symbol of degree $k-N$.

Proof. Term-wise differentiation of the asymptotic expansion $a(x,\xi)=\sum a_i(x,\xi)$. □

Corollary 9.2.1. *$K(x,w)$ is C^∞ on the set $w \neq 0$.*

Now note that by the Fourier inversion formula,

$$a(x,\xi)=\hat{K}(x,\xi).$$

Hence for $\rho \in C_0^\infty(\mathbb{R}^n)$ with support on the set

$$\epsilon < \|\xi\| < \frac{1}{\epsilon},$$

the Schwartz kernel of $A\rho(\hbar D)$ is

$$\hbar^{-n}\int a\left(x,\frac{\xi}{\hbar}\right)\rho(\xi)e^{i\frac{(x-y)\cdot\xi}{\hbar}}d\xi$$

by (9.3). Hence, by (9.6), $A\rho(\hbar D)\in \Psi_{00}(\mathbb{R}^n)$.

More , if $P\in\Psi_{00}(\mathbb{R}^n)$ and $\rho \equiv 1$ on the microsupport of P, then by (8.45)

$$P=\rho(\hbar D)P$$

and hence

$$AP=(A\rho(\hbar D))P\in\Psi_{00}(\mathbb{R}^n).$$

Conclusion: A is a polyhomogeneous pseudo-differential operator in our sense.

Let us now get rid of the assumption that A is an integral operator:

Let X be a manifold and

$$A: C_0^\infty(X)\to C^{-\infty}(X)$$

be a continuous operator with Schwartz kernel $K_A(x,y)$. Pre- and post- multiplying K_A by compactly supported cut-off functions we may assume that

$$K_A\in C_0^{-\infty}(\mathbb{R}^n\times\mathbb{R}^n).$$

Hence by Schwartz's theorem,

$$K_A=\langle D_x\rangle^{2N}\langle D_y\rangle^{2N}K_B$$

where

$$K_B \in C_0^\ell(\mathbb{R}^n \times \mathbb{R}^n)$$

for some positive integers ℓ and N. In other words,

$$B = \langle D\rangle^{-2N} A \langle D\rangle^{-2N}$$

is an integral operator with a C^ℓ kernel. Now $\langle D\rangle^{-2N}$ is a classical pseudo-differential operator with polyhomogeneous symbol

$$(1 + \|\xi\|^2)^{-N}$$

and hence by what we proved above, it is a pseudo-differential operator with polyhomogeneous symbol in our sense. Thus if A is a polyhomogeneous pseudo-differential operator in our sense, so is B. We conclude that B is a polyhomogeneous pseudo-differential operator in the standard sense, i.e., operates on $C_0(\mathbb{R}^n)$ by the recipe

$$f \mapsto \frac{1}{(2\pi)^n} \int b(x,\xi) e^{ix\cdot\xi} \hat{f}(\xi) d\xi$$

where b is a standard polyhomogeneous symbol. Thus A is the classical pseudo-differential operator with polyhomogeneous symbol

$$a(x,\xi) = \langle D_x + \xi\rangle^{2N} b(x,\xi) \langle \xi\rangle^{2N}.$$

A consequence of this computation which will be useful later is

Proposition 9.2.1. *Let $A : C_0^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ be a classical pseudo-differential operator with polyhomogeneous symbol $a(x,\xi)$ of order k. Then $A\langle D\rangle^{-2N}$ is a classical pseudo-differential operator with polyhomogeneous symbol $a(x,\xi)\langle\xi\rangle^{-2N}$. In particular, if $k - 2N < -\ell - n$ then $A\langle D\rangle^{-2N}$ is an integral operator and its kernel is in $C^\ell(\mathbb{R}^n \times \mathbb{R}^n)$.*

As a corollary we obtain

Proposition 9.2.2. *Let $A : C_0^\infty(\mathbb{R}^n) \to C^{-\infty}(\mathbb{R}^n)$ be a classical pseudo-differential operator with polyhomogeneous symbol of order k. Then A maps $C_0^\infty(\mathbb{R}^n)$ into $C^\infty(\mathbb{R}^n)$.*

Proof. For any ℓ pick N so that $k - 2N < n - \ell$ and write $A = B\langle D\rangle^{2N}$ where B is a a classical pseudo-differential operator with polyhomogeneous symbol of order $k - 2N < n - \ell$. Now $\langle D\rangle^{2N}$ maps $C_0^\infty(\mathbb{R}^n)$ into itself and B maps $C_0^\infty(\mathbb{R}^n)$ into $C^\ell(\mathbb{R}^n)$. □

Remarks.

1. Formally, the Schwartz kernel of A is the generalized function

$$K_A(x,y) = \int a(x,\xi) e^{i(x-y)\cdot\xi} d\xi.$$

If we make the change of variables $\xi \mapsto \xi/\hbar$ this becomes

$$\hbar^{-n} \int a\left(x, \frac{\xi}{\hbar}\right) e^{i\frac{(x-y)\cdot\xi}{\hbar}} d\xi.$$

In other words, in semi-classical form, A is the operator

$$C_0^\infty(\mathbb{R}^n) \ni f \to \hbar^{-\frac{n}{2}} \int a(x,\xi,\hbar) e^{i\frac{x\cdot\xi}{\hbar}} (\mathcal{F}_\hbar f)(\xi) d\xi$$

where

$$a(x,\xi,\hbar) = a\left(x, \frac{\xi}{\hbar}\right)$$

and $\mathcal{F}_\hbar$ is the semi-classical Fourier transform.

Now $\mathcal{F}_\hbar \rho(\hbar D) f = \rho(\xi)\mathcal{F}_\hbar f$, so $A\rho(\hbar D)$ is the operator given by

$$\left[A\rho(\hbar D)f\right](x) = \hbar^{n/2} \int a(x,\xi,\hbar)\rho(\xi) e^{i\frac{x\cdot\xi}{\hbar}} (\mathcal{F}_\hbar)(\xi) d\xi. \tag{9.7}$$

2. Let $A : C_0^\infty(X) \to C^\infty(X)$ be a smoothing operator. In other words, assume that A has a Schwartz kernel $K = K_A \in C^\infty(X \times X)$. Then A can be viewed as a classical pseudo-differential operator of order $-\infty$. Hence, for every $P \in \Psi_{00}(X)$, the operator PA belongs to $\Psi_{00}^{-\infty}(X)$. This can also be easily proved by the methods of Chapter 8. Indeed, we may write

$$K(x,y) = K(x,y) e^{\frac{i\phi(x,y)}{\hbar}}$$

where $\phi \equiv 0$. Hence A can be regarded as a semi-classical Fourier integral operator with microsupport on the zero section of $T^*(X \times X)$. So if $P \in \Psi_{00}(X)$, its microsupport does not intersect the microsupport of A, and hence AP is a Fourier integral operator (with microsupport on the zero section of $T^*(X \times X)$) of order $-\infty$. In other words from the microlocal perspective it's the zero operator.

9.3 Semi-classical pseudo-differential operators.

We have seen that an operator

$$A : C_0^\infty(X) \to C^{-\infty}(X)$$

is a classical polyhomogeneous pseudo-differential operator if and only if it has the property

$$AP \in \Psi_{00}(X) \quad \text{for all} \quad P \in \Psi_{00}(X).$$

The condition that $P \in \Psi_{00}(X)$ requires not only that P have compact microsupport, but also that *the microsupport of P is disjoint from the zero section of T^*X*. We will now show that it's important to make this stipulation. We will show that if we impose on A the stronger condition: "$AP \in \Psi_0(X)$ for all $P \in \Psi_0(X)$" then essentially the only operators with this property are differential operators.

To see this, let us assume that $X = \mathbb{R}^n$ and that the Schwartz kernel K_A of A is in $C_0^\ell(\mathbb{R}^n \times \mathbb{R}^n)$ for large ℓ. Let K be defined by

$$K_A(x,y) = K(x, x-y),$$

where $K(x,w) = K_A(x, x-w)$. Let $\rho \in C_0^\infty(\mathbb{R}^n)$ with

$$\rho(\xi) \equiv 1 \quad \text{for } \|\xi\| < \frac{1}{\epsilon}.$$

Then $A\rho(\hbar D)$ has kernel

$$(2\pi)^{n/2} \int \hat{K}\left(x, \frac{\xi}{\hbar}\right) \rho(\xi) e^{i\frac{(x-y)\cdot\xi}{\hbar}} d\xi$$

by (9.3). Thus if $A\rho(\hbar D) \in \Psi_0(X)$, we would have an asymptotic expansion

$$\hat{K}\left(x, \frac{\xi}{\hbar}\right) \sim \hbar^k \sum F_i(x,\xi)\hbar^i$$

for $\|\xi\| < \frac{1}{\epsilon}$, with $k \geq \frac{n}{2} + \ell$. Thus for $\hbar < 1$ we may replace ξ by $\hbar\xi$ in this expansion to get

$$\hat{K}(x,\xi) \sim \hbar^k \sum F_i(x, \hbar\xi)\hbar^i$$

and hence, letting $\hbar \to 0$,

$$\hat{K}(x,\xi) \equiv 0.$$

The situation becomes a lot better if we allow our operators to depend on $\hbar$. More explicitly, let

$$A_\hbar : C_0^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$$

be an operator with Schwartz kernel

$$K_A(x,y,\hbar) \in C^\ell\left(\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}\right)$$

and set

$$K(x,w,\hbar) = K_A(x, x-w, \hbar).$$

Since $K(x,w,\hbar)$ is in C_0^ℓ as function of x, there is a constant C such that

$$\int |D_w^\alpha K(x,w,\hbar)|\, dw \leq C, \qquad \forall \ |\alpha| \leq \ell.$$

So if $\hat{K}$ denotes the Fourier transform of K with respect to w, we have

$$\left|\xi^\alpha \hat{K}(x,\xi,\hbar)\right| \leq C \quad \forall \ |\alpha| \leq \ell. \tag{9.8}$$

We now repeat the argument we gave at the beginning of this section, but keep track of the $\hbar$-dependence: As above, let $\rho = \rho(\xi) \in C_0^\infty(\mathbb{R}^n)$ be supported on the set $\|\xi\| < \frac{1}{\epsilon}$ and be identically 1 on the set $\|\xi\| < \frac{1}{2\epsilon}$. By (9.3), the Schwartz kernel of $A\rho(\hbar D)$ is

$$\hbar^{-n} \int \hat{K}\left(x, \frac{\xi}{\hbar}, \hbar\right) \rho(\xi) e^{i\frac{(x-y)\cdot\xi}{\hbar}} d\xi.$$

For $A\rho(\hbar D)$ to be a semi-classical pseudo-differential operator with compact microsupport for all choices of such ρ, we must have

$$\hat{K}\left(x, \frac{\xi}{\hbar}, \hbar\right) = b(x, \xi, \hbar)$$

for some $b \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R})$. In other words, $\hat{K}$ has to be a function of the form

$$\hat{K}(x, \xi, \hbar) = b(x, \hbar\xi, \hbar). \tag{9.9}$$

We have thus proved:

Theorem 9.3.1. *Let $A : C_0^\infty(\mathbb{R}^n) \to C_0^\ell(\mathbb{R}^n)$ be an operator with Schwartz kernel*

$$K = K(x, y, \hbar) \in C_0^\ell(\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}).$$

Suppose that A has the microlocality property

$$AP \in \Psi_0(\mathbb{R}^n) \ \text{ for all } \ P \in \Psi_0(\mathbb{R}^n).$$

Then the Schwartz kernel of A is given by an oscillatory integral of the form

$$\hbar^{-n} \int b(x, \hbar\xi, \hbar) e^{i\frac{(x-y)\cdot\xi}{\hbar}} d\xi \tag{9.10}$$

where (by (8))

$$|b(x, \xi, \hbar)| \le C\hbar^\ell \langle\xi\rangle^{-\ell}. \tag{9.11}$$

We will devote most of the rest of this section to proving a converse result. Let us first note that (9.10) can be written as

$$\int b(x, \hbar\xi, \hbar) e^{i(x-y)\cdot\xi} d\xi \tag{9.12}$$

by making the change of variables $\xi \mapsto \hbar\xi$. So $A = A_\hbar$ is the operator

$$(Af)(x) = \int b(x, \hbar\xi, \hbar) e^{ix\cdot\xi} \hat{f}(\xi) d\xi \tag{9.13}$$

where $\hat{f}$ is the Fourier transform of f. This operator makes sense under hypotheses much weaker than (9.11). Namely, suppose that

$$|b(x, \xi, \hbar)| \le C\langle\xi\rangle^m \tag{9.14}$$

for some (possibly very large) integer m. We claim:

Theorem 9.3.2. *For b satisfying (9.14) the operator (9.13) is well defined and has the microlocality property*

$$AP \in \Psi_0(\mathbb{R}^n) \quad \textit{if} \quad P \in \Psi_0(\mathbb{R}^n).$$

Proof. Since $f \in C_0^\infty(\mathbb{R}^n)$ we have

$$|\hat{f}(\xi)| \leq C_\ell \langle \xi \rangle^{-\ell}$$

for all ℓ so the operator (9.13) is well defined. Moreover, for $\rho \in C_0^\infty(\mathbb{R}^n)$,

$$(A\rho(\hbar D))(x) = \hbar^{-n/2} \int b(x,\xi,\hbar)\rho(\xi) e^{i\frac{x\cdot\xi}{\hbar}} \mathcal{F}_\hbar f(\xi) d\xi \tag{9.15}$$

so

$$A\rho(\hbar D) \in \Psi(\mathbb{R}^n).$$

□

For the operator A to have other desirable properties, one has to impose some additional conditions on b. For instance, one such desirable property is that the range of A be contained in $C^\infty(\mathbb{R}^n)$. We will show that a sufficient condition for this to be the case is a mild strengthening of (9.15):

Theorem 9.3.3. *Suppose that for every multi-index α there is a $C = C(\alpha)$ and an $N = N(\alpha)$ such that*

$$|D_x^\alpha b(x,\xi,\hbar)| \leq C \langle \xi \rangle^N. \tag{9.16}$$

Then A maps $C_0^\infty(\mathbb{R}^n)$ into $C^\infty(\mathbb{R}^n)$.

Proof. By (9.13)

$$(D_x^\alpha A f)(x) = \int (D_x + \xi)^\alpha b(x, \hbar\xi, \hbar) e^{ix\cdot\xi} \hat{f}(\xi) d\xi$$

and by (9.16) the integral on the right is well defined. □

Another desirable property is "pseudolocality". Recall that if X is a manifold, and $A : C_0^\infty(X) \to C^\infty(X)$ is a linear operator, then A is said to be pseudolocal if, for every pair of functions $\rho_1,\ \rho_2 \in C_0^\infty(X)$ with non-overlapping supports, the operator

$$C_0^\infty(X) \ni f \mapsto \rho_2 A \rho_1 f$$

is a smoothing operator, i.e. an operator of the form

$$f \mapsto \int \rho_2(x) K(x,y) \rho_1(y) dy$$

where K is a C^∞ function on the set $x \neq y$. We claim that we can achieve this property for the operator (9.10) by imposing a condition analogous to (9.16) on the ξ derivatives of $b(x,\xi,\hbar)$:

Theorem 9.3.4. *Suppose that for all multi-indices α there is a constant $C = C(\alpha)$ such that*

$$\left|D_\xi^\alpha b(x,\xi,\hbar)\right| \le C\langle\xi\rangle^{m-|\alpha|}. \tag{9.17}$$

Then the operator (9.10) is pseudolocal.

Proof. For k large,

$$Af = A_{\text{new}}\langle D\rangle^{2k} f$$

where

$$(A_{\text{new}}f)(x) = \int b(x,\xi,\hbar)\langle\xi\rangle^{-2k} e^{ix\cdot\xi}\hat{f}(\xi)d\xi.$$

Since $\langle D\rangle^{2k}$ is pseudolocal, A will be pseudolocal if A_{new} is pseudolocal. Thus replacing A by A_{new}, we may assume that the m in (9.17) is less than $-n-\ell$ for ℓ large. In other words, we can assume that A is an integral operator with Schwartz kernel

$$K_A = \int b(x,\hbar\xi,\hbar)e^{i(x-y)\xi}d\xi$$

in $C^\ell(\mathbb{R}^n \times \mathbb{R}^n)$. Now for any multi-index α we have

$$(y-x)^\alpha \int b(x,\hbar\xi,\hbar)e^{i(x-y)\xi}d\xi = \int b(x,\hbar\xi,\hbar)(-D_\xi)^\alpha e^{i(x-y)\xi}d\xi$$

$$= \int D_\xi^\alpha b(x,\hbar\xi,\hbar)e^{i(x-y)\xi}d\xi$$

by integration by parts. Thus, by (9.17)

$$(y-x)^\alpha K(x,y,\hbar) \in C^{\ell+|\alpha|}(\mathbb{R}^n \times \mathbb{R}^n).$$

Since $|\alpha|$ can be chosen arbitrarily large, this shows that K_A is C^∞ on the set $x \neq y$, and hence that A is pseudolocal. $\square$

The inequalities (9.16) and (9.17) are the motivation for the following definition:

Definition 9.3.1. *A function $b = b(x,\xi,\hbar)$ is said to be in the symbol class S^m if, for every pair of multi-indices α and β, and for every compact subset $W \subset \mathbb{R}^n$, there is a constant $C_{W,\alpha,\beta}$ such that*

$$\left|D_x^\alpha D_\xi^\beta b(x,\xi,\hbar)\right| \le C_{W,\alpha,\beta}\langle\xi\rangle^{m-|\beta|}$$

for all $x \in W$,

From the previous two theorems we conclude that an operator A given by

$$(Af)(x) = \int b(x,\hbar\xi,\hbar)e^{ix\cdot\xi}\hat{f}(\xi)d\xi$$

with

$$b \in S^m$$

maps $C_0^\infty(\mathbb{R}^n)$ into $C^\infty(\mathbb{R}^n)$ and is both pseudolocal and microlocal.

To relate the results of this section to the theorem we proved in the preceding section, we note that a particularly nice subset of S^m is the set of **polyhomogeneous symbols of degree** m given by the following definition:

Definition 9.3.2. *A symbol $b(x,\xi,\hbar)$ is a polyhomogeneous symbol of degree m if there exist, for $i = m, m-1, \dots$ homogeneous functions of degree i in ξ:*

$$b_i(x,\xi,\hbar) \in C^\infty(\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}) \times \mathbb{R})$$

such that for $\rho \in C_0^\infty(\mathbb{R}^n)$ and $r < m$

$$b - (1-\rho)\sum_r^m b_i \in S^{m-r-1}.$$

Operators with symbols of this type we will call semi-classical polyhomogeneous pseudo-differential operators, or SCPHΨDO's for short.

A nice property of these operators is that they can be completely characterized by microlocal properties: More explicitly, let X be a manifold and

$$A_\hbar : C_0^\infty(X) \to C^\infty(X)$$

be a family of polyhomogeneous operators in the sense of §9.2 which depend smoothly on $\hbar$. By this we mean that its restriction to a coordinate patch has a polyhomogeneous symbol (in the sense of §9.2):

$$a(x,\xi,\hbar) \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}).$$

Then $A_\hbar$, viewed as a semi-classical object, i.e.as an operator depending on $\hbar$, is a SCPHΨDO if

$$a(x,\xi,\hbar) = b(x,\hbar\xi,\hbar)$$

and, as we proved above, this is the case if and only if $AP \in \Psi_0(X)$ for $P \in \Psi_0(X)$.

9.4 The symbol calculus.

The "semi-classical pseudo-differential operators with compact microsupport" that we discussed in §8.7 were integral operators

$$(Af)(x) = \int K_A(x,y,\hbar)f(y)dy$$

with kernel of the form

$$K_A(x,y,\hbar) = \hbar^{-n}\int a(x,y,\xi,\hbar)e^{i\frac{(x-y)\cdot\xi}{\hbar}}d\xi.$$

In particular, the symbol, $a(x, y, \xi, \hbar)$ of A was allowed to be a function of both the variable x *and* the variable y. We will show that the same is true of the semi-classical pseudo-differential operators that we introduced in Section 9.3.

We begin by enlarging the class of symbols that we introduced in Section 9.3:

Definition 9.4.1. *A function*

$$a = a(x, y, \xi, \hbar) \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R})$$

is said to be in the **symbol class** S^m *if for all multi-indices* α, β, γ *and all compact subsets* W *of* $\mathbb{R}^n \times \mathbb{R}^n$ *there are constants* $C_{\alpha,\beta,\gamma,W}$ *such that*

$$\left| D_x^\alpha D_y^\gamma D_\xi^\beta a(x, y, \xi, \hbar) \right| \leq C_{\alpha,\beta,\gamma,W} \langle \xi \rangle^{m-|\beta|} \quad \forall\ (x, y) \in W. \tag{9.18}$$

We will show below that operators with symbols of this type are essentially the same operators that we introduced in Section 9.3. For the moment, let us assume that $m < -\ell - n$ with $\ell \gg 0$. Let A be the operator with Schwartz kernel

$$K_A(x, y, \hbar) = \int a(x, y, \hbar\xi, \hbar) e^{i(x-y)\cdot\xi} d\xi. \tag{9.19}$$

From the above estimate we see that

$$\left| D_x^\alpha D_y^\beta K_A(x, y, \hbar) \right| \leq C_W \int \langle \xi \rangle^{m+\ell} d\xi$$

for $|\alpha| + |\beta| \leq \ell$. Since $m + \ell < -n$ the integral on the right converges, and hence $K_A \in C^\ell(\mathbb{R}^n \times \mathbb{R}^n)$.

A similar argument shows that A is pseudolocal: For $1 \leq r \leq n$

$$\begin{aligned}
(x_r - y_r)^N & \int a(x, y, \hbar\xi, \hbar) e^{i(x-y)\cdot\xi} d\xi \\
&= \int a(x, y, \hbar\xi, \hbar) \left(-i \frac{\partial}{\partial \xi_r} \right)^N e^{i(x-y)\cdot\xi} d\xi \\
&= \int \left(i \frac{\partial}{\partial \xi_r} \right)^N a(x, y, \hbar\xi, \hbar) e^{i(x-y)\cdot\xi} d\xi.
\end{aligned}$$

So by (9.18) and the preceding argument,

$$(x_r - y_r)^N K_A(x, y, \hbar) \in C^{\ell+N}(\mathbb{R}^n \times \mathbb{R}^n).$$

In other words, $K_A \in C^{\ell+N}(\mathbb{R}^n \times \mathbb{R}^n)$ for all N on the set $x \neq y$.

Let us now prove that A is a semiclassical pseudo-differential operator with symbol of type S^m in the sense of Section 9.3: Replace $a(x, y, \xi, \hbar)$ by its Taylor expansion in y about the point x:

$$a(x, y, \xi, \hbar) \sim \sum_\alpha \frac{(y - x)^\alpha}{\alpha!} \left(\frac{\partial}{\partial y} \right)^\alpha a(x, x, \xi, \hbar).$$

Plugging this into the right hand side of (9.19) one gets an asymptotic expansion

$$K_A \sim \sum_\alpha K_\alpha(x,y,\hbar) \tag{9.20}$$

where

$$\begin{aligned} K_\alpha(x,y,\hbar) &= \int \left(\frac{\partial}{\partial y}\right)^\alpha a(x,x,\hbar\xi,\hbar)\frac{(y-x)^\alpha}{\alpha!}e^{i(x-y)\cdot\xi}d\xi \\ &= \frac{1}{\alpha!}\int \left(\frac{\partial}{\partial y}\right)^\alpha a(x,x,\hbar\xi,\hbar)\,(-D_\xi)^\alpha\, e^{i(x-y)\cdot\xi}d\xi \\ &= \frac{\hbar^{|\alpha|}}{\alpha!}\int \left(\frac{\partial}{\partial y}\right)^\alpha D_\xi^\alpha a(x,x,\hbar\xi,\hbar)e^{i(x-y)\cdot\xi}d\xi. \end{aligned}$$

Thus the operator with Schwartz kernel K_α is a semi-classical pseudo-differential operator A_α with symbol

$$a_\alpha = \frac{\hbar^{|\alpha|}}{\alpha!}\left(\frac{\partial}{\partial y}\right)^\alpha D_\xi^\alpha a(x,x,\xi,\hbar).$$

Furthermore,

$$a = a^\sharp + r$$

where $a^\sharp$ is in S^m and has an asymptotic expansion

$$a^\sharp(x,\xi,\hbar) = \sum_\alpha \frac{\hbar^{|\alpha|}}{\alpha!}\left(\frac{\partial}{\partial y}\right)^\alpha D_\xi^\alpha a(x,x,\xi,\hbar) \tag{9.21}$$

and $r(x,y,\xi,\hbar)$ is in $S^{-\infty}$ and vanishes to infinite order at $\hbar = 0$.

Letting $A^\sharp$ and R be the operators with these symbols we conclude that

$$A = A^\sharp + R \tag{9.22}$$

where $A^\sharp \in \Psi^m$ and the Schwartz kernel

$$\int r(x,y,\hbar\xi,\hbar)e^{i(x-y)\cdot\xi}d\xi$$

of R is a C^∞ function which vanishes to infinite order at $\hbar = 0$.

One immediate application of this result is

Theorem 9.4.1. *If A is a semi-classical pseudo-differential operator with symbol in S^m then its transpose is a semi-classical pseudo-differential operator with symbol in S^m.*

Proof. If the Schwartz kernel of A is given by (9.19) then the Schwartz kernel $\overline{K_A(y,x)}$ of $A^\dagger$ is given by

$$\int \overline{a(x,y,\hbar\xi,\hbar)}e^{i(x-y)\cdot\xi}d\xi.$$

□

In particular, one can formulate the notion of microlocality in terms of "multiplication on the left" by microlocal cut-offs:

Proposition 9.4.1. *For every semi-classical pseudo-differential operator P of compact microsupport the operator PA is a semi-classical pseudo-differential operator of compact microsupport.*

We will let $\Psi^k(S^m)$ denote the class of elements of Ψ^k whose symbols belong to S^m. If we do not want to specify k we will simply write $\Psi(S^m)$.

9.4.1 Composition.

We will next show that the composition of two pseudo-differential operators $A \in \Psi(S^{m_1})$ and $B \in \Psi(S^{m_2})$ with $m_i \ll n,\ i = 1, 2$ is in $\Psi(S^{m_1+m_2})$.

Indeed, by what we just proved, we may assume that A has a symbol of the form $a(x, \xi, \hbar)$ and that B has a symbol of the form $b(y, \xi, \hbar)$. This implies that the Schwartz kernel of A is of the form

$$K_A(x, y, \hbar) = K(x, x - y, \hbar)$$

where

$$K(x, w, \xi) = \int a(x, \hbar\xi, \hbar)e^{iw\cdot\xi}d\xi.$$

By the Fourier inversion formula

$$a(x, \hbar\xi, \hbar) = (2\pi)^n \hat{K}(x, w, \hbar) \tag{9.23}$$

where $\hat{K}$ is the Fourier transform of K with respect to w.

By the identities above, the Schwartz kernel of AB is given by

$$\int K(x, x - z, \hbar)e^{i(z-y)\cdot\xi}b(y, \hbar\xi, \hbar)dzd\xi.$$

Making the change of variables $z = w + x$ this becomes

$$\int K(x, -w, \hbar)e^{iw\cdot\xi}e^{(x-y)\cdot\xi}b(y, \xi, \hbar)dwd\xi.$$

By the Fourier inversion formula and (9.21) the inner integral is $a(x, \hbar\xi, \hbar)$ so the above expression for the Schwartz kernel of AB becomes

$$\int a(x, \hbar\xi, \hbar)b(y, \hbar\xi, \hbar)e^{i(x-y)\cdot\xi}d\xi.$$

We have proved

Theorem 9.4.2. *Under the above hypotheses, $AB \in \Psi(S^{m_1+m_2})$ and its symbol is*

$$a(x, \xi, \hbar)b(y, \xi, \hbar).$$

9.4.2 Behavior under coordinate change.

The operators we considered in §8.6 were the restrictions to open sets of $\mathbb{R}^n$ of objects which were well defined on manifolds. To prove the same for the operators we are studying in this chapter, we must prove "invariance under coordinate change", and this we can do by exactly the same argument as in §9.1. More explicitly let U and V be open subsets of $\mathbb{R}^n$ and $f : V \to U$ a diffeomorphism. Let $a(x, y, \xi\hbar)$ be a symbol in S^m with $m \ll -n$ and with support in in the set $\{(x, y) \subset U \times U\}$ and let A be the operator with a as symbol. By the argument in Section 9.1, $f^*A(f^{-1})^*$ is a semi-classical pseudo-differential operator with symbol

$$a_f = a(f(x), f(y), (F^\dagger)^{-1}\xi, \hbar)\left|\det f_y F^{-1}(x, y)\right|$$

and, by inspection $a_f \in S^m$.

Our next task is to get rid of assumption, $a \in S^m$, $m < -n - \ell$. One way to do this is by distributional techniques, but, in the spirit of this book we will do this by a more hands-on approach. For $a \in S^m$, $m < -n$, let

$$Ta = a - \frac{\langle D_x + \xi\rangle^{2N} a}{\langle\xi\rangle^{2N}}\,. \tag{9.24}$$

Then Ta is in S^{m-1} and

$$\begin{aligned} &\langle\hbar D_x\rangle^{2N} \int \frac{a(x, y, \hbar\xi, \hbar)}{\langle\hbar\xi\rangle^{2N}} e^{i(x-y)\cdot\xi}\, d\xi \\ &= \int (a - Ta)(x, y, \hbar\xi, \hbar) e^{i(x-y)\cdot\xi}\, d\xi\,. \end{aligned} \tag{9.25}$$

Thus setting

$$b = \frac{a + Ta + \cdots + T^{2N-1}a}{\langle\xi\rangle^{2N}}$$

we have by (9.25)

$$\begin{aligned} &\langle\hbar D_x\rangle^{2N} \int b(x, y, \hbar\xi, \hbar) e^{i(x-y)\cdot\xi}\, ds \\ &= \int (a - T^{2N}a)(x, y, \hbar\xi, \hbar)\, d\xi\,. \end{aligned}$$

Thus the operators, B and C with symbols, b and $c = T^{2N}a$, are in Ψ^{m-2N} and

$$A = \langle\hbar D_x\rangle^{2N} B + C\,. \tag{9.26}$$

Using this formula we can make sense of A for a in S^m when m is large, namely we can choose N with $m - 2N \ll -n$ and then define A by (9.26). Notice also that by taking transposes in (9.26) we get the transpose identity: $A^t = B^t\langle D_x\rangle^{2N} + C^t$. Moreover by Theorem 5 we can replace A, B and C by their transposes in this identity, and by doing so, we get a "left handed" version of (9.26)

$$A = B\langle \hbar D_x \rangle^{2N} + C \tag{9.27}$$

with B and C in Ψ^{m-2N}. One application of these formulas is making sense of the product, $A_1 A_2$ where A_i is in Ψ^{m_i} and the m_i's are large. Letting

$$A_1 = \langle \hbar D_x \rangle^N B_1 + C_1$$

and

$$A_2 = B_2 \langle \hbar D_x \rangle^{2N} + C_2$$

the product becomes

$$\langle \hbar D_x^{2N} \rangle B_1 B_2 \langle \hbar D_x \rangle^{2N} + \langle \hbar D_x \rangle^N B_1 C_2 + C_1 B_2 \langle \hbar D_x \rangle^N + C_1 C_2$$

and for N large $B_1 B_2$, $B_1 C_2$ and $C_1 C_2$ are in Ψ^k for $k = m_1 + m_2 - 4N \ll -n$. We observed in the preceding paragraph that ΨDO's with symbols in S^m, $m \ll -n$ are invariant under coordinate change and hence are intrinsically defined on manifolds. Combining this with (9.26) and (9.27) we can remove the restriction $m \ll -n$. Indeed, these equations imply

Theorem 9.4.3. *The algebra of ΨDO's with symbol in S^m, $-\infty \leq m < \infty$ is invariant under coordinate change and hence intrinsically defined on manifolds.*

The same argument also shows that the principal symbol, $a(x, x, \xi, 0)$, of $a(x, y, \xi, \hbar)$ is intrinsically defined as a function on T^*U. Indeed, for $m < -n$ one can prove this exactly as we did in Section 9.1, and for first order differential operators (i.e.vector fields) the proof is more or less trivial. Hence by (9.26) and the composition formula for symbols described in Theorem 9.4.2, it is easy to remove the restriction $m < -n$.

Our goal in the last part of this chapter will be to explore in more detail symbolic properties of the operators above. In particular three issues we'll be concerned with are:

1. *Canonical forms for symbols.* We've seen above that every $A \in \Psi^m$ has a unique symbol of the form, $a(x, \xi, \hbar)$, i.e., a symbol not depending on y. These symbols we will call *left Kohn–Nirenberg symbols* (or left KN symbols for short). Similarly by taking transposes we get for $A = (A^t)^t$ a unique *right Kohn–Nirenberg symbol* of the form, $a(y, \xi, h)$. An interesting compromise between these extremes are *Weyl* symbols: symbols of the form, $a(\frac{x+y}{2}, \xi, \hbar)$ and, interpolating between these three classes of symbols, *generalized Weyl symbols*: symbols of the form $a((1-t)x + ty, \xi, \hbar)$, $0 \leq t \leq 1$.

2. *Compositions and transposes.* Let Ψ be the union, $\bigcup \Psi^m$. We have shown that this space of operators is closed under composition and transposes. We would like, however, to have a "symbolic calculus" for these operations, (e.g.) a composition law for symbols analogous to (9.3).

3. *Converting symbols of one type into symbols of another type.* From (9.21) one gets formulas relating the various canonical forms in item 1, e.g. formulas for expressing left KN symbols in terms of right KN symbols or expressing right KN symbols in terms of Weyl symbols. One of our goals will be to describe these "conversion" laws in more detail.

The key ingredient in these computations will be

Theorem 9.4.4. *Two symbols $a_1(x,y,\xi,\hbar)$ and $a_2(x,y,\xi,\hbar)$ in S^m define the same pseudo-differential operator A if*

$$a_1 - a_2 = e^{-i\frac{(x-y)\cdot\xi}{\hbar}} \sum_{j=1}^{n} \frac{\partial}{\partial \xi_j} \left(e^{i\frac{(x-y)\cdot\xi}{\hbar}} c_j \right) \tag{9.28}$$

with

$$c_j \in S^{m+1}.$$

Proof. Let us first prove this result under the assumption that $m < -n-\ell$ with $\ell \gg 0$. Let $b = a_1 - a_2$. The Schwartz kernel of the operator defined by b is

$$\hbar^{-n} \int b(x,y,\xi,\hbar) e^{i\frac{(x-y)\cdot\xi}{\hbar}} d\xi$$

and this vanishes if the integrand is a "divergence", as in the right hand side of (9.28).

To prove this theorem in general, notice that

$$\hbar D_{x_i} + \xi_i = e^{-i\frac{(x-y)\cdot\xi}{\hbar}} \circ (\hbar D_{x_i}) \circ e^{i\frac{(x-y)\cdot\xi}{\hbar}}.$$

So if we apply the operator (9.25) to a divergence

$$e^{-i\frac{(x-y)\cdot\xi}{\hbar}} \sum \frac{\partial}{\partial \xi_i} \left(e^{i\frac{(x-y)\cdot\xi}{\hbar}} c_i \right)$$

we again get such a divergence. □

In particular, for $a \in S^m$, the symbols

$$a(x,y,\xi,\hbar)(x-y)^\alpha$$

and

$$(-\hbar D)^\alpha a(x,y,\xi,\hbar)$$

define the same operator. (We already made use of this observation in the course of proving (9.22) for symbols $a \in S^m$ with $m \ll 0$.)

In the next section we will address the issues raised in items 1-4 above by elevating (9.28) to an equivalence relation, and deriving identities between symbols of varying types by purely formal manipulation.

9.5 The formal theory of symbols.

We say that two symbols $a_1(x,y,\xi,\hbar)$ and $a_2(x,y,\xi,\hbar)$ in S^m are **equivalent** if their associated ΨDOs, A_1 and A_2, differ by a ΨDO, B, with symbol $b(x,y,\xi,\hbar)\in\hbar^\infty S^{-\infty}$.

Starting with the relation

$$a(x,y,\xi,\hbar)(x-y)^\alpha \sim (-\hbar D_\xi)^\alpha(a(x,y,\xi,\hbar))$$

we will generalize the formula (9.22) to $a\in S^m$ with m arbitrary. Namely,

$$\begin{aligned}
a(x,y,\xi,\hbar) &\sim \sum \frac{1}{\alpha!}\left(\frac{\partial}{\partial y}\right)^\alpha a(x,y,\xi,\hbar)\Big|_{y=x}(y-x)^\alpha\\
&\sim \sum \frac{1}{\alpha!}(\hbar D_\xi)^\alpha\left(\frac{\partial}{\partial y}\right)^\alpha a(x,y,\xi,\hbar)\Big|_{y=x}\\
&\sim a_R(x,\xi,\hbar)
\end{aligned}$$

where

$$a_R(x,\xi,\hbar)\sim \exp\left(\hbar\frac{\partial}{\partial y}D_\xi\right)a(x,y,\xi,\hbar)\Big|_{y=x} \tag{9.29}$$

is a right Kohn-Nirenberg symbol (i.e., depending only on x).

Notice that if $a_L(y,\xi,\hbar)$ is a *left* Kohn-Nirenberg symbol (depending only on y) then

$$a_R(x,\xi,\hbar)\sim \exp\left(\hbar\frac{\partial}{\partial x}D_\xi\right)a_L(x,\xi,\hbar) \tag{9.30}$$

and hence

$$a_L(y,\xi,\hbar)\sim \exp\left(-\hbar\frac{\partial}{\partial y}D_\xi\right)a_R(y,\xi,\hbar). \tag{9.31}$$

From now on, to avoid confusing x's and y's, we will replace the x and y by a neutral variable z, and express this relation between right and left symbols as

$$a_R(z,\xi,\hbar)\sim \exp\left(\hbar\frac{\partial}{\partial z}D_\xi\right)a_L(z,\xi,\hbar). \tag{9.32}$$

We can generalize right and left symbols by substituting $(1-t)x+ty$ for z in $a(z,\xi,\hbar)$.

This gives the generalized symbol

$$a_{W,t}(z)=a\left((1-t)x+ty,\xi,\hbar\right).$$

This can be converted by (9.29) into a right Kohn-Nirenberg symbol

$$\begin{aligned}
a_R(x,\xi,\hbar) &\sim \exp\left(\hbar\frac{\partial}{\partial y}D_\xi\right)a((1-t)x+ty,\xi,\hbar)\Big|_{y=x}\\
&= \exp\left(t\hbar\frac{\partial}{\partial x}D_\xi\right)a(x,\xi,\hbar).
\end{aligned}$$

Reverting to our neutral variable z this becomes

$$a_R(z,\xi,\hbar) = \exp\left(t\hbar\frac{\partial}{\partial z}D_\xi\right)a_{W,t}(z,\xi,\hbar) \tag{9.33}$$

and

$$a_{W,t}(z,\xi,\hbar) = \exp\left(-t\hbar\frac{\partial}{\partial z}D_\xi\right)a_R(z,\xi,\hbar) \tag{9.34}$$

9.5.1 Multiplication properties of symbols.

We start with Theorem 9.4.2: If A is a ΨDO with right Kohn-Nirenberg symbol $a(x,\xi,\hbar)$ and B is a ΨDO with left Kohn-Nirenberg symbol $b(y,\xi,\hbar)$ then the symbol of AB is $a(x,\xi,\hbar)b(y,\xi,\hbar)$. (We proved this in Section 9.4 for symbols of large negative degree.) But by (9.26) and(9.27) this extends to symbols of arbitrary degree.)

Let us now convert this, using (9.29) into a right Kohn-Nirenberg symbol: We obtain

$$\begin{aligned}
&\sum\frac{1}{\alpha!}\hbar D_\xi^\alpha\left(a(x,\xi,\hbar)\partial_y^\alpha b(y,\xi,\hbar)\right)\Big|_{y=x}\\
=\;&\sum\frac{1}{\alpha!}\sum_{\beta+\gamma=\alpha}\frac{\alpha!}{\beta!\gamma!}\left(\hbar D_\xi\right)^\beta a(x,\xi,\hbar)\left(\hbar D_\xi\right)^\gamma\partial_x^\alpha b(x,\xi,\hbar)\\
=\;&\sum_{\beta,\gamma}\frac{1}{\beta!}\left(\hbar D_\xi\right)^\beta a(x,\xi,\hbar)\frac{1}{\gamma!}\left(\hbar D_\xi\right)^\gamma\partial_x^\gamma\left(\partial_x^\beta b(x,\xi,\hbar)\right)\\
=\;&\sum_{\beta}\frac{1}{\beta!}\left(\hbar D_\xi\right)^\beta a(x,\xi,\hbar)\partial_x^\beta\exp\left(\hbar D_\xi\partial_x\right)b(x,\xi,\hbar).
\end{aligned}$$

If

$$b(y,\xi,\hbar) = b_L(y,\xi,\hbar) = \exp\left(-\hbar D_\xi\partial_x\right)b_R(x,\xi,\hbar)|_{x=y}$$

this formula simplifies to

$$\sum_\beta\frac{1}{\beta!}\left(\hbar D_\xi\right)^\beta a(x,\xi,\hbar)\partial_x^\beta b_R(x,\xi,\hbar). \tag{9.35}$$

In other words, let a_R and b_R be two right Kohn-Nirenberg symbols and let A and B be the corresponding ΨDO's. Then up to equivalence, the right Kohn-Nirenberg symbol of AB is given by (9.35). This generalizes a formula that we proved in Chapter 8 for ΨDO's of compact microsupport.

There is a more compact version of (9.35): We can write

$$\sum_\beta\frac{1}{\beta!}\left(\hbar D_{\xi_1}\right)^\beta a_R(z_1,\xi_1,\hbar)\partial_{z_2}^\beta b_R(z_2,\xi_2,\hbar)$$

as

$$\exp\left(\hbar D_{\xi_1}\frac{\partial}{\partial z_2}\right)a_R(z_1,\xi_1,\hbar)b_R(z_2,\xi_2,\hbar).$$

We then get (9.35) by setting $z = z_1 = z_2$ and $\xi = \xi_1 = \xi_2$. In other words, the symbol of AB is given by

$$\exp\left(\hbar D_{\xi_1}\frac{\partial}{\partial z_2}\right) a_R(z_1,\xi_1,\hbar)b_R(z_2,\xi_2,\hbar)\Big|_{z=z_1=z_2,\ \xi=\xi_1=\xi_2}. \tag{9.36}$$

Our next task will be to derive an analogue of this formula for symbols of type (W,t). First we show how a product symbol of the form $a(x,\xi,\hbar)b(y,\xi,\hbar)$ can be converted into such a generalized Weyl symbol: Let

$$z = sx + ty, \quad s = 1 - t$$

so that

$$x = z + t(x-y), \qquad y = z - s(x-y).$$

By Taylor's expansion

$$\begin{aligned} a(x,\xi,\hbar)b(y,\xi,\hbar) &= \sum_{\beta,\gamma}\frac{t^\beta}{\beta!}\partial_z^\beta a(z,\xi,\hbar)\frac{(-s)^\gamma}{\gamma!}\partial_z^\gamma b(z,\xi,\hbar)(x-y)^{\beta+\gamma} \\ &= \sum \frac{1}{\alpha!}\left(t\frac{\partial}{\partial u} - s\frac{\partial}{\partial v}\right)^\alpha a(u,\xi,\hbar)b(v,\xi,\hbar)\Big|_{u=v=z}(x-y)^\alpha \\ &\sim \sum \frac{1}{\alpha!}\left(s\frac{\partial}{\partial v} - t\frac{\partial}{\partial u}\right)^\alpha (\hbar D_\xi)^\alpha\left(a(u,\xi,\hbar)b(v,\xi,\hbar)\right)\Big|_{u=v=z} \end{aligned}$$

We can simplify this further: Replace

$$\frac{1}{\alpha!}\left(s\frac{\partial}{\partial v} - t\frac{\partial}{\partial u}\right)^\alpha (\hbar D_\xi)^\alpha\left(a(u,\xi,\hbar)b(v,\xi,\hbar)\right)$$

by the sum

$$\sum_{\mu+\nu=\alpha}(\hbar D_\xi)^\mu\left(s\frac{\partial}{\partial v} - t\frac{\partial}{\partial u}\right)^\mu \frac{1}{\nu!}(\hbar D_\eta)^\nu\left(s\frac{\partial}{\partial v} - t\frac{\partial}{\partial u}\right)^\nu a(u,\xi,\hbar)b(v,\eta,\hbar)$$

evaluated at $\xi = \eta$. Summing this over α then yields

$$\exp\left(\hbar D_\xi\left(s\frac{\partial}{\partial y} - t\frac{\partial}{\partial x}\right) + \hbar D_\eta\left(s\frac{\partial}{\partial y} - t\frac{\partial}{\partial x}\right)\right) a(x,\xi,\hbar)b(y,\eta,\hbar)\Big|_{x=y=z,\ \xi=\eta}. \tag{9.37}$$

Now let $a(z,\xi,\hbar)$ and $b(z,\xi,\hbar)$ be symbols of type (W,t), and let

$$\begin{aligned} a_1 &= \exp\left(t\frac{\partial}{\partial x}\hbar D_\xi\right) a(x,\xi,\hbar) \\ b_1 &= \exp\left(-s\frac{\partial}{\partial y}\hbar D_\eta\right) b(y,\eta,\hbar) \end{aligned}$$

be the corresponding right and left Kohn-Nirenberg symbols so that their symbolic product is $a_1(x,\xi,\hbar)b_1(x,\xi,\hbar)$. We plug this into (9.37). The "exp" part of the formula becomes

$$\exp\left(\hbar D_\xi\left(s\frac{\partial}{\partial y}-t\frac{\partial}{\partial x}\right)+t\hbar D_\xi\frac{\partial}{\partial x}+\hbar D_\eta\left(s\frac{\partial}{\partial y}-t\frac{\partial}{\partial x}\right)-s\hbar D_\eta\frac{\partial}{\partial y}\right)$$
$$=\exp\hbar\left(sD_\xi\frac{\partial}{\partial y}-tD_\eta\frac{\partial}{\partial x}\right).$$

So we have proved:

Theorem 9.5.1. *Let $a(z,\xi,\hbar)$ and $b(z,\xi,\hbar)$ be symbols of type (W,t). Their symbolic product is*

$$\exp\hbar\left(sD_\xi\frac{\partial}{\partial y}-tD_\eta\frac{\partial}{\partial x}\right)a(x,\xi,\hbar)b(y,\eta,\hbar) \tag{9.38}$$

evaluated at $\xi=\eta$ and $x=y=z$.

9.6 The Weyl calculus.

In this section we discuss special properties of symbols of type $(W,\frac{1}{2})$ which we shall simply call Weyl symbols.

For the case $s=t=\frac{1}{2}$ formula (9.38) takes the more symmetric form

$$\exp\frac{\hbar}{2}\left(D_\xi\frac{\partial}{\partial y}-D_\eta\frac{\partial}{\partial x}\right)a(x,\xi,\hbar)b(y,\eta,\hbar). \tag{9.39}$$

Here is another important property of Weyl symbols: The ΨDO A associated to a Weyl symbol $a(z,\xi,\hbar)$ has Schwartz kernel

$$K_A(x,y)=\hbar^{-n}\int a\left(\frac{x+y}{2},\xi,\hbar\right)e^{i\frac{(x-y)\cdot\xi}{2}}d\xi.$$

See the discussion in Chapter 16 of kernels of this type from the point of view of physics and of group theory.

The Schwartz kernel of the formal adjoint of A is the operator with Schwartz kernel $\overline{K_A(y,x)}$ which is

$$\hbar^{-n}\int \overline{a}\left(\frac{x+y}{2},\xi,\hbar\right)e^{i\frac{(x-y)\cdot\xi}{2}}d\xi.$$

So if a is real valued, A is formally self-adjoint.

An important consequence of this is the following: Let a and b be real Weyl symbols and A and B their corresponding ΨDO's which are therefore formally self-adjoint. Consider their commutator: $[A,B]=AB-BA$. The adjoint of this commutator is $BA-AB=-[A,B]$ hence the symbol of $[A,B]$ is purely imaginary. This means that in the symbolic expansion for this commutator all *even* powers of h have to be zero. This can also be seen directly from (9.39) by interchanging a and b and subtracting. This has the consequence that computations with Weyl symbols are usually "twice as fast" as the corresponding computations with Kohn–Nirenberg symbols.

9.7 The structure of $I(X,\Lambda)$ as a module over the ring of semi-classical pseudo-differential operators.

Let X be a manifold and Λ a Lagrangian submanifold of T^*X. In ¶8.9.1 we pointed out that the space of oscillatory $\frac{1}{2}$-densites $I(X,\Lambda)$ is a module over the ring of "semi-classical pseudo-differential operators" where, in Chaper 8, "semi-classical" meant "semi-classical with compact micro-support". We also pointed out in ¶8.10 that $I(X,\Lambda)$ is a module over the ring of differential operators. Both these rings sit inside the ring $\tilde{\Psi}$ of ΨDO's with symbols in S^m, $-\infty \leq m < \infty$. It is easy to extend the results of ¶9.8-8.10 to this more general setting:

Theorem 9.7.1. *Let $P \in \tilde{\Psi}^k(X)$ be a semi-classical ΨDO with a symbol of type S^m. If $\gamma \in I^\ell(X,\Lambda)$ then $P\gamma \in I^{k+\ell}(X,\Lambda)$. Moreover if γ is given locally on an open set $U \subset \mathbb{R}^n$ by the expression (8.49):*

$$\gamma = \hbar^{\ell-\frac{n}{2}} \int b(\xi,\hbar) e^{i\frac{x\cdot\xi-\phi(\xi)}{\hbar}} d\xi$$

where $x\cdot\xi - \phi(\xi)$ is a generating function for Λ with respect to the cotangent fibration $T^*U \ni (x,\xi) \mapsto x \in U$ then

$$P\gamma = \hbar^{k+\ell-\frac{n}{2}} \int a(x,\xi,\hbar) b(\xi,\hbar) e^{i\frac{x\cdot\xi-\phi(\xi)}{\hbar}} d\xi \tag{9.40}$$

where $a(x,\xi,\hbar)$ is the right Kohn-Nirenberg symbol of P.

Proof. If $b(\xi)$ is supported on the set $\|\xi\| \leq N$ and ρ is a compactly supported C^∞ function of ξ which is identically one on this set, then

$$\begin{aligned} \rho(\hbar D)\gamma &= \hbar^{\ell-\frac{n}{2}} \int \rho(\xi) b(\xi,\hbar) e^{i\frac{x\cdot\xi-\phi(\xi)}{\hbar}} d\xi \\ &= \hbar^{\ell-\frac{n}{2}} \int b(\xi,\hbar) e^{i\frac{x\cdot\xi-\phi(\xi)}{\hbar}} d\xi \end{aligned}$$

and hence

$$P\gamma = P\rho(\hbar D)\gamma.$$

so in view of (8.49) and (8.50) the right hand side is given by (9.40). □

Chapter 10

Trace invariants.

10.1 Functions of pseudo-differential operators.

Let $P: C_0^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ be a semi-classical pseudo-differential operator of order zero with right Kohn-Nirenberg symbol $p(x,\xi,\hbar) \in S^m$ with leading symbol $p_0(x,\xi) = p(x,\xi,0)$ and Weyl symbol

$$p^W(x,\xi,\hbar) = \exp(-\frac{\hbar}{2}D_\xi\partial_x)p(x,\xi,\hbar).$$

We showed in ¶9.5 that if p^W is real valued then P is formally self-adjoint. But much more is true: under the above assumption, for sufficiently small values of $\hbar$, the operators $P = P_\hbar$ can be extended to a self adjoint operator with a dense domain $D(P) \subset L^2(\mathbb{R}^n)$. See Chapter 13 for a sketch of how this goes. Hence, by the spectral theorem for self-adjoint operators, one can define the operator $f(P)$ for any bounded continuous or (even measurable) function f on $\mathbb{R}$. (See Chapter 13.)

Moreover, if $f \in C_0^\infty(\mathbb{R})$ then $f(P)$ is itself a semi-classical pseudo-differential operator. A nice exposition of this result based on ideas of Dynkin, Helffer and Sjöstrand can be found in the book [DiSj], Chapter 8. We will give a brief account of the exposition in the paragraphs below. A somewhat more extended description will be given in Chapter 13.

Given $f \in C_0^\infty(\mathbb{R})$, an **almost analytic extension** of f is a function $\tilde{f} \in C_0^\infty(\mathbb{C})$ with the property that

$$\left|\frac{\partial \tilde{f}}{\partial \bar{z}}(x+iy)\right| \leq C_N|y|^N$$

for all $N \in \mathbb{N}$. It is easy to show that almost analytic extensions exist. See , for example [DiSj] or [Davies] - or Chapter 13.

Here is a variant of Cauchy's integral theorem valid for a smooth function g

of compact support in the plane:

$$\frac{1}{\pi}\int_{\mathbb{C}}\frac{\partial g}{\partial \overline{z}}\cdot\frac{1}{z-w}dxdy=-g(w). \tag{10.1}$$

Proof. The integral on the left is the limit of the integral over $\mathbb{C}\setminus D_\delta$ where D_δ is a disk of radius δ centered at w. Since g has compact support, and since

$$\frac{\partial}{\partial\overline{z}}\left(\frac{1}{z-w}\right)=0,$$

we may write the integral on the left as

$$-\frac{1}{2\pi i}\int_{\partial D_\delta}\frac{g(z)}{z-w}dz=-\frac{1}{2\pi}\int_0^{2\pi}\frac{g(w+\delta e^{i\theta})}{\delta}\delta d\theta\to -g(w).$$

□

Suppose now that P is a self-adjoint operator on a Hilbert space $\mathfrak{H}$. A standard theorem in Hilbert space theory (see Chapter 13, for example) says that the resolvent $R(z,P)=(zI-P)^{-1}$ exists as a bounded operator for $\operatorname{Im} z\neq 0$ and its norm blows up as $|\operatorname{Im} z|^{-1}$ as $\operatorname{Im} z\to 0$. Hence from (10.1) one is tempted to believe that

$$f(P):=-\frac{1}{\pi}\int_{\mathbb{C}}\frac{\partial\tilde{f}}{\partial\overline{z}}R(z,P)dxdy, \tag{10.2}$$

where $\tilde{f}$ is an (any) almost holomorphic extension of f. Indeed this formula, due to Helffer and Sjóstrand is true. For a proof see [DiSj] or Chapter 13. In fact, Davies [?] gives a beautiful proof of the spectral theorem starting with (10.2) as a putative formula for $f(P)$.

If P is a semi-classical pseudo-differential operator of order zero one can use the Helffer-Sjóstrand formula (10.2) to prove that $f(P)$ is a semi-classical pseudo-differential operator by reducing this assertion to the assertion that $R(z,P)$ is a a semi-classical pseudo-differential operator, a fact which is much easier to prove.

In addition, one gets from (10.2) a formula for the symbol of $f(P)$: Indeed, using the Weyl calculus, one can solve the equation

$$(z-p^W)\sharp q=1+O(\hbar^\infty)$$

and use this to get a symbolic expansion of $R(z,P)=(zI-P)^{-1}$ and then plug this into (10.2) to get a symbolic expansion for $f(P)$. (Again, see Chapter 13 for more details.)

In this chapter we will develop a functional calculus on a much more modest scale.: Let $\rho\in C_0^\infty(\mathbb{R})$. We will make sense of the expression

$$e^{itP}\rho(\hbar D),\quad -\infty<t<\infty \tag{10.3}$$

mod $O(\hbar^\infty)$ as a semi-classical pseudo-differential operator and then define

$$f(P)\rho(\hbar D)$$

mod $O(\hbar^\infty)$ by Fourier inversion:

$$f(P)\rho(\hbar D), = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \hat{f}(t)e^{itP}dt\rho(\hbar D). \tag{10.4}$$

We will then show that weak "ellipticity type" assumptions allow us to remove the $\rho(\hbar D)$ in (10.4) and so define $f(P)$ itself (again only mod $O(\hbar^\infty)$) as a semi-classical pseudo-differential operator.

A somewhat stronger ellipticity hypothesis enables one not only to define $f(P)$ mod $O(\hbar^\infty)$ but also to conclude that it is of trace class mod $O(\hbar^\infty)$. Namely, suppose that for some compact interval $[a,b]$, $p_0^{-1}([a,b])$ is compact. Then the operator P has discrete spectrum on the interval $[a,b]$. In fact,

$$\mathrm{spec}(P)\cap(a,b) = \{\lambda_i(\hbar), i = 1,\dots,N(\hbar)\}$$

where

$$N(\hbar) \sim (2\pi\hbar)^{-n}\mathrm{Vol}\ \{a \le p_0(x,\xi) \le b\}, \tag{10.5}$$

and hence for $f \in C_0^\infty((a,b))$

$$\mathrm{tr}\, f(P) = \sum f(\lambda_i(\hbar)). \tag{10.6}$$

Hence (10.4) will give, in this case, an asymptotic expansion of (10.6) as $\hbar \to 0$. We will sketch a proof of this fact following an argument of Dimassi-Sjöstrand in Chapter 13. The prototye of this theorem is a well known theorem of Friedrichs [Fr] which asserts that that if the potential is non-negative and $\to \infty$ as $x \to \infty$ then the Schrödinger operator has discrete spectrum.

We now give a brief summary of the contents of this chapter:

In Section 10.2 we will prove that the wave equation

$$\frac{1}{\sqrt{-1}}\frac{\partial}{\partial t}U(t) = U(t)$$

with the initial data

$$U(0) = \rho(\hbar D)$$

is solvable mod $O(\hbar^\infty)$ by the symbol calculus techniques we developed in Chaper 9. This will give us via (10.4) a symbolic expansion for $f(P)\rho(\hbar D)$, and, when when we remove the cutoff, a symbolic expansion for $f(P)$ itself. We will then examine the asymptotics of (10.4) and in particular, prove the Weyl law (10.5).

This wave trace approach to the asymptotics of (10.4) has the virtue that it is relatively easy to implement computationally. We will illustrate this by working through the details for a few simple cases like the Schrödinger operator on the real line and the Schrödinger operator on $\mathbb{R}^n$ with radially symmetric electro-magnetic potential.

The results described above involve operators on $\mathbb{R}^n$. But it is easy to modify this approach so that it applies to operators on manifolds. This we will do in

Section ? We will also point out in that section that the theory developed in this chapter is closely related to a branch of spectral theory that is some sixty years old: the heat trace theory developed by Minakshisundaran-Pleijel in the 1950's and since then generalized and applied to numerous problems in analysis and differential geometry.

10.2 The wave operator for semi-classical pseudo-differential operators.

Let $P \in \Psi^0(S^m(\mathbb{R}^n))$ be a zeroth order semi-classical pseudo-differential operator with right Kohn-Nirenberg symbol $p(x,\xi,\hbar)$ and Weyl symbol $p^W(x,\xi,\hbar)$ which we assume to be real as in the preceding section, so that P is formally self-adjoint. Let $\mathrm{p}_0(x,\xi) = p(x,\xi,0)$ be the leading symbol of P. Let

$$\sum_k \mathrm{p}_k(x,\xi)\hbar^k \tag{10.7}$$

be the Taylor expansion of p in $\hbar$ at 0.

Our goal in this section is to find a family $U(t)$ of semi-classical pseudo-differential operators depending differentiabyy on t for $-\infty < t < \infty$ which satisfies the differential equation

$$\frac{1}{i}\frac{\partial}{\partial t}U(t) = PU(t) \tag{10.8}$$

with the initial condition

$$U(0) = \rho(\hbar D). \tag{10.9}$$

In principle we could solve these equations by the transport equation method of Chapter 8. But a more direct and elementary approach is the following:

Let $\mu(x,y,t,\hbar)$ be the (desired) Schwartz kernel of $U(t)$. We wish this to belong to $I^{-n}(X \times X, \Delta_X)$ for each fixed t. So we want μ to have the form

$$\mu(x,y,t,\hbar) = (2\pi\hbar)^{-n}\int a(x,\xi,t,\hbar)e^{i\frac{(x-y)\cdot\xi}{\hbar}}d\xi. \tag{10.10}$$

Our initial condition (10.9) says that

$$a(x,\xi,0,\hbar) = \rho(\xi). \tag{10.11}$$

Set

$$a(x,\xi,t,\hbar) = e^{it\mathrm{p}_0(x,\xi)}b(x,\xi,t,\hbar)\rho(\xi).$$

So (10.11) becomes

$$b(x,\xi,0,\hbar) \equiv 1 \tag{10.12}$$

while (10.8) (for all ρ) yields

$$\frac{1}{i}\frac{\partial}{\partial t}\left(e^{it\mathrm{p}_0(x,\xi)}b(x,\xi,t,\hbar)\right) = \mathrm{p}(x,\xi,\hbar) \star \left(e^{it\mathrm{p}_0(x,\xi)}b(x,\xi,t,\hbar)\right). \tag{10.13}$$

We can expand (10.13) out as

$$e^{it\mathrm{p}}\left(\frac{1}{i}\frac{\partial b}{\partial t}+\mathrm{p}_0 b\right)=\sum_\alpha \frac{\hbar^\alpha}{\alpha!}D_\xi^\alpha \mathrm{p}\partial_x^\alpha\left(e^{it\mathrm{p}_0}b\right).$$

Write

$$\partial_x^\alpha\left(e^{it\mathrm{p}_0}b\right)=e^{it\mathrm{p}_0}\left(e^{-it\mathrm{p}_0}\partial_x^\alpha e^{it\mathrm{p}_0}\right)b$$

and cancel the factor $e^{it\mathrm{p}_0}$ from both sides of the preceding equation to get

$$\frac{1}{i}\frac{\partial b}{\partial t}+\mathrm{p}_0 b=\sum_\alpha \frac{\hbar^\alpha}{\alpha!}D_\xi^\alpha \mathrm{p}Q^\alpha(b)$$

where

$$Q=\left(\partial_x+it\frac{\partial \mathrm{p}_0}{\partial x}\right). \tag{10.14}$$

Since $Q^0=I$, we can remove the term $\mathrm{p}_0 b$ from both sides of the preceding equation to obtain

$$\frac{1}{i}\frac{\partial b}{\partial t}=\sum_{|\alpha|\geq 1}\hbar^\alpha D_\xi^\alpha \mathrm{p}Q^\alpha(b)+(p-\mathrm{p}_0). \tag{10.15}$$

Let us expand b and p in powers of $\hbar$,

$$b=\sum_k b_k(x,\xi,t)h^k,\qquad \mathrm{p}=\sum_k \mathrm{p}_k\hbar^k,$$

and equate powers of $\hbar$ in (10.15). We get the series of equations

$$\frac{1}{i}\frac{\partial b_m}{\partial t}=\sum_{|\alpha|\geq 1}\sum_{j+k+|\alpha|=m}D_\xi^\alpha \mathrm{p}_j Q^\alpha b_k+\sum_{j\geq 1}\mathrm{p}_j b_{m-j} \tag{10.16}$$

with initial conditions

$$b_0(x,\xi,0)\equiv 1,\quad b_m(x,\xi,0)\equiv 0 \text{ for } m\geq 1.$$

We can solve these equations recursively by integration. In particular, $b_0(x,\xi,t)\equiv 1$.

Proposition 10.2.1. *$b_m(x,\xi,t)$ is a polynomial in t of degree at most $2m$.*

Proof by induction. We know this for $m=0$. For $j+k+|\alpha|=m$, we know by induction that $Q^\alpha b_k$ is a polynomial in t of degree at most $|\alpha|+2k=m-j+k\leq m+k<2m$ so integration shows that b_m is a polynomial in t of degree at most $2m$. □

So we have found a solution mod $\hbar^\infty$ to our wave equation problem.

10.3 The functional calculus modulo $O(\hbar^\infty)$.

Sticking (10.10) into (10.4) we get the following expression for the Schwartz kernel of $f(P)\rho(\hbar D)$:

$$\frac{1}{\sqrt{2\pi}}\int \mu(x,y,t,\hbar)\hat{f}(t)dt \sim \frac{1}{\sqrt{2\pi}}\sum \hbar^{k-n}\sum_{\ell\le 2k}\int \mu_{k,\ell}(x,y,t)\hat{f}(t)dt \tag{10.17}$$

where

$$\frac{1}{\sqrt{2\pi}}\int \mu_{k,\ell}(x,y,t)\hat{f}(t)dt$$

$$= \int b_{k,\ell}(x,\xi)\rho(\xi)e^{i\frac{(x-y)\cdot\xi}{\hbar}}\left(\frac{1}{\sqrt{2\pi}}\int t^{\ell}\hat{f}(t)e^{i\mathrm{p}_0}dt\right)d\xi$$

$$= \int b_{k,\ell}(x,\xi)\rho(\xi)e^{i\frac{(x-y)\cdot\xi}{\hbar}}\left(\left(\frac{1}{i}\frac{d}{ds}\right)^{\ell}f\right)(\mathrm{p}_0(x,\xi))d\xi.$$

Thus the Schwartz kernel of $f(P)\rho(\hbar D)$ has an asymptotic expansion

$$(2\pi\hbar)^{-n}\sum_k \hbar^k\sum_{\ell\le 2k}\int b_{k,\ell}(x,\xi)\rho(\xi))e^{i\frac{(x-y)\cdot\xi}{\hbar}}\left(\left(\frac{1}{i}\frac{d}{ds}\right)^{\ell}f\right)(\mathrm{p}_0(x,\xi))d\xi. \tag{10.18}$$

This shows that $f(P)\rho(\hbar D)\in\Psi^0(\mathbb{R}^n)$ and has left Kohn-Nirenberg symbol

$$b_f(x,\xi,\hbar)\rho(\xi)$$

where

$$b_f(x,\xi,\hbar)\sim\sum_k \hbar^k\left(\sum_{\ell\le 2k} b_{k,\ell}(x,\xi)\left(\left(\frac{1}{i}\frac{d}{ds}\right)^{\ell}f\right)(\mathrm{p}_0(x,\xi))\right). \tag{10.19}$$

In particular, since $b_{0,0}(x,\xi)\equiv 1$, we have

$$b_f(x,\xi,0)=f(\mathrm{p}_0(\mathrm{x},\xi)). \tag{10.20}$$

Now let us show that if one imposes a mild "ellipticity type" assumption on $p_0(x,\xi)$ one can remove the cut-off ρ from the above formula.

We have been assuming that the symbol p of P is in S^m and hence, in particular, that $p(x,\xi)$ satisfies

$$|p_0(x,\xi)|\le C_K\langle\xi\rangle^m$$

as x ranges over a compact set K.

In the cases we are interested in m is positive, so we can impose on p_0 the "ellipticity type" condition

$$|p_0(x,\xi)|\ge C_k\|\xi\|^k+o(\|\xi\|^k) \tag{10.21}$$

for some $0 \leq k \leq m$ and positive constant C_k.

Since f is compactly supported, this assumption tells us that

$$\left(\frac{d}{ds}\right)^{\ell}(f)(p_0(x,\xi))$$

is compactly supported in ξ. Hence, if we choose the cutoff function $\rho(\xi)$ to be equal to 1 on a neighborhood of this support, we can eliminate ρ from (10.18) to get the simpler result

Theorem 10.3.1. *Under the above ellipticity assumptions, if $f \in C_0^\infty(\mathbb{R})$ the operator $f(P)$ is a semi-classical pseudo-differential operator and its Shwartz kernel has the asymptotic expansion*

$$(2\pi\hbar)^{-n}\sum \hbar^k \sum_{\ell\leq 2k}\int b_{k,\ell}(x,\xi)e^{\frac{i(x-y)\cdot\xi}{\hbar}}\frac{1}{i^\ell}f^{(\ell)}(p_0(x,\xi))d\xi.$$

10.4 The trace formula.

Suppose that for some interval $[a,b]$ the set $p_0^{-1}([a,b])$ is compact. Then for $f \in C_0^\infty((a,b))$ the functions $f^{(\ell)}(p_0(x,\xi))$ are compactly supported and hence by the expression for $f(P)$ given in Theorem 10.3.1, the operator $f(P)$ is of trace class modulo $O(\hbar^\infty)$. In Chapter 13 we will show that the "modulo $O(\hbar^\infty)$ proviso can be removed, i.e. that $f(P)$ itself is of trace class and hence that $\mathrm{spec}(P)\cap(a,b)$ is discrete. Assuming this, let $[c,d]$ be a finite subinterval of (a,b), and let $\lambda_i(\hbar)$, $i = 1,2,\cdots$ be the eigenvalues of P lying in $[c,d]$. If we choose our f to be non-negative and $f \equiv 1$ on $[c,d]$ we see that

$$\sum(\lambda_i(\hbar)) \leq \sum f(\lambda_i(\hbar)) \leq \operatorname{tr} f(P) < \infty.$$

We conclude that

Proposition 10.4.1. *For any $[c,d] \subset (a,b)$ the number of eigenvalues of P on $[c,d]$ is finite.*

From Theorem 10.3.1 we have the asymptotic expansion

$$\sum f(\lambda_i(\hbar)) \sim (2\pi\hbar)^{-n}\sum_{k,\ell}\hbar^k\int b_{k,\ell}(x,\xi)\frac{1}{i^\ell}f^{(\ell)}(p_0(x,\xi))dxd\xi. \tag{10.22}$$

Since $b_{0,0} \equiv 1$, the leading term on the right is

$$(2\pi\hbar)^{-n}\int f\left(p_0(x,\xi)\right)dxd\xi. \tag{10.23}$$

If $0 \leq f \leq 1$ and is supported on the interval $(c-\epsilon, d+\epsilon)$ with $f \equiv 1$ on $[c,d]$, then (10.22) and (10.23) imply that

$$\#\{\lambda_i(\hbar) \in [c,d]\} \leq (2\pi\hbar)^{-n}\left(\mathrm{Vol}(c \leq p_0(x,\xi) \leq d) + O(\epsilon)\right).$$

In the opposite direction, if $0 \leq f \leq 1$ with f supported on $[c,d]$ and $\equiv 1$ on $[c+\epsilon, d-\epsilon]$ we get the estimate

$$\#\{\lambda_i(\hbar) \in [c,d]\} \geq (2\pi\hbar)^{-n} \left(\mathrm{Vol}(c \leq p_0(x,\xi) \leq d) + O(\epsilon)\right).$$

Putting these together we get the "Weyl law"

$$\#\{\lambda_i(\hbar) \in [c,d]\} \sim (2\pi\hbar)^{-n} \left(\mathrm{Vol}(c \leq p_0(x,\xi) \leq d)\right) + o(1). \tag{10.24}$$

Let us return to (10.22). The summands on the right, namely

$$\int \sum_{\ell \leq 2k} b_{k,\ell}(x,\xi) \frac{1}{i^\ell} f^{(\ell)}\left(p_0(x,\xi)\right) dx d\xi \tag{10.25}$$

are clearly spectral invariants of P. In the next few sections we will compute the first few of these invariants for the Schrödinger operator

$$S_\hbar = \frac{\hbar^2}{2} \sum_i D_{x_i}^2 + V \tag{10.26}$$

and the Schrödinger operator with vector potential $A = (a_1, \ldots, a_n)$:

$$S_{\hbar,A} = \frac{\hbar^2}{2} \sum_i (D_{x_i} + a_i)^2 + V. \tag{10.27}$$

We will also show how, in one dimension, these invariants serve to determine V in some cases.

The material in the next few sections is taken from the paper [GW].

10.5 Spectral invariants for the Schrödinger operator.

For the Schrödinger operator (10.26), we have

$$p(x,\xi,\hbar) = p_0(x,\xi) =$$

$$p(x,\xi) := \frac{\|\xi\|^2}{2} + V(x). \tag{10.28}$$

Hence the set $a \leq p_0(x,\xi) \leq b$ is compact if and only if the set $a \leq V(x) \leq b$ is compact. For the rest of this chapter let us assume that this is the case. We now compute the trace invariants (10.25) for $S_\hbar$: The first trace invariant is

$$\int f(p(x,\xi)) dx d\xi$$

as we have seen above.

To compute the next trace invariant we observe that the operator Q of (10.14) is given as

$$Q = \partial_x + it\frac{\partial V}{\partial x} \tag{10.29}$$

for the case of the Schrödinger operator (10.26). Since p is quadratic in ξ, equations ((10.16) become

$$\begin{aligned}\frac{1}{i}\frac{\partial b_m}{\partial t} &= \sum_{|\alpha|\geq 1}\sum_{k+|\alpha|=m} D_\xi^\alpha p Q^\alpha b_k \\ &= \sum_k \frac{\xi_k}{i}\left(\frac{\partial}{\partial x_k} + it\frac{\partial V}{\partial x_k}\right) b_{m-1} - \frac{1}{2}\sum_k \left(\frac{\partial}{\partial x_k} + it\frac{\partial V}{\partial x_k}\right)^2 b_{m-2}.\end{aligned}$$

Since $b_0(x,\xi,t) = 1$ and $b_1(x,\xi,0) = 0$, we have

$$b_1(x,\xi,t) = \frac{it^2}{2}\sum_l \xi_l \frac{\partial V}{\partial x_l},$$

and thus

$$\begin{aligned}\frac{1}{i}\frac{\partial b_2}{\partial t} &= \sum_k \frac{\xi_k}{i}\left(\frac{\partial}{\partial_{x_k}} + it\frac{\partial V}{\partial x_k}\right)\left(\frac{it^2}{2}\sum_l \xi_l\frac{\partial V}{\partial x_l}\right) - \frac{1}{2}\sum_k\left(\frac{\partial}{\partial x_k} + it\frac{\partial V}{\partial x_k}\right)^2 (1) \\ &= \frac{t^2}{2}\sum_{k,l}\xi_k\xi_l\left(\frac{\partial^2 V}{\partial x_k \partial x_l} + it\frac{\partial V}{\partial x_k}\frac{\partial V}{\partial x_l}\right) - \frac{1}{2}\sum_k\left(it\frac{\partial^2 V}{\partial x_k^2} - t^2\frac{\partial V}{\partial x_k}\frac{\partial V}{\partial x_k}\right).\end{aligned}$$

It follows that

$$b_2(x,\xi,t) = \frac{t^2}{4}\sum_k \frac{\partial^2 V}{\partial x_k^2} + \frac{it^3}{6}\left(\sum_k\left(\frac{\partial V}{\partial x_k}\right)^2 + \sum_{k,l}\xi_k\xi_l\frac{\partial^2 V}{\partial x_k\partial x_l}\right) - \frac{t^4}{8}\sum_{k,l}\xi_k\xi_l\frac{\partial V}{\partial x_k}\frac{\partial V}{\partial x_l}. \tag{10.30}$$

Thus the next trace invariant will be the integral

$$\begin{aligned}\int -\frac{1}{4}\sum_k \frac{\partial^2 V}{\partial x_k^2} f''\left(\frac{\xi^2}{2} + V(x)\right) - \frac{1}{6}\sum_k\left(\frac{\partial V}{\partial x_k}\right)^2 f^{(3)}\left(\frac{\xi^2}{2} + V(x)\right) \\ -\frac{1}{6}\sum_{k,l}\xi_k\xi_l\frac{\partial^2 V}{\partial x_k \partial x_l} f^{(3)}\left(\frac{\xi^2}{2} + V(x)\right) - \frac{1}{8}\sum_{k,l}\xi_k\xi_l\frac{\partial V}{\partial x_k}\frac{\partial V}{\partial x_l} f^{(4)}\left(\frac{\xi^2}{2} + V(x)\right)\, dx d\xi.\end{aligned} \tag{10.31}$$

We can apply to these expressions the integration by parts formula,

$$\int \frac{\partial A}{\partial x_k} B\left(\frac{\xi^2}{2} + V(x)\right)\, dx d\xi = -\int A(x)\frac{\partial V}{\partial x_k} B'\left(\frac{\xi^2}{2} + V(x)\right)\, dx d\xi \tag{10.32}$$

and

$$\int \xi_k\xi_l A(x) B'\left(\frac{\xi^2}{2} + V(x)\right)\, dx d\xi = -\int \delta_k^l A(x) B\left(\frac{\xi^2}{2} + V(x)\right)\, dx d\xi. \tag{10.33}$$

Applying (10.32) to the first term in (10.31) we get

$$\int \frac{1}{4} \sum_k (\frac{\partial V}{\partial x_k})^2 f^{(3)}(\frac{\xi^2}{2} + V(x))\ dxd\xi,$$

and by applying (10.33) the fourth term in (10.31) becomes

$$\int \frac{1}{8} \sum_k (\frac{\partial V}{\partial x_k})^2 f^{(3)}(\frac{\xi^2}{2} + V(x))\ dxd\xi.$$

Finally applying both (10.33) and (10.32) the third term in (10.31) becomes

$$\int -\frac{1}{6} \sum_k (\frac{\partial V}{\partial x_k})^2 f^{(3)}(\frac{\xi^2}{2} + V(x))\ dxd\xi.$$

So the integral (10.31) can be simplified to

$$\frac{1}{24} \int \sum_k (\frac{\partial V}{\partial x_k})^2 f^{(3)}(\frac{\xi^2}{2} + V(x))\ dxd\xi.$$

We conclude

Theorem 10.5.1. *The first two terms of (10.22) are*

$$\operatorname{tr} f(S_\hbar) = \int f\left(\frac{\xi^2}{2} + V(x)\right)\ dxd\xi + \frac{1}{24}\hbar^2 \int \sum_k (\frac{\partial V}{\partial x_k})^2 f^{(3)}(\frac{\xi^2}{2} + V(x))\ dxd\xi + O(\hbar^4). \tag{10.34}$$

In deriving (10.34) we have assumed that f is compactly supported. However, if we change our compactness hypothesis slightly, and assume that V is bounded from below and that the set $V(x) \leq a$ is compact for some a, the left and right hand sides of (10.34) are unchanged if we replace the "f" in (10.34) by *any* function, f, with support on $(-\infty, a)$, and, as a consequence of this remark, it is easy to see that the following two integrals,

$$\int_{\frac{\xi^2}{2}+V(x)\leq\lambda} dxd\xi \tag{10.35}$$

and

$$\int_{\frac{\xi^2}{2}+V(x)\leq\lambda} \sum_k (\frac{\partial V}{\partial x_k})^2 dxd\xi \tag{10.36}$$

are spectrally determined by the spectrum (**??**) on the interval $[0, a]$. Moreover, from (10.34), one reads off the Weyl law: For $0 < \lambda < a$,

$$\#\{\lambda_i(\hbar) \leq \lambda\} = (2\pi\hbar)^{-n} \left(\operatorname{Vol}(\frac{\xi^2}{2} + V(x) \leq \lambda) + o(1)\right). \tag{10.37}$$

We also note that the second term in the formula (10.34) can, by (10.33), be written in the form

$$\frac{1}{24}\hbar^2 \int \sum_k \frac{\partial^2 V}{\partial x_k^2} f^{(2)}(\frac{\xi^2}{2} + V(x))\ dxd\xi$$

and from this one can deduce an $\hbar^2$-order "cumulative shift to the left" correction to the Weyl law.

We won't attempt to compute the invariants (10.25) explicitly. However we will show that they can be written in the form

$$\nu_k(f) = \int \sum_{j=[\frac{k}{2}+1]}^{k} f^{(2j)} \left(\frac{\xi^2}{2} + V(x)\right) p_{k,j}(DV, \cdots, D^{2k}V) dxd\xi \tag{10.38}$$

where $p_{k,j}$ are universal polynomials, and $D^k V$ the k^{th} partial derivatives of V.

Proof of (10.38). Notice that for m even, the lowest degree term in the polynomial b_m is of degree $\frac{m}{2}+1$, thus we can write

$$b_m = \sum_{l=-\frac{m}{2}+1}^{m} b_{m,l} t^{m+l}.$$

Putting this into the the iteration formula, we will get

$$\begin{aligned}\frac{m+l}{i} b_{m,l} = & \sum \frac{\xi_k}{i} \frac{\partial b_{m-1,l}}{\partial x_k} + \sum \xi_k \frac{\partial V}{\partial x_k} b_{m-1,l-1} - \frac{1}{2} \sum \frac{\partial^2 b_{m-2,l+1}}{\partial x_k^2} \\ & - \frac{i}{2}(\frac{\partial}{\partial x_k}\frac{\partial V}{\partial x_k} + \frac{\partial V}{\partial x_k}\frac{\partial}{\partial x_k}) b_{m-2,l} + \frac{1}{2}\sum (\frac{\partial V}{\partial x_k})^2 b_{m-2,l-1},\end{aligned}$$

from which one can easily conclude that for $l \geq 0$,

$$b_{m,l} = \sum \xi^\alpha (\frac{\partial V}{\partial x})^\beta p_{\alpha,\beta}(DV, \cdots, D^m V) \tag{10.39}$$

where $p_{\alpha,\beta}$ is a polynomial, and $|\alpha| + |\beta| \geq 2l - 1$. Moreover, by integration by parts,

$$\int \xi^\mu \xi_i f^{(r} \left(\frac{\|\xi\}^2}{2} + V(x)\right) d\xi = -\int \left(\frac{\partial}{\partial \xi_i}\xi^\mu\right) i f^{(r-1} \left(\frac{\|\xi\}^2}{2} + V(x)\right) d\xi.$$

It follows from this formula and (10.32) and (10.33), all the $f^{(m+l)}$, $l \geq 0$, in the integrand of the $\hbar^m$th term in the expansion (10.25) can be replaced by $f^{(r)}$'s with $r \leq m$. In other words, only derivatives of f of degree $\leq 2k$ figure in the expression for $\nu_k(f)$. For those terms involving derivatives of order less than $2k$, one can also use integration by parts to show that each $f^{(m)}$ can be replaced by a $f^{(m+1)}$ and a $f^{(m-1)}$. In particular, we can replace all the odd derivatives by even derivatives. This proves (10.38). $\square$

10.6 An Inverse Spectral Result: Recovering the Potential Well

Let us now consider the one dimensional case. Suppose V is a "potential well", i.e. has a unique nondegenerate critical point at $x = 0$ with minimal value $V(0) = 0$, and that V is increasing for x positive, and decreasing for x negative. For simplicity assume in addition that

$$-V'(-x) > V'(x) \tag{10.40}$$

holds for all x. We will show how to use the spectral invariants (10.35) and (10.36) to recover the potential function $V(x)$ on the interval $|x| < a$.

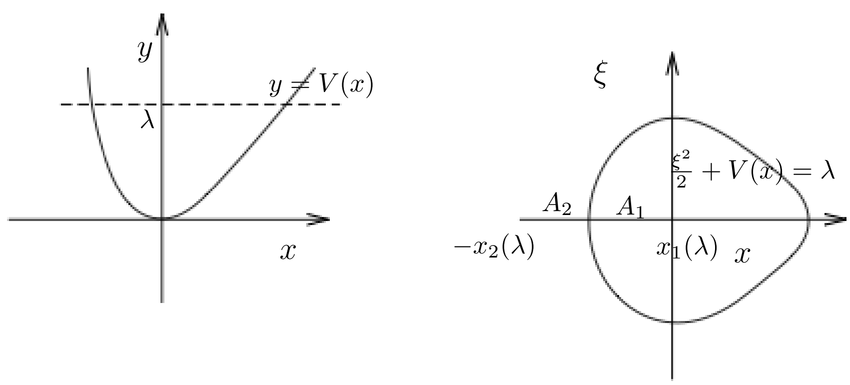


Figure 10.1: Single Well Potential

For $0 < \lambda < a$ we let $-x_2(\lambda) < 0 < x_1(\lambda)$ be the intersection of the curve $\frac{\xi^2}{2} + V(x) = \lambda$ with the x-axis on the $x - \xi$ plane. We will denote by A_1 the region in the first quadrant bounded by this curve, and by A_2 the region in the second quadrant bounded by this curve. Then from (10.35) and (10.36) we can determine

$$\int_{A_1} + \int_{A_2} dx d\xi \tag{10.41}$$

and

$$\int_{A_1} + \int_{A_2} V'(x)^2 dx d\xi. \tag{10.42}$$

Let $x = f_1(s)$ be the inverse function of $s = V(x), x \in (0, a)$. Then

$$\begin{aligned}\int_{A_1} V'(x)^2 \, dx d\xi &= \int_0^{x_1(\lambda)} V'(x)^2 \int_0^{\sqrt{2(\lambda - V(x))}} d\xi dx \\ &= \int_0^{x_1(\lambda)} V'(x)^2 \sqrt{2\lambda - 2V(x)} \, dx \\ &= \int_0^{\lambda} \sqrt{2\lambda - 2s} V'(f_1(s)) \, ds \\ &= \int_0^{\lambda} \sqrt{2\lambda - 2s} \left(\frac{df_1}{ds} \right)^{-1} ds.\end{aligned}$$

Similarly

$$\int_{A_2} V'(x)^2 \, dx d\xi = \int_0^{\lambda} \sqrt{2\lambda - 2s} \left(\frac{df_2}{ds} \right)^{-1} ds,$$

where $x = f_2(s)$ is the inverse function of $s = V(-x), x \in (0, a)$. So the spectrum of $S_\hbar$ determines

$$\int_0^{\lambda} \sqrt{\lambda - s} \left((\frac{df_1}{ds})^{-1} + (\frac{df_2}{ds})^{-1} \right) ds. \tag{10.43}$$

Similarly the knowledge of the integral (10.41) amounts to the knowledge of

$$\int_0^{\lambda} \sqrt{\lambda - s} \left(\frac{df_1}{ds} + \frac{df_2}{ds} \right) ds. \tag{10.44}$$

Recall now that the fractional integration operation of Abel,

$$J^a g(\lambda) = \frac{1}{\Gamma(a)} \int_0^{\lambda} (\lambda - t)^{a-1} g(t) \, dt \tag{10.45}$$

for $a > 0$ satisfies $J^a J^b = J^{a+b}$. Hence if we apply $J^{1/2}$ to the expression (10.44) and (10.43) and then differentiate by λ two times we recover $\frac{df_1}{ds} + \frac{df_2}{ds}$ and $(\frac{df_1}{ds})^{-1} + (\frac{df_2}{ds})^{-1}$ from the spectral data. In other words, we can determine f_1' and f_2' up to the ambiguity $f_1' \leftrightarrow f_2'$.

However, by (10.40), $f_1' > f_2'$. So we can from the above determine f_1' and f_2', and hence $f_i, i = 1, 2$. So we conclude

Theorem 10.6.1. *Suppose the potential function V is a potential well, then the semi-classical spectrum of $S_\hbar$ modulo $o(\hbar^2)$ determines V near 0 up to $V(x) \leftrightarrow V(-x)$.*

Remarks, 1. We will show in Section 10.9 that the hypothesis (10.40) or some "asymmetry" condition similar to it is necessary for the theorem above to be true.

2. The formula (10.44) can be used to construct lots of Zoll potentials, i.e. potentials for which the Hamiltonian flow v_H associated with $H = \xi^2 + V(x)$ is periodic of period 2π. It's clear that the potential $V(x) = x^2$ has this property and is the only even potential with this property. However, by (10.44) and the area-period relation (See Proposition 6.1) every single-well potential V for which

$$f_1(s) + f_2(s) = 2s^{1/2}$$

has this property.

10.7 Semiclassical Spectral Invariants for Schrödinger Operators with Magnetic Fields

In this section we will show how the results in §10.5 can be extended to Schrödinger operators with magnetic fields. Recall that a semi-classical Schrödinger operator with magnetic field on $\mathbb{R}^n$ has the form

$$S_{\hbar}^m := \frac{1}{2}\sum_j \left(\frac{\hbar}{i}\frac{\partial}{\partial x_j} + a_j(x)\right)^2 + V(x) \tag{10.46}$$

where $a_k \in C^\infty(\mathbb{R}^n)$ are smooth functions defining a magnetic field B, which, in dimension 3 is given by $\vec{B} = \vec{\nabla} \times \vec{a}$, and in arbitrary dimension by the 2-form $B = d(\sum a_k dx_k)$. We will assume that the vector potential $\vec{a}$ satisfies the Coulomb gauge condition,

$$\nabla \cdot \vec{a} = \sum_j \frac{\partial a_j}{\partial x_j} = 0. \tag{10.47}$$

(In view of the definition of B, one can always choose such a Coulomb vector potential.) In this case, the Kohn-Nirenberg symbol of the operator (10.46) is given by

$$p(x, \xi, \hbar) = \frac{1}{2}\sum_j (\xi_j + a_j(x))^2 + V(x). \tag{10.48}$$

Recall that

$$Q_\alpha = \frac{1}{\alpha!}\prod_k \left(\frac{\partial}{\partial x_k} + it\frac{\partial p}{\partial x_k}\right)^{\alpha_k}, \tag{10.49}$$

so the iteration formula (**??**) becomes

$$\frac{1}{i}\frac{\partial b_m}{\partial t} = \sum_k \frac{1}{i}\frac{\partial p}{\partial \xi_k}\left(\frac{\partial}{\partial x_k} + it\frac{\partial p}{\partial x_k}\right)b_{m-1} - \frac{1}{2}\sum_k \left(\frac{\partial}{\partial x_k} + it\frac{\partial p}{\partial x_k}\right)^2 b_{m-2}. \tag{10.50}$$

from which it is easy to see that

$$b_1(x, \xi, t) = \sum_k \frac{\partial p}{\partial \xi_k}\frac{\partial p}{\partial x_k}\frac{it^2}{2}. \tag{10.51}$$

Thus the "first" spectral invariant is

$$\int \sum_k (\xi_k + a_k(x)) \frac{\partial p}{\partial x_k} f^{(2)}(p)\ dx d\xi = -\int \sum_k \frac{\partial a_k}{\partial x_k} f'(p) dx d\xi = 0,$$

where we used the fact $\sum \frac{\partial a_k}{\partial x_k} = 0$.

With a little more effort we get for the next term

$$\begin{aligned} b_2(x,\xi,t) =& \frac{t^2}{4} \sum_k \frac{\partial^2 p}{\partial x_k^2} \\ &+ \frac{it^3}{6} \left(\sum_{k,l} \frac{\partial p}{\partial \xi_k} \frac{\partial a_l}{\partial x_k} \frac{\partial p}{\partial x_l} + \sum_{k,l} \frac{\partial p}{\partial \xi_k} \frac{\partial p}{\partial \xi_l} \frac{\partial^2 p}{\partial x_k \partial x_l} + \sum_k (\frac{\partial p}{\partial x_k})^2 \right) \\ &+ \frac{-t^4}{8} \sum_{k,l} \frac{\partial p}{\partial \xi_k} \frac{\partial p}{\partial x_k} \frac{\partial p}{\partial \xi_l} \frac{\partial p}{\partial x_l}. \end{aligned}$$

and, by integration by parts, the spectral invariant

$$I_\lambda = -\frac{1}{24} \int \left(\sum_k \frac{\partial^2 p}{\partial x_k^2} - \sum_{k,l} \frac{\partial a_k}{\partial x_l} \frac{\partial a_l}{\partial x_k} \right) f^{(2)}(p(x,\xi)) dx d\xi. \tag{10.52}$$

Notice that

$$\frac{\partial^2 p}{\partial x_k^2} = \sum_j \frac{\partial^2 a_j}{\partial x_k^2} \frac{\partial p}{\partial \xi_j} + \sum_j (\frac{\partial a_j}{\partial x_k})^2 + \frac{\partial^2 V}{\partial x_k^2}$$

and

$$\|B\|^2 = \mathrm{tr} B^2 = 2 \sum_{j,k} \frac{\partial a_k}{\partial x_j} \frac{\partial a_j}{\partial x_k} - 2 \sum_{j,k} (\frac{\partial a_k}{\partial x_j})^2$$

So the subprincipal term is given by

$$\frac{1}{48} \int f^{(2)}(p(x,\xi)) \left(\|B\|^2 - 2 \sum_k \frac{\partial^2 V}{\partial x_k^2} \right)\ dx\ d\xi.$$

Finally Since the spectral invariants have to be gauge invariant by definition, and since any magnetic field has by gauge change a coulomb vector potential representation, the integral

$$\int_{p<\lambda} \left(\|B\|^2 - 2 \sum_k \frac{\partial^2 V}{\partial x_k^2} \right)\ dx\ d\xi$$

is spectrally determined for an arbitrary vector potential. Thus we proved

Theorem 10.7.1. *For the semiclassical Schrödinger operator (10.46) with magnetic field B, the spectral measure $\nu(f) = \mathrm{trace} f(S_\hbar^m)$ for $f \in C_0^\infty(\mathbb{R})$ has an asymptotic expansion*

$$\nu^m(f) \sim (2\pi\hbar)^{-n} \sum \nu_r^m(f) \hbar^{2r},$$

where

$$\nu_0^m(f) = \int f(p(x,\xi,\hbar))dxd\xi$$

and

$$\nu_1^m(f) = \frac{1}{48}\int f^{(2)}(p(x,\xi,\hbar))(\|B\|^2 - 2\sum \frac{\partial^2 V}{\partial x_i^2}).$$

10.8 An Inverse Result for The Schrödinger Operator with A Magnetic Field

Making the change of coordinates $(x,\xi) \to (x, \xi + a(x))$, the expressions (10.7.1) and (10.8) simplify to

$$\nu_0^m(f) = \int f(\xi^2 + V)dxd\xi$$

and

$$\nu_1^m(f) = \frac{1}{48}\int f^{(2)}(\xi^2 + V)(\|B\|^2 - 2\sum \frac{\partial^2 V}{\partial x_i^2})dxd\xi.$$

In other words, for all λ, the integrals

$$\mathrm{I}_\lambda = \int_{\xi^2 + V(x) < \lambda} dxd\xi$$

and

$$\mathrm{II}_\lambda = \int_{\xi^2 + V(x) < \lambda} (\|B\|^2 - 2\sum \frac{\partial^2 V}{\partial x_i^2})dxd\xi$$

are spectrally determined.

Now assume that the dimension is 2, so that the magnetic field B is actually a scalar $B = Bdx_1 \wedge dx_2$. Moreover, assume that V is a radially symmetric potential well, and the magnetic field B is also radially symmetric. Introducing polar coordinates

$$x_1^2 + x_2^2 = s,\ dx_1 \wedge dx_2 = \frac{1}{2}ds \wedge d\theta$$

$$\xi_1^2 + \xi_2^2 = t,\ d\xi_1 \wedge d\xi_2 = \frac{1}{2}dt \wedge d\psi$$

we can rewrite the integral I_λ as

$$\mathrm{I}_\lambda = \pi^2 \int_0^{s(\lambda)} (\lambda - V(s))ds,$$

where $V(s(\lambda)) = \lambda$. Making the coordinate change $V(s) = x \Leftrightarrow s = f(x)$ as before, we get

$$\mathrm{I}_\lambda = \pi^2 \int_0^{\lambda} (\lambda - x)\frac{df}{dx}dx.$$

A similar argument shows

$$\mathrm{II}_\lambda = \pi^2 \int_0^\lambda (\lambda - x) H(f(x)) \frac{df}{dx} dx,$$

where

$$H(s) = B(s)^2 - 4sV''(s) - 2V'(s).$$

It follows that from the spectral data, we can determine

$$f'(\lambda) = \frac{1}{\pi^2} \frac{d^2}{d\lambda^2} \mathrm{I}_\lambda$$

and

$$H(f(\lambda)) f'(\lambda) = \frac{1}{\pi^2} \frac{d^2}{d\lambda^2} \mathrm{II}_\lambda.$$

So if we normalize $V(0) = 0$ as before, we can recover V from the first equation and B from the second equation.

Remark. In higher dimensions, one can show by a similar (but slightly more complicated) argument that V and $\|B\|$ are both spectrally determined if they are radially symmetric.

10.9 Counterexamples.

Let $V \in C^\infty(\mathbb{R}^n)$ be a potential well - that is a potential with $V(0) = 0$, $V(x) > 0$ for $x \neq 0$ and $V(x) \to +\infty$ as $|x| \to +\infty$. Then, by Proposition 10.4.1, the spectrum of the Shrödinger operator (10.26) is discrete. The question : "to what extent does this spectrum determine V?" is still an open question; however we will show in this section that in dimension one there exist uncountable families of potentials for which the spectral invariants (10.25) are the same and that in dimension greater than one there even exist infinite parameter families of potentials for which these invariants are the same.

We first observe that if $A : \mathbb{R}^n \to \mathbb{R}^n$ is an orthogonal transformation, i.e., $A \in O(n)$ then

$$A^*(S_n)(A^{-1})^* = S_n^A$$

where

$$S_n^A = \frac{\hbar^2}{2} \Delta + V^A(x)$$

and $V^A(x) = V(Ax)$. Thus if $K_f(x, y, \hbar)$ is the Schwartz kernel of the operator $f(S_\hbar)$, then $K_f(Ax, Ay, \hbar)$ is the Schwartz kernel of the operator, $f(S_\hbar^A)$ and, by (10.18), $K_f(Ax, AX)$ has an asymptotic expansion of the form

$$(2\pi\hbar)^{-n} \sum_k \hbar^k \sum_{\ell \le 2k} \int b_{k,\ell}(Ax, \xi) \rho(\xi)) e^{i \frac{(x-y)\cdot \xi}{\hbar}} \left(\left(\frac{1}{i} \frac{d}{ds} \right)^\ell f \right) \left(\frac{\|\xi\|^2}{2} + V(Ax,) \right) d\xi.$$

In particular since the function, $b_{k,\ell}(x,\xi)$ in the expansion (10.19) has the form

$$b_{k,\ell} = \sum \xi^\alpha P_{\alpha,k,\ell}(DV,\dots,d^{2k}V) \tag{10.53}$$

by (39) the corresponding functions for S_n^A have the form

$$b_{k,\ell}^A = \sum \xi^\alpha p_{\alpha,k,\ell}(DV^A,\dots,D^{2k}V^A) \tag{10.54}$$

and hence in particular

$$b_{k,\ell}(\xi, Ax) = \sum \xi^\alpha p_{\alpha,k,\ell}(DV^A,\dots D^{2k}V^A) \tag{10.55}$$

for all $x \in \mathbb{R}^n$.

Now choose V to be rotationally symmetric and let $\rho_i(x)$ be a non-negative C^∞ function with support on the set

$$i < |x| < i+1\,, \qquad x_1 > 0, \dots, x_n > 0$$

with $\rho_i = 0$ for i odd and $\rho_i \neq 0$ for i even. Then, fixing a sequence of rotations,

$$A = \{A_i \in)(n) \quad i = 1,2,3,\dots\}$$

the potentials

$$V_A = V(x) + \sum \rho_i(A_i x)$$

have the same spectral invariants (10.38) for all sequences, A, as can be seen by writing

$$\int b_{k,\ell}(\xi, DV_A,\dots,D^{2k}V_A) f^\ell(\frac{\xi^2}{2} + V_A)\,dx\,d\xi$$

$$= \sum \int_{i \le |x| \le i+1} b_{k,\ell}(\xi, DV_A,\dots,D^{2k}V_A) f^\ell(\frac{\xi^2}{2} + V_A)\,dx\,d\xi$$

$$= \sum \int_{i \le |x| \le i+1} b_{k,\ell}(\xi, D(V+\rho_i)^{A_i},\dots,D^{2k}(V+\rho_i)^{A_i}) f^\ell\left(\frac{\xi^2}{2} + (V+\rho_i)^{A_i}\right)\,dx\,d\xi$$

and observing that this is equal to

$$\sum \int_{i \le |x| \le i+1} b_{k,\ell}(\xi, D(V+\rho_i)) f^\ell\left(\frac{\xi^2}{2} + V + \rho_i\right)\,dx\,d\xi$$

by equation (10.55).

In dimension one this construction doesn't give us an infinite parameter family of potentials with the same spectral invariants (10.38) but it's easy to see that it does give us uncountable family of potentials for which these invariants are the same. Namely for every $K \in [0,1)$ let

$$\alpha = \alpha_1\alpha_2\alpha_3\dots$$

be the binary expansion of α and choose A_{2i} to be the symmetry, $x \to -x$, if α_i is one and $x \to x$ if $\alpha_i$is 0.

This example (which is a slightly modified version of a counterexample by Colin de Verdiere in [Col]) shows why the assumption (40) (or some asymmetry condition similar to (40)) is necessary in the hypotheses of Theorem 10.6.1.

10.10 The functional calculus on manifolds.

Let X^n be a compact manifold and $P_\hbar \in \Psi^0 S^m(X)$ a self-adjoint zero$^{\text{th}}$ order semi-classical pseudodifferential operator with leading symbol $P_0(x,\xi) \in S^m(X)$ satisfying an elliptic estimate of the form

$$P_0(x,\xi) \geq C|\xi|^m \tag{10.56}$$

on every coordinate patch. We will show below how to extend the results of §§10.2–10.3 to manifolds, i.e., how to define $f(P_\hbar)$, modulo $O(\hbar^\infty)$, as a zero$^{\text{th}}$ order semi-classical pseudodifferential operator on X with compact microsupport for all $f \in \mathcal{C}_0^\infty(\mathbb{R})$.

Let V_i, $i = 1,\dots,N$, be a covering of X by coordinate patches, let $\varphi_i \in \mathcal{C}_0^\infty(V_i)$, $i = 1,\dots,N$ be a partition of unity subordinate to this cover, and for each i, let $\psi_i \in \mathcal{C}_0^\infty(V_i)$ be a function which is equal to 1 on a neighborhood of $\operatorname{Supp}\varphi_i$. We can, as in §10.2, construct a family of semi-classical pseudodifferential operators, $U_i(t) : \mathcal{C}_0^\infty(V_i) \to \mathcal{C}^\infty(V_i)$, such that modulo $O(\hbar^\infty)$

$$\begin{aligned} \frac{1}{\sqrt{-1}}\frac{d}{dt}U_i(t) &\equiv P_\hbar U_i(t) \\ U_i(0) &= \rho(\hbar D)\,. \end{aligned}$$

Thus the sum

$$U(t) = \sum \psi_i U_i(t)\varphi_i$$

is a zero$^{\text{th}}$ order semi-classical pseudodifferential operator on X satisfying

$$\begin{aligned} \frac{1}{\sqrt{-1}}\frac{d}{dt}U(t) &\equiv \sum \psi_i P_\hbar U_i(t)\varphi_i \\ &\equiv \sum P_\hbar \psi_i U_i(t)\varphi_i \\ &\equiv P_\hbar U(t) \end{aligned}$$

modulo $O(\hbar^\infty)$ with initial data

$$U(0) = \sum \psi_i \rho(\hbar D)\varphi_i \stackrel{\text{def}}{=} Q_\rho\,, \tag{10.57}$$

i.e., modulo $O(\hbar^\infty)$

$$U(t) = (\exp itP_\hbar)Q_\rho\,. \tag{10.58}$$

Thus for $f \in \mathcal{C}_0^\infty(\mathbb{R})$

$$f(P_\hbar)Q_\rho \equiv \sum \psi_i \left(\int U_i(t)\hat{f}(t)\,dt \right) \varphi_i \tag{10.59}$$

$\operatorname{mod} O(\hbar^\infty)$ where each of the expressions in parentheses has a Schwartz kernel of the form (10.18). Thus by the ellipticity condition (10.56) we can, exactly as in §10.3, remove the cut-off, ρ, to get an asymptotic expansion for the Schwartz kernel of $f(P_h)$ itself of the form,

$$(2\pi h)^{-n} \sum_{i=1}^{N} \sum_{k,\ell} h^k \psi_i \int b_{k,\ell}^i(x,\xi) e^{\dfrac{i(x,y)\cdot\xi}{h}} \left(\frac{i}{i} \frac{d}{ds} \right)^\ell f(p_0(x,\xi))\,d\xi \varphi_i(y)$$

and from this expansion a trace formula of the form (10.22). More explicitly since X is compact the ellipticity conditions (10.56) insure that the spectrum of P_h is discrete and for fixed $\hbar$ consists of a sequence of eigenvalues, $\lambda_i(\hbar)$, $i = 1, 2, \ldots$, which tend to $+\infty$ as i tends to infinity. Hence from the asymptotic expansion above for the Schwartz kernel of $f(P_h)$ one gets an asymptotic expansion for $\sum f(\lambda_i(h))$ of the form

$$(2\pi\hbar)^{-n} \sum_{y=1}^{N} \int \sum_{k,\ell} \hbar^k b_{k,\ell}^j(x,\xi) \varphi_j(x) \left(\frac{1}{i} \frac{d}{ds} \right)^\ell f(P_0(x,\xi))\,dx\,d\xi\,. \tag{10.60}$$

In particular, as we showed in §10.2, $b_{0,0}^i = 1$ so the leading term in this expansion gives the Weyl estimate

$$\sum f(\lambda_i(h)) \sim (2\pi h)^{-n} \left(\int f(P_0(x,\xi) \frac{\omega^n}{n!} + 0()1) \right) \tag{10.61}$$

where $P_0 : T^x X \to \mathbb{R}$ is the intrinsic leading symbol of $P_\hbar$ and $\omega = \sum dx_i \wedge d\xi_i$ is the intrinsic symplectic form on T^*X.

There is an interesting tie-in between this result and the classical "heat-trace" theorem for Riemannian manifolds: Suppose X is a Riemannian manifold and $\Delta : \mathcal{C}^\infty(X) \to \mathcal{C}^\infty(X)$ its Laplace operator. The Minakshisundaram-Plejjel theorem asserts that as $t \to 0+$ one has an asymptotic expansion

$$\operatorname{Tr}(\exp(-t\Delta)) \sim (4\pi t)^{-n/2} \sum a_i t^i \tag{10.62}$$

with $a_0 = \operatorname{vol}(X)$. This is easily deduced from the formula (10.60) by letting $t = h^2$, $P_h = h\sqrt{\Delta}$ and the f in (10.60) a sequence of f's which tend in the Schwartz space norm to e^{-x^2}.

Chapter 11

Fourier Integral operators.

11.1 Semi-classical Fourier integral operators.

As in Chapter 9 one can extend the theory of Fourier integral operators to classes of operators having symbols, $a(x, y, \xi, \hbar)$, which are not compactly supported in ξ; i.e., with "compact support in ξ" replaced by growth conditions in ξ similar to those we discussed for pseudodifferential operators in Chapter 9. We won't, however, attempt to do so here; and, in fact, we will continue to confine ourselves in this chapter to the type of Fourier integral operator we discussed in Chapter 8. We have already seen, however, that these include a lot of interesting real-world examples. For instance, given a C^∞ mapping between manifolds $f : X \to Y$, the pull-back operation, $f^* : \mathcal{C}^\infty(Y) \to \mathcal{C}^\infty(X)$ is microlocally an F.I.O. in the sense that for every semi-classical pseudodifferential operator, $Q : \mathcal{C}^\infty(Y) \to \mathcal{C}^\infty(Y)$, with compact microsupport, f^*Q is a semi-classical F.I.O. Moreover if f is a fiber mapping a similar assertion is true for the push-forward operation, f_*. Given the results of Chapter 9 we can add to this list a lot of other examples such as the operators, $f^*P_\hbar\rho(D)$ and $f_*P_\hbar\rho(D)$ where $P_\hbar$ is in $\Psi^k S^m$ and $\rho = \rho(\xi_1, \ldots, \xi_n)$ is compactly supported. In addition an example about which we will have a lot to say at the end of this chapter is the operator, $\exp(\frac{i\,t}{\hbar}P_\hbar)f(P_\hbar)$, $\quad f \in C_0^\infty(\mathbb{R})$, where $P_\hbar$ is a self-adjoint elliptic operator in $\Psi^0 S^m(X)$. This operator looks suspiciously like the operator, $\exp i\,tP_\hbar$, which we studied in detail in the last chapter, but the presence of the factor "$1/\hbar$" in the exponent gives it a completely different character. In particular we will show that, like the other examples above, it is microlocally an F.I.O. What follows is a brief table of contents for this chapter.

I. Let X be a compact manifold, let $M = T^*X$ and let $\Gamma : T^*X \twoheadrightarrow T^*X$ be a canonical relation which is transversal to Δ_M. We will show in §11.2 that if $F_\hbar$ is a k^{th} order Fourier integral operator with compact microsupport quantizing Γ then one has an asymptotic expansion

$$\text{trace } F_\hbar \sim \hbar^F \sum a_p(\hbar) e^{\frac{i\pi}{4}\sigma_p} e^{\frac{iT_p^*}{\hbar}} \tag{11.1}$$

summed over $p \in \Gamma \cap \Delta_M$ where $a_p(\hbar) = \sum_{i=0}^{\infty} a_{p,i}\hbar^i$ is a formal power series in $\hbar$, σ_p a Maslov factor and the T_p's are symplectic invariants of Γ.

II. In §§11.3 and 11.4 we will show how to compute these invariants when Γ is the graph of a symplectomorphism, and in particular we will show in §11.4 that they have a simple geometric interpretation as the "period spectrum" of a dynamical system living on the mapping torus of f.

III. The second half of this chapter will focus on the two main wave-trace formulas of semi-classical analysis: the Gutzwiller formula and density of states. Let X be $\mathbb{R}^n$ (or, alternatively, let X be a compact manifold) and let $P_h : \mathcal{C}_0^\infty(X) \to \mathcal{C}^\infty(X)$ be a self-adjoint zeroth order semi-classical pseudodifferential operator. We will denote by $H : T^*X \to \mathbb{R}$ its leading symbol and by v_H the Hamiltonian vector field associated with H. We will show in §11.5 that if H is proper the operator, $\exp i\frac{t}{\hbar}P$ is microlocally a semi-classical Fourier integral operator quantizing the symplectomorphism, $\exp tv_H$ and we will show that for cut-offs, ψ and f in $\mathcal{C}_0^\infty(\mathbb{R})$ the trace of the operator

$$\hat{\psi}\left(\frac{P_\hbar}{\hbar}\right) f(P_\hbar) = \int \psi(t) e^{\frac{i\,t}{\hbar}P_\hbar}\, dt f(P_\hbar)$$

has nice asymptotic properties if the flow of v_H on the energy surface, $H = 0$ has non-degenerate periodic trajectories. In particular there is a trace formula

$$\text{trace } \hat{\psi}\left(\frac{P_\hbar}{\hbar}\right) f(P_\hbar) \sim \hbar^{-\frac{n}{2}} \sum_\gamma e^{\frac{iS_\gamma}{\hbar}} \sum_{i=0}^{\infty} a_{\gamma,i}\hbar^i \tag{11.2}$$

similar to (11.1) where the sum is over the periodic trajectories of v_H on $H = 0$ and the S_γ's are the classical "actions" associated with these trajectories:

$$S_\gamma = \int_\gamma \sum \xi_i \, dx_i \,. \tag{11.3}$$

Replacing $P_\hbar$ by $P_\hbar - E$, for any $E \in \mathbb{R}$ one gets an analogous result for the periodic trajectories of v_H on the energy surface $H = E$; so among many other things this result tells us that the classical actions, S_γ, are spectral invariants of $P_\hbar$.

IV. In assuming that the periodic trajectories of v_H on the energy surface of $H = 0$ are non-degenerate we are ruling out the case where a periodic trajectory consists simply of a fixed point for the flow, $\exp tv_H$; i.e., a zero, p, of the vector field v_H. However, if there are a finite number of isolated zeros of $_H$ on $H = 0$ and they are all non-degenerate the density of states formula asserts that for $|t|$ small

$$\text{trace } \exp \frac{i\,t}{\hbar} P_\hbar f(P_\hbar) \sim \sum_p \hbar^{-n/2} e^{\frac{iT_p}{\hbar}} a_p(t,\hbar) \tag{11.4}$$

where $a_p(t,\hbar) \sim \sum_{i=0}^{\infty} a_{p,i}(t)\hbar^i$, and the T_p's are the symplectic invariants figuring in (11.1).

V. In the last section of this chapter we will discuss some applications of the results of this chapter to "heat trace invariants" in Riemannian geometry. Let X be as in §10.11 a compact Riemannian manifold and $g : X \to X$ an isometry of X. In the 1970's Harold Donnelly generalized the heat trace formula that we described in §10.11 by showing that one has an asymptotic expansion

$$\text{trace } g^* e^{-t\Delta_X} \sim \sum_Z (4\pi t)^{-d_Z/2} \sum_{k=0}^{\infty} b_{k,Z} t^k \tag{11.5}$$

where the Z's are the connected components of the fixed point set of g and d_Z is the dimension of Z. Moreover since g is an isometry the eigenvalues of the map $(dg)_p : N_p(Z) \to N_p(Z)$ at $p \in Z$ don't depend on p, and denoting these eigenvalues by $\lambda_{i,Z}$,i=1, $\ldots n - d_\Gamma$ he shows that

$$b_{o,Z} = \text{vol}(Z)(\Pi(1-\lambda_{i,Z}))^{-1}\,. \tag{11.6}$$

If f is the identity map this heat trace expansion is just the Minakshisundaran-Pleijel formula (10.62) and as we pointed out in §10.11 this expansion can be thought of semi-classically as a trace formula for $f(P_\hbar)$ where $P_\hbar = \hbar\sqrt{\Delta}$. In §11.6 we will show that the same is true of the formula (11.5). In fact we will show more generally that if $P_\hbar$ is a self-adjoint semi-classical elliptic pseudodifferential operator of order zero and $g : X \to X$ is a diffeomorphism of X whose graph intersects Δ_X in a finite number of fixed point components, Z, then one has an analogue of the expansion (11.5) for the trace of $g^* f(P_\hbar)$ and that (11.5) can be viewed as a special case of this expansion.

A key ingredient in the proof of all these results is the lemma of stationary phase. A detailed account of the lemma of stationary phase (with a host of applications) can be found in Chapter 15. However, in the next section we will give a brief acount of the manifold version of this lemma, the version that we will need for the applications below.

11.2 The lemma of stationary phase.

Let X be an n-dimensional manifold. A C^∞ function ϕ on X is said to be a **Bott-Morse** function if

- Its critical set $C_\phi := \{x \in X | d\phi_x = 0\}$ is a smooth submanifold and
- For every $p \in C_\phi$ the Hessian $d^2\phi_p : T_pX \to \mathbb{R}$ is non-degenerate on the normal space $N_pC_\phi = T_pX/T_pC_\phi$.

To state the lemma of stationary phase we need to recall some differential invariants which are intrinsically attached to such a function:

Let $W_r,\ r = 1,\dots,N$ be the connected components of C_ϕ, so that ϕ is constant, say identically equal to γ_r on W_r. Similarly, the signature $\mathrm{sgn} d^2\phi$ is constant on each W_r. Let $p \in W_r$ and $w_1,\dots,w_k$ be a basis of N_pW_r. Consider

$$|\det(d^2\phi_p(w_i,w_j))|^{\frac{1}{2}}.$$

If we replace w_i by Aw_i in this expression, where A is some linear operator on N_pW_r we pullout a factor of $|\det A|$. In other words, the above expression defines a density on N_pW_r. From the exact sequence

$$0 \to T_pW_r \to T_pX \to N_pW_r \to 0$$

we know from (6.7) that we have an isomorphism

$$|T_pX| \simeq |T_pW| \otimes |N_pW|.$$

Thus, the above density on N_pW_r together with a given density on T_pX determines a density on T_pW_r.

For example, if $X = \mathbb{R}^n$ with density $dx_1 \dots dx_n$ and ϕ has an isolated nondegenerate fixed point at 0, then the induced "density", which is a number, is

$$\frac{1}{|\det(\partial^2\phi/\partial x_i \partial x_j)(0)|}.$$

In short, a density μ on X determines a density, call it ν_r on each W_r. The lemma of stationary phase says that for μ of compact support we have

$$\int_X e^{\frac{i\phi}{\hbar}}\mu = \sum_r (2\pi\hbar^{\frac{d_r}{2}})\left(e^{\frac{i\gamma_r}{\hbar}} e^{i\frac{\mathrm{sgn}_{W_r}}{4}} \int_{W_r} \nu_r + O(\hbar)\right). \tag{11.7}$$

11.3 The trace of a semiclassical Fourier integral operator.

Let X be an n-dimensional manifold, let $M = T^*X$ and let

$$\Gamma : T^*X \twoheadrightarrow T^*X$$

be a canonical relation. Let $\Delta_M \subseteq M \times M$ be the diagonal and let us assume that

$$\Gamma \pitchfork \Delta_M\,.$$

Our goal in this section is to show that if $F \in \mathcal{F}_0^k(\Gamma)$ is a semi-classical Fourier integral operator "quantizing" the canonical relation Γ then one has a trace formula of the form:

$$\mathrm{tr}\, F = \hbar^k \sum a_p(h) e^{\frac{i\pi}{\eta_p}} e^{iT_p^*/\hbar} \tag{11.8}$$

summed over $p \in \Gamma \cap \Delta_M$. In this formula n is the dimension of X, the η_p's are Maslov factors, the T_p^* are symplectic invariants of Γ at $p \cap \Gamma \circ \Delta_M$ which will be defined below, and $a_p(h) \in C^\infty(\mathbb{R})$.

Let $\varsigma : M \to M$ be the involution, $(x, \xi) \to (x, -\xi)$ and let $\Lambda = \varsigma \circ \Gamma$. We will fix a non-vanishing density, dx, on X and denote by

$$\mu = \mu(x, y, \hbar)\, dx^{\frac{1}{2}}\, dy^{\frac{1}{2}} \tag{11.9}$$

the Schwartz kernel of the operator, F. By definition

$$\mu \in I^{k-\frac{n}{2}}(X \times X, \Lambda)$$

and by (11.9) the trace of F is given by the integral

$$\operatorname{tr} F =: \int \mu(x, x, \hbar)\, dx\,. \tag{11.10}$$

To compute this, we can without loss of generality assume that Λ is defined by a generating function, i.e., that there exists a d-dimensional manifold, S, and a function $\varphi(x, y, s) \in C^\infty(X \times X \times S)$ which generates Λ with respect to the fibration, $X \times X \times S \to X \times X$. Let C_φ be the critical set of φ and $\lambda_\varphi : C_\varphi \to \Lambda$ the diffeomorphism of this set onto Λ. Denoting by $\varphi^\sharp$ the restriction of φ to C_φ and by ψ the function, $\varphi^\sharp \circ \lambda_\varphi^{-1}$, we have

$$d\psi = \alpha_\Lambda \tag{11.11}$$

where α_Λ is the restriction to Λ of the canonical one form, α, on $T^*(X \times X)$.

Lets now compute the trace of F. By assumption μ can be expressed as an oscillatory integral

$$(dx)^{\frac{1}{2}}(dy)^{\frac{1}{2}}\left(h^{k\frac{n}{2}-d/2} \int a(x, y, s, h) e^{\frac{i\varphi(x,y,s)}{\hbar}}\, ds\right)$$

and hence by (11.10)

$$\operatorname{tr} F = \hbar^{k-\frac{n}{2}-d/2} \int a(x, x, s, \hbar) e^{i\frac{\varphi(x,x,s)}{\hbar}}\, ds\, dx\,. \tag{11.12}$$

We claim that: *The function*

$$\varphi(x, x, s) : X \times S \to \mathbb{R} \tag{11.13}$$

is a Morse function, with critical points

$$(x, x, s) = \lambda_\varphi^{-1}(p)\,, \quad p \in \Gamma \cap \Delta_M\,. \tag{11.14}$$

Proof. Consider Γ as a morphism

$$\Gamma : \text{pt.} \to M^- \times M \tag{11.15}$$

and Δ_M^t as a morphism

$$\Delta_M^t : M^- \times M \to pt\,. \tag{11.16}$$

The condition that Γ intersects Δ_M transversally can be interpreted as saying that (11.14) and (11.16) are transversally composable. Thus since $\varphi(x,y,s)$ is a generating function for Γ with respect to the fibration

$$X \times X \times S \to X \times X$$

and $\rho(x,y,\xi) = (x-y)\cdot\xi$ is a generating function for Δ_M with respect to the fibration

$$X \times X \times \mathbb{R}^h \to X \times X$$

the function, $\varphi(x,y,s) + \rho(x,y,\xi)$ is a transverse generating function for $\Delta_M^t \circ \Gamma$ with respect to the fibration

$$X \times X \times S \times \mathbb{R}^h \to pt$$

i.e. is just a Morse function on this set. (See §5.6.)

However if we let $\varphi(x,y,s) = \varphi(x,x,s) + (x-y)\cdot h(x,y,s)$ and set $u = x-y$ and $w = \xi + h(x,y,s)$ then, under this change of coordinates, $\varphi + \gamma$ becomes

$$\varphi(x,x,s) + u \cdot w$$

x, s, u and w being independent variables. Since this is a Morse function its two summands are Morse functions with critical points (x,s) and $u = v = 0$ where

$$\frac{\partial \varphi}{\partial x}(x,x,s) \quad = \quad -\frac{\partial d}{\partial y}(x,x,s)$$

and

$$\frac{\partial d}{\partial s}(x,x,s) \quad = \quad 0$$

i.e. x, x, s is given by (11.14).

□

Since the function (11.13) is a Morse function we can evaluate (11.11) by stationary phase obtaining

$$\operatorname{tr} F = \sum h^k a_p(h) e^{i\frac{\pi}{4}\operatorname{sgn}_p} e^{i\psi(p)/\hbar} \tag{11.17}$$

where sgn_p is the signature of $\varphi(x,x,s)$ at the critical point corresponding to p and

$$\psi(p) = \varphi(x,x,s)\,,$$

the value of $\varphi(x,x,s)$ at this point. This gives us the trace formula (11.8) with $T_p^\sharp = \psi(p)$.

Replacing the transverse composition formula for generating function (§5.6) by the analogous clean composition formula (§5.7) one gets a "clean" version of this result. Namely suppose Γ and Δ_M intersect cleanly in a finite number of connected submanifolds W_r, $r = 1 \ldots, N$ of $\dim d = dr$. Then on each of these submanifolds, ψ is constant: $\psi|W_r = \gamma_r$ and

$$\operatorname{tr} F = h^k \sum_r h^{-\frac{dr}{2}} a(h) e^{i\frac{\gamma_r}{h}} . \tag{11.18}$$

11.3.1 Examples.

Let's now describe how to compute these $T_p^\sharp$'s in some examples: Suppose Γ is the graph of a symplectomorphism

$$f : M \to M .$$

Let pr_1 and pr_2 be the projections of $T^*(X \times X) = M \times M$ onto its first and second factors, and let α_X be the canonical one form on T^*X. Then the canonical one form, α, on $T^*(X \times X)$ is

$$(pr_1)^*\alpha_X + (pr_2)^*\alpha_X ,$$

so if we restrict this one form to Λ and then identify Λ with M via the map, $M \to \Lambda$, $p \to (p, \sigma f(p))$, we get from (11.11)

$$\alpha_X - f^*\alpha_X = d\psi \tag{11.19}$$

and $T_p^\sharp$ is the value of ψ at the point, p.

Let's now consider the Fourier integral operator

$$F^m = \overbrace{F \circ \cdots \circ F}$$

and compute its trace. This operator "quantizes" the symplectomorphism f^m, hence if

$$\operatorname{graph} f^m \pitchfork \Delta_M$$

we can compute its trace by (11.8) getting the formula

$$\operatorname{tr} F^m = \hbar^\ell \sum a_{m,p}(\hbar) e^{i\frac{\pi}{4}\sigma_{m,p}} e^{iT^\sharp_{m,p}/\hbar} . \tag{11.20}$$

with $\ell = km$, the sum now being over the fixed points of f^m. As above, the oscillations, $T^\sharp_{m,p}$, are computed by evaluating at p the function, ψ_m, defined by

$$\alpha_X - (f^m)^*\alpha_X = d\psi_m .$$

However,

$$\begin{aligned} \alpha_X - (f^m)^*\alpha_X &= \alpha_X - f^*\alpha_X + \cdots + (f^{m-1})^*\alpha_X - (f^m)^*\alpha , \\ &= d(\psi + f^*\psi + \cdots + (f^{m-1})^*\psi) \end{aligned}$$

where ψ is the function (11.11). Thus at $p = f^m(p)$

$$T^\sharp_{m,p} = \sum_{i=1}^{m-1} \psi(p_i)\,, \quad p_i = f^i(p)\,. \tag{11.21}$$

In other words $T^\sharp_{m,p}$ is the sum of ψ over the periodic trajectory $(p_1, \ldots, p_{m-1})$ of the dynamical system

$$f^k\,, \quad -\infty < k < \infty\,.$$

We refer to the next subsection "The period spectrum of a symplectomorphism" for a proof that the $T^\sharp_{m,p}$'s are *intrinsic* symplectic invariants of this dynamical system, i.e., depend only on the symplectic structure of M not on the canonical one form, α_X. (We will also say more about the "geometric" meaning of these $T^\sharp_{m,p}$'s in Theorem 11.4.1 below.)

Finally, what about the amplitudes, $a_p(h)$, in formula (11.8)? There are many ways to quantize the symplectomorphism, f, and no canonical way of choosing such a quantization; however, one condition which one can impose on F is that its symbol be of the form:

$$h^{-n}\nu_\Gamma e^{\frac{i\psi}{h}} e^{i\frac{\pi}{4}\sigma_\varphi}\,, \tag{11.22}$$

in the vicinity of $\Gamma \cap \Delta_M$, where ν_Γ is the $\frac{1}{2}$ density on Γ obtained from the symplectic $\frac{1}{2}$ density, ν_M, on M by the identification, $M \leftrightarrow \Gamma$, $p \to (p, f(p))$. We can then compute the symbol of $a_p(h) \in I^0(pt)$ by pairing the $\frac{1}{2}$ densities, ν_M and ν_Γ at $p \in \Gamma \cap \Delta_M$ as in (7.14) obtaining

$$a_p(0) = |\det(I - df_p)|^{-\frac{1}{2}}\,. \tag{11.23}$$

Remark. The condition (11.22) on the symbol of F can be interpreted as a "unitarity" condition. It says that "microlocally" near the fixed points of f:

$$FF^t = I + O(h)\,.$$

11.3.2 The period spectrum of a symplectomorphism.

Let (M, ω) be a symplectic manifold. We will assume that the cohomology class of ω is zero; i.e., that ω is exact, and we will also assume that M is connected and that

$$H^1(M, \mathbb{R}) = 0\,. \tag{*}$$

Let $f : M \to M$ be a symplectomorphism and let $\omega = d\alpha$. We claim that $\alpha - f^*\alpha$ s exact. Indeed $d\alpha - f^*d\alpha = \omega - f^*\omega = 0$, and hence by (*) $\alpha - f^*\alpha$ s exact. Let

$$\alpha - f^*\alpha = d\psi$$

for $\psi \in C^\infty(M)$. This function is only unique up to an additive constant; however, there are many ways to normalize this constant. For instance if W is a connected subset of the set of fixed points of f, and $j : W \to M$ is the inclusion map, then $f \circ j = j$; so

$$j^* \, d\psi = j^*\alpha - j^* f^* \alpha = 0$$

and hence ψ is constant on W. Thus one can normalize ψ by requiring it to be zero on W.

Example. Let Ω be a smooth convex compact domain in $\mathbb{R}^n$, let X be its boundary, let U be the set of points, (x,ξ), $|\xi| < 1$, in T^*X. If $B : U \to U$ is the billiard map and α the canonical one form on T^*X one can take for $\psi = \psi(x,\xi)$ the function

$$\psi(x,\xi) = |x-y| + C$$

where $(y,n) = B(x,\xi)$. B has no fixed points on U, but it extends continuously to a mapping of $\bar{U}$ on $\bar{U}$ leaving the boundary, W, of U fixed and we can normalize ψ by requiring that $\psi = 0$ on W, i.e., that $\psi(x,\xi) = |x-y|$.

Now let

$$\gamma = p_1, \ldots, p_{k+1}$$

be a periodic trajectory of f, i.e.,

$$f(p_i) = p_{i+1} \qquad i = 1, \ldots k$$

and $p_{k+1} = p_1$. We define the *period* of γ to be the sum

$$p(\gamma) = \sum_{i=1}^{k} \psi(p_i)\,.$$

Claim: *$P(\gamma)$ is independent of the choice of α and ψ. In other words it is a* symplectic invariant *of f.*

Proof. Suppose $\omega = d\alpha - d\alpha'$. Then $d(\alpha - \alpha') = 0$; so, by (*), $\alpha' - \alpha = dh$ for some function, $h \in C^\infty(M)$. Now suppose $\alpha - f^*\alpha = d\psi$ and $\alpha' - f^*\alpha' = d\psi'$ with $\psi = \psi'$ on the set of fixed points, W. Then

$$d\psi' - d\psi = d(f^*h - h)$$

and since $f^* = 0$ on W

$$\psi' - \psi = f^*h - h\,.$$

Thus

$$\begin{aligned} \sum_{i=1}^{k} \psi'(p_i) - \psi(p_i) &= \sum_{i=1}^{k} h(f(p_i)) - h(p_i) \\ &= \sum_{i=1}^{k} h(p_{i+1}) - h(p_i) \\ &= 0\,. \end{aligned}$$

Hence replacing ψ by ψ' doesn't change the definition of $P(\gamma)$. □

Example Let $p_i = (x_i, \xi_i)$ $i = 1, \ldots, k+1$ be a periodic trajectory of the billiard map. Then its period is the sum

$$\sum_{i=1}^{k} |x_{i+1} - x_i| \,,$$

i.e., is the *perimeter* of the polygon with vertices at $x_1, \ldots, x_k$. (It's far from obvious that this is a symplectic invariant of B.)

11.4 The mapping torus of a symplectic mapping.

We'll give below a geometric interpretation of the oscillations, $T^{\sharp}_{m,p}$, occurring in the trace formula (11.20). First, however, we'll discuss a construction used in dynamical systems to convert "discrete time" dynamical systems to "continuous time" dynamical systems. Let M be a manifold and $f : M \to M$ a diffeomorphism. From f one gets a diffeomorphism

$$g : M \times \mathbb{R} \to M \times \mathbb{R} \,, \quad g(p,q) = (f(p), q+1)$$

and hence an action

$$\mathbb{Z} \to Diff(M \times \mathbb{R}) \,, \quad k \to g^k \,, \tag{11.24}$$

of the group, $\mathbb{Z}$ on $M \times \mathbb{R}$. This action is free and properly discontinuous so the quotient

$$Y = M \times \mathbb{R}/\mathbb{Z}$$

is a smooth manifold. The manifold is called the *mapping torus* of f. Now notice that the translations

$$\tau_t : M \times \mathbb{R} \to M \times \mathbb{R} \,, \, (p,q) \to (p, q+t) \,, \tag{11.25}$$

commute with the action (11.24), and hence induce on Y a one parameter group of translations

$$\tau^{\sharp}_t : Y \to Y \,, \, -\infty < t < \infty \,. \tag{11.26}$$

Thus the mapping torus construction converts a "discrete time" dynamical system, the "discrete" one-parameter group of diffeomorphisms, $f^k : M \to M$, $-\infty < k < \infty$, into a "continuous time" one parameter group of diffeomorphisms (11.26).

To go back and reconstruct f from the one-parameter group (11.26) we note that the map

$$\iota : M = M \times \{0\} \to M \times \mathbb{R} \to (M \times \mathbb{R})/Z$$

imbeds M into Y as a global cross-section, M_0, of the flow (11.26) and for $p \in M_0$ $\gamma_t(p) \in M_0$ at $t = 1$ and via the identification $M_0 \to M$, the map,

$p \to \gamma_1(p)$, is just the map, f. In other words, $f : M \to M$ is the "first return map" associated with the flow (11.26).

We'll now describe how to "symplecticize" this construction. Let $\omega \in \Omega^2(M)$ be an exact symplectic form and $f : M \to M$ a symplectomorphism. For $\alpha \in \Omega^1(M)$ with $d\alpha = \omega$ let

$$\alpha - f^*\alpha = d\varphi \tag{11.27}$$

and lets assume that φ is bounded from below by a positive constant. Let

$$g : M \times \mathbb{R} \to M \times \mathbb{R}$$

be the map

$$g(p, q) = (p, q + \varphi(x)) . \tag{11.28}$$

As above one gets from g a free properly discontinuous action, $k \to g^k$, of $\mathbb{Z}$ on $M \times \mathbb{R}$ and hence one can form the mapping torus

$$Y = (M \times \mathbb{R})/\mathbb{Z} .$$

Moreover, as above, the group of translations,

$$\tau_t : M \times \mathbb{R} \to M \times \mathbb{R} \, , \; \tau_t(p, q) = (p, q + t) \, ,$$

commutes with (11.28) and hence induces on Y a one-parameter group of diffeomorphisms

$$\tau_t^\sharp : Y \to Y \, ,$$

just as above. We will show, however, that these are not just diffeomorphisms, they are *contacto-morphisms*. To prove this we note that the one-form,

$$\tilde{\alpha} = \alpha + dt \, ,$$

on $M \times \mathbb{R}$ is a contact one-form. Moreover,

$$\begin{aligned} g^*\tilde{\alpha} &= f^*\alpha + d(\varphi + t) \\ &= \alpha + (f^*\alpha - \alpha) + d\varphi + dt \\ &= \alpha + dt = \tilde{\alpha} \end{aligned}$$

by (11.27) and

$$(\tau_a)^*\tilde{\alpha} = \alpha + d(t + a) = \alpha + dt = \tilde{\alpha}$$

so the action of $\mathbb{Z}$ on $M \times \mathbb{R}$ and the translation action of $\mathbb{R}$ on $M \times \mathbb{R}$ are both actions by groups of contacto-morphisms. Thus, $Y = (M \times \mathbb{R})/\mathbb{Z}$ inherits from $M \times \mathbb{R}$ a contact structure and the one-parameter group of diffeomorphisms, $\tau_t^\sharp$, preserves this contact structure.

Note also that the infinitesimal generator, of the group translations, τ_t, is just the vector field, $\frac{\partial}{\partial t}$, and that this vector field satisfies

$$\iota(\frac{\partial}{\partial t})\tilde{\alpha} = 1$$

and

$$\iota(\frac{\partial}{\partial t})\, d\tilde{\alpha} = 0\,.$$

Thus $\frac{\partial}{\partial t}$ is the contact vector field associated with the contract form $\tilde{\alpha}$, and hence the infinitesimal generator of the one-parameter group, $\tau_t^\sharp : Y \to Y$ is the contact vector field associated with the contract form on Y.

Comments:

1. The construction we've just outlined involves the choice of a one-form, α, on M with $d\alpha = \omega$ and a function, φ, with $\alpha = f^x\alpha = d\varphi$; however, it is easy to see that the contact manifold, Y, and one-parameter group of contacto-morphisms are uniquely determined, up to contacto-morphism, independent of these choices.

2. Just as in the standard mapping torus construction f can be shown to be "first return map" associated with the one-parameter group, $\tau_t^\sharp$.

We can now state the main result of this section, which gives a geometric description of the oscillations, $T_{m,p}^\sharp$, in the trace formula.

Theorem 11.4.1. *The periods of the periodic trajectories of the flow, $\tau_t^\sharp$, $-\infty < t < \infty$, coincide with the "length" spectrum of the symplectomorphism, $f : M \to M$.*

Proof. For $(p, a) \in M \times \mathbb{R}$,

$$g^m(p, a) = (f^m(p), q + \varphi(p) + \varphi(p_1) + \cdots + \varphi(p_{m-1})$$

with $p_i = f^i(p)$. Hence if $p = f^m(p)$

$$g^m(p, a) = \tau_{T^\sharp}(p, a)$$

with

$$T^\sharp = T_{m,p}^\sharp = \sum_{i=1}^{m} \varphi(p_i)\,, \quad p_i = f^i(p)\,.$$

Thus if q is the projection of (p, a) onto Y the trajectory of $\tau^\sharp$ through q is periodic of period $T_{m,p}^\sharp$. □

Via the mapping torus construction one discovers an interesting connection between the trace formula in the preceding section and a trace formula which we described in Section 7.7.4.

Let β be the contact form on Y and let

$$M^\sharp = \{(y, \eta) \in T^*Y\,,\, \eta = t\beta_y\,,\, t \in \mathbb{R}_+\}\,.$$

It's easy to see that $M^\sharp$ is a symplectic submanifold of T^*Y and hence a symplectic manifold in its own right. Let

$$H : M^\sharp \to \mathbb{R}^+$$

be the function $H(y, tB_y) = t$. Then Y can be identified with the level set, $H = 1$ and the Hamiltonian vector field ν_H restricted to this level set coincides with the contact vector field, ν, on Y. Thus the flow, $\tau_t^\sharp$, is just the Hamiltonian flow, $\exp t\nu_H$, restricted to this level set. Let's now compute the "trace" of $\exp t\nu_H$ as an element in the category $\tilde{S}$ (the enhanced symplectic category).

The computation of this trace is essentially identical with the computation we make at the end of Section 7.7.4 and gives as an answer the union of the Lagrangian manifolds

$$\Lambda_{T^\sharp_{m,p}} \subset T^*\mathbb{R}\,,\, m \in \mathbb{Z}\,,$$

where the $T^\sharp$'s are the elements of the period spectrum of ν_H and $\Lambda_{T^\sharp}$ is the cotangent fiber at $t = T$. Moreover, each of these $\Lambda_{T^\sharp}$'s is an element of the enhanced symplectic category, i.e. is equipped with a $\frac{1}{2}$-density $\nu_{T^\sharp_{m,p}}$ which we computed to be

$$\overline{T}^\sharp_{m,p}|I - df_p^m|^{-\frac{1}{2}}|d\tau|^{\frac{1}{2}}\,.$$

$\overline{T}^\sharp_{m,p}$ being the *primitive* period of the period trajectory of f through p (i.e., if $p_i = f^i(p)$ $i = 1, \dots, m$ and $p, p_1, \dots, p_{k-1}$ are all distinct but $p = p_k$ then $\overline{T}^\sharp_{m,p} = \overline{T}^\sharp_{k,p}$) . Thus these expressions are just the symbols of the oscillatory integrals

$$\hbar^{-1}a_{m,p}e^{I\imath\overline{T}^\sharp_{p,m}t/\hbar}$$

with $a_{m,p} = \overline{T}^\sharp_{m,p}|I - df_p^m|^{\frac{1}{2}}$.

11.5 The Gutzwiller formula.

Let X be a smooth manifold and $P \in \Psi^0 S^n(X)$ a self-adjoint semi-classical pseudo-differential operator with leading symbol $p(x, \xi($. As in ¶10.4, we will assume that for some real interval $[a, b]$, $p^{-1}([a, b])$ is compact. Our goal in this section is to show that for $f \in C_0^\infty(a, b)$ the operator

$$\exp it\frac{P}{\hbar} \cdot f(P) \qquad (11.29)$$

is compact, and to compute its trace. At first glance it would appear that the techniques of Chapter 10, where we derived a trace formula for the operator $\exp itP\rho(\hbar D)$ would translate more or less verbatim to this setting; i.e. that we should be able to solve the equation

$$\hbar\frac{1}{i}\frac{\partial}{\partial t}U(t) - PU = 0 \qquad (11.30)$$

with initial condition $U(0) = f(P)$ by using the local symbol calculus of ΨDO's as in ¶10.2, and then patch there together to get a manifold result as in ¶10.10.

Unfortunately, however, since the operator $\frac{1}{\hbar}P_\hbar$ is, semi-classically, a *first* order ΨDO, functions of it are no longer ΨDO's, so this result no longer works.

What one *can* do, however, is to solve (11.30) modulo $O(\hbar^\infty)$ by the transport equation techniques of ¶8.7.5 and then use "variation of constants" to get rid of the $O(\hbar^\infty)$. Here are the details:

Let $u(x, y, t, \hbar)$ be the desired Schwartz kernel of $U(t)$. To solve

$$\left(\hbar \frac{1}{i}\frac{\partial}{\partial t} - P(x, D_x, \hbar)\right) u(x, y, t, \hbar) = 0 \tag{11.31}$$

modulo $O(\hbar^\infty)$ with the given initial data, let

$$H = \tau - p(x, \xi)$$

be the leading symbol on the left of (11.31) and let Λ_0 be the set of points

$$(x, \xi, y, \eta, t, \tau) \in T^*(X \times X \times \mathbb{R})$$

where

$$(x, \xi) = (y, \eta), \quad t = 0, \quad H(x, \xi, 0, \tau) = 0, \quad \text{and} \quad (x, \xi) \in p^{-1}(a, b).$$

Since $p^{-1}([a, b])$ is compact and invariant under the flow of the Hamiltonian vector field v_p, the set of points $(x, \xi, y, \eta, t, \tau) \in T^*(X \times X \times \mathbb{R})$ with

$$(x, \xi) = (\exp tv_p)\,(y, \eta), \;\; \tau = p(x, \xi), \;\text{ and } \;\; (y, \eta) \in \Lambda_0$$

is well defined for all t and is an embedded Lagrangian submanifold of $T^*(X \times X \times \mathbb{R})$ on which H is equal to zero. Moreover, the Hamiltonian flow of $v_H = \frac{\partial}{\partial t} - v_p$ is transverse to Λ_0, so we can solve (11.31) modulo $O(\hbar^k)$ for all k be induction on k, at each stage of the induction solving a transport equation for v_H. We can also prescribe arbitrarily the initial value of this solution on the surface Λ_0 and we can choose the initial values inductively so that $\mu(x, y, 0, \hbar)$ is the Schwartz kernel modulo $O(\hbar^\infty)$. Furthermore, for fixed t, the microsupport of $\mu(x, y, t, \hbar)$ is the graph of the symplectomorphism $\exp tv_p$ and hence $\mu(x, y, t, \hbar)$ is the Schwartz kernel of a Fourier integral operator of order zero quantizing this symplectomorphism. We have achieved our first goal, namely the construction of a solution to (11.31) modulo $O(\hbar^\infty)$.

To get rid of the $O(\hbar^\infty)$, we will briefly recall how the method of "variation of constants" works, and show that it is applicable to our set-up:

Let $\mathfrak{H}$ be a a Hilbert space, Q a self-adjoint operator on $\mathfrak{H}$ and $V(t)$ a family of bounded operators on $\mathfrak{H}$ which satisfy

$$\frac{1}{i}\frac{d}{dt}V(t) = QV(t) + R(t) \tag{11.32}$$

and

$$V(0) = A. \tag{11.33}$$

To convert $V(t)$ into a solution of

$$\frac{1}{i}\frac{d}{dt}U(t) = QU(t) \tag{11.34}$$

with intial data

$$U(0) = A, \tag{11.35}$$

we note that by Stone's theorem (see ¶12.3) Q generates a one-parrameter group $\exp itQ$ of unitary operators. Using this fact, set

$$W(t) := \exp itQ \int_0^t \exp(-isQ)R(s)ds. \tag{11.36}$$

Then

$$\frac{1}{i}\frac{d}{dt}W(t) = QW(t) + R(t)$$

and $W(0) = 0$. Then $V - W$ satisfies (11.35) and 11.36).

Let us apply the formula (11.36) to our solution mod $O(\hbar^\infty)$ of (11.30) with $Q = \frac{1}{\hbar}P$. For each t this solution has microsupport in the set $p^{-1}([c,d])$ where $[c,d] \subset (a,b)$ so we can choose functions g and h in $C_0^\infty(a,b)$ such that $g = h \equiv 1$ on $[c,d]$ and $h \equiv 1$ on $\mathrm{Supp}(g)$. Multiplying the solutionwe obtained above fore and aft by $g(P)$ and $h(P)$ we get a new solution of (11.30) mod $O(\hbar^\infty)$ with the same initial data as before, namely $U(0) = f(P)$, but the remainder is now of the form $g(P)R(t)h(P)$ and the $W(t)$ in (11.36) has the form

$$\int_0^t g(P)\left(\exp i(t-s)Q\right)R(s)h(P)ds. \tag{11.37}$$

Now note that $g(P)$ and $h(P)$ are smoothing operat ors and that $\exp i(t-s)Q$ is unitary map of $L^2(X)$ into itself. Moreover, by Proposition 10.4.1, $g(P)$ and $h(P)$ have Schwartz kernels of the form

$$\sum g(\lambda_i(\hbar))\psi_i(x,\hbar)\overline{\psi}_i(y,\hbar)$$

and

$$\sum h(\lambda_i(\hbar))\psi_i(x,\hbar)\overline{\psi}_i(y,\hbar)$$

where the Ψ_i are semi-classical L^2 eigenfunctions of P.

Thus, since the Schwartz kernel of R has compact support, the expression (11.37) is well defined. Moreover, since $R(s,\hbar)$ is $O(\hbar^\infty)$ and the "$\exp i(t-s)R(s)$" factor in the integrand of (11.37) is multiplied fore and aft by operators which are smoothing and smooth as functions of $\hbar$, the integral (11.37) also has this property. This justifies our application of variation of constants.

To recapitulate: We have prove the following (main theorem) of this section:

Theorem 11.5.1. *For $f \in C_0^\infty(a,b)$ the Schwartz kernel of the operator $\left(\exp it\frac{P}{\hbar}\right)f(P)$ is an element of $I^{-n}(X \times X, Lambda)$. In particular, for all t this operator is a semi-classical Fourier integral operator quantizing the symplectomorphism $\exp tv_p$.*

Let ψ be a C^∞ function on $\mathbb{R}$ whose Fourier transform is in $C_0^\infty(\mathbb{R})$. In the next section we will compute

$$\begin{aligned}
&\operatorname{tr} \frac{1}{\sqrt{2\pi}} \int \left(\exp i \frac{t}{\hbar} P\right) f(P) \hat{\psi}(t) dt \\
&= \frac{1}{\sqrt{2\pi}} \sum_k \int e^{i \frac{t \lambda_k}{\hbar}} \hat{\psi}(t) dt f(\lambda_k) \\
&= \sum_k \psi\left(\frac{\lambda_k(\hbar)}{\hbar}\right) f(\lambda_k(\hbar)).
\end{aligned}$$

We will find that the above expression has a very interesting asymptotic expansion involving the periodic trajectories of the vector field v_{p} on the energy surface $\mathrm{p} = 0$. For this we will need to be more explicit about the phase function (in the sense of Chapter 4) of our flowout manifold.

11.5.1 The phase function for the flowout.

Let $M = T^*X$, $\alpha = \alpha_X$ the canonical one form on M. Let

$$\tilde{\alpha} = -\operatorname{pr}_1^* \alpha + + \operatorname{pr}_2^* \alpha + \tau dt$$

be the canonical one form on $M^- \times M \times T^*\mathbb{R}$. We compute the restriction of $\tilde{\alpha}$ to Λ:

Let $\iota_\Lambda : M \times \mathbb{R} \to M^- \times M \times T^*\mathbb{R}$ be the map

$$(x, \xi, t) \mapsto ((x, \xi), \exp t v_{\mathrm{p}}(x, \xi), -t, \tau).$$

This maps $M \times \mathbb{R}$ diffeomorphically onto Λ. We claim that

$$\iota_\Lambda^* \tilde{\alpha} = -\alpha + (\exp t v_{\mathrm{p}})^* \alpha + (\exp t v_{\mathrm{p}})^* \iota(v_{\mathrm{p}}) \alpha dt - \mathrm{p} dt. \tag{11.38}$$

Proof. Holding t fixed, the restriction of $\iota_\Lambda^* \tilde{\alpha}$ to $M \times \{t_0\} \sim M$ is

$$-\iota_\Lambda^* \operatorname{pr}_1^* \alpha + \iota_\Lambda^* \operatorname{pr}_2^* \alpha$$

by the definition of $\tilde{\alpha}$. But

$$\operatorname{pr}_1 \circ \iota_\Lambda = \mathrm{id}_M \quad \text{and} \quad \operatorname{pr}_2 \circ \iota_\Lambda = \exp t v_{\mathrm{p}}.$$

So the preceding expression becomes the sum of the first two terms on the right hand side of (11.38). So to verify (11.38) we need only check the value of $\tilde{\alpha}$ on the tangent vector to the flowout curve

$$t \mapsto (q, \exp t v_{\mathrm{p}}(q), -t, \mathrm{p}(q)).$$

This tangent vector is

$$\left(0, v_{\mathrm{p}}\left(\exp t v_{\mathrm{p}}(q)\right), -\frac{\partial}{\partial t}, 0\right)$$

and this accounts for the second two terms on the right hand side of (11.38). □

Now define the function $\phi \in C^\infty(M \times \mathbb{R})$ by

$$\phi := \int_0^t (\exp s v_{\mathrm{p}})^* \, \iota(v_{\mathrm{p}})\alpha \; ds - t\mathrm{p}. \tag{11.39}$$

We will now show that

$$\iota_\Lambda^* \tilde{\alpha} = d\phi. \tag{11.40}$$

Proof. $-\alpha + (\exp t v_{\mathrm{p}})^* \, \alpha = \int_0^t \frac{d}{ds} (\exp s v_{\mathrm{p}})^* \, \alpha ds$

$$\begin{aligned}
&= \int_0^t (\exp s v_{\mathrm{p}})^* \, L_{v_{\mathrm{p}}} \alpha ds \\
&= \int_0^t (\exp s v_{\mathrm{p}})^* \, d_M \iota(v_{\mathrm{p}})\alpha ds + \int_0^t (\exp s v_{\mathrm{p}})^* \, \iota(v_{\mathrm{p}}) d_M \alpha ds \\
&= \int_0^t (\exp s v_{\mathrm{p}})^* \, d_M \iota(v_{\mathrm{p}})\alpha ds + \int_0^t (\exp s v_{\mathrm{p}})^* \, (-d\mathrm{p}) ds \\
&= \int_0^t (\exp s v_{\mathrm{p}})^* \, d_M \iota(v_{\mathrm{p}})\alpha ds - d\mathrm{p} \int_0^t ds \\
&= \int_0^t (\exp s v_{\mathrm{p}})^* \, d_M \iota(v_{\mathrm{p}})\alpha ds - d\mathrm{p} \int_0^t ds \\
&= d_{M\times\mathbb{R}} \int_0^t (\exp s v_{\mathrm{p}})^* \, d_M \iota(v_{\mathrm{p}})\alpha ds \\
&\quad - \left(\frac{d}{dt} \int_0^t (\exp s v_{\mathrm{p}})^* \, \iota(v_{\mathrm{p}})\alpha ds \right) dt - t d\mathrm{p} \\
&= d_{M\times\mathbb{R}} \int_0^t (\exp s v_{\mathrm{p}})^* \, \iota(v_{\mathrm{p}})\alpha ds - ((\exp t v_{\mathrm{p}})^* \iota(v_{\mathrm{p}})\alpha) \, dt - t d\mathrm{p} \\
&= d_{M\times\mathbb{R}} \phi - ((\exp t v_{\mathrm{p}})^* \iota(v_{\mathrm{p}})\alpha) \, dt + \mathrm{p} dt
\end{aligned}$$

proving (11.40). □

11.5.2 Periodic trajectories of v_{p}.

Suppose that $t \mapsto \gamma(t)$ is a periodic trajectory of v_{p} with (least) period T so that

$$q := \gamma(0) = \gamma(T).$$

Then q is a fixed point of the map $\exp T v_{\mathrm{p}} : M \to M$. The differential of this map, i.e.

$$d \exp T v_{\mathrm{p}} : T_q M \to T_q M$$

maps the subspace $W_q \subset T_q M$ determined by $d\mathrm{p}_q = 0$ into itself and maps $v_{\mathrm{p}}(q)$, which is an element of this subspace into itself. So we get a map, the (reduced) **Poincaré map**

$$P_\gamma : W_q / \{v_{\mathrm{p}}\} \to S_q / \{v_{\mathrm{p}}\}.$$

The trajectory γ is called **non-degenerate** if

$$\det(I - P_\gamma) \neq 0.$$

Let us define

$$S_\gamma := \int_0^T \gamma^* \alpha. \tag{11.41}$$

11.5.3 The trace of the operator (11.29).

Suppose that there are only finitely many periodic trajectories, $\gamma_1, \ldots, \gamma_N$ of v_{p} lying on the energy surface $\mathrm{p} = 0$ whose periods $T_1, \ldots, T_N$ lie in the interval (a, b) and that they are all non-degenerate.

Let $\hat{\psi} \in C_0^\infty((a, b))$. The Gutzwiller trace formula asserts that the trace of the operator

$$\int_{\mathbb{R}} \hat{\psi}(t) \exp i \frac{tP}{\hbar} f(P) dt$$

has an asymptotic expansion

$$\hbar^{\frac{n}{2}} \sum_{r=1}^{N} e^{i \frac{S_\gamma}{\hbar}} \sum_{i=0}^{\infty} a_{r,i} \hbar^i.$$

Proof. Write this trace as $\int \hat{\psi}(t) \mu(x, x, t, \hbar) dt$, where

$$\hat{\psi}(t) \mu(x, y, t, \hbar) \in I^{-\frac{n}{2}}(X \times X \times \mathbb{R}, \Lambda\phi).$$

In other words, it is the integral of $\hat{\psi}(t) \mu(x, y, t, \hbar)$ over the submanifold

$$Y := \Delta_X \times \mathbb{R}$$

of $X \times X \times \mathbb{R}$. The conormal bundle of Γ of Y in $M^- \times M \times T^*\mathbb{R}$ is the set of points

$$(x, \xi, y, \eta, t, \tau)$$

satisfying

$$x = y, \quad \xi = \eta, \quad \tau = 0.$$

This intersects Λ in the set of points $(x, \xi, y, \eta, t, \tau)$ where

$$(\exp t v_{\mathrm{p}})(x, \xi) = (x, \xi), \quad \mathrm{p} = \tau = 0.$$

For $a < t < b$ this is exactly the union of the points on the periodic orbits in this interval. The non-degeneracy condition implies that Γ intersects Λ cleanly. So the Gutzwiller formula above is a special case of our abstract lemma of stationary phase, see Section 8.14.

If we write the trace of the operator $\int_{\mathbb{R}} \hat{\psi}(t) \exp i \frac{tP}{\hbar} f(P) dt$ as

$$\sum_k \int \hat{\psi}(t) e^{it \frac{\lambda_k}{\hbar}} dt \cdot f(\lambda_k(\hbar))$$

and use the Fourier inversion formula this becomes

$$\sqrt{2\pi}\sum \psi\left(\frac{\lambda_k}{\hbar}\right) f(\lambda_k(\hbar))$$

where ψ is the inverse Fourier transform of $\hat{\psi}$.

In this argument, there was nothing special about the zero level set of p We can replace P by $P - E$. So the spectrum of P determines the integrals S_γ for all non-degenerate period trajectories of v_{p}.

11.5.4 Density of states.

The **density of states formula** is a kind of degenerate version of the Gutzwiller formula. It replaces the periodic trajectories of the bicharacteristic flow $f_t := \exp tv_{\mathrm{p}}$ by fixed points of this flow:

More explicitly, let $M = T^*X$, and suppose that for $q = (x, \xi) \in M$ we have $v_{\mathrm{p}}(q) = 0$, so that q is a fixed point of f_t for all t. Let us suppose that for all $0 < t < t_0$ this fixed point is non-degenerate in the sense of Section 11.3. In other words, we assume that the graph of f_t intersects the diagonal $\Delta_M \subset M \times M$, which is equivalent to the condition that the map

$$I - (df_t)_q : \ TM_q \to TM_q$$

is bijective. Let us also suppose that q is the only fixed point of f_t on the energy surface

$$\mathrm{p} = c \quad \text{where} \quad c := \mathrm{p}(q).$$

We can apply the results of Section 11.3 to the Fourier integral operator

$$F_t = \exp \frac{itP}{\hbar}\rho(P), \tag{11.42}$$

where $\rho \in C_0^\infty)\mathbb{R})$ is supported on a small neighborhood of c and is identically one on a still smaller neighborhood. For this choice ρ, the microsupport of F_t intersects Δ_M only at q, Since this intersection is transversal, there is only one summand in (11.8) so (11.8) gives the asymptotic expansion

$$\mathrm{tr}\left(\exp \frac{itP_\hbar}{\hbar}\rho(P_\hbar)\right) = \hbar^{\frac{n}{2}} e^{\frac{i\pi}{\sigma_q}} a_q(\hbar, t) e^{\frac{iT_q^\sharp}{\hbar}} \tag{11.43}$$

where σ_q is a Maslov factor and $T_q^\sharp = \psi(q, t)$ where ψ is defined by the identity (11.19):

$$\alpha_X - f_t^* \alpha_X = d\psi.$$

Since $v_{\mathrm{p}}(q) = 0$, we read off from (11.39) and (11.40) that

$$\psi(q, t) = -t\mathrm{p}(q).$$

Hence from (11.41) we obtain, for $0 < t < t_0$ the **density of states formula**

$$\operatorname{tr}\left(\exp \frac{itP_\hbar}{\hbar}\rho(P_\hbar)\right) = \hbar^{\frac{n}{2}} e^{\frac{i\pi}{\sigma_q}} a_q(\hbar,t)e^{-it\mathrm{p}(q)}. \tag{11.44}$$

Moreover, by (11.23)

$$a_q(0,t) = |\det(I-(df_t)_q|^{\frac{1}{2}}\,. \tag{11.45}$$

We also note that since the left hand side of (11.44) depends smoothly on t, so does $a_q(\hbar,t)$.

11.6 The Donnelly theorem.

Let X be a compact manifold, M its cotangent bundle and P_h a zero$^{\text{th}}$ order self-adjoint elliptic pseudodifferential operator on X. Then for $\rho \in \mathcal{C}_0^\infty$, $\rho(P_h)$ is a zero$^{\text{th}}$ order pseudodifferential operator with compact microsupport. Hence, given a $\mathcal{C}^\infty$ mapping, $f: X \to X$ the operator

$$F = f^*\rho(P_h)$$

is, as we showed in §8.10, a semi-classical Fourier integral operator quantizing the canonical relation, Γ_f, where

$$(x,\xi,y,n) \in \Gamma_f \Leftrightarrow y = f(x) \;\text{ and }\; \xi = df_x^t n\,. \tag{11.46}$$

Therefore if Γ_f intersects Δ_M cleanly we get for the trace of F an asymptotic expansion of the form (11.17). This expansion can also be derived more directly by simply applying stationary phase to the integral (11.10). (Moreover, this approach gives one a lot more information about the individual terms in this asymptotic expansion.)

The details: Let $p(x,\xi)$ be the leading symbol of P. Then the Schwartz kernel of $\rho(P_n)$ is given locally by an oscillatory integral having an asymptotic expansion in powers of h:

$$(2\pi h)^{-d}\sum_{k=0}^{\infty} h^k \int a_{\rho,k}(x,\xi)e^{\frac{i(x-y)\cdot\xi}{h}}\,d\xi \tag{11.47}$$

where

$$a_{\rho,k}(x,\xi) = \sum_{\ell\le 2k} b_{k,\ell}(x,\xi)\left(\left(\frac{d}{ds}\right)^{\ell}\rho\right)(p(x,\xi)) \tag{11.48}$$

and the leading order term in (11.48) is given by $a_{\rho,0} = \rho(p(x,\xi))$. Hence

$$\operatorname{tr} f^*\rho(P_h) \sim (2\pi h)^{-d}\int a_\rho(f(x),\xi,h)e^{\frac{i(f(x)-x)\cdot\xi}{h}}\,dx\,d\xi\,. \tag{11.49}$$

Let's now apply the lemma of stationary phase to the integral (11.49) with phase function

$$\psi(x,\xi) = (f(x)-x)\cdot\xi\,. \tag{11.50}$$

To do so we have to compute C_ψ. But

$$\frac{\partial \psi}{\partial \xi} = 0 \Leftrightarrow x = f(x) \quad \text{and} \quad \frac{\partial \psi}{\partial x} = 0 \Leftrightarrow (df_x - I) \cdot \xi = 0\,. \tag{11.51}$$

Thus C_ψ is just the set (11.46). The method of stationary phase requires that C_ψ be a submanifold of T^*X and that, for $(x,\xi) \in C_\psi$, the Hessian

$$(d^2\psi)_{x,\xi}|N_{x,\xi}C_\psi$$

be non-degenerate, and it is easy to see that these conditions are satisfied if the fixed point set X_f of f is a submanifold of X and if the restriction map $T^*X|X_f \to T^*X_f$ maps (11.46) bijectively onto T^*X_f. Finally, to compute the leading order term in the asymptotic expansion of (11.49) using stationary phase, one has to compute the determinant of the quadratic form (11.46). But

$$\frac{\partial^2 \psi}{\partial \xi^2} = 0 \quad \text{and} \quad \frac{\partial^2 \psi}{\partial \xi \partial x} = \frac{\partial f}{\partial x} - I\ ,$$

so

$$d^2\psi_{x,\xi} = \begin{bmatrix} 0 & \frac{\partial f}{\partial x} - I \\ \frac{\partial f}{\partial x} - I & \cdots \end{bmatrix}$$

and hence

$$\det(d^2\psi_{x,\xi}|N_{x,\xi}C_\psi) = -\left(\det\left(\frac{\partial f}{\partial x} - I\right) \mid N_x X_f\right)^2 .$$

Note also that $\operatorname{sgn} d^2\psi_{x,\xi}|N_{x,\xi}C_\psi = 0$ and $\psi|C_\psi = ((f(x) - x) \cdot \xi)|C_\psi = 0$ by (11.49). Feeding these data into the stationary phase expansion of the integral (11.49) and noting that $a_\rho(x,\xi,0) = \rho(p(x,\xi))$, we get the following variant of Donnelly's theorem.

Theorem 11.6.1. *Let X_i, $i = 1, \ldots, N$, be the connected components of X_f and let $d_i = \dim X_i$. Then*

$$\operatorname{trace} f^*\rho(P_h) \sim \sum (2\pi h)^{-d_i} \sum_{k=0}^{\infty} a_{k,i} h^k.$$

Moreover,

$$a_{0,i} = \int_{T^*X_i} \rho(p(x,\xi)|D(x)|^{-1})\, dx\, d\xi$$

where $dx\, d\xi$ is the symplectic volume form and $D(x) = \det(df_x - I|N_x X_i)$.

Remark:

If we take P_h to be $-h^2\Delta_X$ and $\rho(s)$ to be the function e^{-s}, $s > 0$, (which takes a little justifying since this ρ is not in $\mathcal{C}_0^\infty$), then this theorem reduces to Donnelly's theorem (with h^2 playing the role of t).

Chapter 12

Integrality in semi-classical analysis.

12.1 Introduction.

The semi-classical objects that we have been studying in the last four chapters can be thought of from the symplectic perspective as the quantizations of objects and morphisms in the *exact* symplectic "category". Recall that in this category an object is an exact symplectic manifold, which is a manifold M with a one form α such that $\omega = -d\alpha$ is symplectic. The point morphisms pt. $\to M$, associated with this object are pairs (Λ, ϕ) where $\Lambda \subset M$ is a Lagrangian submanifold and ϕ is a C^∞ function on Λ such that

$$\iota_\Lambda^* \alpha = d\phi. \tag{12.1}$$

If M_1 and M_2 are exact symplectic manifolds, a morphism of M_1 into M_2 is a point morphism of pt. to $M_1^- \times M_2$.

We discussed these categorical issues in Chapter 4. In particular, we showed in §4.13.5 that this category sits inside a slightly larger category: the integral symplectic category. In this "category" the objects are the same as above, but the point morphisms pt. $\to M$ are pairs (Λ, f) where $f : \Lambda \to S^1$ is a C^∞ map that satisfies, as a substitute for (12.1) the equation

$$\iota_\Lambda^* \alpha = \frac{1}{2\pi i} \frac{df}{f}. \tag{12.2}$$

We can view (12.1) as a special case of (12.2) by setting $f = e^{2\pi i \phi}$.

One can show that if our exact symplectic manifolds are cotangent bundles, then most of the semi-classical results that we obtained in the last four chapters can be formulated as results in this larger category. Namely the functions f can always be written locally as

$$f = e^{2\pi i \phi}, \quad \phi \in C^\infty(\Lambda). \tag{12.3}$$

Therefore, since the functions and operators that we have been dealing with in the last four chapters have been defined by first defining them locally, and then extending the local definitions into global definitions via partitions of unity, we can do exactly the same same thing with exact Lagrangian manifolds replaced by integral Lagrangian manifolds. But there is a hitch: the function ϕ in (12.3) is not unique. It is only defined up to an additive constant $c \in \mathbb{Z}$. So if we attach to ϕ oscillatory integrals with phase factor $e^{\frac{2\pi i \phi}{\hbar}}$, these oscillatory integrals will only be well defined modulo factors of the form $e^{\frac{2\pi i c}{\hbar}}$. There is a simple way out - namely, to impose on $\hbar$ the constraint

$$\hbar = \frac{1}{m}, \quad m \in \mathbb{Z}. \tag{12.4}$$

This is the approach we will take below. Note for such $\hbar$, we have $e^{\frac{2\pi i c}{\hbar}} = 1$.

Our motives for introducing these integrality complications into semi-classical analysis will become clearer later in this chapter. We will see in the discussion of concrete examples, that the functions and operators we will use only become well defined if we impose the integrality condition (12.4). What we can say at this point, however, is that these examples, for the most part, have to do with actions of Lie groups on manifolds.

For instance, suppose that X is a manifold and $\pi : P \to X$ a circle bundle. We will show that if A is a classical pseudo-differential operator on P which commutes with the action of S^1 on $C^\infty(P)$, then one can think of A as a semi-classical operator $A_\hbar$ on X, but this operator is only well-defined if $\hbar$ satisfies (12.4).

Or, to cite a second example, suppose that G is a compact Lie group and ρ_m the irreducible representations of G with highest weight $m\beta$. We will show that if $\gamma_m \in C^\infty(G)$ is the character of this representation, the γ_m's define an oscillatory function'$\gamma_\hbar, \hbar = 1/m$ living micro-locally on $\Lambda_O \subset T^*G$ where O is the co-adjoint orbit in $\mathfrak{g}^*$ containing β and Λ_O its character Lagrangian. Thus, in this example too, $\gamma_\hbar$ is only defined when $\hbar$ satisfies (12.4)

Here, as a road map, is a brief outline of the contents of this chapter:

In §12.2 we review standard facts about line bundles and connections. We will need this material in order to explain in detail the correspondence between classical and semi-classical pseudodifferential operators in the example we alluded to above.

In §12.3 we will discuss "integrality" in De Rham theory. In particular we will describe its implications for cohomology classes $[c]$ in $H^*_{DR}(X)$ in dimension s one and two. For instance, we will show that if Λ is a Lagrangian submanifold of the exact symplectic manifold (M, α), the integrality of $\iota^*_\Lambda \alpha$ in the DeRham theoretic sense in just the integrality condition (12.2).

In §12.4 we will review the results of §4.13.5 on integrality in symplectic geometry and discuss some examples of integral Lagrangian submanifolds that we will encounter later in the chapter.

In §12.5 and §12.6 we will develop the symplectic machinery that we will need for applications to group actions that we alluded to above. In particular we study the notion of "symplectic reduction" of the "moment Lagrangian" and of the "character Lagrangian". The first two of these topics were briefly discussed in Chapter 4. We will discuss them in more detail here.

These five sections constitute the "symplectic half" of the this chapter. In the remaining six sections we discuss the semi-classical applications of this material.

In §12.7 we will amplify on what we said above about the semi-classical oscillatory functions and operators associated with integral Lagrangian submanifolds.

In §12.8 and §12.9 we discuss our semi-classical formulation of the theory of characters for representations of compact Lie groups. Our goal in these two sections will be to show that the two classical character formulas for compact Lie groups: the Weyl character formula and the Kirillov character formula are special cases of a more general result, a character formula due to Gross-Kostant-Ramond-Sternberg and to show that the machinery of semi-classical analysis: half-densities, Maslov factors, etc. makes these formulas more transparent.

In §10 we will elaborate on the remark above abut classical pseudodifferential operators on a circle bunlde $P \to X$, i.e. that such operators can be viewed as semi-classical pseudodifferential operators on X

In §11 and §12 we will state and prove the main result of this chapter: an equivariant version of the trace formula that we proved in Chapter 10. In §11 we will prove the S^1 version of this theorem and in §12 use the "character theorems" of §12.8 and §12.9 to extend this result to arbitrary compact Lie groups.

12.2 Line bundles and connections.

Connections, connection forms, and curvature.

Let $\mathbb{L} \to X$ be a complex line bundle over a smooth real manifold. A linear first order differential operator $\nabla : C^\infty(\mathbb{L}) \to C^\infty \mathbb{L} \otimes T^*X)$ is called a **connection** if it satisfies

$$\nabla(fs) = f\nabla s + s \otimes df, \quad \forall s \in C^\infty(\mathbb{L}),\ f \in C^\infty(M). \tag{12.5}$$

If $U \subset X$ is open, and $s : U \to \mathbb{L}$ vanishes nowhere, define the one form $\alpha(s)$ by

$$\alpha(s) := \frac{1}{2\pi i}\frac{\nabla(s)}{s}. \tag{12.6}$$

By (12.5) we have

$$\alpha(fs) = \alpha(s) + \frac{1}{2\pi i}\frac{df}{f} \tag{12.7}$$

for non-vanishing functions f. It follows from (12.7) that ω defined by

$$\omega_U := d\alpha(s) \tag{12.8}$$

is independent of the choice of s and hence is globally defined. From its definition it is clear that $d\omega = 0$. ω is called the **curvature form** of ∇. Its cohomology class is independent of the choice of ∇ and is called the **Chern class** of $\mathbb{L}$.

The condition that the curvature form be real valued.

In general ω could be complex valued, but we suppose that we make the assumption that ω is real valued. It follows from (12.8) that $\mathrm{Im}\,(\alpha(s))$ is closed, and hence if U is simply connected that there is a real valued function h on U with $\mathrm{Im}\,\alpha(s) = dh$. By (12.7)

$$\alpha(e^{-2\pi h}s) = \alpha(s) - \frac{1}{i}\mathrm{Im}\,\alpha(s)$$

which is real. So (with a change in notation) we may assume that all our trivializing sections have the property that $\alpha(s)$ is real. We now examine some consequences of this property.

Let $\mathbb{U} = \{U_\iota\}$ be a good cover (meaning that all intersections are contractible) with trivializing sections s_ι such that all the $\alpha(s_\iota)$ are real. If

$$f_{jk} \in C^\infty(U_j \cap U_k) \tag{$*$}$$

are such that

$$s_j = f_{jk}s_k,$$

then it follows from the reality of the $\alpha(s_\iota)$ and (12.7) that

$$d\left(\mathrm{Im}\,\left(\frac{1}{2\pi i}\log(f_{jk})\right)\right) = 0$$

and since $U_j \cap U_k$ is contractable, that

$$c_{jk} := \mathrm{Im}\,\left(\frac{1}{2\pi i}\log(f_{jk})\right)$$

are constants. Since $f_{jk}f_{k\ell}f_{\ell j}s_j = s_j$ on $U_{jk\ell} := U_j \cap U_k \cap U_\ell$ it follows that $f_{jk}f_{k\ell}f_{\ell j} \equiv 1$ on $U_{jk\ell}$. Hence

$$c_{jk} + c_{k\ell} + c_{\ell j} = \mathrm{Re}\,\log(f_{jk}f_{k\ell}f_{\ell j}) = \mathrm{Re}\,\log 1 = 0 \quad \text{ on } U_{jk\ell}.$$

Thus the c_{jk} define a Čech one cycle.

If X is simply connected, this cocycle is a coboundary, so that there exist constants c_j such that if $U_{jk} := U_j \cap U_k \neq \emptyset$,

$$c_{jk} = c_j - c_k.$$

So if we modify our trivializing sections by replacing s_j by $e^{2\pi c_j} s_j$, we see that we obtain trivializing sections such that the corresponding transition functions satisify

$$|f_{jk}| \equiv 1. \tag{12.9}$$

This allows us to define a Hermitian inner product on $\mathbb{L}$ by defining, for any section s of $\mathbb{L}$ and any U_j

$$\langle s, s\rangle|_{U_j} := |s/s_j|^2. \tag{12.10}$$

Suppose that s is a non-vanishing section of $\mathbb{L}$ such that $\alpha(s)$ is real and we have a Hermitian metric such that $\langle s, s\rangle \equiv 1$. Thus $d\langle s, s\rangle = 0$. On the other hand,

$$\langle \nabla s, s\rangle + \langle s, \nabla s\rangle = (2\pi i)(\alpha(s) - \overline{\alpha(s)})$$

so if $\alpha(s)$ is real, we have

$$\langle \nabla s, s\rangle + \langle s, \nabla s\rangle = d\langle s, s\rangle,$$

since both sides vanish. By (12.7) this equality extends to all sections. Indeed, if we have a section of the form fs then $\langle fs, fs\rangle = |f|^2$ so $d\langle fs, fs\rangle = fd\overline{f} + \overline{f}df$. On the other hand, from (12.7) we have

$$\langle \nabla(fs), fs\rangle + \langle fs, \nabla(fs)\rangle = |f|^2(\langle \nabla s, s\rangle + \langle s, \nabla s\rangle) + \overline{f}df + f\overline{df} = d|f|^2.$$

So we have

$$d\langle u, u\rangle = \langle \nabla u, u\rangle + \langle u, \nabla u\rangle \tag{12.11}$$

for any section u of the form fs. Conversely, suppose that there is a Hermitian metric on $\mathbb{L}$ for which (12.11) holds for all sections. We may choose our trivializing sections s_j to satisfy $\langle s_j, s_j\rangle \equiv 1$, and then conclude that the $\alpha(s_j)$ are real. Of course, if we have trivializing sections such that all the $\alpha(s_j)$ are real, then it follows from (12.8) that the curvature ω is real.

In the case that X is simply connected, and our trivializing sections s_j all have the property that $\alpha_(s_j)$ is real, then for the Hermitian metric given by (12.10), equation (12.11) holds for all sections of $\mathbb{L}$.

The meaning of $\omega = 0$.

Having examined the implications of "ω is real valued" we next examine the implications of the much stronger assumption "$\omega = 0$". This assumption implies that for every trivializing section, $s : U \to \mathbb{L}$, $\alpha(s)$ is closed. Hence if U is simply connected $\alpha(s) = -dh$ for some function $h \in \mathcal{C}^\infty(U)$, and if we replace s by $e^{2\pi i h_s}$ this modified trivializing section satisfies $\triangledown s = 0$. In other words s is an "autoparallel" section of $\mathbb{L}|U$. Suppose now that, as above, $\mathbb{U} = \{U_i\,,\ i = 1, 2, \ldots\}$ is a good cover of X and $s_i : U_i \to \mathbb{L}$ trivializing autoparallel sections of $\mathbb{L}|U_i$. Then the transition functions that we defines above are constants, and, as above, the constants

$$c_{i,j} = \frac{1}{2\pi\sqrt{-1}} \log f_{i,j}$$

define a Čech cocycle in $\check{C}^1(\mathbb{U},\mathbb{R})$. Thus if this cocycle is a coboundary, i.e. if $c_{i,j} = c_i - c_j$ then

$$e^{-2\pi\sqrt{-1}c_i} s_i = e^{-2\pi\sqrt{-1}c_j} s_j \,.$$

In other words these manifold sections patch together to give a global trivializing section of $\mathbb{L}$ with the property, $\nabla s = 0$. Thus, to summarize, we've proved

Theorem 12.2.1. *If X is simply connected and* $\operatorname{curv}(\nabla) = 0$ *there exists a global trivializing section, s, of $\mathbb{L}$ with $\nabla s = 0$.*

Functorial properties of line bundles and connections.

Recall that if Y is a manifold and $\gamma : Y \to X$ a $\mathcal{C}^\infty$ map then one can define a line bundle $\gamma^*\mathbb{L}$ on Y by defining its fiber $\gamma^*\mathbb{L}$ at every point $p \in Y$ to be the fiber, $\mathbb{L}_q$ of $\mathbb{L}$ at the image point $\gamma(p) = q$. Thus if $s : X \to \mathbb{L}$ is a section of $\mathbb{L}$ the composite, $s \circ \gamma$, of the maps, $\gamma : Y \to X$ and $s : X \to \mathbb{L}$ can be viewed as a section of $\gamma^*\mathbb{L}$ and this give one a pull-back operation

$$\gamma^* : \mathcal{C}^\infty(\mathbb{L}) \to \mathcal{C}^\infty(\gamma^*\mathbb{L})\,.$$

By combining this with the pull-back operation on forms: $\gamma^* : \Omega^1(X) \to \Omega^1(Y)$ we get a pull-back operation

$$\gamma^* : \mathcal{C}^\infty(\mathbb{L} \otimes T^*X) \to \mathcal{C}^\infty(\gamma^*\mathbb{L} \otimes T^*Y)$$

and it is easily checked that there is a unique connection, $\gamma^*\Delta$, on $\gamma^*\mathbb{L}$ which is compatible with these two pull-back operations, i.e. satisfies

$$\gamma^*(\nabla s) = \gamma^*\Delta(\gamma^* s)\,. \tag{12.12}$$

Moreover by (12.8) the curvature form of this connection is

$$f^*\omega. \tag{12.13}$$

One elementary application of these functioriality remarks is the following. Suppose Y is just an open subinterval of the real line. Then $\gamma^*\omega = 0$, so by the theorem above the line bundle $\gamma^*\mathbb{L}$ has an autoparallel trivialization. In particular for $a, b \in I$, elements of $\mathbb{L}_p$ at $p = \gamma(a)$ can be identified with elements of $\mathbb{L}_q$ at $q = \gamma(b)$ by "parallel transport along γ".

More generally, if Y is any simply-connected manifold and $\gamma^*\omega = 0$, then the same is true for it: $\gamma^*\mathbb{L}$ has a canonical parallel trivialization. (For instance this is the case if ω is a symplectic form, Y a Lagrangian submanifold of X, and $\gamma : Y \to X$ the inclusion map.)

Line bundles and circle bundles.

We'll conclude this brief review of the theory of connections by describing an alternative way of thinking about line bundle–connection pairs. Let's assume that ω is real and that $\mathbb{L}$ has an intrinsic "autoparallel" Hermitian inner product $\langle\, , \rangle$. If $U \subset X$ is a simply connected open set and $s : U \to \mathbb{L}$ a trivializing section we can assume without loss of generality, that $\langle s, s\rangle = 1$ on U. Thus if we let $P \subseteq \mathbb{L}$ be the circle bundle

$$\{(p, v)\,;\, p \in X\,,\, v \in \mathbb{L}_p\,,\, \langle v, v\rangle_p = 1\}\,,$$

we can view s as being a trivialization section

$$s : U \to P\,. \tag{12.14}$$

Now let $\frac{\partial}{\partial\theta}$ be the infinitesimal generator of the circle action on P. We claim

Theorem 12.2.2. *There exists a unique real-valued one-form, $\alpha \in \Omega^1(P)$, such that*

(i) $\alpha\left(\frac{\partial}{\partial\theta}\right) = \frac{1}{2\pi}$
and

(ii) For all sections, (12.14), of P, α has the reproducing property

$$s^*\alpha = \alpha(s) \tag{12.15}$$

Proof. The trivializing section (12.14) gives one a bundle isomorphism

$$P \simeq U \times S^1$$

and if α has this property it's clear that it has to correspond to the one-form: $\alpha(s) + \frac{d\theta}{2\pi}$. Thus, if an α exists, it has to be unique, and to show that it exists it suffices to show that the form above: $\alpha(s) + \frac{d\theta}{2\pi}$, has properties (i) and (ii) on $U \times S^1$. However if we replace the section (5) by $\tilde{s} = e^{2\pi i h}s$, h being any real-valued $\mathcal{C}^\infty$ function, then

$$(\tilde{s})(\alpha(s) + \frac{d\theta}{2\pi}) = \alpha(s) + dh = \alpha(\tilde{s})$$

by 3.

□

Remarks

1. Since the form $\alpha(s) + \frac{d\theta}{2\pi}$ is S^1 invariant and the identification, $P|U \simeq U \times S^1$ is an S^1-equivariant identification. The form α itself is an S^1 invariant form. In particular,

$$\iota\left(\frac{\partial}{\partial\theta}\right) d\alpha = L_{\frac{\partial}{\partial\theta}}\alpha - d\left(\frac{\partial}{\partial\theta}\right)\alpha = 0\,. \tag{12.16}$$

2. From property (ii) one gets the identity

$$\nabla s = \sqrt{-1} s \otimes s^* \alpha \tag{12.17}$$

which can be viewed as an alternative way of defining ∇ in terms of α.

3. Let π be the projection, $P \to X$. Using the identity (12.15) one can rewrite the identity: $ds^*\alpha = d\alpha(s) = \omega$, more intrinsically in the form

$$\pi^*\omega = d\alpha\,. \tag{12.18}$$

(Notation: We will henceforth refer to α as the connection form of the connection, ∇.)

4. Of particular interest for us will be examples of line bundle–connection pairs, $(\mathbb{L}, \nabla)$ for which the curvature form, ω, is symplectic, i.e. for which (X, ω) is a symplectic manifold and $(\mathbb{L}, \nabla)$ is a "pre-quantization" of this manifold. In this case α is a contact form on P, i.e. for $2m = \dim X$ the $2m + 1$-form, $\alpha \wedge (d\alpha)^m$ is nowhere vanishing. Moreover, one gets from α an exact symplectic form

$$\omega^\# = d(t\alpha)\,, \quad t \in \mathbb{R}^+ \tag{12.19}$$

on the product, $P \times \mathbb{R}^+$. Denoting by $\mathbb{L}^\#$ the complement of the zero section in $\mathbb{L}$ one gets a natural identification

$$P \times \mathbb{R}^+ \simeq \mathbb{L}^\#\,,\ (x, v, t) \mapsto (x, tv)$$

via which we can think of $\omega^\#$ as being an exact symplectic form, $\omega^\# = d\alpha^\#$, $\alpha^\# = t\alpha$, on $\mathbb{L}^\#$. In particular, $\mathbb{L}^\#$ is the *symplectic cone* associated with the contact manifold, (P, α).

12.3 Integrality in DeRham theory.

A coholomogy class, $c \in H^k(X, \mathbb{R})$ is *integral* if it is in the image of the map $H^k(X, \mathbb{Z}) \to H^k(X, \mathbb{R})$, mapping cohomology classes with *integer* coefficients into cohomology classes with *real* coefficients. In this section we will describe the implications of this integrality property in degrees $k = 1$ and $k = 2$.

We begin with the case $k = 1$: Suppose $\alpha \in \Omega^1(X)$ is a closed one-form with $[\alpha] = c$. Let $\mathbb{U} = \{U_i, i = 1, 2, \dots\}$ be a good cover of X. Then, for every U_i, there exists a function, $h_i \in C^\infty(U_i)$ with the property

$$dh_i = \alpha \tag{12.20}$$

and hence on overlaps, $U_i \cap U_j$, there exists constants $c_{i,j}$ satisfying

$$h_i|U_i \cap U_j - h_j|U_i \cap U_j = c_{i,j}\,. \tag{12.21}$$

Moreover if $U_i \cap U_j \cap U_k$ is non-empty

$$c_{i,j} + c_{j,k} + c_{k,i} = 0$$

and hence the $c_{i,j}$'s define a Čech cocycle $\check{c} \in \check{C}^1(\mathbb{U}, \mathbb{R})$; and the correspondence, $\alpha \to \check{c}$, gives rise, at the level of cohomology, to the standard isomorphism, $H^1_{DR}(X) \to H^1(X, \mathbb{R})$. Suppose now that c is an integral Čech cocycle, i.e. $c_{i,j\in\mathbb{Z}}$. Then by equation (12.21)

$$e^{2\pi\sqrt{-1}h_i} = e^{2\pi\sqrt{-1}h_j}$$

on $U_i \cap U_j$, so these functions define a map $f : X \to S^1$ whose restriction to U_i is $e^{2\pi\sqrt{-1}h_i}$ and hence by (12.20)

$$\alpha = \frac{1}{2\pi i} \frac{df}{f} . \tag{12.22}$$

In other words we've proved (most of) the following assertion.

Theorem 12.3.1. *A cohomology class, $c \in H^1(X, \mathbb{R})$ is integral iff it has a DeRham representative of the form*

$$\alpha = \frac{1}{2\pi} f^* \, d\theta \tag{12.23}$$

where f is a map of X into S^1 and θ the standard angle variable on S^1.

Let us now turn to to the slightly more complicated problem of deciphering the implications of integrality for cohomology classes, c, in $H^2(X, \mathbb{Z})$. If $\omega \in \Omega^2(X)$ is a closed 2-form representing this class, the Čech cocyle corresponding to ω can be constructed by a sequence of operations similar to (12.20)–(12.22). Namely let

$$\omega | U_i = d\alpha_i \, , \quad \alpha_i \in \Omega^1(U_i) \, , \tag{12.24}$$

and on $U_i \cap U_j$ let

$$\alpha_i = \alpha_j = dh_{i,j} \tag{12.25}$$

where $h_{i,j} = -h_{j,i}$ is in $\mathcal{C}^\infty(U_i \cap U_j)$. Then by (12.25)

$$d(h_{i,j} + h_{j,k} + h_{k,i}) \quad = \quad \alpha_i - \alpha_j + \alpha_j - \alpha_k + \alpha_k - \alpha_i = 0 \, .$$

so

$$c_{i,j,k} \quad = \quad h_{i,j} + h_{j,k} + h_{k,i} \tag{12.26}$$

is a constant. Moreover from this identity it is easy to see that the Čech cochain, $\check{c} \in \check{C}^2(\mathbb{U}, \mathbb{R})$, defined by the $c_{i,j,k}$'s satisfies $\delta c(i,j,k,\ell) = c(j,k,\ell) - c(i,k,\ell) + c(i,j,\ell) - c(i,j,k) = 0$ and hence is a cocycle. Moreover, as above the correspondence $\omega \to \check{c}$, defines, at the level of cohomology, the standard isomorphism, $H^2_{DR}(X) \to H^2(X, \mathbb{R})$.

Suppose now that $\check{c}$ is an integral cocyle, i.e. the $c_{i,j,k}$'s are integers. Then, letting $f_{i,j} = e^{2\pi\sqrt{-1}h_{i,j}}$, one gets from (12.26) the identities

$$f_{i,j}f_{j,k}f_{k,i} = 1\,, \tag{12.27}$$

and it is easy to see from these identities that the $f_{i,j}$'s are transition functions for a line bundle $\mathbb{L} \to X$. Indeed this line bundle can be defined explicitly as the union:

$$\tilde{\mathbb{L}} = \sqcup_i U_i \times \mathbb{C} \tag{12.28}$$

modulo the identifications:

$$(x, c_i) \sim (x, c_j) \Leftrightarrow c_i = f_{i,j}(x)c_j \tag{12.29}$$

for x's on the overlap $U_i \cap U_j$. Moreover the maps

$$\tilde{s}_i : U_i \to \tilde{\mathbb{L}}\,, \quad x \to (x, 1)$$

define trivializing sections, s_i of $\mathbb{L}$, and these have the $f_{i,j}$'s as their associated transition functions. In addition one can define a connection, ∇, on $\mathbb{L}$ by setting

$$\frac{1}{2\pi\sqrt{-1}}\,\frac{\nabla s_i}{s_i} = \alpha_i \tag{12.30}$$

where the α_i's are the α_i's in (12.24)–(12.25) and by (12.24) the curvature form of this connection is ω. Thus we've proved (most of) the following assertion.

Theorem 12.3.2. *If $c \in H^2(X, \mathbb{R})$ is an integral cohomology class there exists a line bundle connection pair $\mathbb{L}, \nabla$ with $c = [\mathrm{curv}(\nabla)]$.*

Remarks

1. One can define a Hermitian inner product on $\mathbb{L}$ by requiring that the s_i's above satisfy $\langle s_i, s_j \rangle \equiv 1$ on U_i.

2. This theorem is a key ingredient in the proof of the following purely topological result.

Theorem 12.3.3. *There is a bijection between $H^2(X, \mathbb{Z})$ and the set of equivalence classes of complex line bundles on X.*

We won't prove this result here but a nice proof of it can be found in [Weil].

12.4 Integrality in symplectic geometry.

In Chapter 4 we defined an *exact symplectic manifold* to be a pair (M, α) consisting of a symplectic manifold (M, ω) and a one-form, α, for which $\omega = d\alpha$. We also defined an *exact Lagrangian submanifold* of (M, α) to be a pair, (Λ, φ) consisting of a Lagrangian submanifold, Λ of M and a real-valued function

$\varphi \in \mathcal{C}^\infty(\Lambda)$ for which $\iota_\Lambda^*\alpha = d\varphi$. These were the building blocks of the "exact symplectic category" that we discussed in §4.13. In this category the (M,α)'s played the role of objects, the categorical points of (M,α) were its exact Lagrangian submanifolds; and given two objects (M_1,α_1) and (M_2,α_2) we defined the morphisms between them to be the categorical points of the product manifold

$$M \quad = \quad M_1^- \times M_2 \tag{12.31}$$

equipped with the one-form

$$\alpha \quad = \quad -(pr_1)^*\alpha_1 + pr_2^*\alpha_2\,. \tag{12.32}$$

Recall from §4.13.5 that this category sits inside a slightly larger category which, for lack of a better term, we called the *integral* symplectic category. In this category the objects are the same as above: Exact symplectic manifolds: (M,α). However morphisms between two objects (M,α_1) and (M_2,α_2) are be pairs (Γ, f) where Γ is a Lagrangian submanifold of the product (1) and f a $\mathcal{C}^\infty$ map, $\Gamma \to S^1$ satisfying

$$\iota_\Gamma^*\alpha = \frac{1}{2\pi i}\,\frac{df}{f}\,. \tag{12.33}$$

Thus if (Γ,α) is a morphism in the exact symplectic category we can convert it into a morphism in this category by setting $f = e^{2\pi\varphi i}$. Note that the forms $\iota_\Gamma^*\alpha$ are *integral* one-forms (this being our reason for calling this the "integral" symplectic category). Also as in Chapter 4 the term, category, continues to mean "category-in-quotations marks". To compose morphisms $\Gamma_1 M_2 \to M_3$ and $\Gamma_2 : M_2 \to M_3$ we will have to assume that they are cleanly composible in the sense of §4.2 and in particular that the map defined by :

$$\kappa : \Gamma_2 * \Gamma_1 \to \Gamma_2 \circ \Gamma_1$$

is a smooth fibration with connected fibers. Assuming this we defined the composition operation for morphisms (Γ_1, f_1) and (Γ_2, f_2) in this new category in more or less the same way as in §4.12. We will simply replaced the composition law in the exact symplectic category by the composition law

$$(\Gamma_1, f_1) \circ (\Gamma_2, f_2) = (\Gamma, f) \tag{12.34}$$

if $\Gamma = \Gamma_2 \circ \Gamma_1$ and

$$K^* f = \rho_1^* f_1 \rho_2^* f_2\,. \tag{12.35}$$

Thus by this composition law our recipe for converting an exact canonical relation, (Γ,α), into an integral canonical relation, (Γ, f) by letting $f = e^{2\pi i\varphi}$, defines an imbedding of the exact symplectic category into the integral symplectic category.

Given an exact symplectic manifold (M,α) its "categorical point": the morphisms, $pt \to M$, are by definition pairs, (Λ, f) where Λ is a Lagrangian submanifold of M and f a map of M into S^1 satisfying $\iota_\Lambda^*\alpha = \frac{1}{2\pi i}\,\frac{df}{f}$. We'll devote the rest of this section to describing some examples of such point-morphisms (example which will resurface in the last couple of sections of this chapter).

Example 1.

At the end of §12.2 we showed that if (X, ω) is an integral symplectic manifold and $(\mathbb{L}, \nabla)$ a pre-quantization of X, we get an exact symplectic manifold $(L^{\#}, \alpha^{\#})$ by deleting the zero section from $\mathbb{L}$. Moreover, if P is the unit circle bundle in $\mathbb{L}$ and $\alpha \in \Omega^1(P)$ the connection form then, via the identification $\mathbb{L}^{\#} = P \times \mathbb{R}^+$, $\alpha^{\#}$ becomes the one-form, $t\alpha$. Now let $\Lambda \subset X$ be a Lagrangian submanifold and $\iota_\Lambda : \Lambda \to X$ the inclusion map. Using the functorial properties of line bundles described in §12.2 one gets a line bundle with connection on Λ

$$\mathbb{L}_\Lambda = \iota_\Lambda^* \mathbb{L} \quad \text{and} \quad \nabla_\Lambda = \iota_\Lambda^* \nabla \,.$$

Moreover, by the functorial property (12.13) of the curvature form

$$\operatorname{curv}(\nabla_\Lambda) = \iota_\Lambda^* \omega = 0$$

since Λ is Lagrangian. Thus if π is the projection map of P onto X, and $\Lambda^{\#} = \pi^{-1}(\Lambda)$ then

$$\iota_{\Lambda^\sharp}^* \, d\alpha = \pi^* \iota_\Lambda^* \omega = 0$$

so $\iota_\Lambda^* \alpha$ is closed.

Definition 12.4.1. *Λ satisfies the* **Bohr–Sommerfeld condition** *if this closed form is integral.*

There are a number of other formulations of this condition, the one of most relevance for us being the following:

Proposition 12.4.1. *Let s be a trivializing section of $\mathbb{L}_\Lambda$. Then Λ satisfies Bohr–Sommerfeld iff $\operatorname{Re} \alpha(s)$ is integral.*

Proof. Replacing s by $\langle s_0, s \rangle^{-\frac{1}{2}} s$ we can convert s into a trivializing section of $\Lambda^{\#} = P|\Lambda$, giving us identifications

$$\Lambda^{\#} \quad = \quad \Lambda \times S^+$$

and

$$\iota_{\Lambda^\sharp}^* \alpha \quad = \quad \alpha(s) + \frac{d\theta}{2\pi} \,.$$

Therefore α is integral if and only if $\alpha(s)$ is integral.

□

Example 2.
In example 1 replace X by $X^- \times X$ and $\mathbb{L}$ by $\mathbb{L}^* \boxtimes \mathbb{L}$, and let $f : X \to X$ be a

symplectomorphism. We will say that f is *pre-quantizable* if there exists a line bundle automorphism

$$\mathbb{L} \simeq f^*\mathbb{L} \tag{12.36}$$

satisfying

$$\nabla f^* s = f^* \nabla s \tag{12.37}$$

$$f^* \langle s, s \rangle = \langle f^* s, f^* s \rangle \tag{12.38}$$

for all $s \in C^\infty(\mathbb{L})$. Now let the Λ in example 1 be the graph of f viewed as a Lagrangian manifold of $X^- \times X$. The conditions (12.36)–(12.38) can be reformulated as saying that $\mathbb{L}^* \boxtimes \mathbb{L}|\Lambda$ had a canonical autoparallel trivializing section. Hence by the proposition above Λ satisfies Bohr–Sommerfeld and the $\Lambda^\#$ sitting above it in $(\mathbb{L}^* \boxtimes \mathbb{L})^\#$ is integral.

Example 3. The character Lagrangian.
Let G be an n-dimensional torus and

$$\chi : G \to \mathrm{Hom}(V)$$

an irreducible unitary representation of G. For χ to be irreducible and unitary the vector space V has to be one dimensional and $\chi(g)$ has to be multiplication by an element, $f(g)$ of S^+, hence such a representation is basically a homomorphism, $f : G \to S^+$, and this homomorphism, is by definition the character of χ. As for the *character Lagrangian*, this is by definition the graph in T^*G of the one-form, $\alpha = \frac{1}{2\pi} i \frac{df}{f}$ and hence is an integral Lagrangian submanifold of T^*G. We will show in the next section how to define an analogue of this object for G non-abelian, and at the end of this chapter discuss some semi-classical results in which it plays an important role.

We recalled at the beginning of this section that one way to generalize the notion of "morphism" in the exact symplectic category was by replacing " exactness" by "integrality". As we pointed out in §4.13.5, one can go in the opposite direction and define a class of morphisms which are much more restrictive than the exact morphisms but which play a prominent role in the applications we've just alluded to. Let (M_i, α_i), $i = 1, 2$, be exact symplectic manifolds and $\Gamma \subseteq M_1^- \times M_2$ a canonical relation. If M_1 and M_2 are cotangent bundles so is $M_1 \times M_2$. Thus $M_1 \times M_2$, with its zero section deleted, is a symplectic cone, and we will say that Γ is **conormal** if it is a conic submanifold of this cone, A simple condition for this to be the case is that for α the one-form (12.32) to satisfy

$$\iota_\Gamma^* \alpha = 0\,, \tag{12.39}$$

and this motivates the following:

Definition 12.4.2. *Let (M_i, α_i), $i = 1, 2$, be exact symplectic manifolds and $\Gamma \subseteq M_1^* \times M_2$ a canonical relations. We will say that Γ is* **conormal** *if it satisfies the condition (12.39).*

12.5 Symplectic reduction and the moment map.

Let M be a symplectic manifold, G a Lie group and $\tau : G \to \mathrm{Diff}(M)$ a Hamiltonian action of G on M. From this action we get a moment map

$$\phi : M \to \mathfrak{g}^* \tag{12.40}$$

with the defining property:

$$\iota(v_M)\omega = d\langle \phi, v\rangle \tag{12.41}$$

for all $v \in \mathfrak{g}$. (This identity only defines the $\langle \phi, v\rangle$'s up to additive constants; however, the cases we will be interested in, we can choose these constants so that the map (12.40) is G-equivariant. For instance suppose (M, α) is an exact symplectic manifold and α is G invariant. Then $L_{v_M}\alpha = 0$, so

$$\iota(v_M)\alpha = -d\iota(v_M)\alpha\,, \tag{12.42}$$

so one can take as one's definition of ϕ

$$\langle \phi, v\rangle = -\iota(v_M)\alpha \tag{12.43}$$

giving one a "ϕ" that is patently G-equivariant.)

The identity (12.41), evaluated at $p \in M$, says that

$$d\langle \phi, v\rangle_p = \iota(v_M(p))\omega_p\,. \tag{12.44}$$

Therefore, since ω_p is non-degenerate $\langle d\phi_p, v\rangle = 0$ if and only if $v_M(p) = 0$, i.e.

$$\text{Image } (d\phi_p : T_pM \to \mathfrak{g}^*) = \mathfrak{g}_p^\perp \tag{12.45}$$

where

$$\mathfrak{g}_p = \{v \in \mathfrak{g}\,,\ v_M(p) = 0\}\,. \tag{12.46}$$

From this one gets the following pertinent fact:

Proposition 12.5.1. *A point, a, of $\mathfrak{g}^*$ is a regular value of ϕ iff for every $p \in \phi^{-1}(a)$ $\mathfrak{g}_p = 0$; in other words iff the action of G at p is locally free.*

In particular, because of the G-equivariance of ϕ, the set $Z = \phi^{-1}(0)$ is a G-invariant closed subset of M, and if 0 is a regular value, is a G-invariant submanifold on which G acts in a locally free fashion. Therefore, if we assume in addition that G acts freely the quotient

$$B = Z/G \tag{12.47}$$

is a manifold and the projection $\pi : Z \to B$ makes Z into a principal G-bundle over B. Moreover the identity (12.44) tells us that at $p \in Z$

$$\langle (d\phi \circ \iota_Z)_p\,,\ v\rangle = \iota(v_Z(p))\iota_Z^*\omega_p\,.$$

However $\phi \circ \iota_Z = 0$, so

$$\iota(v_Z)\iota_Z^*\omega = 0\,. \tag{12.48}$$

This together with the fact that $\iota_Z^*\omega$ is G-invariant tells us that $\iota_Z^*\omega$ is *basic* with respect to the fibration, $\pi : Z \to B$. In other words there exists a unique two-form, ω_B, on B satisfying

$$\pi^*\omega_B = \iota_Z^*\omega\,. \tag{12.49}$$

A simple computation shows that ω_B is symplectic, and hence (12.49) implies that Z is a coisotropic submanifold of M. From Section 4.6 we know that this corresponds to a reduction morphism in the symplectic category. We recall how this goes:

Let Γ be the graph of π. By definition this sits in $Z \times B$; but, via the inclusion, $Z \to M$, we can think of Γ as a submanifold of $M \times B$, and the identity (12.49) can be interpreted as saying that Γ is a *Lagrangian submanifold* of $M^- \times B$ i.e. a *canonical relation*

$$\Gamma \in \mathrm{Morph}(M, B)$$

which is a reduction in the categorical sense. We will call this canonical relation the *reduction morphism* associated with the action τ, and the pair (B, ω_B) is called the *symplectic reduction of M with respect to the action*, τ.

Suppose now that M is an exact symplectic manifold and that $\omega = d\alpha$, $\alpha \in \Omega^1(M)^G$. Then, as we saw above the moment map associated with τ is given by (12.43) and hence, for $p \in Z$, $\iota(v_M)\alpha_p = 0$. This together with the G-invariance of α tells us that $\iota_Z^*\alpha$ is basic, and hence that there exists a one-form, $\alpha_B \in \Omega^1(B)$, satisfying

$$\omega_B \quad = \quad d\alpha_B \tag{12.50}$$

and

$$\pi^*\alpha_B = \iota_Z^*\alpha\,. \tag{12.51}$$

These two identities, however, simply say that the canonical relation, Γ, is conormal in the sense of Section 4.13.5. In other words:

Theorem 12.5.1. *In the exact symplectic category the reduction morphism*

$$\Gamma : M \to B$$

is an conormal canonical relation.

Example: Let $M = T^*G$. Then from the right action of G on T^*G one gets a trivialization $T^*G = G \times \mathfrak{g}^*$ which is invariant with respect to the left action of G on T^*G and the moment map associated with this left action is the map, $(x, \xi) \in G \times \mathfrak{g} \to -\xi$. Thus, in this example, Z is the zero section in T^*G, Z/G is the point manifold, "pt.", and $\Gamma^\dagger$ the point morphism, pt. $\to Z$.

As an application of these ideas we will come back to a notion that we discussed in Chapter 4, the notion of "moment Lagrangian" and provide an alternative perspective on it in terms of symplectic reduction: Consider the product action of G on $M \times T^*G$. Its moment map is the map

$$(x, g, \xi) \to \phi(x) - \xi \tag{12.52}$$

hence the zero level set of this moment map: the set, Z, in the discussion above, can be identified with $M \times G$ via the identification

$$(x, g) \to (x, g, \phi(x)) . \tag{12.53}$$

Thus G acts freely on Z, a global cross-section for this action being given by $M \times \{e\}$. Moreover the restriction to this cross-section of the product symplectic form on $M \times T^*G$ is the standard symplectic form on M so the symplectic reduction of $M \times T^*G$ by the product action of G is M itself. As for the canonical relation, Γ, associated with this reduction: this is by definition the graph of the fibration $\pi : Z \to M$; therefore, identifying M with the cross-section, $M \times \{e\}$ we see that the fiber above (x, e) in Z is the G orbit through (x, e) i.e. the set, $\{(g \times g^{-1}), g \in G\}$ and hence the graph of Γ is the set of all pairs $(p, \pi(p))$ where $p = (x, g)$ and $\pi(p) = gx$. Hence if we imbed Z into $M \times T^*G$ via the map (12.53) Γ becomes the set of points

$$(x, gx\,;\, g, \phi(x)) , \quad (x, g) \in M \times G \tag{12.54}$$

in $M \times T^*G$, which by comparison with the description for Γ_τ in §4.10.1 is seen to be Weinstein's moment Lagrangian. In other words the *moment Lagrangian, Γ_τ, is just the reduction morphism associated* with the action of G on $M \times T^*G$.

One consequence of this is that if M is an exact symplectic manifold Γ_τ is an *conormal* canonical relation.

In particular, suppose that M is the cotangent bundle T^*X of an n-dimensional manifold X, and that τ is the lift to M of an action

$$X \times G \to X \tag{12.55}$$

of G on X.

As we explained in §4.7, τ_X defines a morphism

$$\Gamma_{\tau_X} : T^*X \to T^*(X \times G) \tag{12.56}$$

i.e. a Lagrangian submanifold of

$$(T^*X)^- \times T^*X \times T^*G.$$

Claim:

Theorem 12.5.2. *The Γ_{τ_X} defined by (12.56) is identical with the moment Lagrangian (12.54).*

Proof. Recall that for any map $f : Y \to X$, the Lagrangian manifold Γ_f consists of the set of pairs $((y,\eta),(x,\xi))$ such that

$$x = f(y) \ \text{ and } \eta = df_y^*\xi.$$

So we have to check that for $Y = X \times G$ and $f = \tau$ this set coincides with (12.54).

This follows from the following lemma:

Lemma 12.5.1. *The moment map $\phi : T^*X \to \mathfrak{g}^*$ of the lifting of τ_X to T^*X is given by*

$$\langle \phi(x,\xi), v\rangle = -\langle \xi, v_X(x)\rangle, \quad v \in \mathfrak{g}. \tag{12.57}$$

The proof of the lemma follows from the identity (12.43) and the fact that at the point $p = (x,\xi) \in T^*X$ the right hand side of (12.43) is $\langle \xi, v_X(x)\rangle$ by the defining property of the canonical one form α on T^*X □

Thus, in this example, Γ_τ is not only conormal, but is, in fact, just the conormal bundle of the graph of τ_X.

Let us return to the general formula (12.54): By rearranging factors we can think of Γ as a morphism

$$\Gamma : M^- \times M \to T^*G\,.$$

If this morphism is composable with the diagonal, Δ, in $M^- \times M$ we get another object that we studied in Chapter 4 the character Lagrangian, $\Gamma_\tau \circ \Delta$, in T^*G. One consequence of the composition theorem that we proved in §112.3.4 is that if Γ_τ is conormal and Δ is an integral Lagrangian submanifold of $M^- \times M$ then the character Lagrangian is an integral Lagrangian submanifold.

An example of this which we will encounter later in this chapter is the following. Let (X,ω) be a (not-necessarily-exact) symplectic manifold and τ an action of G on X. Suppose X is pre-quantizable and let $\mathbb{L}$ be its pre-quantum line bundle and ∇ and $\langle\,,\,\rangle$ the pre-quantum connection and Hermitian inner product on $\mathbb{L}$. We will say that τ is pre-quantizable if it lifts to an action of G on $\mathbb{L}$ that preserves ∇ and $\langle\,,\,\rangle$. In this case it is easy to see that τ has to be a Hamiltonian action. In fact to see this let P be the unit circle bundle in $\mathbb{L}$ and $\alpha \in \Omega^1(P)$ its connection form. Then the action of G on P satisfies, for all $v \in \mathfrak{g}$

$$d\iota(v_P)\alpha = -\iota(v_P)\, d\alpha = -\iota(v_P)\pi^*\omega\,. \tag{12.58}$$

But $\iota(v_P)\pi^*\omega = \pi^*\iota(v_X)\omega$ and $\iota(v_P)\alpha$ is an S^1-invariant $\mathcal{C}^\infty$ function on M and hence is the pull-back by π of a $\mathcal{C}^\infty$ function $-\langle\phi, v\rangle$ on X. Thus we can rewrite the identity above in the form, $d\langle\phi, v\rangle = \iota(v_X)\omega$. Q.E.D.

Let $M = P \times \mathbb{R}^+ = (\mathbb{L})^\#$ be the symplectic cone associated with (P,α) and $\alpha_M = t\alpha$ its associated one-form. From the Hamiltonian action of S^1 on M we get a Hamiltonian action of the two-torus $T = S^1 \times S^1$ on $M^- \times M$ and associated with this action a reduction morphism

$$\Gamma : M^- \times M \to X^- \times X\,.$$

Moreover, this morphism can be factored into a product of two simpler morphisms: From the line bundle, $\mathbb{L}$, we get a pre-quantum line bundle, $\mathbb{L}^* \boxtimes \mathbb{L}$ over $X^- \times X$ and this comes equipped with a product connection and product Hermitian structure. Let Q be the circle sub-bundle of this product bundle and β the connection form on Q and let

$$W = (\mathbb{L}^* \boxtimes \mathbb{L})^{\#} = Q \times \mathbb{R}^+$$

be the symplectic cone associated with Q and β. In terms of these data the factorization of Γ that we alluded to above is the following. Factor the torus, T, as a product, $T_1 \times T_2$, where T_1 is the group of pairs, $(e^{i\theta}, e^{-i\theta})$, $e^{i\theta} \in S^1$ and T_2 the group of pairs $(e^{i\theta}, e^{i\theta})$. Then if we reduce $M^- \times M$ by the action of T^1 we get a reduction morphism

$$\Gamma_1 : M^- \times M \to W$$

and if we reduce W by the action of T_2 we get a reduction morphism

$$\Gamma_2 : W \to X^- \times X$$

and this "reduction in stages" factors Γ into a composite reduction $\Gamma = \Gamma_2 \circ \Gamma_1$. Now let Δ_X be the diagonal in $X^- \times X$. Then $\Delta^{\#} = \Gamma_2^t \circ \Delta_X$ is just the pre-image of Δ in Q and hence, as we showed in §112.3.4 is an integral Lagrangian submanifold of W. Moreover $\Gamma_1^t \circ \Delta^{\#} = \Gamma \circ \Delta$; so it is just the diagonal Δ_M in $M^- \times M$, and hence Δ_M is integral. Finally it is easy to check that the moment Lagrangian associated with the action of G on M is just the composition of the morphisms

$$\Gamma : M^- \times M \Rightarrow X^- \times X$$

and

$$\Gamma_\tau : X^- \times X \Rightarrow T^*G\,.$$

Hence, as we showed above, this composite morphism is an *conormal* canonical relation. Moreover the identity, "$\Delta_M = \Gamma^t \circ \Delta_X$", can be interpreted as saying that $\Delta_X = \Gamma \circ \Delta_M$. Therefore if Γ_τ and Δ_X are cleanly composable so are $\Gamma_\tau \circ \Gamma \circ \Delta_M$; and

$$\Gamma_\tau \circ \Delta_X = \Gamma_\tau \circ \Gamma \circ \Delta_M\,.$$

In other words the character Lagrangian associated with the action of G on M, coincides with the character Lagrangian associated with the action of G on X, and hence since the first of these is integral so is the second.

We conclude this discussion of symplectic reduction by pointing out that the reduction morphism $\Gamma \in \mathrm{Morph}(M, B)$ can be equipped with a canonical half density and hence converted into a morphism in the *enhanced* symplectic "category".

Namely, by identifying Γ with the zero level set Z of the moment map, we can think of it as a principal G - bundle $Z \to B$. From Haar measure on G and the symplectic volume form on B one gets a non-vanishing smooth density on Z whose square root is a non-vanishing half density σ. Thus via the identification of Z with Γ this becomes a half density on Γ and hence the pair (Γ, σ) is a morphism in the enhancedd symplectic "category".

In particular, this remark applies to the moment Lagrangian

$$\Gamma_\tau : M^- \times M \to T^*G.$$

Moreover, from the sympletcic volume form on M, and the identification of M with the diagonal in $M^- \times M$ one gets a volume form on Δ whose square root in a non-vanishing half-density, μ. Thus, if Γ_τ and Δ are cleanly composible, the composition law for morphisms in the enhanced symplectic "category" converts $\Gamma_\tau \circ \Delta$ into an enhanced Lagrangian $(\Gamma_\tau \circ \Delta, \sigma \circ \mu)$, i.e. equips $\Gamma_\tau \circ \Delta$ with a canonical half-density $\sigma \circ \mu$.

12.6 Coadjoint orbits.

To extend the character formula (3.6.7) to non-abelian groups we will have to describe the analogues for these groups of the elements, α, of the weight lattice of G and this will require a brief review of the theory of co-adjoint orbits. As above let G be a connected Lie group, $\mathfrak{g}$ its Lie algebra and $(\mathrm{Ad})^* : G \to \mathrm{Aut}(\mathfrak{g}^*)$ the co-adjoint action of G on $\mathfrak{g}^*$. Let O be an orbit of G in $\mathfrak{g}^*$ and f a point on this orbit. We claim that one can define an alternating bilinear form on T_fO by setting

$$\omega_f(v_O, w_O) = \langle f, [v, w]\rangle \tag{12.59}$$

for $v, w \in \mathfrak{g}$. To show that the left hand side is well-defined we note

$$\langle f, [v, w]\rangle = \langle f, gd(v)w\rangle = \langle \mathrm{ad}(v)^* f, w\rangle \tag{12.60}$$

and $\mathrm{ad}(v)^* f = 0$ if and only if $v_O(f) = 0$ so the expression on the right only depends on v_O and, with the roles of v and w reversed, only depends on w_O.

Suppose in addition that $\omega_f(v_), w_O) = 0$ for all w_O. Then by (12.60), $(\mathrm{ad}\, v)^* f = 0$ and hence $v_)(f) = 0$. Therefore since the vectors $v_0(f)$, $v \in \mathfrak{g}$ span the tangent space to O at f, the bilinear form (12.59) is non-degenerate.

Let ω_O be the 2-form on O defined by the assignment, $f \to \omega_f$. It is clear that this form is G-invariant. Moreover, if one lets $\phi_v \in \mathcal{C}^\infty(O)$ be the function, $\phi_v(f) = \langle f, v\rangle$ then

$$\begin{aligned} d\phi_v(w_O)_f &= -d\phi_v((\mathrm{ad})^*(w)f) \\ &= -\langle \mathrm{ad}(w)^* f, v\rangle \\ &= -\langle f, \mathrm{ad}(w)v\rangle = \langle f, [v, w]\rangle \\ &= \iota(v_O)\omega_f(w_O) \end{aligned}$$

and hence

$$\iota(v_O)\omega_O = d\phi_v \, . \tag{12.61}$$

From this one easily deduces that the following are true:

1. The two-form, ω_O, is closed (and hence symplectic.)
2. The action of G on O is Hamiltonian.
3. The moment map associated with this action is the inclusion map, $O \hookrightarrow \mathfrak{g}^*$.

Proof. Since ω_O is G-invariant

$$0 = Lv_0\omega_O = \iota(v_O)\, d\omega_O + d\iota(v_O)\omega_)$$

and by (12.61) the second summand on the right vanishes. Thus for all $v \in \mathfrak{g}$, $\iota(v_O\, d\omega_O = 0)$ and since the $v_O(f)$'s span the tangent space to O at each point, f, of O, this implies that $d\omega_O = 0$. Moreover if we denote by ϕ the inclusion map of O into $\mathfrak{g}^*$ we can rewrite (12.61) in the form

$$\iota(v_O)\omega_0 = d\langle \phi, v\rangle \, . \tag{12.62}$$

□

The next issue we'll address is the question of whether O can be equipped with G equivariant pre-quantum structure. Let $\mathbb{L}$ be a line bundle on O and suppose that the action of G on O can be lifted to an action of G on $\mathbb{L}$ by bundle morphisms, that this action preserves a connection, $\triangledown$, and a Hermitian inner product $\langle\, ,\, \rangle$ and, finally, that $\mathrm{curv}(\triangledown) = \omega$. Then equivalently, the G-action on $\mathbb{L}$ preserves the circle subbundle, P, of $\mathbb{L}$ defined by $\langle\, ,\, \rangle$ commutes with the S^1 action on this bundle and preserves the connection form, $\alpha \in \Omega^1(P)$. Now let's fix a point, f, of O and let $G_f \subset G$ be the stabilizer group of f in G. From the action of G on $\mathbb{L}$ we get the representation of G_f on $\mathbb{L}_f$. Moreover, for $v \in \mathfrak{g}$

$$\alpha(v_P) = \langle \phi, v\rangle(f) = \langle f, v\rangle \tag{12.63}$$

by (12.62). Also, if $v \in \mathfrak{g}_f$, then v_P is tangent to the fiber, P_f of P above f. However, on this fiber, $\alpha = d\theta$, so for every $v \in \mathfrak{g}_f$

$$\langle d\theta, v_P\rangle = \langle f, v\rangle \, . \tag{12.64}$$

Thus the character of the representation of G_f on $\mathbb{L}_f$ is just the map

$$\exp v \in G_f \to e^{2\pi i\langle f, v\rangle} \, . \tag{12.65}$$

In other words the weight of this representation is f. This proves

Theorem 12.6.1. *If the action of G on O is prequantizable, then for $f \in O$, $f|\mathfrak{g}_f$ is in the weight lattice of the group G_f.*

Remark For the groups we'll be interested in: connected compact groups the converse of this result is true. (See for instance Kostant *Unitary representation ...*)

We will henceforth call a coadjoint orbit, O, **integral** if it has this property.

From this result one gets a description of the line bundle, $\mathbb{L}$, as the quotient

$$\mathbb{L} \simeq G \times \mathbb{C}/G_f$$

where the action of G_f on $G \times \mathbb{C}$ is the product of its right action on G and the action (12.63) on $\mathbb{C}$. Moreover the connection is determined as well by these data. Namely the connection form α on P satisfies

$$\alpha\left(\frac{\partial}{\partial\theta}\right) \quad = \quad 1$$

and

$$\alpha(v_P)_f = \langle \phi, v\rangle_f \quad = \quad \langle f, v\rangle$$

for all $f \in O$ and since the v_P's and $\frac{\partial}{\partial\theta}$ span the tangent space of P at each of its points these conditions completely determine α.

We will next compute the character Lagrangian for the action of G on O. By definition a point $(g, f) \in G \times \mathfrak{g}^*$ is in this character Lagrangian if there exists a point $x \in O$ such that $gx = x$ and $\phi(x) = f$. However, since ϕ is just the inclusion map of O into $\mathfrak{g}^*$, this character Lagrangian, which we will denote by Λ_O, is the set

$$\Lambda_O = \{(g, f) \in G \times O\,, \quad \mathrm{Ad}(g)^* f = f\}\,. \tag{12.66}$$

Thus the projection

$$\Lambda_O \to O\,, \quad (g, f) \to f \tag{12.67}$$

is just a fiber mapping with fiber, G_f, above f. However the projection

$$\Lambda_O \to G\,, \quad (g, f) \to g \tag{12.68}$$

is a lot more complicated. Over generic points of G the set of f's in O for which $\mathrm{Ad}(g)^* f = f$ is finite and over these generic points (12.68) is just a covering map. However, if g is, for instance, the identity element of G then the fiber above g is all of O.

We will give a much more detailed description of this map in §12.8.

We conclude this discussion of coadjoint orbits by describing a generalization, involving coadjoint orbits, of the symplectic reduction operation that we discussed in §12.5. This generalization will play an important role in the analytic applications of integrability the we will take up in Sections **??** and **??**.

Let M be a Hamiltonian G-manifold and $\Phi : M \to \mathfrak{g}^*$ its moment map. If $O \subset \mathfrak{g}^*$ is a coadjoint orbit, its pre-image

$$\Sigma := \Phi^{-1}(O)$$

is a G-invariant subset of M. If G acts freely on this set, the by (12.45) the map $|Phi$ is transversal to O, and hence Σ is a submanifold of M of codimension equal to the dimension of O. Moreover, since G acts freely on Σ, the quotient $B = \Sigma/G$ is a manifold of dimension

$$\dim M - \dim O - \dim G$$

and the projection

$$\pi : \Sigma \to B$$

makes Σ into a principal G-bundle over B.

We call B the **symplectic reduction of M with respect to O**. To justify this nomenclature, we show that B has an intrinsic symplectic structure. To see this, we note that B has an alternative description: Consider the product action of G on the symplectic manifold $M \times O^-$. This is a Hamiltonian G-action with moment map

$$\Psi_O : M \times O^- \to \mathfrak{g}, \quad (p, \ell) \mapsto \Phi(p) - \ell.$$

The zero level set of Ψ_O is the set of $(p, \ell) \in M \times O^-$ such that $\Phi(p) = \ell$. So it can be identified with Σ via the map $\Sigma \ni p \mapsto (p, \Phi(p))$. This identification is G-equivariant, so as G acts freely on Σ, it acts freely on this zero level set. Hence the symplectic reduction

$$\Psi_O^{-1}(0)/G$$

of $M \times O^0-$ that we defined in §12.5 canbe identified with B. This allows us to equip B with an intrinsic symplectic structure.

12.7 Integrality in semi-classical analysis

In chapter 8 we showed that if one is given a manifold, X, and an exact Lagrangian submanifold, (Λ, φ) of T^*X, then one can attach to these data a space of oscillatory half-densities $I(\Lambda; X)$. Let's briefly recall the role of the function, φ, in the definition of this space. Given any Λ one can find, at least locally, a fiber bundle, $\pi : Z \to X$ and a generating function for Λ, $\psi \in \mathcal{C}^\infty(Z)$ whose defining property is that critical set of ψ with respect to the fibration, π, is mapped diffeomorphically onto Λ by the map

$$\gamma_\psi : C_\psi \to \Lambda\,,\; z \to d_X\psi\,. \tag{12.69}$$

Given ψ one then defines the space $I(\Lambda, \psi)$ to be the set of oscillatory functions,

$$\pi_* a(z,h) e^{\frac{i\psi}{\hbar}}\,,\; a \in \mathcal{C}^\infty(Z \times \mathbb{R}^*)\,. \tag{12.70}$$

One problem with this definition however is that there is an unspecified additive constant involved in the choice of ψ: for every $c \in \mathbb{R}$, $\psi + c$ doesn't change either the critical set C_ψ or the parametrization of Λ. It does however have a discernible effect on the oscillatory behavior of the oscillatory integral

(12.70), i.e. it multiplies it by the factor $e^{i\frac{c}{\hbar}}$. (This situation becomes ever worse if one tries to define elements of $I(X;\Lambda)$ by patching together contributions from N local parametrizations of Λ in which case the elements of $I(X;\Lambda)$ become expressions of the form

$$\sum_{k=1}^{N}(\pi_k)_*\left(a_k e^{i\frac{\psi_k}{\hbar}}\right)e^{i\frac{c_k}{\hbar}} \tag{12.71}$$

which for the c_k's arbitrary and N large can be made to have more or less random oscillatory behavior.) The role of the φ in the exact pair (Λ,φ) is to avoid these complications by requiring that the generating function ψ satisfy

$$\gamma_\psi^*\varphi = \psi|C_\psi \tag{12.72}$$

and as we showed in chapter 8 this *does* avoid these complications and give one a satisfactory global theory of oscillatory functions.

Suppose now that $\Lambda = (\Lambda, f)$ is an integral Lagrangian submanifold of T^*X. In this case one can still to a certain extent avoid these complications by replacing (12.72) by

$$\gamma_\psi^* f = e^{i\psi}\,. \tag{12.73}$$

This does not entirely get rid of the ambiguity of an additive constant in the definition of ψ but does force this constant to be of the form $2\pi n$, n a positive integer. Thus "random sums" like the expression (12.71) can be eliminated by the simple expedient of requiring that $1/\hbar$ be an integer. In fact one can show that if one imposes this condition the results that we proved in chapter 8 all extend, more or less verbatim, to Lagrangian manifolds and canonical relations which are integral. Moreover we can now define some objects which we weren't able to fit into our theory before:

Example: Let G be an n-torus and $f : G \to S^1$ a function of the form $f(x) = e^{2\pi i\alpha(x)}$, where $\alpha \in \mathbb{Z}_G^* \subset \mathfrak{g}^*$ is an element of the weight lattice of G. Then the function

$$f^m = e^{\frac{2\pi i\alpha(x)}{\hbar}}\,, \quad \hbar = \frac{1}{m} \tag{12.74}$$

can be regarded as an element of $I^n(G;\Lambda_\alpha)$ where Λ_α is the character Lagrangian associated with f, and f^m is the character of the representation of G with weight $m\alpha$.

This example turns out to be a special case of a larger class of examples involving characters of representations of Lie groups, and we'll discuss these examples in the next three sections.

12.8 The Weyl character formula.

In this section we will assume that G is a compact simply connected Lie group and that for every $\beta \in O$ the isotropy group, G_β, is a subtorus of G, i.e. as a homogeneous space, O, is the quotient, G/T, of G by the Cartan subgroup,

T, of G. In addition we will assume that O is an integral coadjoint orbit: $O = \mathrm{Ad}^*(G)\beta_0$ where β_0 is an integer lattice vector in the interior of the positive Weyl chamber, t^*_+, of t^*. Let γ_m be the character of the irreducible representation of G with highest weight, $m\beta_0$. We will show that the sequence of functions, γ_m, $m = 1, 2, \ldots$, define, for $\hbar = 1/m$, an element $\gamma(g, \hbar)$ of $I^0(\Lambda_O; G)$ and that its symbol is $\chi_\rho m|\nu_O|^{\frac{1}{2}}$ where $|\nu_O|^{\frac{1}{2}}$ is the canonical $\frac{1}{2}$-density on Λ_O that we defined in §12.5, m is a Maslov factor and χ_ρ a conversion factor which effectively converts $|\nu_O|^{\frac{1}{2}}$ into a "$\frac{1}{2}$-form". (For more about "$\frac{1}{2}$-forms" and their relation to $\frac{1}{2}$-densities see [GS], chapter V, §4.)

We will in fact prove a stronger result. We will show that, with this "$\frac{1}{2}$-form" correction, the recipe we give in chapter 8 for associating to $m|\nu_O|^{\frac{1}{2}}$ an oscillatory $\frac{1}{2}$-density turns out to give, even for $\hbar = 1$ an exact formula for $\gamma(g, \hbar)$ (not, as one would expect, a formula that's asymptotic in $\hbar$). We will verify this assertion by computing this $\frac{1}{2}$-density at regular points of the group, G, and comparing it with the Weyl character formula for γ_n. This computation will require our reviewing a few basic facts about roots and weights, but in principle is fairly easy since the projection, $\Lambda_O \to G$, is just a finite-to-one covering over the set of regular points in G. However, we will also show in the next section that our recipe for quantizing $\xi_\rho m\sigma^{\frac{1}{2}}$ gives an exact answer in a neighborhood of the identity element where the projection, $\Lambda_O \to G$ is highly singular. (This will again be a proof by observational mathematics. We'll show that the recipe for computing the oscillatory function associated with $\xi_\rho m\sigma_O^{\frac{1}{2}}$ by generating functions coincides with the Kirillov formula for γ_n.) We will also say a few words about the computation of γ_n at arbitrary points of G, (in which case the methods of chapter 8 turn out to give a generalized Kirillov formula due to Gross, Kostant, Ramond, Sternberg[GKRS]).

We'll start by describing a few elementary properties of the manifold Λ_O and of the fibration, $\Lambda_O \to G$.

Proposition 12.8.1. *There is a canonical diffeomorphism of G spaces*

$$k_O : \Lambda_O \to O \times T\,. \tag{12.75}$$

Proof. Λ_O is the subset

$$\{(g, \beta) \in G \times O\,,\, (\mathrm{Ad})^*(g)\beta = \beta\} \tag{12.76}$$

of $G \times O$; so the projection, p of Λ_O onto O is a fibration with fiber,

$$G_\beta = \{h \in G\,,\, (\mathrm{Ad})^*(h)\beta = \beta\} \tag{12.77}$$

above β. Thus if $\beta = g\beta_0$, $(\mathrm{Ad})^*(g^{-1}hg)\beta = \beta_0$ and since β_0 is the interior of the positive Weyl chamber $g^{-1}hg \in T$. Thus the map

$$(h, \beta) \in \Lambda_O \to (\beta, g^{-1}hg) \tag{12.78}$$

is a G-equivariant diffeomorphism of Λ_O onto $O \times T$.

□

Another slightly more complicated description of Λ_O is in terms of the projection

$$\pi : \Lambda_O \to G\,,\ (g,\beta) \to g\,. \tag{12.79}$$

For this projection

$$\pi^{-1}(g) = \{\beta \in O\,, (\mathrm{Ad})^*(g)\beta = \beta\} \tag{12.80}$$

so in particular if $K = \{a \in G, a^{-1}ga = g\}$ is the centralizer of g in G then for every $\beta \in \pi^{-1}(g)$ the $(\mathrm{Ad})^*$ orbit of K through β is in $\pi^{-1}(g)$. We claim, in fact, that $\pi^{-1}(g)$ consists of a finite number of K orbits. To see this, we can without loss of generality assume that g is in T. Let $N(T)$ be the normalizer of T in G and $W = N(T)/T$ the Weyl group. We claim that for g in T

$$\pi^{-1}(g) = \bigcup K w \beta_0\,, \quad w \in W\,. \tag{12.81}$$

Proof. Let $\beta = h\beta_0$, $h \in G$, be an element of $\pi^{-1}(g)$. Then $\mathrm{Ad}^*(g)\beta = \beta$, so $\mathrm{Ad}^*(gh)\beta_0 = (\mathrm{Ad})^*(h)\beta_0$ and hence since β_0 is in Int $\mathfrak{t}^*$, $h^{-1}gh$ is in T. Therefore, $h^{-1}gh = aga^{-1}$ for some $a \in N(T)$ and hence ah is in K i.e. h is in wK where w is the image of a^{-1} in $N(T)/T = W$. □

Let G_{reg} be the set of regular elements of G: elements whose centralizers are maximal tori. As a corollary of the result above we get the following:

Proposition 12.8.2. *Over G_{reg} the map $\pi : \Lambda_O \to G$ is an N to 1 covering map where N is the cardinality of W.*

Proof. It suffices to verify this for $g \in T_{\mathrm{reg}}$ in which case $K = T$ and hence by (12.81):

$$\pi^{-1}(g) = \{w\beta_0\,, w \in W\}\,. \tag{12.82}$$

□

Thus over T_{reg}, Λ_O is the disjoint union of the Lagrangian manifolds

$$\Lambda_w = \quad \text{graph}\ \left(\frac{1}{2\pi i}\,\frac{df_w}{f}\right)\,, \quad w \in W \tag{12.83}$$

where $f_w(t) = e^{2\pi i\langle w\beta_0, t\rangle}$. Moreover, the complement of G_{reg} in G is an algebraic subvariety of G of codimension ≥ 2. Therefore since G is simply connected, G_{reg} is simply connected, and the covering map, $\Lambda_{\mathrm{reg}} \to G_{reg}$ is a trivial covering map mapping the connected components,

$$\{(g,\beta) \in G_{\mathrm{reg}} \times O\,,\ \beta = \mathrm{Ad}^*(g)w\beta_O\}$$

of Λ_O bijectively onto G_{reg}.

Now let dg and dt be the standard Haar measure on G and T and μ_O the symplectic volume form on O. As we explained in Section 12.5 one gets from μ_O a canonical $\frac{1}{2}$-density on the character Lagrangian, Λ_O and a simple computation (which we'll spare the reader) shows that the square of this $\frac{1}{2}$-density is given by

$$\nu_O = k_O^*(\mu_O \otimes dt) \tag{12.84}$$

where k_O is the mapping (12.75) and $\mu_O \otimes dt$ is the product on $O \times T$ of the densities, μ_O, and dt.

Let's now come back to the goal of this section as enunciated above: to show that if γ_m is in the character of the irreducible representation of G with highest weight, $m\beta_0$, the oscillatory function

$$\gamma(g,\hbar) = \gamma_m(g)\,, \quad \hbar = 1/m \tag{12.85}$$

defines an element, $\gamma(g,\hbar)|dg|^{\frac{1}{2}}$, in $I^0(\Lambda_O;G)$ and that its symbol is a $\frac{1}{2}$-density on Λ_O of the form, $\xi_\rho m|\nu_O|^{\frac{1}{2}}$ where ξ_ρ is a "$\frac{1}{2}$-density-to-$\frac{1}{2}$-form" conversion factor and m a Maslov factor (both of which will be defined shortly). Let $\Lambda_{\text{reg}} = \pi^{-1}(G_{\text{reg}})$. Then $\pi : \Lambda_{\text{reg}} \to G_{\text{reg}}$ is an N to 1 covering map which splits over T_{reg} into the union of Lagrangian manifolds, Λ_w. Let $f_w : T \to S^1$, be the function, (12.83), and f the unique G-invariant function on Λ_{reg} whose restriction to Λ_w is $\pi^* f_w$. We will prove that on Λ_{reg} (where the mapping π is locally a diffeomorphism at every point)

$$\pi_*(f^m \xi_\rho m|\nu_O|^{\frac{1}{2}}) = \gamma(g,\hbar)|dg|^{\frac{1}{2}}$$

by explicitly computing the push-forward on the left hand side and comparing it with the expression for $\gamma_m(g)$ given by the Weyl character formula. To perform this computation we will first review a few elementary facts about the adjoint representation of T on the Lie algebra, g of G.

Under this representations, $\mathfrak{g} \otimes \mathbb{C}$, splits into T-invariant complex subspaces

$$n \oplus \bar{n} \oplus t \otimes \mathbb{C} \tag{12.86}$$

where n is a nilpotent Lie subalgebra of $\mathfrak{g} \otimes \mathbb{C}$. Moreover, n and $\bar{n}$ split into direct sums of one-dimensional subspaces

$$n \;=\; \oplus \mathfrak{g}_\alpha\,, \quad \alpha = \alpha_k\,,\, k = 1,\ldots,d \tag{12.87}$$

and

$$\bar{n} \;=\; \oplus \bar{\mathfrak{g}}_\alpha\,, \quad \alpha = -\alpha_k\,,\, k = 1,\ldots,d \tag{12.88}$$

where $d = \dim G/T$, and α is the weight of the representation of T on $\mathfrak{g}_\alpha$. The α_k's are by definition the *positive roots* of $\mathfrak{g}$ and the $-\alpha_k$'s the *negative roots*. We'll denote the set of these roots by ϕ and the subset of positive roots by ϕ^+. For $\alpha \in \phi^+$ let Z_α be a basis vector for $\mathfrak{g}_\alpha$ and $Z_{-\alpha} = \bar{Z}_\alpha$ the corresponding basis vector for $\mathfrak{g}_{-\alpha} = \bar{\mathfrak{g}}_\alpha$. Then for $X \in T$

$$[X, Z_\alpha] = 2\pi i \alpha(X) Z_\alpha \tag{12.89}$$

and hence by Jacobi's identity

$$[X,[Z_\alpha, Z_\beta]] = 2\pi i(\alpha+\beta)(X)[Z_\alpha, Z_\beta]\,. \tag{12.90}$$

Hence either $[Z_\alpha, Z_\beta] = 0$ or $\alpha + \beta$ is again a root or, for $\beta = -\alpha$, $[Z_{\alpha, Z_\beta}]$ is in $t \otimes \mathbb{C}$. The sum

$$\rho = \frac{1}{2} \sum \alpha_k \,, \quad \alpha_k \in \phi^+ \tag{12.91}$$

will play an important role in the computations below as will the identities

$$e^{2\pi i\rho} \prod (1 - e^{-2\pi i \alpha_k}) \quad = \quad \prod (e^{\pi i \alpha_k} - e^{-\pi i \alpha_k}) \tag{12.92}$$

and

$$\sum_{w \in W} (-1)^w e^{2\pi i w\rho} \quad = \quad e^{2\pi i \rho} \prod (1 - e^{-2\pi i \alpha_k}) \,. \tag{12.93}$$

(The first identity is obvious and the second a consequence of the fact that

$$w\phi^+ = \{\pm\alpha_1, \ldots, \pm\alpha_k\}$$

and that all possible combinations of plus and minus signs can occur.)

Now fix an element, h of T. We will begin our computation of the left hand side of the character formula (12.85) by computing the derivative of the mapping

$$\gamma_h : G/T \to G \,,\; gT \to h^{-1} g^{-1} hg \tag{12.94}$$

at the identity coset, $p_0 = eT$, of G/T. If we identify $T_{p_0} \otimes \mathbb{C} = \mathfrak{g}/t \otimes \mathbb{C}$ with $n + \bar{n}$ and let $h = \exp X$ we get

$$\begin{aligned}(d\gamma_h)_{p_0}(Z_\alpha) &= \frac{d}{dt} h^{-1} (\exp -tZ_\alpha) h (\exp tZ_\alpha)|_{t=0} \\ &= \mathrm{Ad}(h) Z_\alpha - Z_\alpha\end{aligned}$$

and hence by (12.88):

$$(d\gamma_h)_{p_0}(Z_\alpha) = (e^{2\pi i \alpha(X)} - 1) Z_\alpha \,. \tag{12.95}$$

Next consider the mapping

$$\gamma : G/T \times T \to G \,,\; (gT, h) \to g^{-1} hg \,.$$

If we let $TG = G \times \mathfrak{g}$ be the right invariant trivialization of TG and identify the complexified tangent spaces to $G/T \times T$ at (p_0, h) and to G at h with $n \oplus \bar{n} + t \otimes \mathbb{C}$ the determinant of $(d\gamma)_{p_0}$ is equal, by (12.95) to $|D(h)|^2$ where $D(h)$ is the Weyl product

$$D(h) = e^{2\pi i \rho(X)} \prod (1 - e^{-2\pi 1 \alpha_k(X)}) \,. \tag{12.96}$$

Hence at (p_0, h)

$$\gamma^* \, dg = |D(h)|^2 \mu_{G/T} \otimes dt \tag{12.97}$$

where $\mu_{G/T}$ is the unique G-invariant density on G/T whose integral over G/T is 1. Thus if we make the trivial identifications, $G/T = O$ and $\mu_{G/T} = \mu_O$, note that $\gamma \centerdot k_O = \pi$ and recall that by definition, $\nu_O = k_O^*(\mu_O \otimes dt)$ we obtain from (12.97) the formula

$$\pi^* \, dg = |D(h)|^2 \nu_O \tag{12.98}$$

at points on Λ_O above h. Therefore at *regular* points, h, of T

$$\pi_* \nu_O = |W||D(h)|^{-2} \, dg \tag{12.99}$$

since there are exactly $N = |W|$ preimage points of h in Λ_O. Thus if we take the square root of (12.99) at each of these points we also get, for the $\frac{1}{2}$-density, $|\nu_O|^{\frac{1}{2}}$,

$$\pi_* |\nu_O| = |W||D(h)|^{-1} |dg|^{\frac{1}{2}} \tag{12.100}$$

at regular points, h, of T.

Now let m be the function on $\pi^{-1}(T_{\mathrm{reg}})$ whose restriction to Λ_w is the pull-back to Λ_w of the function

$$\frac{1}{|W|} \frac{|D(h)|}{D(h)} (-1)^w \tag{12.101}$$

and let ξ_ρ be the function on $\pi^{-1}(T_{\mathrm{reg}})$ whose restriction to Λ_w is the pull-back to Λ_w of the function

$$e^{2\pi i \langle w\rho, X\rangle} \, . \tag{12.102}$$

These functions extend to G-invariant functions on Λ_{reg} and by (12.100)–(12.102) we get for $\pi_* \xi_\rho m f^m |\nu_O|^{\frac{1}{2}}$ the expression

$$D(h)^{-1} \sum (-1)^w e^{2\pi i \langle w(\rho + m\beta_O, X\rangle)} |dg|^{\frac{1}{2}} \tag{12.103}$$

at points, $h = \exp X$ in T_{reg}; and by the Weyl character formula the expression (12.103) is $\gamma_m |dg|^{\frac{1}{2}}$.

Remarks

1. Another corollary of the formula (12.97) is the integration theorem which asserts that for $f \in L^1(G)$

$$\int f(g) \, dg = \frac{1}{|W|} \int_{G/T} f(g^{-1} t g) \, d\mu_{G/T} |\Delta(t)|^2 \, dt \tag{12.104}$$

   and, in fact, one can give a simple direct proof of the Weyl character formula itself based solely on this identity and the identity (12.93). (See Remark 4, below.)

2. Moreover the identity (12.93) has a nice interpretation in terms of the Weyl character formula, It says that with $\beta_0 = 0$ the expression (12.102) is equal to 1, i.e. the character of the trivial representation of G is 1.

3. We will briefly explain what the function (12.102) has to do with Maslov indices: At $X \in t_{\text{reg}}$, $\Delta(h)/|\Delta(h)|$ is equal to the product

$$(i)^d \prod \frac{\sin \frac{\alpha_k}{z}(X)}{|\sin \frac{\alpha_k}{2}(X)|} = u^d (-1)^{\sigma(X)} \tag{12.105}$$

where

$$\sigma(X) = \#\{k, \alpha_k(X) > O\} - \#\{k, \alpha_k(X) < O\}$$

and we will see in the next section that the function

$$X \in t_{\text{reg}} \to i^d (-1)^{\sigma(X)}$$

can be interpreted as a section of the Maslov line bundle on $\Lambda_O|T$.

4. **Proof of the Weyl character formula.** Assume that $\pi \in \widehat{G}$ is an irreducible representation of the highest weight $\lambda \in A^+$. From the orthogonormality of characters, we have orthogonality of character, we have

$$\int_G |\chi_\lambda(g)|^2 \, dg = 1 \, .$$

Since χ_λ is a class-function, the integration formula yields

$$\int_T |\chi_\lambda(t)\Delta(t)|^2 \, dt = |W| \, .$$

We now analyze the integrand $\chi_\lambda \Delta$. First note that since $Q \simeq N_G(T)/T$, there exists $n \in N_G(T)$ such that $\pi(n) : V(\lambda) \to V(\omega\lambda)$. In particular, $\dim V(\mu) = \dim(\omega\mu)$. Hence, if $n_\mu = \dim V(\mu)$, then

$$\chi_\lambda|_T = \sum n_\mu e^\mu \;\text{ and }\; n_{\omega\mu} \, , \quad \forall_\mu \in \mathcal{E}(\pi) \, , \, \forall_\omega \in W \, .$$

On the other hand since Δ is W-skew symmetric $\chi_\lambda \Delta$ is W-skew symmetric. This means that if we write

$$\chi_\lambda \Delta = \left(\sum_\mu n_\mu e^\mu \right) \left(\sum_{\omega - \in W} \epsilon(\omega) e^{\omega\rho} \right) = \sum c(\beta) e^\beta$$

where $c(\beta) \in \mathbb{Z}$ are the coefficients of the various $\beta = \omega\rho + \mu$ after we open the parenthesis, then

$$c(\omega\beta) = \epsilon(\omega) c(\beta) \, , \quad \forall \omega \in W \, .$$

Since $c(\lambda + \rho) = 1$, we have $c(\omega(\lambda + \rho)) = \epsilon(\omega))$, $\forall \omega \in W$.

On the other hand, the Parseval identity on the torus

$$\int_T |\xi_\lambda \Delta|^2 = \sum_\beta |c(\beta)|^2$$

implies

$$\begin{aligned}|W| &= \sum_{\beta}|c(\beta)|^2 = \sum_{\omega\in W}|c(\omega(\lambda+\rho))|^2 + \sum_{\beta\notin W\cdot(\lambda+\rho)}|c(\beta)|^2\\ &= 1+1+\cdots+1(|W|\text{times}) + \sum_{\beta\notin W\cdot(\lambda+\rho)}|c(\beta)|^2\,.\end{aligned}$$

Hence $c(\beta)=0$ when $\beta\notin W\cdot(\lambda+\rho)$ and $\chi_\lambda\Delta=\sum_{\omega\in W}\epsilon(\omega)e^{\Omega(\lambda+\rho)}$.

12.9 The Kirillov character formula.

The fibration, $\pi:\Lambda_O\to G$ is just a finite-to-one covering map over points of G_{reg}, so locally, at any point, $g\in G_{\text{reg}}$, each sheet, Λ_w, of this covering map is the graph of a one-form, $d\varphi_w$, and this φ_w can be taken to be the generating function for Λ_O in a neighborhood of g. However over the identity element, π degenerates and the pre-image of e becomes the whole orbit, O so this naive recipe no longer works. Nonetheless, there is still a simple description of Λ_O at e in terms of generating functions.

Theorem 12.9.1. *Let $\varphi:O\times\mathfrak{g}\to\mathbb{R}$ be the function $\varphi(\beta,X)=\beta(X)$. Then via the identification*

$$O\times\mathfrak{g}\to O\times G\,,\ (\beta,X)\to(\beta,\exp X) \tag{12.106}$$

φ becomes a generating function for Λ_O, locally near e, with respect to the fibration, $O\times G\to G$.

Remark

The qualification "locally near e" is necessary because exp is only a diffeomorphism in a neighborhood of e; however the open set on which this theorem is true turns out, in fact, to be a rather large open neighborhood of e.

To prove this result fix an $X\in\mathfrak{g}$ and let $\ell_X:O\to\mathbb{R}$ be the function, $\ell_X(\beta)=\varphi(\beta,X)=\beta(X)$. We will first prove

Lemma 12.9.1. *ℓ_X is a Bott–Morse function whose critical set is the set*

$$\{\beta\in O\,,\quad \text{ad}(X)^*\beta=0\}\,. \tag{12.107}$$

Proof. $(d\ell_X)_\beta=0$ iff, for all $Y\in\mathfrak{g}$

$$\text{ad}(Y)^*\beta(X)=0\,. \tag{12.108}$$

But $\text{ad}(Y)^*\beta(X)=-\beta([Y,X])$ and $-\beta([Y,X])=\text{ad}(X)^*\beta(Y)$. This proves the lemma.

To prove the theorem let C_φ be the critical set of this generating function. Then C_φ intersects the fiber above $g = \exp X$ in the critical set of ℓ_X which, by the lemma is just the set

$$\{\beta \in O\,, \quad (\mathrm{Ad})^*(g)\beta = \beta\}\,. \tag{12.109}$$

Hence by (12.80) the inclusion map $C_\varphi \to T^*G$ maps C_φ onto Λ_O.

□

Let μ_O be the symplectic volume form on O. Then since $\lambda_m(g) = \gamma(g, \hbar)$, $\hbar = 1/m$, is in the space of oscillatory functions, $I^0(\Lambda_O, G)$ there exists an amplitude, $a(\beta, X, \hbar)$, defined locally near $X = 0$ such that

$$\gamma_m(\exp X) = \int a(\beta, X, \hbar) e^{2\pi i m \varphi(\beta, X)} \mu_O\,. \tag{12.110}$$

Kirillov's theorem ([Ki]) is the following explicit formula for this amplitude.

Let $\gamma_\rho : O \times \mathfrak{g} \to S^1$ be the function

$$\gamma_\rho(\beta, X) = e^{2\pi i \langle \mathrm{Ad}(g)^* \rho, X\rangle} \tag{12.111}$$

where the "g" in the expression on the right is the unique element of $G \bmod T$ satisfying $\beta = \mathrm{Ad}(g)^*\beta_O$. Also let $v(\alpha)$, for $\alpha \in t^*$, be the symplectic volume of the coadjoint orbit through α and let $j(X)$ be the square root of the Jacobian at X of the exponential map, $\mathfrak{g} \to G$. Then for $\hbar = 1/m$

$$a(\beta, X, \hbar) = j(X)^{-1} \frac{v(\rho + m\beta_0)}{v(\beta_0)} \gamma_\rho(\beta, X)\,. \tag{12.112}$$

Note by the way that

$$\frac{v(\rho + m\beta_0)}{v(\beta_0)} = \hbar^{-d}(1 + O(h)) \tag{12.113}$$

where $2d = \dim O$ and hence by (8.1) the oscillatory integral (12.110) is in fact in $I^0(\Lambda_O; G)$. We won't attempt to prove this result but we will show how to get from it a concrete description of the Maslov factor in the symbol of $\gamma(g, \hbar)$ on Λ_O.

We first note that for $X \in t_{\mathrm{reg}}$, the critical points of ℓ_X are, by (12.82) and the lemma, just the points, $w\beta_0$, $w \in W$. Identifying the tangent space to O at β_0 with $\mathfrak{g}/t$ we will prove

Lemma 12.9.2. *The Hessian, $(d^2\ell_X)_{\beta_0}$, of ℓ_X at β_0 is the bilinear form*

$$(Y, Z) \in \mathfrak{g}/t \to \beta_0([Y, [Z, X]])\,. \tag{12.114}$$

Remark

Since $\mathrm{Ad}(X)^*\beta_0 = 0$

$$\begin{aligned} 0 &= \beta_0([Y,Z],X) \\ &= \beta_0([Y,[Z,X]]) - \beta_0([Z,[Y,X]]) \end{aligned}$$

so the bilinear form (12.114) is symmetric.

Proof of the lemma: By definition

$$\begin{aligned} (d^2\ell_X)_{\beta_0}(Y,Z) &= (\mathrm{ad}(Y)^*\,\mathrm{ad}(Z)^*\beta_0)(X) \\ &= \beta_0(\mathrm{ad}(Z)(\mathrm{ad}\,Y)X) \\ &= \beta_0([Z,[Y,\Lambda]])\,. \end{aligned}$$

□

By (12.86) we can identify $\mathfrak{g}/t \otimes \mathbb{C}$ with $n \oplus \bar{n}$ and take as basis vectors of n the vectors, Z_α, $\alpha \in \phi^+$. We then get by (12.86)

$$(d^2\ell_X)_{\beta_0}(Z_\alpha, \bar{Z}_\beta) = 0 \tag{12.115}$$

if $\alpha \neq \beta$ and

$$(d^2\ell_X)_{\beta_0}(Z_\alpha, \bar{Z}_\alpha) = 2\pi\alpha(X)\beta_0(X_\alpha) \tag{12.116}$$

where

$$X_\alpha = \sqrt{-1}[Z_\alpha, \bar{Z}_\alpha] \in t\,. \tag{12.117}$$

However (see for instance [FH])

$$\beta_0 \in \mathrm{Int}_+^* \Leftrightarrow \beta_0(X_\alpha) > 0 \;\text{ for all }\quad \alpha \in \phi^*\,.$$

Hence by (12.116) we get for the signature of $(d^2\ell_X)_{\beta_0}$ the expression

$$2(\#\{\alpha \in \phi^+, \alpha(X) > 0\} - \#\{\alpha \in \phi^2, \alpha(X) < 0\}) \tag{12.118}$$

and hence

$$\exp i\frac{\Pi}{4}\,\mathrm{sgn}(d^2\ell_X)_{\beta_0} = i^d\,\frac{\Pi \sin 2\pi\alpha(X)}{\pi|\sin 2\pi\alpha(X)|} \tag{12.119}$$

for points $X \in t_{reg}$ close to $X = 0$. But for $g = \exp X$, the right hand side is $D(g)/|D(g)|$ where $D(g)$ is the Weyl denominator (12.96). Thus finally

$$\exp\frac{i\pi}{4}\,\mathrm{sgn}(d^2\ell_X)_{\beta_0} = \frac{D(g)}{|D(g)|}\,. \tag{12.120}$$

A similar computation shows that

$$\exp\frac{i\pi}{4}\,\mathrm{sgn}(d^2\ell_X)_{w\beta_0} = \frac{D(g)}{|D(g)|}(-1)^w\,. \tag{12.121}$$

Thus the right hand side of (12.121) is just the value of the function m (in our formula in §12.8) for the symbol of $\gamma(g,h)|dg|^{\frac{1}{2}})$ at the points (g,β_0) of Λ_O above $g \in T_{\mathrm{reg}}$ and the left hand side is the formula for the Maslov factor in this symbol at these points as defined in § refsec8.5

12.10 The GKRS character formula.

We will next show that Kirillov's theorem gives a generating function description of Λ_O at arbitrary points of Λ_O. To see this let k_0 be an element of G (which, without loss of generality we can assume to be in T) and let K be its centralizer in G. Then T is contained in K and the normalizer, $N_K(T)$ of T in K is contained in the normalizer, $N(T)$, of T in G; so one gets an inclusion of Weyl groups:

$$W_K = N_K(T)/T \to N(T)/T = W$$

and to each right coset, $W_K w$, in W a K-orbit

$$O_K^w = Kw\beta_0 \tag{12.122}$$

in O. As we saw in §12.8 the union of these K orbits is the preimage of k_0 in Λ_O. We will, for the moment, view (12.122) as sitting inside $\mathfrak{k}^*$ and apply (a slightly modified version of) the Kirillov theorem to it. More explicitly: the mapping, $X \in \mathfrak{k} \to (\exp X)k_0 \in K$, is a diffeomorphism of a neighborhood of 0 in $\mathfrak{k}$ onto a neighborhood, U_0 of k_0 in K, and since k_0 is in the center of K the function

$$\phi^w : O_K^w \times U_0 \to \mathbb{R} \tag{12.123}$$

defined by the pairing

$$\phi^w(\beta, k) = \langle \beta, \exp^{-1}(kk_0^{-1})\rangle \tag{12.124}$$

is a generating function for the character Lagrangian, $\Lambda_{O_K^w} \to K$ over the neighborhood, U_0 of k_0. Now let $C(k_0)$ be the conjugacy xlass of k_0 in G and for each $g \in C(k_0)$ let K_g be the group gKg^{-1}, let $O_g^w = \mathrm{Ad}(g)^* O_K^w$ be the coadjoint orbit of K_g corresponding to O_K^w , let $U_g = gU_0 g^{-1}$, let $Z_g^w = O_g^w \times U_g$ and let

$$\varphi_g^w : Z_g^w \to \mathbb{R} \tag{12.125}$$

be the function, $\varphi_g^w(\beta, u) = \varphi^w(\mathrm{Ad}^*(g^{-1})\beta, g^{-1}ug)$. Then φ_g^w is a generating function for the character Lagrangian of O_g^w with respect to the fibration

$$Z_g^w = O_g^w \times U_g \xrightarrow{\pi^w} U_g\,. \tag{12.126}$$

One can easily amalgamate all these data into a single set of generating data for Λ_O on a neighborhood, U, in G of $C(k_0)$. Namely let Z^w be the disjoint union of the Z_g^w's, let U be the disjoint union of the U_g's, let $\pi^w : Z^w \to U$ be the fiber mapping whose restriction to Z_g^w is the projection (12.126) and let $\phi^w : Z^w \to \mathbb{R}$ be the function whose restriction to Z_g^w is the function (12.125). We claim that

Theorem 12.10.1. *φ^w is a generating function for the component of Λ_O above U containing $w\beta_0$.*

Proof. This is an immediate consequence of the fact that, restricted to the set, $\pi^{-1}(U_g) = O_g^w \times U_g$, φ^w is a generating function for the character Lagrangian of the coadjoint orbit, O_g^w, in $\mathfrak{k}_g$. □

Example

If we take k_0 to be an element of T_{reg}, $K = T$, $O_K^w = w\beta_0$, and the description of Λ_0 that we get from this theorem is just our description of Λ_{reg} in §12.8.

This result can be viewed as a semi-classical formulation of a well-known result of Gross–Kostant–Ramond–Sternberg. (See [GKRS] and [Ko].) To describe their result and its connection with the construction above, we will begin by making a careful choice of the representative, "w" in the right coset, $W_K w$, of $W_K \backslash W$; i.e. the w involved in the definition of the coadjoint K-orbit (12.122). If w_0 is any element of this coset, then there exists a unique $w_1 \in W_K$ such that $w_1 w_0 \beta_0$ is a dominant weight of the group, K, i.e. sits inside the interior of the positive Weyl chamber $(t_K^*)^+$ of t^*. Thus letting $w = w_1 w_0$, there exists a unique w in the coset $W_K w$ such that $w\beta_0$ is a dominant weight of K. In fact the same is true for the weights

$$m w\beta_0 + w\rho - \rho_K \tag{12.127}$$

where $2\rho_K$ is the sum of the positive roots of K and $m = 1/h$ is a positive integer. Let $\gamma_K^w(k,\hbar)$ be the character of the irreducible representation of K with weight (12.127). Then the GKRS theorem asserts that for $k \in T$ the character of the irreducible representation of G with highest weight, $m\beta_0$, is expressible in terms of these characters by the simple identity

$$\gamma(k,\hbar) = \frac{1}{\Delta} \sum (-1)^w \gamma_K^w(k,\hbar) \tag{12.128}$$

where

$$\Delta = \prod e^{\pi i\alpha} - e^{-\pi i\alpha}\,, \quad \alpha \in \Phi^+ \tag{12.129}$$

and ϕ^+ is the set of positive roots of G that are not positive roots of K. Thus, locally near $k = k_0$ in T, the summands in (12.128) are given by oscillatory integrals associated with the fibration (12.123) and the generating functions (12.112), and the amplitudes in the oscillatory integrals are given by K analogues of the amplitude (12.112) in the Kirillov formula.

12.11 The pseudodifferential operators on line bundles

In their article, "Sur la formule des traces",[PU] Thierry Paul and Alejandro Uribe develop an approach to the theory of semi-classical pseudodifferential operators which involves identifying the algebra of semi-classical pseudodifferential operators on a manifold, X, with the algebra of S^1-invariant classical pseudodifferential operators on $X \times S^1$. Their idea is the following: Let X, for simplicity, be $\mathbb{R}^n$ and let $A\left(x, \frac{\partial}{\partial x}, \frac{\partial}{\partial\theta}\right)$ be an invariant m^{th} order pseudodifferential differential operator having, as in §9.2, a polyhomogeneous symbol

$$a(x,\xi,\tau) = \sum a_j(x,\xi,\tau)\,,\ \infty < j < m\,, \tag{12.130}$$

τ being the dual variable to the angle variable, θ on S^1. Then, for functions of the form, $f(x)e^{ik\theta}$

$$\hbar^m A(fe^{ik\theta}) = \left(A_\hbar\left(x, \frac{\partial}{\partial x}\right) f\right) e^{ik\theta} \tag{12.131}$$

where $\hbar = 1/k$ and

$$A_\hbar\left(x, \frac{\partial}{\partial x}\right) f = \hbar^m \left(\frac{1}{2\pi h}\right)^n \int a(x,\xi(\hbar,1/\hbar)e^{\frac{i(x-y)\cdot\xi}{\hbar}} f(y)\,dy\,d\xi \tag{12.132}$$

is a zeroth order semi-classical pseudodifferential operator with leading symbol

$$a_m(x,\xi,1) \tag{12.133}$$

where $a_m(x,\xi,\tau)$ is the leading symbol of A. The definitions (12.131) and (12.132) set up a correspondence between classical pseudodifferential operators on $\mathbb{R}^n \times S^1$ and semi-classical pseudodifferential operators on $\mathbb{R}^n$, and in [PU], Paul and Uribe use this correspondence to give a classical proof of the semi-classical trace formula that we discussed in §11.5.3.

We will show below that their approach adapts nicely to the theory of pseudodifferential operators on line bundles: Let $\mathbb{L} \to X$ be a complex line bundle on X, $\langle\,,\,\rangle : \mathbb{L} \to \mathbb{R}$, a Hermitian inner product on $\mathbb{L}$, and $P \subset \mathbb{L}$ the unit circle bundle associated with $\langle\,,\,\rangle$. Let $\Gamma(\mathbb{L})$ denote the space of smooth sections of $\mathbb{L}$. Then the correspondence (12.131)–(12.132) can be converted into a correspondence which associates to an S^1-invariant classical pseudodifferential operator

$$A : \Gamma(\mathbb{L}) \to \Gamma(\mathbb{L}) \tag{12.134}$$

(a family of classical pseudodifferential operators

$$A_k :\ \Gamma(\mathbb{L}^k) \to \Gamma(\mathbb{L}^k), \tag{12.135}$$

and these, in turn, can be viewed as a semi-classical pseudo-differential operator $A_\hbar,\ \hbar = 1/k$.

To see this, we will begin by identifying $\Gamma(\mathbb{L}^k)$ with the space $C_k^\infty(P)$ of functions on P which have the transformation properties

$$f(e^{i\theta}p) = e^{ik\theta} f(p)\,,. \tag{12.136}$$

Now let A be an m^{th} order S^1-invariant classical pseudodifferential operator on P and define $A_\hbar$ to be the operator

$$A_\hbar = \hbar^m A|\mathcal{C}_k^\infty(P)\,, \quad k = 1/\hbar\,. \tag{12.137}$$

Locally the operators, A and $A_\hbar$ look like the operators (12.131) and (12.132). Namely let U be an open subset of X and $P|U = U \times S^1$ a trivialization of P over U. Then on U, A is a classical pseudodifferential operator of the form (12.131), $A_\hbar$ is the operator (12.131), and its symbol is defined by the expression (**??**). The global definition of its symbol, however, is a little trickier: From the action of S^1 on P one gets a Hamiltonian action of S^1 on T^*P with moment map

$$(p,\eta) \in T^*P \xrightarrow{\phi} \langle \eta, \left(\frac{\partial}{\partial\theta}\right)_p \rangle \,. \tag{12.138}$$

Let

$$(T^*P)_{\text{red}} = \phi^{-1}(1)/S^1 \tag{12.139}$$

be the symplectic reduction of T^*P at $\phi = 1$. Then since A is S^1 invariant its leading symbol, $\sigma(A) : T^*P \to \mathbb{C}$, is also S^1 invariant so the restriction

$$\sigma(A)|\phi^{-1}(1) = \sigma(A_\hbar) \tag{12.140}$$

is in fact a function on $(T^*P)_{\text{red}}$ and this we will define to be the symbol of $A_\hbar$. (Note that if $P_U = U \times S^1$ is a trivialization of P then by (12.138) , $(T^*P_U)_{\text{red}} = T^*U$ and the definition, (12.140) coincides with the definition (12.132).)

This correspondence between A and $A_\hbar$ is particularly easy to describe if A is a differential operator. In this case the restriction of A to U is of the form

$$A = \sum_{|\mu|+r=m} a_{\mu,r}(x)\left(\frac{1}{i}\frac{\partial}{\partial\theta}\right)^r D_x^\mu \tag{12.141}$$

and $A_\hbar$ is the operator

$$\sum_{j=0}^m h^i \sum_{|\mu|+r+j=m} a_{\mu,r}(hD_X)^\mu \,. \tag{12.142}$$

One can get a more intrinsic description of these operators by equipping $\mathbb{L}$ with a connection

$$\nabla : \mathcal{C}^\infty(\mathbb{L}) \to \mathcal{C}^\infty(\mathbb{L} \otimes T^*X) \,. \tag{12.143}$$

This connection extends to a connection

$$\nabla : \mathcal{C}^\infty(\mathbb{L}^k) \to \mathcal{C}^\infty(\mathbb{L}^k \otimes T^*X) \tag{12.144}$$

with the property: $\nabla s^k = ks^{k-1}\nabla s$, and in particular if $s : U \to P$ is a trivializing section of $\mathbb{L}$ and v a vector field on U, the operator

$$\frac{1}{i}\nabla_v : \mathcal{C}^\infty(\mathbb{L}^k) \to \mathcal{C}^\infty(\mathbb{L}^k) \tag{12.145}$$

is given locally on U by the expression:

$$\frac{\hbar}{i}\nabla_v f s^k = \frac{\hbar}{i}(L_v f + k a_v f)s^k \tag{12.146}$$

where by (12.2) and (12.7)

$$a_v = \frac{1}{i}\langle \frac{\nabla s}{s}, v\rangle = 2\pi\langle s^*\alpha, v\rangle\,. \tag{12.147}$$

More generally, *every* semi-classical differential operator of order m

$$A_\hbar : \mathcal{C}^\infty(\mathbb{L}^k) \to \mathcal{C}^\infty(\mathbb{L}^k)\,, \quad k = 1/h$$

can be written, intrinsically, or a coordinate patch, U as an operator of this form:

$$A_\hbar = \sum_{j=0}^{m} \hbar^j \sum_{\mu+j+r=m} a_{\mu,r}(x)\left(\frac{\hbar}{i}\nabla\partial/\partial x\right)^\mu \tag{12.148}$$

as one can see by letting $s : U \to P$ be a local trivialization of P and comparing the operator

$$s^{-k}A_\hbar s^k = \sum_{j=0}^{m}\hbar^j \sum_{|\mu|+j+r=m} a_{\mu,r}(\hbar D_X + \langle s^*\alpha, \partial/\partial x\rangle)^\mu \tag{12.149}$$

with the operator (12.142).

We have seen that the symbols of these semi-classical operators live globally on $(T^*P)_{\text{red}}$; however, we will show below that these symbols can be thought of as living on the usual tangent bundle of X. However, the price we will have to pay for this is that we will have to equip this tangent bundle with a non-standard symplectic form. We first observe that the zero level set of the moment mapping (12.138) is just the pull-back, π^*T^*X, of T^*X with respect to the fibration, $\pi : P \to X$, i.e. each point (p,η) on this level set is of the form, $\eta = (d\pi)_p^*\xi$ for some $\xi \in T^*_{\pi(p)}$. Thus the reduced space $\phi^{-1}(0)/S^1$ can be canonically identified with T^*X.

Now let α be the connection form on P, let $\beta = 2\pi\alpha$ and let

$$\gamma_\beta : T^*P \to T^*P \tag{12.150}$$

be the map,

$$(p,\eta) \to (p, \eta + \beta_p)\,.$$

Since $\langle\beta, \partial/\partial\theta\rangle = 1$ this map maps the zero level set of the moment map (12.138) onto the level set, $\phi = 1$. Moreover if ω is the symplectic form on T^*P

$$\gamma_\beta^*\omega = \omega + \pi_P^*\, d\beta \tag{12.151}$$

where $\pi_P : T^*P \to P$ is the cotangent fibration. Thus if $\text{curv}(\nabla)$ is the curvature form of the connection, ∇, and $\nu_X = 2\pi\,\text{curv}(\nabla)$, $d\beta = \pi^*\nu_X$ and

$$\gamma_\beta^*\omega = \omega + (\pi\circ\pi_P)^*\nu_X. \tag{12.152}$$

Moreover, since γ_β is S^1 invariant and maps the level set $\phi^{-1}(0)$ onto the level set $\phi^{-1}(1)$, it induces a map of $\phi^{-1}(0)/S^1$ onto $\phi^{-1}(1)/S^1$, i.e. a diffeomprphism

$$\rho_\beta : T^*X \to (T^*P)_{\text{red}} \tag{12.153}$$

and by (12.152) this satisfies

$$\rho_\beta^* \omega_{\text{red}} = \omega_X + \pi_X^* \nu_X \tag{12.154}$$

where ω_X is the standard symplectic form on T^*X and $\pi_X : T^*X \to X$ is the cotangent fibration of X. In other words $(T^*P)_{\text{red}}$ with its natural "reduced" symplectic form is symplectomorphic to T^*X with its "α-twisted" symplectic form (12.154). Via this isomorphism we can think of the symbol of a semi-classical pseudodifferential operator of type (12.141) as being a function on T^*X; however if we want to compute the Hamiltonian flow associated with this symbol we will have to do so with respect to the symplectic form, (12.154), *not* with respect to the usual symplectic form on T^*X.

12.12 Spectral properties of the operators, $A_\hbar$

In the last two sections of this chapter we will describe some applications of the results of earlier sections to spectral theory. In this section we will show how to extend the trace formula of chapter 10 to operators of the form (12.137) and in the next section show how to reformulate this result as a theorem in "equivariant" spectral theory for circle actions on manifolds. We will then make use of the semi-classical version of the Weyl character formula that we proved in §12.8 to generalize this theorem to arbitrary compact Lie groups. As above let $A : \mathcal{C}^\infty(P) \to \mathcal{C}^\infty(P)$ be a classical m^{th} order pseudodifferential operator. We will assume in this section that A is selfadjoint and elliptic and we will also assume, for simplicity, that X is compact. Since A is selfadjoint its symbol is real valued and ellipticity implies that, for fixed x, $|\sigma(A(x,\xi))| \to +\infty$ as ξ tends to infinity. Therefore if $P_U = U \times S$ then on U, the leading terms, $a_0(x,\xi,\tau)$, in (12.130) satisfies

$$|a_0(x,\xi,\tau)| \geq C(|\xi|^2 + |\tau|^2)^{m/2} \tag{12.155}$$

for some positive constant C. Hence since the operator

$$A_\hbar = \hbar^m k A | \mathcal{C}_k^\infty(P) \tag{12.156}$$

is a standard semi-classical pseudodifferential operator of the form (12.132) on U with symbol $a_0(x,\xi,1)$ its symbol satisfies an estimate of the form

$$|\sigma(A_\hbar)(x,\xi| \geq C(|\xi|^2 + 1)^{m/2} \tag{12.157}$$

on U. Finally, the assumption that X is compact implies that A has discrete spectrum: there is an orthonormal basis, φ_j, of $L^2(X)$ with $\varphi_j \in \mathcal{C}^\infty(X)$ and

$$A\varphi_j = \lambda_j \varphi_j \,. \tag{12.158}$$

the λ_j's tending to infinity as j tends to infinity. Thus, if f is in $\mathcal{C}_0^\infty(\mathbb{R})$ the operator $f(A)$ is the finite rank smoothing operator

$$f(A)\varphi_k = f(\lambda_k)\varphi_k \,. \tag{12.159}$$

It follows that similar assertions are true for the restriction of A to $\mathcal{C}_k^\infty(P)$ and hence for the semi-classical operator

$$A_\hbar = \hbar^m A|\mathcal{C}_k^\infty(P)\,, \quad \hbar = 1/k\,.$$

In particular the operator

$$f(A_\hbar) = f(\hbar^m A)|\mathcal{C}_k^\infty(P) \tag{12.160}$$

is a finite rank smoothing operator. Moreover, the restriction to U of $f(A_\hbar)$ has to coincide with the operator $f(A_\hbar|U)$ so by the results of §10, its Schwartz kernel is of the form

$$\left(\frac{1}{2\pi\hbar}\right)^n \int f(\sigma(A(x,\xi)))a_U(x,\xi,h)e^{\frac{i(x-y)\cdot\xi}{\hbar}}\,d\xi \tag{12.161}$$

with

$$a_U(x,\xi,\hbar) \sim \sum_{\ell=0}^{\infty} a_{\ell,U}(x,\xi)\hbar^\ell \tag{12.162}$$

and $a_U(x,\xi,0) = 1$.

Now let U_j, $j = 1,\dots,N$ be an open cover of X by coordinate patches such that, for each j, $P|U_j \simeq U_j \times S^1$ and let ρ_j and χ_j be functions in $\mathcal{C}_0^\infty(U_j)$ with the property, $\sum \rho_j = 1$, and $\chi_j \equiv 1$ on the support of ρ_j. Then, by pseudolocality,

$$f(A_\hbar) = \sum \chi_j f(A_\hbar|U_j)\rho_j \tag{12.163}$$

$\operatorname{mod} O(\hbar^\infty)$; so modulo $O(\hbar^\infty)$, the trace of $f(A)$ is given by the sum

$$(2\pi h)^{-n}\sum_j \int f(\sigma(A)(x,\xi))\rho_j(x)a_{U_j}(x,\xi)\,dx\,d\xi \tag{12.164}$$

and hence admits an asymptotic expansion

$$\operatorname{trace} f(A_\hbar) \sim (2\pi\hbar)^{-n}\sum_{r=0}^{\infty} C_r\hbar^{ir} \tag{12.165}$$

with leading term

$$c_0 = \int_{(T^*P)_{\mathrm{red}}} f(\sigma(A))\mu \tag{12.166}$$

where μ is the symplectic volume form on $(T^*P)_{\mathrm{red}}$. In particular one easily deduces from this the Weyl estimate

$$N_h(I) \sim (2\pi h)^{-n}\operatorname{vol}(\sigma(A)^{-1}(I)) \tag{12.167}$$

where I is any bounded sub-interval of $\mathbb{R}$, and $N_h(I)$ the number of eigenvalues of $A_\hbar$ on I. Translating this back into an assertion about A this gives us the estimate (12.167) for the number of eigenvalues of the operator, $A|\mathcal{C}_k^\infty(P)$ lying on the interval, $k^m I$.

12.13 Equivariant spectral problems in semi-classical analysis

Let X be a manifold, G a compact connected Lie group and $\tau : G \to Diff(X)$ a $\mathcal{C}^\infty$ action of G on X. Suppose that

$$A : \mathcal{C}_0^\infty(X) \to \mathcal{C}^\infty(X)$$

is a self-adjoint operator, e.g. a classical or semi-classical pseudodifferential operator which commutes with this action, and suppose, for simplicity, that the spectrum of A is discrete. Then for each eigenvalue λ, one gets a representation ρ_λ, of G on the corresponding eigenspace

$$V_\lambda = \{\varphi \in \mathcal{C}^\infty(X)\,, \quad A\varphi = \lambda\varphi\}, \tag{12.168}$$

and the equivariant spectrum of A is, by definition, the set of data

$$\{(\lambda, \rho_\lambda)\,; \quad \lambda \in \mathrm{Spec}(A)\}\,. \tag{12.169}$$

For instance if G is S^1 the equivariant spectrum consists of the eigenvalues of A plus, for each eigenvalue, λ, a list

$$m(\lambda, k)\,, \quad -\infty < k < \infty \tag{12.170}$$

of the multiplicities with which the irreducible representations, $\rho_k = e^{i\theta k}$, of S^1 occur as subrepresentations of ρ_λ. An example is the operator, $A : \mathcal{C}^\infty(P) \to \mathcal{C}^\infty(P)$, in §12.12 whose equivariant spectrum is the spectrum

$$\lambda_1(h)\ \lambda_2(h), \ldots, \quad h = 1/k$$

of the operator $A|\mathcal{C}_k^\infty(P)$, i.e. a formatted version of the usual spectrum of A in which we keep track of the dependence on k.

In equivariant spectral theory one is concerned with the same basic problem as in ordinary spectral theory: to extract geometric information from the data, (12.169); however, one has a larger arsenal of weapons at one's disposal for doing so; for instance, for A a semi-classical pseudodifferential operator of order zero, one has twisted versions:

$$\mathrm{trace}(\tau_g^* e^{-tA_h})\,, \quad g \in G$$

of the heat trace invariants that we discussed in chapter 10, and twisted versions

$$\mathrm{trace}\left(\tau_g^* e^{\frac{iA_h}{h}}\right)\,, \quad g \in G$$

of the wave trace invariants that we discussed in chapter 11. To cite another example: for the operator, A, in §11, one can consider in addition to its usual heat-trace invariants the more sophisticated heat trace invariants (12.165).

The goal of this section will to be to generalize the trace formula (12.165)-(12.166) viewed in this light (i.e. viewed as a theorem about *equivariant* spectra) to groups other than S^1. More explicitly we will let X be a compact G-manifold, $A_h : \mathcal{C}^\infty(X) \to \mathcal{C}^\infty(X)$, a G-equivariant semi-classical pseudodifferential operator of order zero and $\rho_{k\alpha}$ the irreducible representation of G with highest weight, $k\alpha$, and will prove an analogue of the formulas (12.165)-(12.166) involving the spectral data,

$$\lambda, m(k\alpha, \lambda)\,, \quad \lambda \in \operatorname{Spec}(A_\hbar) \tag{12.171}$$

where $h = 1/k$ and $m(k\alpha, \lambda)$ is the multiplicity with which $\rho_{k\alpha}$ occurs in V_λ. Our main result will be a trace formula for the operator

$$\int \tau_g^*(A_\hbar)\overline{\gamma_{k\alpha}(g)}\, dg \tag{12.172}$$

where $\gamma_{k\alpha}$ is the Weyl character of the representation, $\rho_{k\alpha}$. To prove this result we will make crucial use of the fact that $\gamma_{k\alpha}$ can be viewed as an element

$$\gamma_\hbar(g) \in I^0(G, \Lambda_O)\,, \quad \hbar = 1/k$$

where O is the coadjoint orbit through α. To keep the exposition below from getting too unwieldy we will henceforth make the following simplifying assumptions.

1. $A_\hbar$ is self-adjoint as an operator on $L^2(X)$.

2. For some open subinterval, I, of $\mathbb{R}$ $\sigma(A_\hbar)^{-1}(I)$ is compact.

3. O is a generic coadjoint orbit of G, i.e. $\dim O = \dim G - \dim T$.

4. Let $\Phi : T^*X \to \mathfrak{g}$ be the moment map associated with the lifted action of G on T^*X. Then G acts freely on the preimage

$$\Sigma = \Phi^{-1}(O)\,. \tag{12.173}$$

Concerning this last hypothesis we note that if G acts freely on Σ then the reduced space

$$(T^*X)_O = \Sigma/G \tag{12.174}$$

is well-defined. We will denote by μ_O its symplectic volume form and by $\sigma(A)_{\text{red}}$ the reduced symbol of $A_\hbar$: the function on $(T^*X)_O$ defined by

$$\iota_\Sigma^*\sigma(A_\hbar) = \pi_\Sigma^*\sigma(A_h)_{\text{red}} \tag{12.175}$$

where $\iota_{\sum}$ is the inclusion of $\sum$ into T^*X and $\pi_{\sum}$ the projection of $\sum$ onto $(T^*X)^O$. (This is well-defined since $\sigma(A_\hbar)$ is G-invariant.) With this notation we will prove

Theorem 12.13.1. *For $f \in \mathcal{C}_0^\infty(I)$ the trace of the operator (12.171) admits an asymptotic expansion*

$$(2\pi\hbar)^{-m} \sum_{k=0}^{\infty} c_k \hbar^k \tag{12.176}$$

where $m = \dim X - \frac{1}{2}(\dim T + \dim G)$ and

$$c_0 = \int_{(T^*X)_O} f(\sigma(A_\hbar)_{\mathrm{red}})\mu_{\mathrm{red}}\,. \tag{12.177}$$

As a first step in the proof we will prove

Lemma 12.13.1. *Let $Q \in \Psi^0(X)$ be a semi-classical zeroth order pseudodifferential operator with compact microsupport. The the Schwartz kernel of the operator, $\tau_g^* Q$, viewed as an oscillatory function on $X \times X \times G$, is an element of the space $I^{-n}(\Gamma_\tau\,;\, X \times X \times G)$ where Γ_τ is the moment Lagrangian associated with the lifted action of G on T^*X.*

Proof. We recall that if X and Y are manifolds and $f : X \to Y$ a $\mathcal{C}^\infty$ map, this map lifts to a canonical relation

$$\Gamma_f : T^*X \to T^*Y$$

with the defining property: $(x, \xi, y, \eta) \in \Gamma_f$ iff $y = f(x)$ and $\xi = (df_x)^*\nu$. We pointed out in §12.5 that for the map

$$\tau : X \times G \to X\,, \quad (x, g) \to \tau_g(x)$$

Γ_τ is just the moment Lagrangian, and we get the lemma above by applying this observation to $\tau^* Q$.

□

We now turn to the proof of the theorem:

Proof. Let $M = T^*X$, and, by rearranging factors, regard Γ_τ as being the usual moment canonical relation

$$\Gamma_\tau : M^- \times M \to T^*G\,. \tag{12.178}$$

Then by the lemma the operator

$$L_Q : \mathcal{C}^\infty(G) \to \mathcal{C}^\infty(X \times X) \tag{12.179}$$

mapping φ to $\int \tau_g^* Q \varphi(g)\, dg$ is a semi-classical Fourier integral operator quantizing the canonical relation

$$\Gamma^\dagger : T^*G \to M^- \times M\,,$$

and the trace operator

$$\text{trace} : k(x, y) \in \mathcal{C}^\infty(X \times X) \to \int k(x, x)\, dx$$

is a semi-classical Fourier integral operator quantizing the canonical relation,

$$\Delta^\dagger : M \times M \to \text{pt.}\,.$$

Thus with $Q = f(A)$ the expression (12.171) can be interpreted as the operator, trace $\circ L_Q$ applied to $\gamma_\hbar(g) \in I^0(\Lambda_O, G)$. But a point $(p,q) \in M \times M$ is in $\Gamma^t \circ \Lambda_O$ iff

(a) $q = \tau_g^* p$

(b) $\phi(p) \in O$

and

(c) $\text{Ad}(g)^* \phi(p) = \phi(p)$

and such a point is in $\Delta^\dagger : M \times M \to \text{pt.}$ iff, in addition

(d) $p = q$.

However, by (b), p is in $\phi^{-1}(O)$, and since G acts freely on $\phi^{-1}(O)$

(e) $g = e$.

Thus the canonical relations

$$\Gamma^\dagger \circ \Lambda_O : \text{pt.} \to M^- \times M$$

and

$$\Delta^\dagger : M^- \times M \to \text{pt.}$$

compose cleanly and by the clean composition formula of chapter 8 §8.13, the expression

$$\text{trace} \int \tau_g^* Q \hbar(g)\, dg \tag{12.180}$$

is an element of $I^{-m}(\text{pt.})$ i.e. a formal power series

$$c(h) = (2\pi\hbar)^{-m} \sum c_k \hbar^k \tag{12.181}$$

whose leading symbol can be computed by the "clean" symbol calculus of chapter 8, i.e. as a symbolic integral over the fibers of the fibration

$$(\Delta^t) \star (\Gamma^t \circ \Lambda_O) \to \Delta^\dagger \circ \Gamma^\dagger \circ \Lambda_O\,. \tag{12.182}$$

But since $\Delta^t \circ \Gamma^t \circ \Lambda_O = \text{pt.}$, this becomes an integral over the space $\Delta^\dagger \star (\Gamma^t \circ \Lambda_O)$ itself, i.e. over the set

$$\{(p,p) \in M^- \times M\,, \quad p \in \Sigma\}\,. \tag{12.183}$$

In other words the symbol, c_0, of the series (12.181) can be computed by a symbolic computation only involving the symbols of Q restricted to the set, Σ and of $\gamma_\hbar$ restricted to the set

$$\{g \in G\,,\ \tau_g^* p = p\,, \quad \text{for some} \ \ p \in \Sigma\}\,, \tag{12.184}$$

and by condition (e) this is just the set $\{e\}$. Thus this symbolic integral over (12.182) only involves the symbol of $\gamma_\hbar$ restricted to the fiber, O, of Λ_O above $e \in G$; and by the Kirillov formula this is just the symplectic volume form, μ_O, on O; i.e. doesn't involve the complicated Maslov factors in the expression (12.105). From this one easily deduces that the integral over $\Delta^t * (\Gamma^t \circ \Lambda_O)$ in the clean composition formula for symbols that we cited above gives us for the symbol of the expression

$$\text{trace} \int \tau_g^* Q \gamma_h(g)\, dg$$

the integral

$$\int_{\sum} \iota_{\sum}^* \sigma(Q)\, dg (\pi_{\sum})^* \mu_{\text{red}} \tag{12.185}$$

which in the case of $Q = f(A_h)$ reduces the integral (12.177). □

From this result we get the following generalization of the Weyl law (12.167):

Theorem 12.13.2. *Let $\lambda_i(\hbar), i = 1, \ldots, \ell$ be the eigenvalues of $A_\hbar$ lying on the interval I and let V_I be the sum of the corresponding eigenspaces and $N_\hbar(I)$ the multiplicity with which the representation $\rho_{m\alpha}$, $m = 1/\hbar$ occurs as a subrepresentation of the representation of G on V_I. Then*

$$N_\hbar(I) \sim (2\pi\hbar)^{-m} \operatorname{vol}\left(\sigma_{\text{red}}(A)^{-1}(I)\right). \tag{12.186}$$

Chapter 13

Spectral theory and Stone's theorem.

In this chapter we gather various facts from functional analysis that we use, or which motivate our constructions in Chapter 10. All the material we present here is standard, and is available in excellent modern texts such as Davies, Reed-Simon, Hislop-Sigal, Schecter, and in the classical text by Yosida. Our problem is that the results we gather here are scattered among these texts. So we had to steer a course between giving a complete and self-contained presentation of this material (which would involve writing a whole book) and giving a bare boned listing of the results.

We also present some results relating semi-classical analysis to functional analysis on L_2 which allow us to provide the background material for the results of Chapters 9-11. Once again the material is standard and can be found in the texts by Dimassi-Sjöstrand, Evans-Zworski, and Martinez. And once again we steer a course between giving a complete and self-contained presentation of this material giving a bare boned listing of the results.

The key results are:

- **The spectral theorem for self-adjoint operators.** We will recall the somewhat subtle definition of a self-adjoint operator on a Hilbert space below. The spectral theorem then (in functional calculus form) allows the construction of an operator $f(A)$ for any self-adjoint operator A, and for a reasonable class of functions f on $\mathbb{R}$. The map $f \mapsto f(A)$ is to be linear, multiplicative, and take complex conjugation into adjoint, i.e. $\overline{f} \mapsto f(A)^*$. (The map $f \mapsto f(A)$ should be non-trivial and unique in an appropriate sense.) For the full spectral theorem, we want the class of functions to include the bounded Borel measurable functions on $\mathbb{R}$. For our purposes it is enough to have such a functional calculus for functions belonging to the Schwartz space $\mathcal{S}(\mathbb{R})$, or even for smooth functions of compact support.

- **Stone's theorem.** This has two parts: 1) Given any self-adjoint operator

A, the family $U(t) = \exp itA$ is a unitary one parameter group of transformations. This is an immediate consequence of the spectral theorem if the class of functions in the functional calculus includes the functions $x \mapsto e^{itx}$ as is the case for the full spectral theorem. 2) Conversely, given a unitary one parameter group $U(t)$, its infinitesimal generator (see below for the definition) is self-adjoint.

Starting from Stone's theorem, one can get the functional calculus for functions in the Schwartz space $\mathcal{S}(\mathbb{R})$ by a straightforward generalization of the formula for the inverse Fourier transform, namely by setting

$$f(A) = \frac{1}{\sqrt{2\pi}} \int \hat{f}(t)U(t)dt$$

where $\hat{f}$ is the Fourier transform of f. So it is desirable to have a proof (and formulation) of Stone's theorem independent of the spectral theorem. In fact, Stone's theorem is a special case of the Hille-Yosida theorem about one-parameter semi-groups on Frechet spaces and their infiinitesimal generators. So we discuss the Hille-Yosida theorem and its proof below.

One of the main efforts and tools in Chapter 10 is to provide and use a semi-classical version of Stone's theorem.

- **The Dynkin-Helffer-Sjöstrand formula.** We stated this formula, namely
$$f(P) := -\frac{1}{\pi} \int_{\mathbb{C}} \frac{\partial \tilde{f}}{\partial \bar{z}} R(z, P) dxdy, \qquad (10.2)$$
in Chapter 10. In fact, it is an immediate consequence of the multiplication version of the spectral theorem.

  The Dynkin- Helffer-Sjösrand formula allows one to show that if H is a self adjoint operator associated to a pseudo-differential operator with real Weyl symbol p, then for $f \in C_0^\infty(\mathbb{R})$, the operator $f(H)$ provided by the functional calculus is associated to $f(p)$.

- **The Calderon-Vallaincourt theorem.** This says that if P is a semi-classical pseudo-differential operator satisfying appropriate conditions, it extends to a family of bounded operator on L_2 whose L_2 bounds are given in terms of the sup norms of a finite number of derivatives of p.

13.1 Unbounded operators, their domains, their spectra and their resolvents.

13.1.1 Linear operators and their graphs.

Let B and C be Banach spaces. We make $B \oplus C$ into a Banach space via

$$\|\{x, y\}\| = \|x\| + \|y\|.$$

Here we are using $\{x, y\}$ to denote the ordered pair of elements $x \in B$ and $y \in C$ so as to avoid any conflict with our notation for scalar product in a Hilbert space. So $\{x, y\}$ is just another way of writing $x \oplus y$. A subspace

$$\Gamma \subset B \oplus C$$

will be called a **graph** (more precisely a graph of a linear transformation) if

$$\{0, y\} \in \Gamma \quad \Rightarrow \quad y = 0.$$

Another way of saying the same thing is

$$\{x, y_1\} \in \Gamma \ \text{ and } \ \{x, y_2\} \in \Gamma \quad \Rightarrow \quad y_1 = y_2.$$

In other words, if $\{x, y\} \in \Gamma$ then y is determined by x.

In the language of ¶ 3.3.5 Γ is a graph if it co-injective as a relation.

The domain and the map of a graph.

So let

$D(\Gamma)$ denote the set of all $x \in B$ such that there is a $y \in C$ with $\{x, y\} \in \Gamma$.

Then $D(\Gamma)$ is a linear subspace of B, but, and this is very important, $D(\Gamma)$ is *not* necessarily a closed subspace. We have a linear map

$$T(\Gamma): \ D(\Gamma) \to C, \quad Tx = y \ \text{ where } \ \{x, y\} \in \Gamma.$$

The graph of a linear transformation.

Equally well, we could start with the linear transformation: Suppose we are given a (not necessarily closed) subspace $D(T) \subset B$ and a linear transformation

$$T : D(T) \to C.$$

We can then consider its graph $\Gamma(T) \subset B \oplus C$ which consists of all

$$\{x, Tx\}, \quad x \in D(T).$$

Thus the notion of a graph, and the notion of a linear transformation defined only on a subspace of B are logically equivalent. When we start with T (as usually will be the case) we will write $D(T)$ for the domain of T and $\Gamma(T)$ for the corresponding graph.

There is a certain amount of abuse of language here, in that when we write T, we mean to include $D(T)$ and hence $\Gamma(T)$ as part of the definition.

13.1.2 Closed linear transformations.

A linear transformation is said to be **closed** if its graph is a closed subspace of $B \oplus C$.

Let us disentangle what this says for the operator T. It says that if $f_n \in D(T)$ then

$$f_n \to f \text{ and } Tf_n \to g \quad \Rightarrow \quad f \in D(T) \text{ and } Tf = g.$$

This is a much weaker requirement than continuity. Continuity of T would say that $f_n \to f$ alone would imply that Tf_n converges to Tf. Closedness says that if we know that *both*

$$f_n \text{ converges } \textbf{and } g_n = Tf_n \text{ converges to } g$$

then we can conclude that $f = \lim f_n$ lies in $D(T)$ and that $Tf = g$.

13.1.3 The resolvent, the resolvent set and the spectrum.

The resolvent and the resolvent set .

Let $T : B \to B$ be an operator with domain $D = D(T)$. A complex number z is said to belong to the **resolvent set** of T if the operator

$$zI - T$$

maps D onto all of B and has a bounded inverse. We denote this bounded inverse by $R(z, T)$ or $R_z(T)$ or simply by R_z if T is understood. So

$$R(z, T) := (zI - T)^{-1} \text{ maps } B \to D(T)$$

and is bounded. $R(z, T)$ is called the **resolvent** of T at the complex number z.

The spectrum.

The complement of the resolvent set is called the **spectrum** of T and is denoted by $\operatorname{spec}(T)$.

Theorem 13.1.1. *The set* $\operatorname{spec}(T)$ *is a closed subset of* $\mathbb{C}$. *In fact, if* $z \notin \operatorname{spec}(T)$ *and* $c := \|R(z, T)\|$ *then the spectrum does not intersect the disk*

$$\{w \in \mathbb{C} | \ |(w - z)| < c^{-1}\}.$$

For w *in this disk*

$$R(w, T) = \sum_0^\infty (-(w - z))^n R(z, T)^{n+1}$$

and so is an analytic operator valued function of w. *Differentiating this series term by term shows that*

$$\frac{d}{dz} R(z, T) = -R(z, T)^2.$$

Proof, part 1. The series given in the theorem certainly converges in operator norm to a bounded operator for w in the disk. For a fixed w in the disk, let C denote the operator which is the sum of the series. Then

$$C = R(z,T) - (w-z)R(z,T)C.$$

This shows that C maps B to $D(T)$ and has kernel equal to the kernel of $R(z,T)$ which is $\{0\}$. So C is a bounded injective operator mapping B into D. Also

$$C = R(z,T) - (w-z)CR(z,T)$$

which shows that the image of $R(z,T)$ is contained in the image of C and so the image of C is all of D.

Proof, part 2.

$$C := \sum_0^\infty (-(w-z))^n R(z,T)^{n+1}.$$

If $f \in D$ and $g = (zI-T)f$ then $f = R(z,T)g$ and so $Cg = f - (w-z)Cf$ and hence

$$C(zf - Tf) = f - (w-z)Cf$$

or

$$C(-Tf) = f - wCf \quad \text{so} \quad C(wI-T)f = f$$

showing that C is a left inverse for $wI - T$. A similar argument shows that it is a right inverse. So we have proved that the series converges to the resolvent proving that the resolvent set is open and hence that the spectrum is closed. The rest of the theorem is immediate. □

A useful lemma.

Lemma 13.1.1. *If $T : B \to B$ is an operator on a Banach space whose spectrum is not the entire plane then T is closed.*

Proof. Assume that $R = R(z,T)$ exists for some z. Suppose that f_n is a sequence of elements in the domain of T with $f_n \to f$ and $Tf_n \to g$. Set $h_n := (zI-T)f_n$ so

$$h_n \to zf - g.$$

Then $R(zf-g) = \lim Rh_n = \lim f_n = f$. Since R maps B to the domain of T this shows that f lies in this domain. Multiplying $R(zf-g) = f$ by $zI - T$ gives

$$zf - g = zf - Tf$$

showing that $Tf = g$. □

13.1.4 The resolvent identities.

The first resolvent identity.

Let z and w both belong to the resolvent set. We have

$$wI - T = (w - z)I + (zI - T).$$

Multiplying this equation on the left by R_w gives

$$I = (w - z)R_w + R_w(zI - T),$$

and multiplying this on the right by R_z gives

$$R_z - R_w = (w - z)R_w R_z.$$

It follows (interchanging z and w) that $R_z R_w = R_w R_z$, in other words

all resolvents R_z commute with one another.

So we can write the preceding equation as

$$R_z - R_w = (w - z)R_z R_w. \tag{13.1}$$

This equation, known as the **first resolvent equation** (or identity), dates back to the theory of integral equations in the 19th century.

Relation with the Laplace transform.

Let L denote the Laplace transform:

$$L(G)(\lambda) = \int_0^\infty e^{-\lambda t} G(t) dt.$$

Here, say, G is a bounded continuous function with values in a Banach space. So $L(G)(\lambda)$ is defined for Re $\lambda > 0$.

If we take G to be $\mathbb{C}$ valued, given by $G(t) = e^{zt}$ where Re $z \leq 0$ we have

$$L(G)(\lambda) = \frac{1}{\lambda - z}.$$

More generally, suppose that $G(t) = e^{At}$ where A is a bounded operator on a Banach space and e^{At} is given by the usual exponential series. Assume that A is such that e^{At} is uniformly bounded (in the operator norm) in t so that the Laplace transform $L(G)$ is defined for Re $\lambda > 0$. Then

$$(\lambda I - A)L(G)(\lambda) = \int_0^\infty (\lambda I - A)e^{-(\lambda I - A)t} dt = I.$$

In other words,

$$L(G)(\lambda) = R(\lambda, A) \quad \text{for Re } \lambda > 0.$$

One of our tasks will be to generalize this to a broader class of operators.

Let us return to the general Laplace transform.

Integration by parts shows that $L(G')(\lambda) = \lambda L(G)(\lambda) - G(0)$. Apply this to G given by

$$G(t) = \int_0^t e^{-c(t-s)} g(s) ds.$$

Then

$$G'(t) = g(t) - cG(t), \quad G(0) = 0,$$

so $L(g)(\lambda) = L(G')(\lambda) + cL(G)(\lambda) = (c+\lambda)L(G)(\lambda)$. Thus the Laplace transform of G is given by

$$L(G)(\lambda) = \frac{1}{\lambda + c} L(g)(\lambda). \tag{13.2}$$

Let F be the Laplace transform of f. Then we claim that

$$\int_0^\infty \int_0^\infty e^{-\lambda s - \mu t} f(s+t) ds dt = \frac{F(\mu) - F(\lambda)}{\lambda - \mu} \tag{13.3}$$

when $\lambda \neq \mu$.

Proof. We may assume (by analytic continuation) that λ and μ are real, and, without loss of generality, that $\lambda > \mu$. Write the integral with respect to t as $e^{-\lambda s} \int_0^\infty e^{-\mu t} f(s+t) dt$. Make the change of variables $w = s + t$ so that $\int_0^\infty e^{-\mu t} f(s+t) dt$

$$= e^{\mu s} \int_s^\infty e^{-\mu w} f(w) dw = e^{\mu s} F(\mu) - e^{\mu s} \int_0^s e^{-\mu w} f(w) dw.$$

Then apply the Laplace transform with respect to s and use (13.2) with $c = -\mu$ for the second term □

Suppose that f takes values in a Banach algebra. Then (by uniqueness of the Laplace transform) we see that f satisfies the identity

$$f(s+t) = f(s)f(t)$$

if and only if its Laplace transform F satisfies the identity

$$F(\lambda) \cdot F(\mu) = \frac{F(\mu) - F(\lambda)}{\lambda - \mu}.$$

In other words, the first resolvent identity is a reflection of the semigroup property $f(s+t) = f(s)f(t)$ in case $f(s) = e^{sA}$ when e^{sA} is uniformly bounded in s.

The second resolvent identity.

The first resolvent identity relates the resolvents of a fixed operator at two different points in the resolvent set. The second resolvent identity relates the resolvents of two different operators at the same point. Here is how it goes:

Let a and b be operators whose range is the whole space and with bounded inverses. Then

$$a^{-1} - b^{-1} = a^{-1}(b-a)b^{-1}$$

assuming that the right hand side is defined. For example, if A and B are closed operators with $D(B-A) \supset D(A)$ we get

$$R_A(z) - R_B(z) = R_A(z)(B-A)R_B(z). \tag{13.4}$$

This is the **second resolvent identity**. It also dates back to the 19th century.

13.1.5 The adjoint of a densely defined linear operator.

Suppose that we have a linear operator $T : D(T) \to C$ and let us make the hypothesis that

$$D(T) \textit{ is dense in } B.$$

Any element of B^* is then completely determined by its restriction to $D(T)$. Now consider

$$\Gamma(T)^* \subset C^* \oplus B^*$$

defined by

$$\{\ell, m\} \in \Gamma(T)^* \quad \Leftrightarrow \quad \langle \ell, Tx\rangle = \langle m, x\rangle \quad \forall\ x \in D(T). \tag{13.5}$$

Since m is determined by its restriction to $D(T)$, we see that $\Gamma^* = \Gamma(T^*)$ is indeed a graph. (It is easy to check that it is a linear subspace of $C^* \oplus B^*$.) In other words we have defined a linear transformation

$$T^* := T(\Gamma(T)^*)$$

whose domain consists of all $\ell \in C^*$ such that there exists an $m \in B^*$ for which $\langle \ell, Tx\rangle = \langle m, x\rangle \quad \forall\ x \in D(T)$.

The adjoint of a linear transformation is closed.

If $\ell_n \to \ell$ and $m_n \to m$ then the definition of convergence in these spaces implies that for any $x \in D(T)$ we have

$$\langle \ell, Tx\rangle = \lim \langle \ell_n, Tx\rangle = \lim\langle m_n, x\rangle = \langle m, x\rangle.$$

If we let x range over all of $D(T)$ we conclude that Γ^* is a closed subspace of $C^* \oplus B^*$. In other words we have proved

Theorem 13.1.2. *If $T : D(T) \to C$ is a linear transformation whose domain $D(T)$ is dense in B, it has a well defined adjoint T^* whose graph is given by (13.5). Furthermore T^* is a closed operator.*

13.2 Self-adjoint operators on a Hilbert space.

13.2.1 The graph and the adjoint of an operator on a Hilbert space.

Now let us restrict to the case where $B = C = \mathfrak{H}$ is a Hilbert space, so we may identify $B^* = C^* = \mathfrak{H}^*$ with $\mathfrak{H}$ via the Riesz representation theorem which says that the most general continuous linear function on $\mathfrak{H}$ is given by scalar product with an element of $\mathfrak{H}$.

If $T : D(T) \to \mathfrak{H}$ is an operator with $D(T)$ dense in $\mathfrak{H}$ we may identify the graph of T^* as consisting of all $\{g, h\} \in \mathfrak{H} \oplus \mathfrak{H}$ such that

$$(Tx, g) = (x, h) \quad \forall x \in D(T)$$

and then write

$$(Tx, g) = (x, T^*g) \quad \forall \; x \in D(T), \quad g \in D(T^*).$$

Notice that we can describe the graph of T^* as being the orthogonal complement in $\mathfrak{H} \oplus \mathfrak{H}$ of the subspace

$$M := \{\{Tx, -x\} \quad x \in D(T)\}.$$

The domain of the adjoint.

The domain $\mathcal{D}$ of T^* consists of those g such that there is an h with $(Tx, g) = (x, h)$ for all x in the domain of T. We claim that $\mathcal{D}$ is dense in $\mathfrak{H}$. Suppose not. Then there would be some $z \in \mathfrak{H}$ with $(z, g) = 0$ for all $g \in D(T^*)$. Thus $\{z, 0\} \perp M^\perp = D(T^*)$. But $(M^\perp)^\perp$ is the closure $\overline{M}$ of M. This means that there is a sequence $x_n \in D(T)$ such that $Tx_n \to z$ and $x_n \to 0$. So if we assume that T is closed, we conclude that $z = 0$. In short, if T is a closed densely defined operator so is T^*.

13.2.2 Self-adjoint operators.

We now come to the central definition: An operator A defined on a domain $D(A) \subset \mathfrak{H}$ is called **self-adjoint** if

- $D(A)$ is dense in $\mathfrak{H}$,
- $D(A) = D(A^*)$, and
- $Ax = A^*x \; \forall x \in D(A)$.

The conditions about the domain $D(A)$ are rather subtle. For the moment we record one immediate consequence of the theorem of the preceding section:

Proposition 13.2.1. *Any self adjoint operator is closed.*

13.2.3 Symmetric operators.

A densely defined operator S on a Hilbert space is called **symmetric** if

- $D(S) \subset D(S^*)$ and
- $Sx = S^*x \ \ \forall\, x \in D(S)$.

Another way of saying the same thing is: S is symmetric if $D(S)$ is dense and

$$(Sx, y) = (x, Sy) \quad \forall \ \ x, y \in D(S).$$

Every self-adjoint operator is symmetric but not every symmetric operator is self adjoint. This subtle difference will only become clear as we go along.

A sufficient condition for a symmetric operator to be self-adjoint.

Let A be a symmetric operator on a Hilbert space $\mathfrak{H}$. The following theorem will be very useful:

Theorem 13.2.1. *If there is a complex number z such that $A + zI$ and $A + \overline{z}I$ both map $D(A)$ surjectively onto $\mathfrak{H}$ then A is self-adjoint.*

We must show that if ψ and f are such that

$$(f, \phi) = (\psi, A\phi) \quad \forall\ \phi \in D(A)$$

then

$$\psi \in D(A) \ \text{ and } \quad A\psi = f.$$

Once we show that $\psi \in D(A)$ then, since $D(A)$ is assumed to be dense and $(\psi, A\phi) = (A\psi, \phi)$ for $\psi, \phi \in D(A)$ and this equals (f, ψ) by hypothesis, we conclude that $A\psi = f$. So we must prove that $\psi \in D(A)$.

Proof. Choose $w \in D(A)$ such that $(A+\overline{z}I)w = f+\overline{z}\psi$. Then for any $\phi \in D(A)$

$$(\psi, (A + zI)\phi) = (f + \overline{z}\psi, \phi) = (Aw + \overline{z}w, \phi) = (w, A\phi + z\phi).$$

Then choose $\phi \in D(A)$ such that $(A+zI)\phi = \psi - w$. So $(\psi, \psi - w) = (w, \psi - w)$ and hence $\|\psi - w\|^2 = 0$, i.e $\psi = w$, so

$$\psi \in D(A).$$

□

Here is an important application of the theorem we just proved.

Multiplication operators.

Let $(X, \mathcal{F}, \mu)$ be a measure space and let $\mathfrak{H} := L_2(X, \mu)$. Let a be a real valued $\mathcal{F}$ measurable function on X with the property that a is bounded on any measurable subset of X of finite measure. Let

$$\mathcal{D} := \left\{ u \in \mathfrak{H} \,\middle|\, \int_X (1 + a^2)|u|^2 d\mu < \infty \right\}.$$

Notice that $\mathcal{D}$ is dense in $\mathfrak{H}$. Let A be the linear operator

$$u \mapsto au$$

defined on the domain $\mathcal{D}$. Notice that A is symmetric.

Proposition 13.2.2. *The operator A with domain $\mathcal{D}$ is self-adjoint.*

Proof. The operator consisting of multiplication by

$$\frac{1}{i + a}$$

is bounded since $\left|\frac{1}{i+a}\right| \leq 1$ and clearly maps $\mathfrak{H}$ to $\mathcal{D}$. Its inverse is multiplication by $i + a$. Similarly multiplication by $-i + a$ maps $\mathcal{D}$ onto $\mathfrak{H}$. So we may take $z = i$ in Theorem 13.2.1. □

Notice that for any bounded measurable function f on $\mathbb{R}$, we may define the operator $f(A)$ to consist of multiplication by $f(a)$. It is clear that the map $f \mapsto f(A)$ satisfies all the desired properties of the functional calculus. In particular

$$R(z, A) \text{ consists of multiplication by } \frac{1}{z - a} \tag{13.6}$$

when $\operatorname{Im} z \neq 0$.

The Dynkin-Helffer-Sjöstrand formula for multiplication operators.

Recall that if $f \in C_0^\infty(\mathbb{R}$, a function $\tilde{f} \in C_0^\infty(\mathbb{C})$ is called an almost analytic extension of f if

$$\left|\overline{\partial}\tilde{f}\right| \leq C_n |\operatorname{Im} z|^N \ \forall N \in \mathbb{N} \ \text{ and } \ f_{|\mathbb{R}} = f.$$

It is easy to show that almost analytic extensions always exist. For a proof, see Davies or Dimassi-Sjöstrand. We will reproduce the proof from Dimassi-Sjöstrand at the end of this chapter.

Recall also that for any $g \in C_0^\infty(\mathbb{C})$ we have the formula

$$g(w) = -\frac{1}{\pi} \int_{\mathbb{C}} \frac{\partial g}{\partial \overline{z}} \cdot \frac{1}{z - w} dx dy.$$

Applied to the function $\tilde{f}$ and $w \in \mathbb{R}$ we have

$$f(w) = -\frac{1}{\pi}\int_{\mathbb{C}} \frac{\partial \tilde{f}}{\partial \bar{z}} \cdot \frac{1}{z-w} dxdy.$$

Letting $w = a(m),\ m \in X$ we see that the function $f(a)$ is given by

$$f(a) = -\frac{1}{\pi}\int_{\mathbb{C}} \frac{\partial \tilde{f}}{\partial \bar{z}} \cdot \frac{1}{z-a} dxdy.$$

Hence the operator $f(A)$ is given by

$$f(A) := -\frac{1}{\pi}\int_{\mathbb{C}} \frac{\partial \tilde{f}}{\partial \bar{z}} R(z,A) dxdy. \qquad (10.2)$$

This proves the Dynkin-Helffer-Sjöstrand formula (10.2) for the case of multiplication operators. A bit later we will prove the multiplication version of the spectral theorem which says that any self-adjoint operator on a separable Hilbert space is unitarily equivalent to a multiplication operator. This implies that (10.2) is true in general.

Using the Fourier transform.

The Fourier transform is a unitary operator on $L_2(\mathbb{R}^n)$ (Plancherel's theorem) , and carries constant coefficent partial differential operators into multiplication by a polynomial. So

Proposition 13.2.3. *If D is a constant coefficient differential operator which is carried by the Fourier transform into a* real *polynomial, then D is self-adjoint.*

An example is the Laplacian, which goes over into multiplication by $\|k\|^2$ under the Fourier transform. The domain of the Laplacian consists of those $f \in L_2$ whose Fourier transform $\hat{f}$ have the property that $\|k\|^2 \hat{f}(k) \in L_2$.

We shall see below that there is a vast generalization of this fact. Namely for a broad class of real Weyl symbols, p, the associated operators P, (originally defined, say as maps from $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$) in fact define self adjoint operators on $L_2(\mathbb{R}^n)$ when passing to the closure of these operators.

13.2.4 The spectrum of a self-adjoint operator is real.

The following theorem is central. Once we will have stated and proved the spectral theorem, the following theorem will be an immediate consequence. But we will proceed in the opposite direction, first proving the theorem and then using it to prove the spectral theorem:

Theorem 13.2.2. *Let A be a self-adjoint operator on a Hilbert space $\mathfrak{H}$ with domain $D = D(A)$. Let*

$$c = \lambda + i\mu, \quad \mu \neq 0$$

be a complex number with non-zero imaginary part. Then

$$(cI - A) : D(A) \to \mathfrak{H}$$

is bijective. Furthermore the inverse transformation

$$(cI - A)^{-1} : \mathfrak{H} \to D(A)$$

is bounded and in fact

$$\|(cI - A)^{-1}\| \le \frac{1}{|\mu|}. \tag{13.7}$$

We will prove this theorem in stages:

We show that $\|f\|^2 = \|(\lambda I - A)g\|^2 + \mu^2\|g\|^2$ **for** $g \in D(A)$.

Let $g \in D(A)$ and set $f := (cI - A)g = [\lambda I - A]g + i\mu g$. Then $\|f\|^2 = (f, f) =$

$$\|[\lambda I - A]g\|^2 + \mu^2\|g\|^2 + ([\lambda I - A]g, i\mu g) + (i\mu g, [\lambda I - A]g).$$

The last two terms cancel: Indeed, since $g \in D(A)$ and A is self adjoint we have

$$(\mu g, [\lambda I - A]g) = (\mu[\lambda I - A]g, g) = ([\lambda I - A]g, \mu g)$$

since μ is real. Hence

$$([\lambda I - A]g, i\mu g) = -i(\mu g, [\lambda I - A]g).$$

We have thus proved that

$$\|f\|^2 = \|(\lambda I - A)g\|^2 + \mu^2\|g\|^2. \tag{13.8}$$

We show that $\|(cI - A)^{-1}\| \le \frac{1}{|\mu|}$.

It follows from (13.8) that

$$\|f\|^2 \ge \mu^2\|g\|^2$$

for all $g \in D(A)$. Since $|\mu| > 0$, we see that $f = 0 \Rightarrow g = 0$ so $(cI - A)$ is injective on $D(A)$, and furthermore that $(cI - A)^{-1}$ (which is defined on the image of $(cI - A)$) satisfies

$$\|(cI - A)^{-1}\| \le \frac{1}{|\mu|}.$$

We must show that the image of $(cI - A)$ is all of $\mathfrak{H}$.

We show the image of $(cI - A)$ is dense in $\mathfrak{H}$.

For this it is enough to show that there is no $h \neq 0 \in \mathfrak{H}$ which is orthogonal to im $(cI - A)$. So suppose that

$$([cI - A]g, h) = 0 \quad \forall g \in D(A).$$

Then

$$(g, \overline{c}h) = (cg, h) = (Ag, h) \ \ \forall g \in D(A)$$

which says that $h \in D(A^*)$ and $A^*h = \overline{c}h$. But A is self adjoint so $h \in D(A)$ and $Ah = \overline{c}h$. Thus

$$\overline{c}(h, h) = (\overline{c}h, h) = (Ah, h) = (h, Ah) = (h, \overline{c}h) = c(h, h).$$

Since $c \neq \overline{c}$ this is impossible unless $h = 0$. We have now established that the image of $cI - A$ is dense in $\mathfrak{H}$.

We show that image of $(cI - A)$ is all of $\mathfrak{H}$, completing the proof of the theorem.

Let $f \in \mathfrak{H}$. We know that we can find

$$f_n = (cI - A)g_n, \quad g_n \in D(A) \quad \text{with } f_n \to f.$$

The sequence f_n is convergent, hence Cauchy, and from

$$\|(cI - A)^{-1}\| \leq \frac{1}{|\mu|} \quad (13.7)$$

applied to elements of im $D(A)$ we know that

$$\|g_m - g_n\| \leq |\mu|^{-1}\|f_n - f_m\|.$$

Hence the sequence $\{g_n\}$ is Cauchy, so $g_n \to g$ for some $g \in \mathfrak{H}$. But we know that A is a closed operator. Hence $g \in D(A)$ and $(cI - A)g = f$. $\square$

13.3 Stone's theorem.

As indicated in the introduction to this chapter, we will present a generalization of Stone's theorem due to Hille and Yosida. The setting will be the study of a one parameter semi-group on a Frechet space. A Frechet space $\mathbf{F}$ is a vector space with a topology defined by a sequence of semi-norms and which is complete. An important example is the Schwartz space $\mathcal{S}$. Let $\mathbf{F}$ be such a space.

13.3.1 Equibounded continuous semi-groups.

We want to consider a one parameter family of operators T_t on $\mathbf{F}$ defined for all $t \geq 0$ and which satisfy the following conditions:

- $T_0 = I$
- $T_t \circ T_s = T_{t+s}$
- $\lim_{t\to t_0} T_t x = T_{t_0} x \ \forall t_0 \geq 0$ and $x \in \mathbf{F}$.
- For any defining seminorm p there is a defining seminorm q and a constant K such that $p(T_t x) \leq K q(x)$ for all $t \geq 0$ and all $x \in \mathbf{F}$.

We call such a family an **equibounded continuous semigroup**. We will usually drop the adjective "continuous" and even "equibounded" since we will not be considering any other kind of semigroup.

The treatment here will essentially follow that of Yosida, *Functional Analysis* especially Chapter IX.

13.3.2 The infinitesimal generator.

We are going to begin by showing that every such semigroup has an "infinitesimal generator", i.e. can be written in some sense as $T_t = e^{At}$.

The definition of A.

We define the operator A as

$$Ax = \lim_{t \searrow 0} \frac{1}{t}(T_t - I)x.$$

That is, A is the operator so defined on the domain $D(A)$ consisting of those x for which the limit exists.

Our first task is to show that $D(A)$ is dense in $\mathbf{F}$. For this we begin with a "putative resolvent"

$$R(z) := \int_0^\infty e^{-zt} T_t dt \tag{13.9}$$

which is defined (by the boundedness and continuity properties of T_t) for all z with Re $z > 0$.

One of our tasks will be to show that $R(z)$ as defined in (13.9) is in fact the resolvent of A. We begin by checking that every element of im $R(z)$ belongs to $D(A)$ and that $(zI - A)R(z) = I$: We have

$$\frac{1}{h}(T_h - I)R(z)x = \frac{1}{h}\int_0^\infty e^{-zt} T_{t+h} x dt - \frac{1}{h}\int_0^\infty e^{-zt} T_t x dt =$$

$$\frac{1}{h}\int_h^\infty e^{-z(r-h)} T_r x dr - \frac{1}{h}\int_0^\infty e^{-zt} T_t x dt$$

$$= \frac{e^{zh}-1}{h}\int_h^\infty e^{-zt}T_t x dt - \frac{1}{h}\int_0^h e^{-zt}T_t x dt$$
$$= \frac{e^{zh}-1}{h}\left[R(z)x - \int_0^h e^{-zt}T_t dt\right] - \frac{1}{h}\int_0^h e^{-zt}T_t x dt.$$

If we now let $h \to 0$, the integral inside the bracket tends to zero, and the expression on the right tends to x since $T_0 = I$. We thus see that

$$R(z)x \in D(A)$$

and

$$AR(z) = zR(z) - I,$$

or, rewriting this in a more familiar form,

$$(zI - A)R(z) = I. \tag{13.10}$$

This equation says that $R(z)$ is a right inverse for $zI - A$. It will require a lot more work to show that it is also a left inverse.

We show that $D(A)$ is dense in $\mathbf{F}$.

We will prove that $D(A)$ is dense in $\mathbf{F}$ by showing that, taking s to be real, that

$$\lim_{s\to\infty} sR(s)x = x \quad \forall\ x \in \mathbf{F}. \tag{13.11}$$

Indeed,

$$\int_0^\infty se^{-st}dt = 1$$

for any $s > 0$. So we can write

$$sR(s)x - x = s\int_0^\infty e^{-st}[T_t x - x]dt.$$

Applying any seminorm p we obtain

$$p(sR(s)x - x) \le s\int_0^\infty e^{-st}p(T_t x - x)dt.$$

For any $\epsilon > 0$ we can, by the continuity of T_t, find a $\delta > 0$ such that

$$p(T_t x - x) < \epsilon \ \ \forall \ \ 0 \le t \le \delta.$$

Now let us write

$$s\int_0^\infty e^{-st}p(T_t x - x)dt = s\int_0^\delta e^{-st}p(T_t x - x)dt + s\int_\delta^\infty e^{-st}p(T_t x - x)dt.$$

The first integral is bounded by

$$\epsilon s \int_0^\delta e^{-st} dt \le \epsilon s \int_0^\infty e^{-st} dt = \epsilon.$$

As to the second integral, let M be a bound for $p(T_t x) + p(x)$ which exists by the uniform boundedness of T_t. The triangle inequality says that $p(T_t x - x) \le p(T_t x) + p(x)$ so the second integral is bounded by

$$M \int_\delta^\infty s e^{-st} dt = M e^{-s\delta}.$$

This tends to 0 as $s \to \infty$, completing the proof that $sR(s)x \to x$ and hence that $D(A)$ is dense in $\mathbf{F}$.

The differential equation.

Theorem 13.3.1. *If $x \in D(A)$ then for any $t > 0$*

$$\lim_{h \to 0} \frac{1}{h}[T_{t+h} - T_t]x = AT_t x = T_t Ax.$$

In colloquial terms, we can formulate the theorem as saying that

$$\frac{d}{dt} T_t = AT_t = T_t A$$

in the sense that the appropriate limits exist when applied to $x \in D(A)$.

Proof. Since T_t is continuous in t, we have

$$T_t Ax = T_t \lim_{h \searrow 0} \frac{1}{h}[T_h - I]x = \lim_{h \searrow 0} \frac{1}{h}[T_t T_h - T_t]x =$$

$$\lim_{h \searrow 0} \frac{1}{h}[T_{t+h} - T_t]x = \lim_{h \searrow 0} \frac{1}{h}[T_h - I]T_t x$$

for $x \in D(A)$. This shows that $T_t x \in D(A)$ and

$$\lim_{h \searrow 0} \frac{1}{h}[T_{t+h} - T_t]x = AT_t x = T_t Ax.$$

To prove the theorem we must show that we can replace $h \searrow 0$ by $h \to 0$. Our strategy is to show that with the information that we already have about the existence of right handed derivatives, we can conclude that

$$T_t x - x = \int_0^t T_s Ax ds.$$

Since $t \mapsto T_t$ is continuous, this is enough to give the desired result. In order to establish the above equality, it is enough, by the Hahn-Banach theorem to prove that for any $\ell \in \mathbf{F}^*$ we have

$$\ell(T_t x) - \ell(x) = \int_0^t \ell(T_s Ax) ds.$$

In turn, it is enough to prove this equality for the real and imaginary parts of ℓ.

So it all boils down to a lemma in the theory of functions of a real variable:

A lemma in the theory of functions of a real variable.

Lemma 13.3.1. *Suppose that f is a continuous real valued function of t with the property that the right hand derivative*

$$\frac{d^+}{dt} f := \lim_{h \searrow 0} \frac{f(t+h) - f(t)}{h} = g(t)$$

exists for all t and $g(t)$ is continuous. Then f is differentiable with $f' = g$.

Proof of the lemma. We first prove that $\frac{d^+}{dt} f \geq 0$ on an interval $[a, b]$ implies that $f(b) \geq f(a)$. Suppose not. Then there exists an $\epsilon > 0$ such that

$$f(b) - f(a) < -\epsilon(b - a).$$

Set

$$F(t) := f(t) - f(a) + \epsilon(t - a).$$

Then $F(a) = 0$ and

$$\frac{d^+}{dt} F > 0.$$

At a this implies that there is some $c > a$ near a with $F(c) > 0$. On the other hand, since $F(b) < 0$, and F is continuous, there will be some point $s < b$ with $F(s) = 0$ and $F(t) < 0$ for $s < t \leq b$. This contradicts the fact that $[\frac{d^+}{dt} F](s) > 0$. Thus if $\frac{d^+}{dt} f \geq m$ on an interval $[t_1, t_2]$ we may apply the above result to $f(t) - mt$ to conclude that

$$f(t_2) - f(t_1) \geq m(t_2 - t_1),$$

and if $\frac{d^+}{dt} f(t) \leq M$ we can apply the above result to $Mt - f(t)$ to conclude that $f(t_2) - f(t_1) \leq M(t_2 - t_1)$. So if $m = \min g(t) = \min \frac{d^+}{dt} f$ on the interval $[t_1, t_2]$ and M is the maximum, we have

$$m \leq \frac{f(t_2) - f(t_1)}{t_2 - t_1} \leq M.$$

Since we are assuming that g is continuous, this is enough to prove that f is indeed differentiable with derivative g. $\square$.

13.3.3 The resolvent of the infinitesimal generator.

We have already verified that

$$R(z) = \int_0^\infty e^{-zt} T_t dt$$

maps $\mathbf{F}$ into $D(A)$ and satisfies

$$(zI - A)R(z) = I$$

for all z with Re $z > 0$, cf (13.10).

We shall now show that for this range of z

$$(zI - A)x = 0 \quad \Rightarrow \; x = 0 \quad \forall\, x \in D(A)$$

so that $(zI - A)^{-1}$ exists, and that it is given by $R(z)$:

Suppose that

$$Ax = zx \qquad x \in D(A)$$

and choose $\ell \in \mathbf{F}^*$ with $\ell(x) = 1$. Consider

$$\phi(t) := \ell(T_t x).$$

By Theorem 13.3.1 we know that ϕ is a differentiable function of t and satisfies the differential equation

$$\phi'(t) = \ell(T_t Ax) = \ell(T_t zx) = z\ell(T_t x) = z\phi(t), \quad \phi(0) = 1.$$

So

$$\phi(t) = e^{zt}$$

which is impossible since $\phi(t)$ is a bounded function of t and the right hand side of the above equation is not bounded for $t \geq 0$ since the real part of z is positive.

We have from (13.10) that

$$(zI - A)R(z)(zI - A)x = (zI - A)x$$

and we know that $R(z)(zI - A)x \in D(A)$. From the injectivity of $zI - A$ we conclude that $R(z)(zI - A)x = x$.

From $(zI - A)R(z) = I$ we see that $zI - A$ maps $\operatorname{im} R(z) \subset D(A)$ onto $\mathbf{F}$ so certainly $zI - A$ maps $D(A)$ onto $\mathbf{F}$ bijectively. Hence

$$\operatorname{im}(R(z)) = D(A), \quad \operatorname{im}(zI - A) = \mathbf{F}$$

and

$$R(z) = (zI - A)^{-1}.$$

Summary of where we are.

The resolvent $R(z) = R(z, A) := \int_0^\infty e^{-zt} T_t dt$ is defined as a strong limit for Re $z > 0$ and, for this range of z:

$$\begin{aligned} D(A) &= \operatorname{im}(R(z, A)) && (13.12) \\ AR(z, A)x &= && \\ R(z, A)Ax &= (zR(z, A) - I)x, \ x \in D(A) && (13.13) \\ AR(z, A)x &= (zR(z, A) - I)x, \ \forall \ x \in \mathbf{F} && (13.14) \\ \lim_{z \nearrow \infty} zR(z, A)x &= x \text{ for } z \text{ real } \ \forall x \in \mathbf{F}. && (13.15) \end{aligned}$$

The operator A is closed.

We claim that

Theorem 13.3.2. *The operator A is closed.*

Proof. Suppose that $x_n \in D(A)$, $x_n \to x$ and $y_n \to y$ where $y_n = Ax_n$. We must show that $x \in D(A)$ and $Ax = y$. Set

$$z_n := (I - A)x_n \quad \text{so} \quad z_n \to x - y.$$

Since $R(1, A) = (I - A)^{-1}$ is a bounded operator, we conclude that

$$x = \lim x_n = \lim (I - A)^{-1} z_n = (I - A)^{-1}(x - y).$$

From (13.12) we see that $x \in D(A)$ and from the preceding equation that $(I - A)x = x - y$ so $Ax = y$. □

13.3.4 Application to Stone's theorem.

We now have enough information to prove one half of Stone's theorem, namely that any continuous one parameter group of unitary transformations on a Hilbert space has an infinitesimal generator which is skew adjoint:

Suppose that $U(t)$ is a one-parameter group of unitary transformations on a Hilbert space $\mathfrak{H}$. We have $(U(t)x, y) = (x, U(t)^{-1}y) = (x, U(-t)y)$ and so differentiating at the origin shows that the infinitesimal generator A, which we know to be closed, is skew-symmetric:

$$(Ax, y) = -(x, Ay) \quad \forall \ x, y \in D(A).$$

Also the resolvents $(zI - A)^{-1}$ exist for all z which are not purely imaginary, and $(zI - A)$ maps $D(A)$ onto the whole Hilbert space $\mathfrak{H}$.

Writing $A = iT$ we see that T is symmetric and that $\pm iI + T$ is surjective. Hence T is self-adjoint. This proves that every one parameter group of unitary transformations is of the form e^{iTt} with T self-adjoint.

We now want to turn to the other half of Stone's theorem: We want to start with a self-adjoint operator T, and construct a (unique) one parameter group of unitary operators $U(t)$ whose infinitesimal generator is iT. As mentioned in the introduction to this chapter, this fact is an immediate consequence of the spectral theorem. But we want to derive the spectral theorem from Stone's theorem, so we need to provide a proof of this half of Stone's theorem which is independent of the spectral theorem. We will state and prove the Hille-Yosida theorem and find that this other half of Stone's theorem is a special case.

13.3.5 The exponential series and sufficient conditions for it to converge.

In finite dimensions we have the formula

$$e^{tB} = \sum_0^\infty \frac{t^k}{k!} B^k$$

with convergence guaranteed as a result of the convergence of the usual exponential series in one variable. (There are serious problems with this definition from the point of view of numerical implementation which we will not discuss here.)

In infinite dimensional spaces some additional assumptions have to be placed on an operator B before we can conclude that the above series converges. Here is a very stringent condition which nevertheless suffices for our purposes:

Let $\mathbf{F}$ be a Frechet space and B a continuous map of $\mathbf{F} \to \mathbf{F}$. We will assume that the B^k are **equibounded** in the sense that for any defining semi-norm p there is a constant K and a defining semi-norm q such that

$$p(B^k x) \leq K q(x) \quad \forall\ k = 1, 2, \ldots \quad \forall\ x \in \mathbf{F}.$$

Here the K and q are required to be independent of k and x.

Then

$$p(\sum_m^n \frac{t^k}{k!} B^k x) \leq \sum_m^n \frac{t^k}{k!} p(B^k x) \leq K q(x) \sum_n^n \frac{t^k}{k!}$$

and so

$$\sum_0^n \frac{t^k}{k!} B^k x$$

is a Cauchy sequence for each fixed t and x (and uniformly in any compact interval of t). It therefore converges to a limit. We will denote the map $x \mapsto \sum_0^\infty \frac{t^k}{k!} B^k x$ by

$$\exp(tB).$$

This map is linear, and the computation above shows that

$$p(\exp(tB)x) \leq K \exp(t) q(x).$$

The usual proof (using the binomial formula) shows that $t \mapsto \exp(tB)$ is a one parameter equibounded semi-group. More generally, if B and C are two such operators then if $BC = CB$ then $\exp(t(B+C)) = (\exp tB)(\exp tC)$.

Also, from the power series it follows that the infinitesimal generator of $\exp tB$ is B.

13.3.6 The Hille Yosida theorem.

Let us now return to the general case of an equibounded semigroup T_t with infinitesimal generator A on a Frechet space $\mathbf{F}$ where we know that the resolvent $R(z, A)$ for Re $z > 0$ is given by

$$R(z, A)x = \int_0^\infty e^{-zt} T_t x dt.$$

This formula shows that $R(z, A)x$ is continuous in z. The resolvent equation

$$R(z, A) - R(w, A) = (w - z)R(z, A)R(w, A)$$

then shows that $R(z, A)x$ is complex differentiable in z with derivative $-R(z, A)^2 x$. It then follows that $R(z, A)x$ has complex derivatives of all orders given by

$$\frac{d^n R(z, A)x}{dz^n} = (-1)^n n! R(z, A)^{n+1} x.$$

On the other hand, differentiating the integral formula for the resolvent n- times gives

$$\frac{d^n R(z, A)x}{dz^n} = \int_0^\infty e^{-zt}(-t)^n T_t dt$$

where differentiation under the integral sign is justified by the fact that the T_t are equicontinuous in t.

Putting the previous two equations together gives

$$(zR(z, A))^{n+1} x = \frac{z^{n+1}}{n!} \int_0^\infty e^{-zt} t^n T_t x dt.$$

This implies that for any semi-norm p we have

$$p((zR(z, A))^{n+1} x) \leq \frac{z^{n+1}}{n!} \int_0^\infty e^{-zt} t^n \sup_{t \geq 0} p(T_t x) dt = \sup_{t \geq 0} p(T_t x)$$

since

$$\int_0^\infty e^{-zt} t^n dt = \frac{n!}{z^{n+1}}.$$

Since the T_t are equibounded by hypothesis, we conclude

Proposition 13.3.1. *The family of operators $\{(zR(z, A))^n\}$ is equibounded in Re $z > 0$ and $n = 0, 1, 2, \ldots$.*

Statement of the Hille-Yosida theorem.

Theorem 13.3.3. [Hille -Yosida.] *Let A be an operator with dense domain $D(A)$, and such that the resolvents*

$$R(n, A) = (nI - A)^{-1}$$

exist and are bounded operators for $n = 1, 2, \ldots$. Then A is the infinitesimal generator of a uniquely determined equibounded semigroup if and only if the operators

$$\{(I - n^{-1}A)^{-m}\}$$

are equibounded in $m = 0, 1, 2 \ldots$ and $n = 1, 2, \ldots$.

If A is the infinitesimal generator of an equibounded semi-group then we know that the $\{(I - n^{-1}A)^{-m}\}$ are equibounded by virtue of the preceding proposition. So we must prove the converse. Our proof of the converse will be in several stages:

The definition of J_n.

Set

$$J_n = (I - n^{-1}A)^{-1}$$

so $J_n = n(nI - A)^{-1}$ and so for $x \in D(A)$ we have

$$J_n(nI - A)x = nx$$

or

$$J_n Ax = n(J_n - I)x.$$

Similarly $(nI - A)J_n = nI$ so $AJ_n = n(J_n - I)$. Thus we have

$$AJ_n x = J_n Ax = n(J_n - I)x \quad \forall\ x \in D(A). \tag{13.16}$$

Idea of the proof.

The idea of the proof is now this: By the results of the preceding section on the exponential series, we can construct the one parameter semigroup $s \mapsto \exp(sJ_n)$. Set $s = nt$. We can then form $e^{-nt}\exp(ntJ_n)$ which we can write as $\exp(tn(J_n - I)) = \exp(tAJ_n)$ by virtue of (13.16). We expect from

$$\lim_{s\to\infty} sR(s)x = x \quad \forall\ x \in \mathbf{F}$$

that

$$\lim_{n\to\infty} J_n x = x \quad \forall \quad x \in \mathbf{F}. \tag{13.17}$$

This then suggests that the limit of the $\exp(tAJ_n)$ be the desired semi-group.

Proof that $\lim_{n\to\infty} J_n x = x \quad \forall \quad x \in \mathbf{F}. \quad (13.17).$

So we begin by proving (13.17). We first prove it for $x \in D(A)$. For such x we have $(J_n - I)x = n^{-1}J_n Ax$ by (13.16) and this approaches zero since the J_n are equibounded. But since $D(A)$ is dense in $\mathbf{F}$ and the J_n are equibounded we conclude that (13.17) holds for all $x \in \mathbf{F}$.

Now define

$$T_t^{(n)} = \exp(tAJ_n) := \exp(nt(J_n - I)) = e^{-nt}\exp(ntJ_n).$$

We know from our study of the exponential series that

$$p(\exp(ntJ_n)x) \le \sum \frac{(nt)^k}{k!} p(J_n^k x) \le e^{nt}Kq(x)$$

which implies that

$$p(T_t^{(n)}x) \le Kq(x). \tag{13.18}$$

Thus the family of operators $\{T_t^{(n)}\}$ is equibounded for all $t \ge 0$ and $n = 1, 2, \ldots$.

The $\{T_t^{(n)}\}$ converge as $n \to \infty$ uniformly on each compact interval of t.

We next want to prove that the $\{T_t^{(n)}\}$ converge as $n \to \infty$ uniformly on each compact interval of t: The J_n commute with one another by their definition, and hence J_n commutes with $T_t^{(m)}$. By the semi-group property we have

$$\frac{d}{dt}T_t^m x = AJ_m T_t^{(m)}x = T_t^{(m)}AJ_m x$$

so

$$T_t^{(n)}x - T_t^{(m)}x = \int_0^t \frac{d}{ds}(T_{t-s}^{(m)}T_s^{(n)})x ds = \int_0^t T_{t-s}^{(m)}(AJ_n - AJ_m)T_s^{(n)}x ds.$$

Applying the semi-norm p and using the equiboundedness we see that

$$p(T_t^{(n)}x - T_t^{(m)}x) \le Ktq((J_n - J_m)Ax).$$

From (13.17) this implies that the $T_t^{(n)}x$ converge (uniformly in every compact interval of t) for $x \in D(A)$, and hence since $D(A)$ is dense and the $T_t^{(n)}$ are equicontinuous for all $x \in \mathbf{F}$. The limiting family of operators T_t are equicontinuous and form a semi-group because the $T_t^{(n)}$ have this property.

We show that the infinitesimal generator of this semi-group is A.

Let us temporarily denote the infinitesimal generator of this semi-group by B, so that we want to prove that $A = B$. Let $x \in D(A)$.

We know that
$$p(T_t^{(n)}x) \le Kq(x). \quad (13.18).$$
We claim that
$$\lim_{n\to\infty} T_t^{(n)} AJ_n x = T_t A x \tag{13.19}$$
uniformly in in any compact interval of t. Indeed, for any semi-norm p we have
$$\begin{aligned} p(T_tAx - T_t^{(n)}AJ_nx) &\le p(T_tAx - T_t^{(n)}Ax) + p(T_t^{(n)}Ax - T_t^{(n)}AJ_nx) \\ &\le p((T_t - T_t^{(n)})Ax) + Kq(Ax - J_nAx) \end{aligned}$$
where we have used (13.18) to get from the second line to the third. The second term on the right tends to zero as $n \to \infty$ and we have already proved that the first term converges to zero uniformly on every compact interval of t. This establishes (13.19).

$$\begin{aligned} T_tx - x &= \lim_{n\to\infty}(T_t^{(n)}x - x) \\ &= \lim_{n\to\infty}\int_0^t T_s^{(n)}AJ_nxds \\ &= \int_0^t (\lim_{n\to\infty} T_s^{(n)}AJ_nx)ds \\ &= \int_0^t T_sAxds \end{aligned}$$
where the passage of the limit under the integral sign is justified by the uniform convergence in t on compact sets. It follows from $T_tx - x = \int_0^t T_sAxds$ that x is in the domain of the infinitesimal operator B of T_t and that $Bx = Ax$. So B is an extension of A in the sense that $D(B) \supset D(A)$ and $Bx = Ax$ on $D(A)$.

But since B is the infinitesimal generator of an equibounded semi-group, we know that $(I - B)$ maps $D(B)$ onto $\mathbf{F}$ bijectively, and we are assuming that $(I - A)$ maps $D(A)$ onto $\mathbf{F}$ bijectively. Hence $D(A) = D(B)$.

This concludes the proof of the Hille-Yosida theorem.

13.3.7 The case of a Banach space.

In case $\mathbf{F}$ is a Banach space, so there is a single norm $p = \|\ \|$, the hypotheses of the theorem read: $D(A)$ is dense in $\mathbf{F}$, the resolvents $R(n, A)$ exist for all integers $n = 1, 2, \ldots$ and there is a constant K independent of n and m such that
$$\|(I - n^{-1}A)^{-m}\| \le K \quad \forall\ n = 1, 2, \ldots,\ m = 1, 2, \ldots. \tag{13.20}$$

Contraction semigroups.

In particular, if A satisfies
$$\|(I - n^{-1}A)^{-1}\| \le 1 \tag{13.21}$$

condition (13.20) is satisfied, and such an A then generates a semi-group. Under this stronger hypothesis we can draw a stronger conclusion: In (13.18) we now have $p = q = \|\cdot\|$ and $K = 1$. Since $\lim_{n\to\infty} T_t^n x = T_t x$ we see that under the hypothesis (13.21) we can conclude that

$$\|T_t\| \leq 1 \quad \forall\, t \geq 0.$$

A semi-group T_t satisfying this condition is called a **contraction semi-group**.

13.3.8 The other half of Stone's theorem.

We have already given a direct proof that if S is a self-adjoint operator on a Hilbert space then the resolvent exists for all non-real z and satisfies

$$\|R(z,S)\| \leq \frac{1}{|\mathrm{Im}\ (z)|}.$$

This implies (13.21) for $A = iS$ and $-iS$ giving us a proof of the existence of $U(t) = \exp(iSt)$ for any self-adjoint operator S, a proof which is independent of the spectral theorem.

13.4 The spectral theorem.

13.4.1 The functional calculus for functions in $\mathcal{S}$.

The Fourier inversion formula for functions f whose Fourier transform $\hat{f}$ belongs to L_1 (say for $f \in \mathcal{S}$, for example) says that

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(t) e^{itx} dt.$$

If we replace x by A and write $U(t)$ instead of e^{itA} this suggests that we define

$$f(A) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(t) U(t) dt. \tag{13.22}$$

Checking that $(fg)(A) = f(A)g(A)$.

To check that $(fg)(A) = f(A)g(A)$ we use the fact that $(fg)\hat{} = \hat{f} \star \hat{g}$ so

$$\begin{aligned}(fg)(A) &= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \hat{f}(t-s)\hat{g}(s) U(t) ds dt \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \hat{f}(r)\hat{g}(s) U(r+s) dr ds \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \hat{f}(r)\hat{g}(s) U(r) U(s) dr ds \\ &= f(A)g(A).\end{aligned}$$

Checking that the map $f \mapsto f(A)$ sends $\overline{f} \mapsto (f(A))^*$.

For the standard Fourier we know that the Fourier transform of $\overline{f}$ is given by

$$\hat{\overline{f}}(\xi) = \overline{\hat{f}(-\xi)}.$$

Substituting this into the right hand side of (13.22) gives

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \overline{\hat{f}(-t)} U(t) dt = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \overline{\hat{f}(-t)} U^*(-t) dt = \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(-t) U(-t) dt \right)^*$$
$$= (f(A))^*$$

by making the change of variables $s = -t$.

Checking that $\|f(A)\| \leq \|f\|_\infty$.

Let $\|f\|_\infty$ denote the sup norm of f, and let $c > \|f\|_\infty$. Define g by

$$g(s) := c - \sqrt{c^2 - |f(s)|^2}.$$

So g is a real element of $\mathcal{S}$ and

$$\begin{aligned} g^2 &= c^2 - 2c\, s\sqrt{c^2 - |f|^2} + c^2 - |f|^2 \\ &= 2cg - f\overline{f} \\ \text{so} & \\ f\overline{f} - 2cg + g^2 &= 0. \end{aligned}$$

So by our previous results,

$$f(A)^* f(A) - cg(A) - c(g(A)^* + g(A)^* g(A) = 0$$

i.e.

$$f(A)^* f(A) + (c - g(A))^* (c - g(A) = c^2.$$

So for any $v \in \mathfrak{H}$ we have

$$\|f(A)\|^2 \leq \|f(A)v\|^2 + \|(c - g(A))v\|^2 = c^2 \|v\|^2$$

proving that

$$\|f(A)\| \leq \|f\|_\infty. \tag{13.23}$$

Enlarging the functional calculus to continuous functions vanishing at infinity.

Equation (13.23) allows us to extend the functional calculus to all continuous functions vanishing at infinity. Indeed if $\hat{f}$ is an element of L_1 so that its inverse Fourier transform f is continuous and vanishes at infinity (by Riemann-Lebesgue) the formula (13.22) applies to f.

We will denote the space of continuous functions vanishing at infinity by $C_0(\mathbb{R})$.

Checking that (13.22) is non-trivial and unique.

We checked above that for z not real the function r_z given by

$$r_z(x) = \frac{1}{z-x}$$

has the property that

$$r_z(A) = R(z,A) = (zI - A)^{-1}$$

is given by an integral of the type (13.22). This involved some heavy lifting but not the spectral theorem. This shows that (13.22), is not trivial. Once we know that $r_z(A) = R(z,a)$ the Stone-Weierstrass theorem gives uniqueness.

Still missing one important item.

We still need to prove:

Proposition 13.4.1. *If* $\mathrm{Supp}(f) \cap \mathrm{spec}(A) = \emptyset$ *then* $f(A) = 0$.

We will derive this from the multiplication version of the spectral theorem.

13.4.2 The multiplication version of the spectral theorem.

In this section we follow the treatment in Davies.

The cyclic case.

A vector $v \in \mathfrak{H}$ is called **cyclic** for A if the linear combinations of all the vectors $R(z,A)v$ as z ranges over all non-real complex numbers is dense in $\mathfrak{H}$. Of course there might not be any cyclic vectors.

But suppose that v is a cyclic vector. Consider the continuous linear function ℓ on $C_0(\mathbb{R})$ given by

$$\ell(f) := (f(A)v, v).$$

If f is real valued and non-negative, then $\ell(f) = (f^{\frac{1}{2}}(A)v, f^{\frac{1}{2}}(A)v) \geq 0$. In other words, ℓ is a non-negative continuous linear functional. The Riesz representation theorem then says that there is a non-negative, finite, countably additive measure μ on $\mathbb{R}$ such that

$$\ell(f) = \int_{\mathbb{R}} f d\mu.$$

In fact, from its definition, the total measure $\mu(\mathbb{R}) \leq \|v\|^2$.

Let us consider $C_0(\mathbb{R})$ as a (dense) subset of $L_2(\mathbb{R},\mu)$, and let $(\cdot,\cdot)_2$ denote the scalar product on this L_2 space. Then for $f, g \in C_0(\mathbb{R})$ we have

$$(f,g)_2 = \ell(\overline{g}f) = (g(A)^* f(A)v, v) = (f(A)v, g(A)v),$$

(where the last two scalar products are in $\mathfrak{H}$). This shows that the map

$$f \mapsto f(A)v$$

is an isometry from $C_0(\mathbb{R})$ to the subspace of $\mathfrak{H}$ consisting of vectors of the form $f(A)v$. The space of vectors of the form $f(A)v$ is dense in $\mathfrak{H}$ by our assumption of cyclicity (since already the linear combinations of vectors of the form $r_z(A)$, $z \notin \mathbb{R}$ are dense). The space $C_0(\mathbb{R})$ is dense in $L_2(\mathbb{R})$. So the map above extends to a unitary map from $L_2(\mathbb{R}, \mu)$ to $\mathfrak{H}$ whose inverse we will denote by U.

So $U : \mathfrak{H} \to L_2(\mathbb{R}, \mu)$ is a unitary isomorphism such that

$$U(f(A)) = f, \quad \forall\, f \in C_0(\mathbb{R}).$$

Now let $f, g, h \in C_0(\mathbb{R})$ and set

$$\phi := g(A)v, \qquad \psi := h(A)v.$$

Then

$$(f(A)\phi, \psi) = \int_{\mathbb{R}} fg\overline{h}d\mu = (fU(\phi), U(\psi))_2$$

where, in this last term, the f denotes the operator of multiplication by f.

In other words,

$$Uf(A)U^{-1}$$

is the operator of multiplication by f on $L_2(\mathbb{R}, \mu)$. In particular, U of the image of the operator $f(A)$ is the image of multiplication by f in L_2.

Let us apply this last fact to the function $f = r_z,\ z \notin \mathbb{R}$, i.e.

$$r_z(x) = \frac{1}{z - x}.$$

We know that the resolvent $r_z(A)$ maps $\mathfrak{H}$ onto the domain $D(A)$, and that multiplication by r_z, which is the resolvent of the operator on L_2 maps L_2 to the domain of the operator of multiplication by x. This latter domain is the set of $k \in L_2$ such that $xk(x) \in L_2$. Now $(zI - A)r_z(A) = I$, so

$$Ar_z(A) = zr_z(A) - I.$$

Applied to $U^{-1}g, \ \ g \in L_2(\mathbb{R}, \mu)$ this gives .

$$Ar_z(A)U^{-1}g = zr_z(A)U^{-1}g - U^{-1}g.$$

So

$$AU^{-1}Ur_z(A)U^{-1}g = zU^{-1}Ur_z(A)U^{-1}g - U^{-1}g,$$

and multiplying by U gives

$$UAU^{-1}r_z \cdot g = zr_z \cdot g - g.$$

So if we set $h = r_z \cdot g$ so $zr_z \cdot g - g = xh$ we see that

$$UAU^{-1}h = x \cdot h. \tag{13.24}$$

If $y \notin \mathrm{Supp}(\mu)$ then multiplication by r_y is bounded on $L_2(\mathbb{R}, \mu)$ and conversely. So the support of μ is exactly the spectrum of A.

The general case.

Now for a general separable Hilbert space $\mathfrak{H}$ with a self-adjoint operator A we can decompose $\mathfrak{H}$ into a direct sum of Hilbert spaces each of which has a cyclic vector. Here is a sketch of how this goes. Start with a countable dense subset $\{x_1, x_2, \dots\}$ of $\mathfrak{H}$. Let $\mathfrak{L}_1$ be the cyclic subspace generated by x_1, i.e. $\mathfrak{L}_1$ is the smallest (closed) cyclic subspace containing x_1. Let $m(1)$ be the smallest integer such that $x_{m(1)} \notin \mathfrak{L}_1$. Let $y_{m(1)}$ be the component of $x_{m(1)}$ orthogonal to $\mathfrak{L}_1$, and let $\mathfrak{L}_2$ be the cyclic subspace generated by $y_{m(1)}$. Proceeding inductively, suppose that we have constructed the cyclic subspaces $\mathfrak{L}_i, \quad i = 1, \dots, n$ and let $m(n)$ be the smallest integer for which $x_{m(n)}$ does not belong to the (Hilbert space direct sum) $\mathfrak{L}_1 \oplus \mathfrak{L}_2 \oplus \cdots \oplus \mathfrak{L}_n$. Let $y_{m(n)}$ be the component of $x_{m(n)}$ orthogonal to this direct sum and let $\mathfrak{L}_{n+1}$ by the cyclic subpace generated by $y(m)$. At each stage of the induction there are two possibilities: If no $m(n)$ exists, the $\mathfrak{H}$ is the finite direct sum $\mathfrak{L}_1 \oplus \mathfrak{L}_2 \oplus \cdots \oplus \mathfrak{L}_n$. If the induction continues indefinitely, then the closure of the infinite Hilbert space direct sum $\mathfrak{L}_1 \oplus \mathfrak{L}_2 \oplus \cdots \oplus \mathfrak{L}_n \oplus \cdots$ contains all the x_i and so coincides with $\mathfrak{H}$.

By construction, each of the spaces $\mathfrak{L}_i$ is invariant under all the $R(z, A)$ so we can apply the results of the cyclic case to each of the $\mathfrak{L}_i$. Let us choose the cyclic vector $v_i \in \mathfrak{L}_i$ to have norm 2^{-n} so that the total measure of $\mathbb{R}$ under the corresponding measure μ_i is 2^{-2n}. Recall that S denotes the spectrum of A and each of the measures μ_i is supported on S. So we put a measure μ on $S \times \mathbb{N}$ so that the restriction of μ to $S \times \{n\}$ is μ_n. Then combine the U_n given above in the obvious way.

We obtain the following theorem:

Theorem 13.4.1. *Let A be a self-adjoint operator on a separable Hilbert space $\mathfrak{H}$ and let $S = \operatorname{spec}(A)$. There exists a finite measure μ on $S \times \mathbb{N}$ and a unitary isomorphism*

$$U : \mathfrak{H} \to L_2(S \times \mathbb{N}, \mu)$$

such that UAU^{-1} is multiplication by the function $a(s, n) = s$. More precisely, U takes the domain of A to the set of functions $h \in L_2$ such that $ah \in L_2$ and for for all such functions h we have

$$UAU^{-1}h = ah.$$

For any $f \in C_0(\mathbb{R})$ we have

$$Uf(A)U^{-1} = \text{multiplication by } f.$$

In particular, if supp$(f) \cap S = \emptyset$ then $f(A) = 0$.

The general version of the Dynkin-Helffer-Sjöstrand formula is true.

As a corollary of the preceding theorem, we conclude, as mentioned above, that the Dynkin-Helffer-Sjöstrand formula (10.2) is true in general.

Enlarging the functional calculus to bounded Borel functions.

We can now use the preceding theorem to define $f(A)$ where f is an arbitrary bounded Borel function, in such a way that it extends the preceding functional calculus. Here is how it goes: Let $\mathcal{B}$ denote the space of bounded Borel functions on $\mathbb{R}$. We say that $f_n \in \mathcal{B}$ increases monotonically to $f \in \mathcal{B}$ if $f_n(x)$ increases monotonically to $f(x)$ for every $x \in \mathbb{R}$. In particular the

$$\|f_n\| = \|f_n\|_0 = \sup_{x \in \mathbb{R}} |f_n(x)|$$

are uniformly bounded.

Theorem 13.4.2. *There exists a map from $\mathcal{B}$ to to bounded operators on $\mathfrak{H}$, $f \mapsto f(A)$ extending the map defined in Section 13.4.1 on $\mathcal{S}$ having all of the same properties (including the property that if $\mathrm{Supp}(f) \cap S = \emptyset$ then $f(A) = 0$.) This map is unique subject to the additional condition that whenever $f_n \in \mathcal{B}$ converges monotonically to $f \in \mathcal{B}$ then*

$$f_n(A) \to f(A)$$

in the sense of strong limits.

Proof. We may identify $\mathfrak{H}$ with $L_2(S \times \mathbb{N}, \mu)$ and A with the multiplication operator by a where $a(s,n) = s$ by the preceding theorem. Then for any $f \in \mathcal{B}$ define $f(A)$ to be multiplication by $f \circ a$. This has all the desired properties. The monotone convergence property is a consequence of the monotone convergence theorem in measure theory. This establishes the existence of the extension of the map $f \mapsto f(A)$ to $\mathcal{B}$.

For the uniqueness we use a monotone class argument. We have the uniqueness of the extension to $C_0(\mathbb{R})$. So let $\mathcal{C}$ denote the class on which two putative extension agree. Then $\mathcal{C}$ is mbonotone class containing $C_0(\mathbb{R})$. But the smallest such class is $\mathcal{B}$. $\square$

Corollary 13.4.1. *The spectrum of A equals the essential range of a defined as the set of all $\lambda \in \mathbb{R}$ such that*

$$\mu\left(x | a(\{x\} \times \mathbb{N}) - \lambda|\right) < \epsilon > 0$$

for all $\epsilon > 0$. If $\lambda \notin \mathrm{spec}(A)$ then

$$\|(\lambda I - A)^{-1})\| = \|R(\lambda, A)\| = |\mathrm{dist}\ (\lambda, S)|^{-1}.$$

By the multiplicative form of the spectral theorem it is enough to prove this when $A = a$ is a multiplication operator, and we will leave the details in this case to the reader, or refer to Davies [Davies] page 17.

The projection valued measure form of the spectral theorem.

Let us return to Theorem 13.4.2. If B is any Borel subset of $\mathbb{R}$ and $\mathbf{1}_B$ denotes the indicator function of B (i.e. the function which equals 1 on B and zero elsewhere) then $\mathbf{1}_B(A)$ is a self-adjoint projection operator which we will sometimes denote by P_B (where the operator A is understood). We have:

Theorem 13.4.3. *If B is any Borel subset of $\mathbb{R}$ then $\mathbf{1}_B(A)$ is a projection which commutes with A. If B_1 and $B_2 \in \mathcal{B}$ and $B_1 \cap B_2 = \emptyset$ then $\mathbf{1}_{B_1 \cup B_2}(A) = \mathbf{1}_{B_1}(A) + \mathbf{1}_{B_2}(A)$. If (a,b) is an open interval and f_n is an increasing sequence of continuous functions which converge to $\mathbf{1}_{(a,b)}$ then $f_n(A)$ converge strongly to the projection $P_{(a,b)} := \mathbf{1}_{(a,b)}$. We have $P_{(a,b)}A = 0 \iff (a,b) \cap S = \emptyset$.*

13.5 The Calderon-Vallaincourt theorem.

indexCalderon-Vallaincourt theorem In Chapters 9-11 we considered operators associated to symbols $a = a(x, \xi, \hbar)$, namely

$$(Op_t a(x, \hbar D, \hbar)u)(x) := \frac{1}{(2\pi\hbar)^n} \int \int e^{\frac{i}{\hbar}(x-y)\cdot\xi} a(tx + (1-t)y, \xi, \hbar) u(y) dy d\xi.$$

If (for each fixed $\hbar$) the function $a(\cdot, \cdot, \hbar)$ belongs to $\mathcal{S}(\mathbb{R}^{2n})$ then this operator is given by a kernel $K = K_\hbar \in \mathcal{S}(\mathbb{R}^{2n})$:

$$(Op_t a u)(x) = \int_{\mathbb{R}^n} K_\hbar(x,y) u(y) dy$$

where

$$K_\hbar(x,y) = \frac{1}{(2\pi\hbar)^n} \int_{\mathbb{R}^n} e^{\frac{i}{\hbar}(x-y)\cdot\xi} a(tx + (1-t)y, \xi, \hbar) d\xi.$$

As an operator, $K_\hbar$ maps $\mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$. At the other extreme, if $a \in \mathcal{S}'(\mathbb{R}^{2n})$ the above formula for $K = K_\hbar$ shows that $K \in \mathcal{S}'(\mathbb{R}^{2n})$. Hence $Op_t a$ is defined as an operator from $\mathcal{S}$ to $\mathcal{S}'$ given by

$$\langle Op_t(a)u, v \rangle = \langle K, u \otimes v \rangle.$$

The Schwartz kernel theorem guarantees that a continuous map from $\mathcal{S} \to \mathcal{S}'$ is in fact given by a kernel $K \in \mathcal{S}'(\mathbb{R}^{2n})$ and the above relation between K and a shows that every such map is of the form $Op_t(a)$ for a unique a.

The Calderon-Vallaincourt theorem imposes conditions on a to guarantee that $Op_t a$ gives a family of bounded operators on L_2. For simplicity we state for the case $t = \frac{1}{2}$, i.e Weyl quantization.

The conditions are: For each α and β there are constants $C_{\alpha,\beta}$ such that

$$\|\partial_x^\alpha \partial_\xi^\beta a\|_\infty \le C_{\alpha,\beta}.$$

Here $\|\dot{\|}_\infty$ denotes the sup norm on $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+$. Then

Theorem 13.5.1. [Calderon-Vallancourt.] *Under the above hypotheses, the operators $Op_\hbar^W a$ are continuous in the L_2 norm and so extend to bounded operators on L_2. Furthermore, there exists a positive constant M_n depending only on n, and a positive constant C_n depending only on n and $\hbar_0$ such that*

$$\|Op_\hbar^W a\|_{L_2} \leq C_n \left(\sum_{|\alpha|+|\beta|\leq M_n} \|\|\partial_x^\alpha \partial_\xi^\beta a\|_\infty \right).$$

The proof of this important theorem can be found in Evans-Zworski, Sjostrand-Dimassi, or in Martinez.

For a fixed $\hbar$ (so that a is now a function of just x and ξ) the change of variables $\xi \mapsto \hbar\xi$ converts $Op_\hbar^W a$ into $Op_1^W a(x, \hbar\xi)$ and

$$\partial_x^\alpha \partial_\xi^\beta a(x, \hbar\xi) = \hbar^{|\beta|}(\partial_x^\alpha \partial_\xi^\beta a)(x, \hbar\xi)$$

which is bounded by $\hbar^\beta \|\partial_x^\alpha \partial_\xi^\beta a\|_\infty$. So as long as $\hbar$ lies in a bounded interval, it is enough to prove the theorem for $\hbar = 1$. In other words, if $a = a(x, \xi)$ is bounded with all its derivatives on $\mathbb{R}^n \times \mathbb{R}^n$ and we define the operator A on $C_0^\infty(\mathbb{R}^n)$ by

$$Au(x) = \int\int a\left(\frac{x+y}{2}, \xi, \hbar\right) u(y) e^{i(x-y)\cdot\xi} dy d\xi$$

then

Theorem 13.5.2. Calderon-Vaillancourt *A is bounded as an operator on L^2 with bound*

$$C_n \left(\sum_{|\alpha|\leq M_n} \|\partial^\alpha a\|_\infty \right)$$

where C_n and M_n depend only on n.

The proof consists of a partition of unity argument followed by an application of a lemma in Hilbert space theory known as the Cotlar-Stein lemma. We refer to Martinez pp. 43-49 for an exceptionally clear presentation of this proof.

13.5.1 Existence of inverses.

In this section we present an important application of the Calderon-Vallaincourt theorem. We follow the exposition in Martinez. We begin by imposing some growth conditions on symbols.

A function $g : \mathbb{R}^{2n} \to \mathbb{R}^+$ is called an **order function** if

$$\partial_z^\alpha g = O(g)$$

for any $\alpha \in \mathbb{N}^{2n}$ and uniformly on $\mathbb{R}^{2n}$. For us, the key examples are

$$g(x, \xi) = \langle\xi\rangle^m = (1 + \|\xi\|^2)^{m/2}$$

and
$$g(x,\xi) = (1 + \|x\|^2 + \|\xi\|^2)^{m/2}$$
for various values of m.

Notice that it follows from Leibnitz's rule that if g is an order function then so is $1/g$.

A function $a = a(x,\xi,\hbar)$ defined on $\mathbb{R}^{2n} \times (0,\hbar_0]$ for some $\hbar_0 > 0$ is said to belong to $S(g)$ if it depends smoothly on (x,ξ) and for any $\alpha \in \mathbb{N}^{2d}$
$$\partial^\alpha a(x,\xi,\hbar) = O(g)$$
uniformly with respect to $(x,\xi,\hbar) \in \mathbb{R}^{2n} \times (0,\hbar_0]$.

For example, if $g = 1$ then $S(1)$ consists of C^∞ functions on $\mathbb{R}^{2n}$ parametrized by $\hbar \in (0,\hbar_0]$ which are uniformly bounded together with all their derivatives.

If $g = \langle\xi\rangle^m$ then the condition for a to belong to $S(g)$ is different from the condition on symbols that we imposed in Chapter 9 in that we are now demanding uniform bounds on all of $\mathbb{R}^{2n}$ whereas in Chapter 9 we allowed the bounds to depend on compact subsets of $\mathbb{R}^n$. On the other hand, in Chapter 9 we imposed the condition that locally $\partial_x^\beta \partial_\xi^\alpha a = O(\langle\xi\rangle^{m-|\alpha|}$ where here we are demanding that $\partial_x^\beta \partial_\xi^\alpha a = O(\langle\xi\rangle^m)$.

Notice that if g_1 and g_2 are order functions then so is $g_1 g_2$, and if $a \in S(g_1)$ and $b \in S(g_2)$ then $ab \in S(g_1 g_2)$.

Here is an unfortunate definition which seems to be standard in the subject: A symbol $a \in S(g)$ is called **elliptic** if there is a positive constant C_0 such that
$$|a| \geq \frac{1}{C_0} g$$
uniformly on $\mathbb{R}^{2n} \times (0,\hbar_0]$ for some $\hbar_0 > 0$.

For example, if
$$a(x,\xi,\hbar) = a_0(x,\xi) + \hbar a_1(x,\xi) + \cdots + \hbar^{N-1} a_{N-1}(x,\xi) + \hbar^N c(x,\xi,\hbar)$$
with $c \in S(g)$, and if there is a constant C_1 such that
$$|a_0| \geq \frac{1}{C_1} g$$
then a is elliptic.

From Leibniz's rule it follows that if $a \in S(g)$ is elliptic, then $1/a \in S(1/g)$. But more is true: using the symbolic calculus of Chapter 9:

Proposition 13.5.1. *Let $a \in S(g)$ be elliptic. Then there exists $b \in S(1/g)$ such that*
$$\begin{aligned} Op_\hbar(a) \circ Op_\hbar(b) &= 1 + Op_\hbar(r) \\ Op_\hbar(b) \circ Op_\hbar(a) &= 1 + Op_\hbar(r') \end{aligned}$$
with $r, r' \in O(\hbar^\infty)$ in $S(1)$.

Proof. Let $b_0 := 1/a$. We know that $b_0 \in S(1/g)$. Looking for $b \sim \sum \hbar^j b_j$ we solve for $b_j \in S(1/g)$ recursively so that

$$a\sharp b = 1 + O(\hbar^\infty) \quad \text{in} \quad S(1).$$

Similarly, find b' such that

$$b'\sharp a = 1 + O(\hbar^\infty) \quad \text{in} \quad S(1).$$

So

$$\begin{aligned} Op_\hbar(a) \circ Op_\hbar(b) &= 1 + Op_\hbar(r) \\ Op_\hbar(b') \circ Op_\hbar(a) &= 1 + Op_\hbar(r') \end{aligned}$$

with $r, r' = O(\hbar^\infty)$ in $S_{3n}(1)$. So

$$(1 + Op_\hbar(r')) \circ Op_\hbar(b) = Op_\hbar(b')(1 + Op_\hbar(r)).$$

Multiplying out gives

$$Op_\hbar(b) = Op_\hbar(b') + Op_\hbar(b') \circ Op_\hbar(r) - Op_\hbar(r') \circ Op_\hbar(b).$$

The last two terms together are of the form $Op_\hbar(r_1)$ with $r_1 = O(\hbar^\infty)$ in $S(1)$.
So

$$\begin{aligned} Op_\hbar(b) \circ Op_\hbar(a) &= Op_\hbar(b') \circ Op_\hbar(a) + Op_\hbar(r_1) \circ Op_\hbar(a) \\ &= 1 + Op_\hbar(r_2) \end{aligned}$$

with $r_2 = O(\hbar^\infty)$ in $S(1)$. So b does the trick. □

Let A be be the (family of) operator(s) $Op_\hbar(a)$. (say defined on $C_0^\infty(\mathbb{R}^n) \subset L_2(\mathbb{R}^n)$) Suppose that $g \geq 1$ so that $1/g \leq 1$ and hence $B = Op_h(b)$ is a family of bounded operators on $L_2 = L_2(\mathbb{R}^n)$ for suffficiently small $\hbar$ by the Calderon-Vallaincourt theorem, and let $R_1 := Op_\hbar(r)$ and $R_2 := Op_\hbar(r')$. Again by the Calderon-Vallaincourt theorem, R_1 and R_2 define bounded operators on L_2 and their norms as L_2 operators satisfy

$$\|R_1\| + \|R_2\| = O(\hbar^\infty).$$

In particular, the Neumann series for $(1+R_2)^{-1}$ converges for $\hbar$ small enough and hence $(1 + R_2)^{-1}B$ is a left inverse for A. (In case A were a bounded operator, so defined on all of L_2 we could similarly construct a right inverse and then the two inverses would coincide.) We wish to know that the inverse we constructed belongs to $S(g^{-1})$.

For this we apply Beal's characterization of operators $C = C_\hbar :\to \mathcal{S} \to \mathcal{S}'$ which are of the form $C = Op_\hbar(c)$ for $c \in S(1)$. Here is a statement of Beal's theorem: First some notation: If $\ell = \ell(x, \xi)$ is a linear function of (x, ξ) we denote the corresponding operators $Op_\hbar(\ell)$ by $\ell(\hbar D)$.

Theorem 13.5.3. [Semi-classical form of Beal's theorem.] *Let $C = C_\hbar : \mathcal{S} \to \mathcal{S}'$ be a continuous (family of) linear operators and so is of the form $Op_h c$ for $c \in \mathcal{S}'(\mathbb{R}^{2n}$. The following conditions are equivalent:*

1. *$c \in S(1)$.*
2. *For every $N \in \mathbb{N}$ and every collection $\ell_1, \dots, \ell_N$ of linear functions on $\mathbb{R}^{2n}$ the operators $\mathrm{ad}(\ell_1(x, \hbar D)) \circ \cdots \circ \mathrm{ad}(\ell_N(x, \hbar D))C$ are bounded in the L_2 operator norm and their operator norm is $O(\hbar^N)$.*

For a proof of Beal's theorem, see Dimassi-Sjöstrand pp. 98-99 or Evans-Zworski (Theorem 8.13).

Let us go back to our construction of the inverse the operator $Op(p)$ corresponding to an elliptic $p \in S(q)$. If we define $\tilde{q} = 1/p$ then our functional calculus tells us that $p\sharp\tilde{q} = 1 - \hbar r$ with $r \in S(1)$. So $1 - \hbar r$ satisfies condition 2 in Beal's theorem. But

$$\mathrm{ad}(\ell)\left((1-\hbar r)^{-1}\right) = -\left((1-\hbar r)^{-1}\right)\left(\mathrm{ad}(\ell)(1-\hbar r)\right)\left((1-\hbar r)^{-1}\right).$$

Repeated application of this identity shows that $(1-\hbar r)^{-1} \in S(1)$ so $q := \tilde{q}\sharp l\left((1-\hbar r)^{-1}\right) \in S(1)$ and the corresponding operator is the inverse of $Op(p)$.

13.6 The functional calculus for Weyl operators.

Let $g \geq 1$ be an order function, and let $p \in S(g)$ be real valued. Let $p^w(x, \hbar D, \hbar)$ be the corresponding Weyl operators, so initially all we know is that $p^w(x, \hbar D, \hbar)$ maps $\mathcal{S} \to \mathcal{S}'$. The main result of this section is that if $g = O((1+\|x\|^2+\|\xi\|^2)^m)$ for some m, then p^w defines an essentially self-adjoint operator on L_2.

We begin by sketching the fact that if $g = O((1+\|x\|^2+\|\xi\|^2)^m)$ for some m, and $a \in S(g)$ then $Op(a) : \mathcal{S} \to \mathcal{S}$. The idea is to use integration by parts to rewrite the operator $Op(a)$ for $a \in \mathcal{S}(\mathbb{R}^{2n}$ using integration by parts, and then to approximate $a \in S(g)$ by elements of $S(g)$. We use the operators

$$L_y := \frac{1 - \hbar\xi \cdot D_y}{1 + \|\xi\|^2} \quad \text{and} \quad L_\xi := \frac{1 + \hbar(x-y) \cdot D_\xi}{1 + \|x-y\|^2}.$$

Both operators satisfy

$$Le^{i\frac{(x-y)\cdot\xi}{\hbar}} = e^{i\frac{(x-y)\cdot\xi}{\hbar}}.$$

Integration by parts p times with respect to y using $L_y$gives

$$\frac{1}{(2\pi\hbar)^n}\int e^{i\frac{(x-y)\cdot\xi}{\hbar}} a(x,y,\xi,\hbar)u(y)d\xi dy$$

$$= \frac{1}{(2\pi\hbar)^n}\int e^{i\frac{(x-y)\cdot\xi}{\hbar}} ({}^tL)^p(au)d\xi dy$$

for $a \in \mathcal{S}(\mathbb{R}^{2n}$. But this last integral makes sense when $m - p < -n$ for $a \in S(g)$, and so, by continuity, we see that $Op(a)$ maps $\mathcal{S}$ into functions, in fact C^∞ functions on $\mathbb{R}^n$ and $Op(a)$ has the above form. We then integrate by parts with respect to ξ using L_ξ to conclude that $x^\alpha \partial_x^\beta (Op(a))u$ lies in $\mathcal{S}$. For details, see Martinez pages 24-25.

We now know that

$$p^w : \mathcal{S} \to \mathcal{S} \subset L_2.$$

We will let $P = p^w$ when thought of as an operator on L_2. We may (initially) consider P as a symmetric operator with domain $\mathcal{S} = \mathcal{S}(\mathbb{R}^n) \subset L_2$.

Also assume that $p \pm i$ is elliptic, so that we can construct their inverses as in the preceding section, and the symbols corresponding to them as above which we shall denote by $(p \pm i)^{-1} \in S(g^{-1})$ for small enough $\hbar$, and the corresponding bounded operators on L_2 which we denote by $(p^w \pm i)^{-1}$.

The following discussion is taken directly from Dimassi-Sjöstrand page 101:

It is easy to check that $(p^w \pm i)^{-1} L_2$ is independent of the choice of $\pm$. We denote it by $\mathcal{D}_P$.

Proposition 13.6.1. *The closure $\overline{P}$ of P has domain $\mathcal{D}_P$ and is self-adjiont.*

Proof. To say that u is in the domain of $\overline{P}$ means that there exists a sequence $u_j \to u$ with $v_j = Pu_j$ converging to some v (both in the L_2 norm). (In particular this converges as elements of $\mathcal{S}'$ and $p^w u = v$ as elements of $\mathcal{S}'$ and hence as elements of L_2). We have

$$(p^w + i)u_j = v_j + iu_j$$

and hence $u_j = (p^w + i)^{-1}(v_j + iu_j)$, and since $(p^w + i)^{-1}$ is a bounded operator on L_2 we conclude that $u = (p^w + i)^{-1}(v + iu) \in \mathcal{D}_P$.

Conversely, suppose that $u \in \mathcal{D}_P$, so that $u = (p^w + i)^{-1} w$ for some $w \in L_2$. Choose $f_j \in \mathcal{S}$ with $f_j \to w$ in L_2, and let $u_j = (p^w + i)^{-1} f_j$. So $u_j \in \mathcal{S}$ and $u_j \to u$. Also $(p^w + i)u_j = f_j$ so $p^w u_j \to v - iu$. This shows that $u \in \mathcal{D}_P$. So we have proved that $\overline{P}$ has domain $\mathcal{D}_P$ and coincides with p^w there.

Suppose that u is in the domain of P^* and $P^* u = v$. From the formal self-adjointness of p^w it follows that $p^w u = v$ as elements of $\mathcal{S}'$ and hence as elements of L_2 and hence that $(p^w + i)u = v + iu$ and therefore $u = (p^w + i)^{-1}(v + iu) \in \mathcal{D}_p$. So we have shown that the domain of P^* is $\mathcal{D}_P$ and $\overline{P} = P^*$. □

In fact we have proved that P has a unique self-adjoint extension (with domain $\mathcal{D}_P$) which we will now write simply as P instead of $\overline{P}$.

For example, consider the operators $\hbar^2 \Delta + V$ where $V \geq 0$ is a real function with $V \in S(\langle x \rangle^m)$ for some m and such that $1 + V$ is an order function. This operator corresponds to the symbol $\|\xi\|^2 + V(x)$ which belongs to $S(1 + \xi^2 + V)$. So the operator $1 + \hbar^2 \Delta + V$ (and hence the operator $\hbar^2 \Delta + V$) is essentially self-adjoint.

(Of course, for the case of the Schrödinger operator, much weaker conditions guarantee that it is essentially self adjoint; for example that the potential be ≥ 0 and locally L_2. See for example, Hislop-Segal page 86.)

We know from the preceding paragraph (via Beal's theorem) that the resolvent $R(z,P)$ is a Weyl operator for $\operatorname{Im} z \neq 0$ (Proposition 8.6 of Dimassi-Sjöstrand). Then using the Dynkin-Helffer-Sjöstrand formula we obtain

Theorem 13.6.1. [Theorem 8.7 of Dimassi-Sjöstrand.] *If $f \in C_0^\infty(\mathbb{R})$ then $f(P) \in Op_h(g^{-k})$ for any $k \in \mathbb{N}$. Furthermore, the two leading terms in the symbol of $f(P)$ are $a_0 = f(p_0)$ and $a_1 = p_1 f'(p_0)$.*

13.6.1 Trace class Weyl operators.

Suppose that $J \subset \mathbb{R}$ is an interval such that $p_0^{-1}(J) = \emptyset$. Then for any smaller interval $I \subset J$ (say with compact closure), the inverse of $rI - P$ exists for all $r \in I$ and sufficiently small $\hbar$. In other words, $\operatorname{spec}(P) \cap I = \emptyset$. So if $f \in C_0^\infty(\mathbb{R}$ has support in I then $f(P) = 0$.

Now suppose only that $p_0^{-1}(J)$ is contained in a compact subset $K \subset \mathbb{R}^{2n}$, and suppose that f has support in I. We will conclude that $f(P)$ is of trace class by the following beautiful argument due to Dimassi-Sjöstrand page 115:

Let $\tilde{p}$ be a real symbol which coincides with p outside some larger compact set and $\tilde{p}$ takes no values in J. So $a := \tilde{p} - p$ compact support and its corresponding operator A is of trace class with

$$\operatorname{tr} A = \frac{1}{(2\pi)^n} \int\int a(x,\xi,\hbar)\,dx\,d\xi$$

as can easily be checked.

Now apply the second resolvent identity to $P = p^w$ and $\tilde{P} = \tilde{p}^w$ which says that

$$R(z,P) = R(z,\tilde{P}) + R(z,P)(P - \tilde{P})R(z,\tilde{P}).$$

Plug this into the Dynkin-Helffer-Sjöstrand formula to obtain

$$f(P) = f(\tilde{P}) - \frac{1}{\pi} + \int \overline{\partial}\tilde{f} R(z,P)(P-\tilde{P})R(z,\tilde{P})\,dz.$$

The first term vanishes since the support of f lies in I and $\tilde{p}^{-1}(J) = \emptyset$. In the second term, the two resolvents blow up to order $|\operatorname{Im} z|^{-1}|$ while $\overline{\partial}\tilde{f}$ vanishes to infinite order in $|\operatorname{Im} z|$. Since $P - \tilde{P}$ is a trace class operator we conclude that $f(P)$ is a trace class operator!

13.7 Kantorovitz's non-commutative Taylor's formula.

13.7.1 A Dynkin-Helffer-Sjöstrand formula for derivatives.

Recall that if $f \in C_0^\infty(\mathbb{R})$ and if $\tilde{f}$ is an almost holomorphic extension of f then for any $w \in \mathbb{R}$ we have

$$f(w) = -\frac{1}{\pi}\int_{\mathbb{C}} \overline{\partial}\tilde{f} \cdot \frac{1}{z-w}\,dx\,dy.$$

The term $\overline{\partial}\tilde{f}$ vanishes to infinite order along the real axis. So we may differentiate under the integral sign as often as we like and conclude that

$$\frac{1}{j!}f^{(j)}(w) = -\frac{1}{\pi}\int_{\mathbb{C}}\overline{\partial}\tilde{f}\cdot\frac{1}{(z-w)^{j+1}}dxdy.$$

We may now apply the multiplication form of the spectral theorem as above to conclude

Proposition 13.7.1. *Let $f \in C_0^\infty(\mathbb{R})$ and $\tilde{f}$ an almost holomorphic extension of f. Then for any self-adjoint operator A we have*

$$\frac{1}{j!}f^{(j)}(A) = -\frac{1}{\pi}\int\overline{\partial}\tilde{f}R(z,A)^{j+1}dxdy. \tag{13.25}$$

Let B be another self-adjoint operator. We will use the above proposition to obtain a formula (due to Kantorovitz) which expresses $f(B)$ in terms of $f(A)$ as a sort of "Taylor expansion" about A.

13.7.2 The exponential formula.

Before proceeding to the general case, we illustrate it in a very important special case. Let $\mathfrak{A}$ be a Banach algebra (say the algebra of bounded operators on a Hilbert space), and let $a, b \in \mathfrak{A}$. The usual formula for the exponential series converges, so we have

$$e^{ta} = I + ta + \frac{1}{2}t^2a^2 + \cdots$$

with a similar formula for e^{tb}. We can regard the exponential formula as an asymptotic series if we like, i.e.

$$e^{ta} = I + ta + \cdots + \frac{1}{n!}t^na^n + O(t^{n+1}).$$

The special case of Kantorovitz's non-commutative Taylor formula that we study in this section expresses e^{tb} in terms of e^{ta} as follows: Define

$$X_0 := I,\ X_1 = b - a,\ X_2 := b^2 - 2ba + a^2,$$

and, in general,

$$X_n := b^n - nb^{n-1}a + \binom{n}{2}b^{n-2}a^2 + \cdots \pm a^n. \tag{13.26}$$

In other words, X_n looks like the binomial expansion of $(b-a)^n$ with all the b's moved to the left and all the a's to the right. The formula we want says that

$$e^{tb} = \left(I + tX_1 + \frac{1}{2}t^2X_2 + \cdots\right)e^{ta}. \tag{13.27}$$

Proof. If a and b commute, this is simply the assertion that $e^{tb} = e^{t(b-a)}e^t a$. But in trying to verify (13.27) all the a's lie to the right of all the b's, and we never move an a past a b, so (13.27) is true in general. □

An asymptotic consequence of (13.27) is

$$e^{tb} = \left(I + tX_1 + \frac{1}{2}t^2 X_2 + \cdots + \frac{1}{n!}t^n X_n\right) e^{ta} + O(t^{n+1}). \tag{13.28}$$

Polterovitch's idea.

Notice that we can obtain the X_n inductively as $X_0 = I$ and

$$X_{n+1} = (b-a)X_n + [a, X_n]. \tag{13.29}$$

Suppose that a and b are themselves asymptotic series in $\hbar$:

$$a \sim a_0 + a_1\hbar + a_2\hbar^2 + \cdots, \quad b = b_0 + b_1\hbar + b_2\hbar^2 + \cdots.$$

Suppose that $a - b = O(\hbar)$ and that bracket by a raises degree, i.e if $Y = O(\hbar^j)$ then $[a, Y] = O(\hbar^{j+1})$. Then it follows from the inductive definition (13.29) that

$$X_n = O(\hbar^n).$$

Polterovich and Hilkin-Polterovitch use this idea to greatly simplify an old formula of Agmon-Kannai about the asymptotics of the resolvents of elliptic operators. See our discussion in Chapter 11.

13.7.3 Kantorovitz's theorem.

We continue with the above notations, so $\mathfrak{A}$ is a Banach algebra and $a, b \in \mathfrak{A}$. We let $\sigma(a),\ \sigma(b)$ denote the spectra of a and b and $R(z,a),\ R(z,b)$ denote the resolvents of a and b.

Let L_a denote the operator of left multiplication by a and R_b denote the operator of right multiplication by b and

$$C(a,b) := L_a - R_b.$$

so

$$C(a,b)x = ax - xb.$$

Since right and left multiplications commute (by the associative law) we have the "binomial formula"

$$C(a,b)^n = L_a^n - nL_a^{n-1}R_b + \binom{n}{2} L_a^{n-2}R_b^2 + \cdots.$$

$\Omega \subset \mathbb{C}$ denotes an open set containing $\sigma(a) \cup \sigma(b)$ and Γ denotes a finite union of closed curves lying in Ω and containing $\sigma(a) \cup \sigma(b)$ in its interior.

Finally, f is a complex function defined and holomorphic on Ω.

Theorem 13.7.1. [Kantorovitch *For* $n = 0, 1, 2, \ldots$

$$f(b) = \sum_{j=0}^{n} (-1)^j f^{(j)}(a)[C(a,b)^j 1]/j! + L_n(f,a,b) \tag{13.30}$$

$$= \sum_{j=0}^{n} [C(b,a)^j 1] \cdot f^{(j)}(a)/j! + R_n(f,a,b) \tag{13.31}$$

where

$$L_n(f,a,b) := (-1)^{n+1} \frac{1}{2\pi i} \int_\Gamma f(z) R(z,a)^{n+1} [C(a,b)^{n+1} 1] \cdot R(z,b) dz \tag{13.32}$$

$$R_n(f,a,b) := \frac{1}{2\pi i} \int_\Gamma f(z) R(z,b) [C(b,a) 1| \cdot R(z,a)^{n+1} dz. \tag{13.33}$$

Example. In (13.31) take $f(x) = e^{tx}$ so that

$$f^{(j)}(a) = t^j e^{ta}$$

(and letting $n = \infty$ and ignoring the remainder) we get the formula of the preceding section for exponentials.

Proof. Let ϕ and ψ be invertible elements of a Banach algebra. Clearly

$$\psi = \phi + \phi(\phi^{-1} - \psi^{-1})\psi.$$

Suppose that z is in the resolvent set of a and b and take

$$\phi = R(z,a) = (zI - a)^{-1}, \qquad \psi = R(z,b) = (zI - b)^1$$

in the above formula. We get

$$R(z,b) = R(z,a) + R(z,a)(b-a)R(z,b).$$

This is our old friend, the second resolvent identity. Now let

$$Q := \phi(\psi^{-1} - \phi^{-1}) = \phi\psi^{-1} - I$$

so

$$(I+Q)^{-1} = \phi\psi^{-1} - I$$

so $\quad I + Q = \phi\psi^{-1} \quad$ is invertible and

$$(I+Q)^{-1} = \psi\phi^{-1}.$$

On the other hand, from high school algebra (the geometric sum) we know that for any integer $n \geq 0$ we have

$$(I+Q)^{-1} = \sum_{j=0}^{n} (-1)^j Q^j + (-1)^{n+1} Q^{n+1} (I+Q)^{-1},$$

as can be verified by multiplying on the right by $I + Q$. Multiplying this geometric sum on the right by ϕ gives

$$\psi = \sum_{j=0}^{n} \left[\phi(\psi^{-1} - \phi^{-1})\right]^{n} \phi + (-1)^{n+1} \left[\phi(\psi^{-1} - \phi^{-1})\right]^{n+1} \psi.$$

Substituting

$$\phi = R(z, a), \quad \psi = R(z, b)$$

gives Kantorovitz's extension of the second resolvent identity:

$$R(z, b) =$$

$$\sum_{j=0}^{n} (-1)^{j} \left[R(z, a)(a - b)\right]^{j} R(z, a) + (-1)^{n+1} \left[R(z, a)(a - b)\right]^{n+1} R(z, b). \quad (13.34)$$

In case a and b commute, the expression

$$\sum_{j=0}^{n} (-1)^{j} \left[R(z, a)(a - b)\right]^{j} R(z, a) + (-1)^{n+1} \left[R(z, a)(a - b)\right]^{n+1} R(z, b)$$

simplifies to

$$\sum_{j=0}^{n} R(z, a)^{j+1}(a - b)^{j} + (-1)^{n+1} R(z, a)^{n+1} R(z, b)(a - b)^{n+1}.$$

Now L_a and R_b always commute and $L_{c^{-1}} = (L_c)^{-1}$ for any invertible c and similarly $R_{c^{-1}} = (R_c)^{-1}$. So the above equation with a replaced by L_a and b replaced by R_b becomes

$$R_{R(z,b)} = \sum_{j=0}^{n} (L_{R(z,a)})^{j+1} C(a, b)^{j} + (-1)^{n+1} (L_{R(z,a)})^{j+1} R_{R(z,b)} C(a, b)^{n+1}.$$

If we apply this operator identity to the element $1 \in A$ we get

$$R(z, b) = \sum_{j=0}^{n} (-1)^{j} R(z, a)^{j+1} C(a, b)^{j} \cdot 1$$

$$+ (-1)^{n+1} R(z, a)^{n+1} \left[C(a, b)^{n+1} \cdot 1\right] R(z, b). \quad (13.35)$$

If we replace a by R_a and b by L_b in

$$\sum_{j=0}^{n} R(z, a)^{j+1}(a - b)^{j} + (-1)^{n+1} R(z, a)^{n+1} R(z, b)(a - b)^{n+1}$$

and apply to I we obtain

$$R(z, b) = \sum_{j=0}^{n} X_j R(z, a)^{j+1} + R(z, b) X_{n+1} R(z, a)^{n+1} \quad (13.36)$$

where, we recall, $X_j = [C(b,a)^j]I$.

The Riesz-Dunford functional calculus (which is basically an extension to Banach algebras of the Cauchy integral formula) says that for a function f analytic in Ω,

$$f(b) = \frac{1}{2\pi i}\int_\Gamma f(z)R(z,b)dz$$

and

$$\frac{1}{j!}f^{(j)}(a) = \frac{1}{2\pi i}\int_\Gamma f(z)R(z,a)^{j+1}dz.$$

Applied to (13.35) this gives (13.30) and (13.32).

A similar argument using (13.36) gives (13.31) and (13.33).

□

13.7.4 Using the extended Dynkin-Helffer-Sjöstrand formula.

For possibly unbounded operators we have to worry about domains. So the operators $C(a,b)^j I$ (where I is the identity operator) will be defined on the domain

$$\mathcal{D}_j := \mathcal{D}\left[(C(a,b)^j I\right] = \bigcap_{k=0}^{j}\mathcal{D}(a^k b^{j-k})$$

and (13.35) holds as an operator with domain $\mathcal{D}_{n+1}$.

If we multiply this equation by $\overline{\partial}\tilde{f}$ and integrate over $\mathbb{C}$ we obtain, as an analogue of Kantorovitz's first formula, for $f \in C_0^\infty(\mathbb{R})$:

$$\begin{aligned} f(b) &= \sum_{j=0}^{n}(-1)^j f^{(j)}(a)[C(a,b)^j 1]/j! + L_n(f,a,b) \\ &= \sum_{j=0}^{n}[C(b,a)^j 1]\cdot f^{(j)}(a)/j! + R_n(f,a,b) \end{aligned}$$

where

$$L_n(f,a,b) = \frac{(-1)^n}{\pi}\int_{\mathbb{C}}\overline{\partial}\tilde{f}(z)R(z,a)^{n+1}[C(a,b)^{n+1}1]\cdot R(z,b)dxdy. \tag{13.37}$$

A similar expression holds for the right remainder.

13.8 Appendix: The existence of almost holomorphic extensions.

We follow the discussion in Dimassi-Sjöstrand.

Let $f \in C_0^\infty(\mathbb{R})$, $\psi \in C_0^\infty(\mathbb{R})$, with $\psi \equiv 1$ on $\mathrm{Supp}(f)$, and $\chi \in C_0^\infty(\mathbb{R})$ with $\chi \equiv 1$ near 0. Define

$$\tilde{f}(x+iy) := \frac{\psi(x)}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{i(x+iy)\xi} \chi(y\xi) \hat{f}(\xi) d\xi,$$

where $\hat{f}$ is the Fourier transform of f. By the Fourier inversion formula

$$\tilde{f}_{|\mathbb{R}} = f. \tag{13.38}$$

With $\overline{\partial} := \frac{1}{2}(\partial_x + i\partial_y)$ we have

$$\overline{\partial}\tilde{f} = \frac{i}{2}\frac{\psi(x)}{\sqrt{2\pi}} \int e^{i(x+iy)\xi)}(-\xi\chi(y\xi) + \chi'(y\xi))\xi\hat{f}(\xi)d\xi$$

$$+\frac{1}{2}\frac{\psi(x)}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{i(x+iy)\xi} i\xi\chi(y\xi)\hat{f}(\xi)d\xi + \frac{1}{2}\frac{\psi'(x)}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{i(x+iy)\xi}\chi(y\xi)\hat{f}(\xi)d\xi$$

$$= \frac{i}{2}\frac{\psi(x)}{\sqrt{2\pi}} \int e^{i(x+iy)\xi)}\chi'(y\xi)\xi\hat{f}(\xi)d\xi + \frac{1}{2}\frac{\psi'(x)}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{i(x+iy)\xi}\chi(y\xi)\hat{f}(\xi)d\xi.$$

Define

$$\chi_N(t) := t^{-N}\chi'(t).$$

We can insert and extract a factor of y^N in the first integral above and write this first integral as

$$y^N \frac{i}{2}\frac{\psi(x)}{\sqrt{2\pi}} \int e^{i(x+iy)\xi)}\chi_N(y\xi)\xi^{N+1}\hat{f}(\xi)d\xi$$

and so get a bound on this first integral of the form

$$C_N|y|^N\|\xi^{N+1}\hat{f}(\xi)\|_{L^1}.$$

For the second integral we put in the expression of $\hat{f}$ as the Fourier transform of f to get

$$\frac{1}{2}\frac{\psi'(x)}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} e^{i(x-r+iy)\xi}\chi(y\xi)f(r)drd\xi.$$

Now $\psi' = 0$ on $\mathrm{Supp}(f)$ so $x - r \neq 0$ on $\mathrm{Supp}(\psi'(x)f(r))$ so this becomes

$$\frac{1}{4\pi}\psi'(x)\int\int D_\xi\left(e^{i(x-r+iy)\xi}\right)\frac{\chi(y\xi)}{x-r+iy}f(r)drd\xi.$$

Integration by parts turns this into

$$\frac{1}{4\pi}\psi'(x)\int\int e^{i(x-r+iy)\xi}\frac{\chi'(y\xi)y}{x-r+iy}f(r)drd\xi.$$

We can insert and extract a factor of y^N and also of $(\xi+i)^2$ so that the double integral becomes

$$y^N \int\int e^{i(x-r+iy)\xi}(\xi+i)^2 \frac{\chi_N(y\xi)y}{(x-r+iy)(\xi+i)^2} f(r)drd\xi$$

$$= y^N \int\int (i-D_r)^2(-D_r)^N \left(e^{i(x-r+iy)\xi}\right) \frac{\chi_N(y\xi)y}{(x-r+iy)} f(r) \frac{1}{(\xi+i)^2} drd\xi.$$

Integration by parts again brings the derivatives over to the the term $\frac{f(r)}{x-r+iy}$ and shows that the second integral is also $O(|y|^N)$. So we have proved that

$$|\overline{\partial}\tilde{f}(z)| \le C_N|\mathrm{Im}z|^N. \tag{13.39}$$

Thus for any $f \in C_0^\infty(\mathbb{R})$ we have produced an "almost holomorphic" extension $\tilde{f}$ satisfying (13.39) and (13.38).

Chapter 14

Differential calculus of forms, Weil's identity and the Moser trick.

The purpose of this chapter is to give a rapid review of the basics of the calculus of differential forms on manifolds. We will give two proofs of Weil's formula for the Lie derivative of a differential form: the first of an algebraic nature and then a more general geometric formulation with a "functorial" proof that we learned from Bott. We then apply this formula to the "Moser trick" and give several applications of this method.

14.1 Superalgebras.

A (commutative associative) **superalgebra** is a vector space

$$A = A_{even} \oplus A_{odd}$$

with a given direct sum decomposition into even and odd pieces, and a map

$$A \times A \to A$$

which is bilinear, satisfies the associative law for multiplication, and

$$\begin{aligned} A_{even} \times A_{even} &\to A_{even} \\ A_{even} \times A_{odd} &\to A_{odd} \\ A_{odd} \times A_{even} &\to A_{odd} \\ A_{odd} \times A_{odd} &\to A_{even} \\ \omega \cdot \sigma &= \sigma \cdot \omega \text{ if either } \omega \text{ or } \sigma \text{ are even,} \\ \omega \cdot \sigma &= -\sigma \cdot \omega \text{ if both } \omega \text{ and } \sigma \text{ are odd.} \end{aligned}$$

We write these last two conditions as

$$\omega \cdot \sigma = (-1)^{\deg \sigma \deg \omega} \sigma \cdot \omega.$$

Here deg $\tau = 0$ if τ is even, and deg $\tau = 1$ (mod 2) if τ is odd.

14.2 Differential forms.

A **linear differential form** on a manifold, M, is a rule which assigns to each $p \in M$ a linear function on TM_p. So a linear differential form, ω, assigns to each p an element of TM_p^*. We will, as usual, only consider linear differential forms which are smooth.

The superalgebra $\Omega(M)$ is the superalgebra generated by smooth functions on M (taken as even) and by the linear differential forms, taken as odd.

Multiplication of differential forms is usually denoted by $\wedge$. The number of differential factors is called the *degree* of the form. So functions have degree zero, linear differential forms have degree one.

In terms of local coordinates, the most general *linear* differential form has an expression as $a_1 dx_1 + \cdots + a_n dx_n$ (where the a_i are functions). Expressions of the form

$$a_{12} dx_1 \wedge dx_2 + a_{13} dx_1 \wedge dx_3 + \cdots + a_{n-1,n} dx_{n-1} \wedge dx_n$$

have degree two (and are even). Notice that the multiplication rules require

$$dx_i \wedge dx_j = -dx_j \wedge dx_i$$

and, in particular, $dx_i \wedge dx_i = 0$. So the most general sum of products of two linear differential forms is a differential form of degree two, and can be brought to the above form, locally, after collections of coefficients. Similarly, the most general differential form of degree $k \leq n$ on an n dimensional manifold is a sum, locally, with function coefficients, of expressions of the form

$$dx_{i_1} \wedge \cdots \wedge dx_{i_k}, \quad i_1 < \cdots < i_k.$$

There are $\begin{pmatrix} n \\ k \end{pmatrix}$ such expressions, and they are all even, if k is even, and odd if k is odd.

14.3 The d operator.

There is a linear operator d acting on differential forms called *exterior* differentiation, which is completely determined by the following rules: It satisfies Leibniz' rule in the "super" form

$$d(\omega \cdot \sigma) = (d\omega) \cdot \sigma + (-1)^{\deg \omega} \, \omega \cdot (d\sigma).$$

On functions it is given by

$$df = \frac{\partial f}{\partial x_1} dx_1 + \cdots + \frac{\partial f}{\partial x_n} dx_n$$

and, finally,

$$d(dx_i) = 0.$$

Since functions and the dx_i generate, this determines d completely. For example, on linear differential forms

$$\omega = a_1 dx_1 + \cdots a_n dx_n$$

we have

$$\begin{aligned} d\omega &= da_1 \wedge dx_1 + \cdots + da_n \wedge dx_n \\ &= \left(\frac{\partial a_1}{\partial x_1} dx_1 + \cdots \frac{\partial a_1}{\partial x_n} dx_n \right) \wedge dx_1 + \cdots \\ &\quad \left(\frac{\partial a_n}{\partial x_1} dx_1 + \cdots + \frac{\partial a_n}{\partial x_n} dx_n \right) \wedge dx_n \\ &= \left(\frac{\partial a_2}{\partial x_1} - \frac{\partial a_1}{\partial x_2} \right) dx_1 \wedge dx_2 + \cdots + \left(\frac{\partial a_n}{\partial x_{n-1}} - \frac{\partial a_{n-1}}{\partial x_n} \right) dx_{n-1} \wedge dx_n. \end{aligned}$$

In particular, equality of mixed derivatives shows that $d^2 f = 0$, and hence that $d^2\omega = 0$ for any differential form. Hence the rules to remember about d are:

$$\begin{aligned} d(\omega \cdot \sigma) &= (d\omega) \cdot \sigma + (-1)^{\deg \omega} \, \omega \cdot (d\sigma) \\ d^2 &= 0 \\ df &= \frac{\partial f}{\partial x_1} dx_1 + \cdots + \frac{\partial f}{\partial x_n} dx_n. \end{aligned}$$

14.4 Derivations.

A linear operator $\ell : A \to A$ is called an *odd derivation* if, like d, it satisfies

$$\ell : A_{even} \to A_{odd}, \quad \ell : A_{odd} \to A_{even}$$

and

$$\ell(\omega \cdot \sigma) = (\ell\omega) \cdot \sigma + (-1)^{\deg \omega} \, \omega \cdot \ell\sigma.$$

A linear map $\ell : A \to A$,

$$\ell : A_{even} \to A_{even}, \quad \ell : A_{odd} \to A_{odd}$$

satisfying

$$\ell(\omega \cdot \sigma) = (\ell\omega) \cdot \sigma + \omega \cdot (\ell\sigma)$$

is called an *even derivation*. So the Leibniz rule for derivations, even or odd, is

$$\ell(\omega \cdot \sigma) = (\ell\omega) \cdot \sigma + (-1)^{\deg \ell \deg \omega} \, \omega \cdot \ell\sigma.$$

Knowing the action of a derivation on a set of generators of a superalgebra determines it completely. For example, the equations

$$d(x_i) = dx_i, \quad d(dx_i) = 0 \ \ \forall i$$

implies that

$$dp = \frac{\partial p}{\partial x_1} dx_1 + \cdots + \frac{\partial p}{\partial x_n} dx_n$$

for any polynomial, and hence determines the value of d on any differential form with polynomial coefficients. The local formula we gave for df where f is any differentiable function, was just the natural extension (by continuity, if you like) of the above formula for polynomials.

The sum of two even derivations is an even derivation, and the sum of two odd derivations is an odd derivation.

The composition of two derivations will not, in general, be a derivation, but an instructive computation from the definitions shows that the *commutator*

$$[\ell_1, \ell_2] := \ell_1 \circ \ell_2 - (-1)^{\text{deg}\ell_1 \text{deg}\ell_2} \ \ell_2 \circ \ell_1$$

is again a derivation which is even if both are even or both are odd, and odd if one is even and the other odd.

A derivation followed by a multiplication is again a derivation: specifically, let ℓ be a derivation (even or odd) and let τ be an even or odd element of A. Consider the map

$$\omega \mapsto \tau \ell \omega.$$

We have

$$\begin{aligned} \tau\ell(\omega\sigma) &= (\tau\ell\omega) \cdot \sigma + (-1)^{\text{deg}\ell \text{deg}\omega} \tau\omega \cdot \ell\sigma \\ &= (\tau\ell\omega) \cdot \sigma + (-1)^{(\text{deg}\ell + \text{deg}\tau)\text{deg}\omega} \omega \cdot (\tau\ell\sigma) \end{aligned}$$

so $\omega \mapsto \tau\ell\omega$ is a derivation whose degree is

$$\text{deg}\tau + \text{deg}\ell.$$

14.5 Pullback.

Let $\phi : M \to N$ be a smooth map. Then the pullback map ϕ^* is a linear map that sends differential forms on N to differential forms on M and satisfies

$$\begin{aligned} \phi^*(\omega \wedge \sigma) &= \phi^*\omega \wedge \phi^*\sigma \\ \phi^* d\omega &= d\phi^*\omega \\ (\phi^* f) &= f \circ \phi. \end{aligned}$$

The first two equations imply that ϕ^* is completely determined by what it does on functions. The last equation says that on functions, ϕ^* is given by "substitution": In terms of local coordinates on M and on N ϕ is given by

$$\begin{aligned}\phi(x^1,\ldots,x^m) &= (y^1,\ldots,y^n)\\ y^i &= \phi^i(x^1,\ldots,x^m)\quad i=1,\ldots,n\end{aligned}$$

where the ϕ_i are smooth functions. The local expression for the pullback of a function $f(y^1,\ldots,y^n)$ is to substitute ϕ^i for the y^is as into the expression for f so as to obtain a function of the x's.

It is important to observe that the pull back on differential forms is defined for any smooth map, not merely for diffeomorphisms. This is the great advantage of the calculus of differential forms.

14.6 Chain rule.

Suppose that $\psi : N \to P$ is a smooth map so that the composition

$$\psi\circ\phi : M \to P$$

is again smooth. Then the *chain rule* says

$$(\psi\circ\phi)^* = \phi^*\circ\psi^*.$$

On functions this is essentially a tautology - it is the associativity of composition: $f\circ(\psi\circ\phi) = (f\circ\psi)\circ\phi$. But since pull-back is completely determined by what it does on functions, the chain rule applies to differential forms of any degree.

14.7 Lie derivative.

Let ϕ_t be a one parameter group of transformations of M. If ω is a differential form, we get a family of differential forms, $\phi_t^*\omega$ depending differentiably on t, and so we can take the derivative at $t = 0$:

$$\frac{d}{dt}\left(\phi_t^*\omega\right)_{|t=0} = \lim_{t=0}\frac{1}{t}\left[\phi_t^*\omega - \omega\right].$$

Since $\phi_t^*(\omega\wedge\sigma) = \phi_t^*\omega\wedge\phi_t^*\sigma$ it follows from the Leibniz argument that

$$\ell_\phi :\ \omega \mapsto \frac{d}{dt}\left(\phi_t^*\omega\right)_{|t=0}$$

is an even derivation. We want a formula for this derivation.

Notice that since $\phi_t^* d = d\phi_t^*$ for all t, it follows by differentiation that

$$\ell_\phi d = d\ell_\phi$$

and hence the formula for ℓ_ϕ is completely determined by how it acts on functions.

Let X be the vector field generating ϕ_t. Recall that the geometrical significance of this vector field is as follows: If we fix a point x, then

$$t \mapsto \phi_t(x)$$

is a curve which passes through the point x at $t = 0$. The tangent to this curve at $t = 0$ is the vector $X(x)$. In terms of local coordinates, X has coordinates $X = (X^1, \dots, X^n)$ where $X^i(x)$ is the derivative of $\phi^i(t, x^1, \dots, x^n)$ with respect to t at $t = 0$. The chain rule then gives, for any function f,

$$\begin{aligned} \ell_\phi f &= \frac{d}{dt} f(\phi^1(t, x^1, \dots, x^n), \dots, \phi_n(t, x^1, \dots, x^n))_{|t=0} \\ &= X^1 \frac{\partial f}{\partial x_1} + \cdots + X^n \frac{\partial f}{\partial x_n}. \end{aligned}$$

For this reason we use the notation

$$X = X^1 \frac{\partial}{\partial x_1} + \cdots + X^n \frac{\partial}{\partial x_n}$$

so that the differential operator

$$f \mapsto Xf$$

gives the action of ℓ_ϕ on functions.

As we mentioned, this action of ℓ_ϕ on functions determines it completely. In particular, ℓ_ϕ depends only on the vector field X, so we may write

$$\ell_\phi = D_X$$

where D_X is the even derivation determined by

$$D_X f = Xf, \quad D_X d = dD_X.$$

14.8 Weil's formula.

But we want a more explicit formula for D_X. For this it is useful to introduce an odd derivation associated to X called the *interior product* and denoted by $i(X)$. It is defined as follows: First consider the case where

$$X = \frac{\partial}{\partial x_j}$$

and define its interior product by

$$i\left(\frac{\partial}{\partial x_j}\right) f = 0$$

for all functions while

$$i\left(\frac{\partial}{\partial x_j}\right) dx_k = 0, \quad k \neq j$$

and

$$i\left(\frac{\partial}{\partial x_j}\right)dx_j = 1.$$

The fact that it is a derivation then gives an easy rule for calculating $i(\partial/\partial x_j)$ when applied to any differential form: Write the differential form as

$$\omega + dx_j \wedge \sigma$$

where the expressions for ω and σ do not involve dx_j. Then

$$i\left(\frac{\partial}{\partial x_j}\right)[\omega + dx_j \wedge \sigma] = \sigma.$$

The operator

$$X^j i\left(\frac{\partial}{\partial x_j}\right)$$

which means first apply $i(\partial/\partial x_j)$ and then multiply by the function X^j is again an odd derivation, and so we can make the definition

$$i(X) := X^1 i\left(\frac{\partial}{\partial x_1}\right) + \cdots + X^n i\left(\frac{\partial}{\partial x_n}\right). \tag{14.1}$$

It is easy to check that this does not depend on the local coordinate system used.

Notice that we can write

$$Xf = i(X)df.$$

In particular we have

$$\begin{aligned} D_X dx_j &= dD_X x_j \\ &= dX_j \\ &= di(X)dx_j. \end{aligned}$$

We can combine these two formulas as follows: Since $i(X)f = 0$ for any function f we have

$$D_X f = di(X)f + i(X)df.$$

Since $ddx_j = 0$ we have

$$D_X dx_j = di(X)dx_j + i(X)ddx_j.$$

Hence

$$D_X = di(X) + i(X)d = [d, i(X)] \tag{14.2}$$

when applied to functions or to the forms dx_j. But the right hand side of the preceding equation is an even derivation, being the commutator of two odd derivations. So if the left and right hand side agree on functions and on the differential forms dx_j they agree everywhere. This equation, (14.2), known as *Weil's formula*, is a basic formula in differential calculus.

We can use the interior product to consider differential forms of degree k as $k-$multilinear functions on the tangent space at each point. To illustrate, let σ be a differential form of degree two. Then for any vector field, X, $i(X)\sigma$ is a linear differential form, and hence can be evaluated on any vector field, Y to produce a function. So we define

$$\sigma(X,Y) := [i(X)\sigma](Y).$$

We can use this to express exterior derivative in terms of ordinary derivative and Lie bracket: If θ is a linear differential form, we have

$$\begin{aligned} d\theta(X,Y) &= [i(X)d\theta](Y) \\ i(X)d\theta &= D_X\theta - d(i(X)\theta) \\ d(i(X)\theta)(Y) &= Y[\theta(X)] \\ [D_X\theta](Y) &= D_X[\theta(Y)] - \theta(D_X(Y)) \\ &= X[\theta(Y)] - \theta([X,Y]) \end{aligned}$$

where we have introduced the notation $D_XY =: [X,Y]$ which is legitimate since on functions we have

$$(D_XY)f = D_X(Yf) - YD_Xf = X(Yf) - Y(Xf)$$

so D_XY as an operator on functions is exactly the commutator of X and Y. (See below for a more detailed geometrical interpretation of D_XY.) Putting the previous pieces together gives

$$d\theta(X,Y) = X\theta(Y) - Y\theta(X) - \theta([X,Y]), \tag{14.3}$$

with similar expressions for differential forms of higher degree.

14.9 Integration.

Let

$$\omega = f dx_1 \wedge \cdots \wedge dx_n$$

be a form of degree n on $\mathbf{R}^n$. (Recall that the most general differential form of degree n is an expression of this type.) Then its integral is defined by

$$\int_M \omega := \int_M f dx_1 \cdots dx_n$$

where M is any (measurable) subset. This, of course is subject to the condition that the right hand side converges if M is unbounded. There is a lot of hidden subtlety built into this definition having to do with the notion of orientation. But for the moment this is a good working definition.

The *change of variables formula* says that if $\phi : M \to \mathbf{R}^n$ is a smooth differentiable map which is one to one whose Jacobian determinant is everywhere positive, then

$$\int_M \phi^*\omega = \int_{\phi(M)} \omega.$$

14.10 Stokes theorem.

Let U be a region in $\mathbf{R}^n$ with a chosen orientation and smooth boundary. We then orient the boundary according to the rule that an outward pointing normal vector, together with the a positive frame on the boundary give a positive frame in $\mathbf{R}^n$. If σ is an $(n-1)-$form, then

$$\int_{\partial U} \sigma = \int_U d\sigma.$$

A manifold is called *orientable* if we can choose an atlas consisting of charts such that the Jacobian of the transition maps $\phi_\alpha \circ \phi_\beta^{-1}$ is always positive. Such a choice of an atlas is called an orientation. (Not all manifolds are orientable.) If we have chosen an orientation, then relative to the charts of our orientation, the transition laws for an $n-$form (where $n = \dim M$) and for a density are the same. In other words, given an orientation, we can identify densities with $n-$forms and $n-$form with densities. Thus we may integrate $n-$forms. The change of variables formula then holds for orientation preserving diffeomorphisms as does Stokes theorem.

14.11 Lie derivatives of vector fields.

Let Y be a vector field and ϕ_t a one parameter group of transformations whose "infinitesimal generator" is some other vector field X. We can consider the "pulled back" vector field ϕ_t^*Y defined by

$$\phi_t^*Y(x) = d\phi_{-t}\{Y(\phi_t x)\}.$$

In words, we evaluate the vector field Y at the point $\phi_t(x)$, obtaining a tangent vector at $\phi_t(x)$, and then apply the differential of the (inverse) map ϕ_{-t} to obtain a tangent vector at x.

If we differentiate the one parameter family of vector fields ϕ_t^*Y with respect to t and set $t = 0$ we get a vector field which we denote by D_XY:

$$D_XY := \frac{d}{dt}\phi_t^*Y_{|t=0}.$$

If ω is a linear differential form, then we may compute $i(Y)\omega$ which is a function whose value at any point is obtained by evaluating the linear function $\omega(x)$ on the tangent vector $Y(x)$. Thus

$$i(\phi_t^*Y)\phi_t^*\omega(x) = \langle (d(\phi_t)_x)^*\omega(\phi_t x), d\phi_{-t}Y(\phi_t x)\rangle = \{i(Y)\omega\}(\phi_t x).$$

In other words,

$$\phi_t^*\{i(Y)\omega\} = i(\phi_t^* Y)\phi_t^*\omega.$$

We have verified this when ω is a differential form of degree one. It is trivially true when ω is a differential form of degree zero, i.e. a function, since then both sides are zero. But then, by the derivation property, we conclude that it is true for forms of all degrees. We may rewrite the result in shorthand form as

$$\phi_t^* \circ i(Y) = i(\phi_t^* Y) \circ \phi_t^*.$$

Since $\phi_t^* d = d\phi_t^*$ we conclude from Weil's formula that

$$\phi_t^* \circ D_Y = D_{\phi_t^* Y} \circ \phi_t^*.$$

Until now the subscript t was superfluous, the formulas being true for any fixed diffeomorphism. Now we differentiate the preceding equations with respect to t and set $t = 0$. We obtain,using Leibniz's rule,

$$D_X \circ i(Y) = i(D_X Y) + i(Y) \circ D_X$$

and

$$D_X \circ D_Y = D_{D_X Y} + D_Y \circ D_X.$$

This last equation says that Lie derivative (on forms) with respect to the vector field $D_X Y$ is just the commutator of D_X with D_Y:

$$D_{D_X Y} = [D_X, D_Y].$$

For this reason we write

$$[X, Y] := D_X Y$$

and call it the Lie bracket (or commutator) of the two vector fields X and Y. The equation for interior product can then be written as

$$i([X, Y]) = [D_X, i(Y)].$$

The Lie bracket is antisymmetric in X and Y. We may multiply Y by a function g to obtain a new vector field gY. Form the definitions we have

$$\phi_t^*(gY) = (\phi_t^* g)\phi_t^* Y.$$

Differentiating at $t = 0$ and using Leibniz's rule we get

$$[X, gY] = (Xg)Y + g[X, Y] \tag{14.4}$$

where we use the alternative notation Xg for $D_X g$. The antisymmetry then implies that for any differentiable function f we have

$$[fX, Y] = -(Yf)X + f[X, Y]. \tag{14.5}$$

From both this equation and from Weil's formula (applied to differential forms of degree greater than zero) we see that the Lie derivative with respect to X at a point x depends on more than the value of the vector field X at x.

14.12 Jacobi's identity.

From the fact that $[X,Y]$ acts as the commutator of X and Y it follows that for any three vector fields X, Y and Z we have

$$[X,[Y,Z]] + [Z,[X,Y]] + [Y,[Z,X]] = 0.$$

This is known as **Jacobi's identity**. We can also derive it from the fact that $[Y,Z]$ is a natural operation and hence for any one parameter group ϕ_t of diffeomorphisms we have

$$\phi_t^*([Y,Z]) = [\phi_t^* Y, \phi_t^* Z].$$

If X is the infinitesimal generator of ϕ_t then differentiating the preceding equation with respect to t at $t = 0$ gives

$$[X,[Y,Z]] = [[X,Y],Z] + [Y,[X,Z]].$$

In other words, X acts as a derivation of the "mutliplication" given by Lie bracket. This is just Jacobi's identity when we use the antisymmetry of the bracket. In the future we we will have occasion to take cyclic sums such as those which arise on the left of Jacobi's identity. So if F is a function of three vector fields (or of three elements of any set) with values in some vector space (for example in the space of vector fields) we will define the cyclic sum $\mathcal{C}yc\ F$ by

$$\mathcal{C}yc\ F(X,Y,Z) := F(X,Y,Z) + F(Y,Z,X) + F(Z,X,Y).$$

With this definition Jacobi's identity becomes

$$\mathcal{C}yc\ [X,[Y,Z]] = 0. \tag{14.6}$$

14.13 A general version of Weil's formula.

Let W and Z be differentiable manifolds, let I denote an interval on the real line containing the origin, and let

$$\phi : W \times I \to Z$$

be a smooth map. We let $\phi_t : W \to Z$ be defined by

$$\phi_t(w) := \phi(w,t).$$

We think of ϕ_t as a one parameter family of maps from W to Z. We let ξ_t denote the tangent vector field along ϕ_t. In more detail:

$$\xi_t : W \to TZ$$

is defined by letting $\xi_t(w)$ be the tangent vector to the curve $s \mapsto \phi(w,s)$ at $s = t$.

If σ is a differential form on Z of degree $k+1$, we let the expression $\phi_t^* i(\xi_t)\sigma$ denote the differential form on W of degree k whose value at tangent vectors $\eta_1, \dots, \eta_k$ at $w \in W$ is given by

$$\phi_t^* i(\xi_t)\sigma(\eta_1, \dots, \eta_k) := (i(\xi_t)(w))\sigma)(d(\phi_t)_w \eta_1, \dots, d(\phi_t)_w \eta_k). \tag{14.7}$$

It is only the combined expression $\phi_t^* i(\xi_t)\sigma$ which will have any sense in general: since ξ_t is not a vector field on Z, the expression $i(\xi_t)\sigma$ will not make sense as a stand alone object (in general).

Let σ_t be a smooth one-parameter family of differential forms on Z. Then

$$\phi_t^* \sigma_t$$

is a smooth one parameter family of forms on W, which we can then differentiate with respect to t. The general form of Weil's formula is:

$$\frac{d}{dt}\phi_t^* \sigma_t = \phi_t^* \frac{d\sigma_t}{dt} + \phi_t^* i(\xi_t) d\sigma + d\phi_t^* i(\xi_t)\sigma. \tag{14.8}$$

Before proving the formula, let us note that it is functorial in the following sense: Suppose that that $F : X \to W$ and $G : Z \to Y$ are smooth maps, and that τ_t is a smooth family of differential forms on Y. Suppose that $\sigma_t = G^* \tau_t$ for all t. We can consider the maps

$$\psi_t : X \to Y, \quad \psi_t := G \circ \phi_t \circ F$$

and then the smooth one parameter familiy of differential forms

$$\psi_t^* \tau_t$$

on X. The tangent vector field ζ_t along ψ_t is given by

$$\zeta_t(x) = dG_{\phi_t(F(x))}\left(\xi_t(F(x))\right).$$

So

$$\psi_t^* i(\zeta_t)\tau_t = F^*\left(\phi_t^* i(\xi_t) G^* \tau_t\right).$$

Therefore, if we know that (14.8) is true for ϕ_t and σ_t, we can conclude that the analogous formula is true for ψ_t and τ_t.

Consider the special case of (14.8) where we take the one parameter family of maps

$$f_t : W \times I \to W \times I, \quad f_t(w, s) = (w, s+t).$$

Let

$$G : W \times I \to Z$$

be the map ϕ, and let

$$F : W \to W \times I$$

be the map

$$F(w) = (w, 0).$$

Then

$$(G \circ f_t \circ F)(w) = \phi_t(w).$$

Thus the functoriality of the formula (14.8) shows that we only have to prove it for the special case $\phi_t = f_t : W \times I \to W \times I$ as given above!

In this case, it is clear that the vector field ξ_t along ψ_t is just the constant vector field $\frac{\partial}{\partial s}$ evaluated at $(x, s+t)$. The most general differential (t-dependent) on $W \times I$ can be written as

$$ds \wedge a + b$$

where a and b are differential forms on W. (In terms of local coordinates $s, x^1, \ldots, x^n$ these forms a and b are sums of terms that have the expression

$$cdx^{i_1} \wedge \cdots \wedge dx^{i_k}$$

where c is a function of s, t and x.) To show the full dependence on the variables we will write

$$\sigma_t = ds \wedge a(x,s,t)dx + b(x,s,t)dx.$$

With this notation it is clear that

$$\phi_t^* \sigma_t = ds \wedge a(x, s+t, t)dx + b(x, s+t, t)dx$$

and therefore

$$\frac{d\phi_t^*\sigma_t}{dt} = ds \wedge \frac{\partial a}{\partial s}(x, s+t, t)dx + \frac{\partial b}{\partial s}(x, s+t, t)dx$$

$$+ds \wedge \frac{\partial a}{\partial t}(x, s+t, t)dx + \frac{\partial b}{\partial t}(x, s+t, t)dx.$$

So

$$\frac{d\phi_t^*\sigma_t}{dt} - \phi_t^*\frac{d\sigma_t}{dt} = ds \wedge \frac{\partial a}{\partial s}(x, s+t, t)dx + \frac{\partial b}{\partial s}(x, s+t, t)dx.$$

Now

$$i\left(\frac{\partial}{\partial s}\right)\sigma_t = adx$$

so

$$\phi_t^* i(\xi_t)\sigma_t = a(x, s+t, t)dx.$$

Therefore

$$d\phi_t^* i(\xi_t)\sigma_t = ds \wedge \frac{\partial a}{\partial s}(x, s+t, t)dx + d_W(a(x, s+t, t)dx).$$

Also

$$d\sigma_t = -ds \wedge d_W(adx) + \frac{\partial b}{\partial s} ds \wedge dx + d_W bdx$$

so

$$i\left(\frac{\partial}{\partial s}\right) d\sigma_t = -d_W(adx) + \frac{\partial b}{\partial s}dx$$

and therefore

$$\phi_t^* i(\xi_t) d\sigma_t = -d_W a(x, s+t, t)dx + \frac{\partial b}{\partial s}(x, s+t, t)dx.$$

So

$$\begin{aligned} d\phi_t^* i(\xi_t)\sigma_t + \phi_t^* i(\xi_t) d\sigma_t &= ds \wedge \frac{\partial a}{\partial s}(x, s+t, t)dx + \frac{\partial b}{\partial s}(x, s+t, t)dx \\ &= \frac{d\phi_t^*\sigma_t}{dt} - \phi_t^* \frac{d\sigma_t}{dt} \end{aligned}$$

proving (14.8).

A special case of (14.8) is the following. Suppose that $W = Z = M$ and ϕ_t is a family of diffeomorphisms $f_t : M \to M$. Then ξ_t is given by

$$\xi_t(p) = v_t(f_t(p))$$

where v_t is the vector field

$$v_t(f(p)) = \frac{d}{dt} f_t(p).$$

In this case $i(v_t)\sigma_t$ makes sense, and so we can write (14.8) as

$$\frac{d\phi_t^*\sigma_t}{dt} = \phi_t^* \frac{d\sigma_t}{dt} + \phi_t^* D_{v_t}\sigma_t. \tag{14.9}$$

14.14 The Moser trick.

Let M be a differentiable manifold and let ω_0 and ω_1 be smooth k-forms on M. Let us examine the following question: does there exist a diffeomorphism $f : M \to M$ such that $f^*\omega_1 = \omega_0$?

Moser answers this kind of question by making it harder! Let ω_t, $0 \leq t \leq 1$ be a family of k-forms with $\omega_t = \omega_0$ at $t = 0$ and $\omega_t = \omega_1$ at $t = 1$. We look for a one parameter family of diffeomorphisms

$$f_t : M \to M, \quad 0 \leq t \leq 1$$

such that

$$f_t^*\omega_t = \omega_0 \tag{14.10}$$

and

$$f_0 = \mathrm{id}\,.$$

Let us differentiate (14.10) with respect to t and apply (14.9). We obtain

$$f_t^*\dot{\omega}_t + f_t^* D_{v_t}\omega_t = 0$$

where we have written $\dot{\omega}_t$ for $\frac{d\omega_t}{dt}$. Since f_t is required to be a diffeomorphism, this becomes the requirement that

$$D_{v_t}\omega_t = -\dot{\omega}_t. \tag{14.11}$$

Moser's method is to use "geometry" to solve this equation for v_t if possible. Once we have found v_t, solve the equations

$$\frac{d}{dt}f_t(p) = v_t(f_t(p), \qquad f_0(p) = p \tag{14.12}$$

for f_t. Notice that for p fixed and $\gamma(t) = f_t(p)$ this is a system of ordinary differential equations

$$\frac{d}{dt}\gamma(t) = v_t(\gamma(t)), \quad \gamma(0) = p.$$

The standard existence theorems for ordinary differential equations guarantees the existence of of a solution depending smoothly on p at least for $|t| < \epsilon$. One then must make some additional hypotheses that guarantee existence for all time (or at least up to $t = 1$). Two such additional hypotheses might be

- M is compact, or
- C is a closed subset of M on which $v_t \equiv 0$. Then for $p \in C$ the solution for all time is $f_t(p) = p$. Hence for p close to C solutions will exist for a long time. Under this condition there will exist a neighborhood U of C and a family of diffeomorphisms

  $$f_t : U \to M$$

  defined for $0 \leq t \leq 1$ such

  $$f_0 = \text{id}, \qquad f_{t|C} = \text{id}\,\forall t$$

  and (14.10) is satisfied.

We now give some illustrations of the Moser trick.

14.14.1 Volume forms.

Let M be a compact oriented connected n-dimensional manifold. Let ω_0 and ω_1 be nowhere vanishing n-forms with the same volume:

$$\int_M \omega_0 = \int_M \omega_1.$$

Moser's theorem asserts that under these conditions there exists a diffeomorphism $f : M \to M$ such that

$$f^*\omega_1 = \omega_0.$$

Moser invented his method for the proof of this theorem.

The first step is to choose the ω_t. Let

$$\omega_t := (1-t)\omega_0 + t\omega_1.$$

Since both ω_0 and ω_1 are nowhere vanishing, and since they yield the same integral (and since M is connected), we know that at every point they are either both positive or both negative relative to the orientation. So ω_t is nowhere vanishing. Clearly $\omega_t = \omega_0$ at $t = 0$ and $\omega_t = \omega_1$ at $t = 1$. Since $d\omega_t = 0$ as ω_t is an n-from on an n-dimensional manifold,

$$D_{v_t}\omega_t = di(v_t)\omega_t$$

by Weil's formula. Also

$$\dot{\omega}_t = \omega_1 - \omega_0.$$

Since $\int_M \omega_0 = \int_M \omega_1$ we know that

$$\omega_0 - \omega_1 = d\nu$$

for some $(n-1)$-form ν. Thus (14.11) becomes

$$di(v_t)\omega_t = d\nu.$$

We will certainly have solved this equation if we solve the harder equation

$$i(v_t)\omega_t = \nu.$$

But this equation has a unique solution since ω_t is no-where vanishing. QED

14.14.2 Variants of the Darboux theorem.

We present these in Chapter 2.

14.14.3 The classical Morse lemma.

Let $M = \mathbb{R}^n$ and $\phi_i \in C^\infty(\mathbb{R}^n),\ i = 0, 1$. Suppose that 0 is a non-degenerate critical point for both ϕ_0 and ϕ_1, suppose that $\phi_0(0) = \phi_1(0) = 0$ and that they have the same Hessian at 0, i.e. suppose that

$$\left(d^2\phi_0\right)(0) = \left(d^2\phi_1\right)(0).$$

The Morse lemma asserts that there exist neighborhoods U_0 and U_1 of 0 in $\mathbb{R}^n$ and a diffeomorphism

$$f : U_0 \to U_1, \quad f(0) = 0$$

such that

$$f^*\phi_1 = \phi_0.$$

Proof. Set

$$\phi_t := (1-t)\phi_0 + t\phi_1.$$

The Moser trick tells us to look for a vector field v_t with

$$v_t(0) = 0, \quad \forall\ t$$

and

$$D_{v_t}\phi_t = -\dot{\phi}_t = \phi_0 - \phi_1.$$

The function ϕ_t has a non-degenerate critical point at zero with the same Hessian as ϕ_0 and ϕ_1 and vanishes at 0. Thus for each fixed t, the functions

$$\frac{\partial \phi_t}{\partial x^i}$$

form a system of coordinates about the origin.

If we expand v_t in terms of the standard coordinates

$$v_t = \sum_j v_j(x,t)\frac{\partial}{\partial x^j}$$

then the condition $v_j(0,t) = 0$ implies that we must be able to write

$$v_j(x,t) = \sum_i v_{ij}(x,t)\frac{\partial \phi_t}{\partial x^i}.$$

for some smooth functions v_{ij}. Thus

$$D_{v_t}\phi_t = \sum_{ij} v_{ij}(x,t)\frac{\partial \phi_t}{\partial x^i}\frac{\partial \phi_t}{\partial x^j}.$$

Similarly, since $-\dot{\phi}_t$ vanishes at the origin together with its first derivatives, we can write

$$-\dot{\phi}_t = \sum_{ij} h_{ij}\frac{\partial \phi_t}{\partial x^i}\frac{\partial \phi_t}{\partial x^j}$$

where the h_{ij} are smooth functions. So the Moser equation $D_{v_t}\phi_t = -\dot{\phi}_t$ is satisfied if we set

$$v_{ij}(x,t) = h_{ij}(x,t).$$

Notice that our method of proof shows that if the ϕ_i depend smoothly on some paramters lying in a compact manifold S then the diffeomorphism f can be chosen so as to depend smoothly on $s \in S$.

In Section 5.11 we give a more refined version of this argument to prove the Hörmander-Morse lemma for generating functions.

In differential topology books the classical Morse lemma is usually stated as follows:

Theorem 14.14.1. *Let M be a manifold and $\phi : M \to \mathbb{R}$ be a smooth function. Suppose that $p \in M$ is a non-degenerate critical point of ϕ and that the signature of $d^2\phi_p$ is $(k, n-k)$. Then there exists a system of coordinates $(U, x_1, \ldots, x_n)$ centered at p such that in this coordinate system*

$$\phi = c + \sum_{i=1}^{k} x_i^2 - \sum_{i=k+1}^{n} x_i^2.$$

Proof. Choose any coordinate system $(W, y_1, \ldots y_n)$ centered about p and apply the previous result to

$$\phi_1 = \phi - c$$

and

$$\phi_0 = \sum h_{ij} y_i y_j$$

where

$$h_{ij} = \frac{\partial^2 \phi}{\partial y_i \partial y_i}(0).$$

This gives a change of coordinates in terms of which $\phi - c$ has become a non-degenerate quadratic form. Now apply Sylvester's theorem in linear algebra which says that a linear change of variables can bring such a non-degenerate quadratic form to the desired diagonal form.

Chapter 15

The method of stationary phase

15.1 Gaussian integrals.

We recall a basic computation in the integral calculus:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2/2}dx = 1. \tag{15.1}$$

This is proved by taking the square of the left hand side and then passing to polar coordinates:

$$\begin{aligned}
\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2/2}dx\right]^2 &= \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}dxdy \\
&= \frac{1}{2\pi}\int_{0}^{2\pi}\int_{0}^{\infty} e^{-r^2/2}rdrd\theta \\
&= \int_{0}^{\infty} e^{-r^2/2}rdr \\
&= 1.
\end{aligned}$$

15.1.1 The Fourier transform of a Gaussian.

Now

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2/2}e^{-\eta x}dx$$

converges for all complex values of η, uniformly in any compact region. Hence it defines an analytic function which may be evaluated by taking η to be real and then using analytic continuation. For real η we complete the square and make a change of variables:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(-\frac{x^2}{2}-x\eta)dx =$$

$$\begin{aligned}
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\frac{1}{2}(-(x+\eta)^2+\eta^2)dx \\
&= \exp(\eta^2/2)\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(-(x^2+\eta^2)/2)dx \\
&= \exp(\eta^2/2).
\end{aligned}$$

As we mentioned, this equation is true for any complex value of η. In particular, setting $\eta = -i\xi$ we get

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(-x^2/2+i\xi x)dx = \exp(-\xi^2/2). \tag{15.2}$$

In short,

15.1.1. *The Fourier transform of the Gaussian function* $x \mapsto \exp(-x^2/2)$ *is* $\xi \mapsto e^{-\xi^2/2}$.

If f is any smooth function vanishing rapidly at infinity, and $\hat{f}$ denotes its Fourier transform, then the Fourier transform of $x \mapsto f(cx)$ is $\xi \mapsto \frac{1}{c}\hat{f}(\xi/c)$. In particular, if we take $\lambda > 0,\ c = \lambda^{\frac{1}{2}}$ we get

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(-\lambda x^2/2+i\xi x)dx = \left(\frac{1}{\lambda}\right)^{\frac{1}{2}}\exp(-\xi^2/2\lambda). \tag{15.3}$$

We proved this formula for λ real and positive. But the integral on the left makes sense for all λ with Re $\lambda > 0$, and hence this formula remains true in the entire open right hand plane Re $\lambda > 0$, provided we interpret the square root occurring on the right as arising by analytic continuation from the positive real axis.

We can say more: The integral on the left converges uniformly (but not absolutely) for λ in any region of the form

$$\text{Re } \lambda \geq 0, \quad |\lambda| > \delta > 0.$$

To see this, observe that for any $S > R > 0$ we have

$$e^{-\lambda x^2/2} = -\frac{1}{\lambda x}\frac{d}{dx}\exp(-\lambda x^2/2) \quad \text{for } R \leq x \leq S$$

so we can apply integration by parts to get

$$\int_R^S e^{-\lambda x^2/2}e^{i\xi x}dx =$$

$$\frac{1}{\lambda}\left(\frac{1}{R}e^{-\lambda R^2/2+i\xi R}-\frac{1}{S}e^{-S^2/2+i\xi S}+\int_R^S e^{-\lambda x^2/2}\frac{d}{dx}\left(\frac{e^{i\xi x}}{x}\right)dx\right)$$

and integrate by parts once more to bound the integral on the right. We conclude that

$$\left|\int_R^S e^{-\lambda x^2/2}e^{i\xi x}dx\right| = O(\frac{1}{|\lambda R|}).$$

15.2 The integral $\int e^{-\lambda x^2/2}h(x)dx$.

This same argument shows that

$$\int e^{-\lambda x^2/2}h(x)dx$$

is convergent for any h with two bounded continuous derivatives. Indeed,

$$\int_R^S e^{-\lambda x^2/2}h(x)dx =$$

$$\begin{aligned}
&= -\frac{1}{\lambda}\int_R^S \frac{h(x)}{x}\frac{d}{dx}e^{-\lambda x^2/2}dx \\
&= -\lambda^{-1}e^{-\lambda x^2/2}(h(x)/x)\Big|_R^S \\
&\quad +\frac{1}{\lambda}\int_R^S e^{-\lambda x^2/2}\frac{d}{dx}(h(x)/x)dx \\
&= -\lambda^{-2}e^{-\lambda x^2/2}[\lambda(h(x)/x)-(1/x)\frac{d}{dx}(h(x)/x)]\Big|_R^S \\
&\quad +\lambda^{-2}\int_R^S e^{-\lambda x^2/2}[(1/x)(h(x)/x)']'dx.
\end{aligned}$$

This last integral is absolutely convergent, and the boundary terms tend to zero as $R\to\infty$.

This argument shows that if M is a bound for h and its first two derivatives, the above expressions can all be estimated purely in terms of M. Thus if h depends on some auxiliary parameters, and is uniformly bounded together with its first two derivatives with respect to these parameters, then the integral $\int_{-\infty}^{\infty} h(x)\exp(-\lambda x^2/2)dx$ converges uniformly with respect to these parameters.

Let us push this argument one step further. Suppose that h has derivatives of all order which are bounded on the entire real axis, and suppose further that $h\equiv 0$ in some neighborhood, $|x|<\epsilon$, of the origin. If we do the integration by parts

$$\int_R^S e^{-\lambda x^2/2}h(x)dx$$

$$= -\lambda^{-1}e^{-\lambda x^2/2}(h(x)/x)\Big|_R^S + \frac{1}{\lambda}\int_R^S e^{-\lambda x^2/2}\frac{d}{dx}(\frac{h(x)}{x})dx,$$

choose $R < \epsilon$ and let $S \to \infty$. We conclude that

$$\int_{-\infty}^{\infty} e^{-\lambda x^2/2}h(x)dx = \frac{1}{\lambda}\int_{-\infty}^{\infty} e^{-\lambda x^2/2}\frac{d}{dx}(h(x)/x)dx.$$

The right hand side is a function of the same sort as h. We conclude that

$$\int_{\mathbb{R}} e^{-\lambda x^2/2}h(x)dx = O(\lambda^{-N})$$

for all N if h vanishes in some neighborhood of the origin has derivatives of all order which are each bounded on the entire line.

15.3 Gaussian integrals in n dimensions.

Getting back to the case $h \equiv 1$, if we take $\lambda = \mp ir, \quad r > 0$ and set $\xi = 0$ in (15.3) then analytic continuation from the positive real axis gives $\lambda^{\frac{1}{2}} = e^{\mp\pi i/4}$ and we obtain

$$\int_{-\infty}^{\infty} e^{\pm irx^2/2}dx = \left(\frac{2\pi}{r}\right)^{\frac{1}{2}} e^{\pm\pi i/4}. \tag{15.4}$$

Doing the same computation in n - dimensions gives

$$\int e^{i\tau Q/2}dy = \left(\frac{2\pi}{\tau}\right)^{\frac{n}{2}} \left(\frac{1}{r_1 \cdot r_2 \cdots r_n}\right)^{\frac{1}{2}} e^{i\,\mathrm{sgn}\,Q\pi/4} \tag{15.5}$$

if

$$Q(y) = \sum \pm r_i(y^i)^2.$$

Now $r_1 \cdot r_2 \cdots r_n = |\det Q|$. So we can rewrite the above equation as

$$\int e^{i\tau Q/2}dy = \left(\frac{2\pi}{\tau}\right)^{\frac{n}{2}} \frac{1}{\sqrt{|\det Q|}} e^{i\,\mathrm{sgn}\,Q\pi/4} \tag{15.6}$$

We proved this formula under the assumption that Q was in diagonal form. But if Q is any non-degenerate quadratic form, we know that there is an orthogonal change of coordinates which brings Q to diagonal form. By this change of variables we see that

15.3.1. *(15.6) is valid for any non-degenerate quadratic form.*

15.4 Using the multiplication formula for the Fourier transform.

Recall that in one dimension this says that if $f, g \in \mathcal{S}(\mathbb{R})$ and $\hat{f}, \hat{g}$ denote their Fourier transforms then

$$\int_{\mathbb{R}} \hat{f}(\xi) g(\xi) d\xi = \int_{\mathbb{R}} f(x) \hat{g}(x) dx.$$

In this formula let us take

$$g(\xi) = e^{-\frac{\xi^2}{2\lambda}}$$

where Re $\lambda > 0$ so that

$$\hat{g}(x) = \lambda^{\frac{1}{2}} e^{-\lambda x^2/2}$$

where the square root is given by the positive square root on the positive axis and extended by analytic continuation. So the multiplication formula yields

$$\int_{\mathbb{R}} \hat{f}(\xi) e^{-\frac{\xi^2}{2\lambda}} d\xi = \lambda^{\frac{1}{2}} \int_{\mathbb{R}} f(x) e^{-\frac{\lambda x^2}{2}} dx.$$

Take

$$\lambda = \epsilon - ia, \quad \epsilon > 0, \quad a \in \mathbb{R} - \{0\}$$

and let $\epsilon \searrow 0$. We get

$$\int_{\mathbb{R}} \hat{f}(\xi) e^{-\frac{i\xi^2}{2a}} = |a|^{\frac{1}{2}} e^{-\frac{\pi i}{4} \operatorname{sgn} a} \int_{\mathbb{R}} f(x) e^{\frac{iax^2}{2}} dx$$

which we can rewrite as

$$\int_{\mathbb{R}} f(x) e^{i\frac{ax^2}{2}} dx = |a|^{-\frac{1}{2}} e^{\frac{\pi i}{4}} \int_{\mathbb{R}} \hat{f}(\xi) e^{-\frac{i\xi^2}{2a}} d\xi.$$

We can pass from this one dimensional formula to an n - dimensional formula as follows: Let $A = (a_{k\ell})$ be a non-singular symmetric $n \times n$ matrix and let sgn A denote the signature of the quadatic form

$$Q(x) = \langle Ax, x \rangle = \sum a_{ij} x_i x_j.$$

Let

$$B := A^{-1}.$$

Then for any $t > 0$ we have

$$\int_{\mathbb{R}^n} f(x) e^{i\frac{t}{2}\langle Ax, x\rangle} dx = t^{-\frac{n}{2}} |\det A|^{-\frac{1}{2}} e^{\frac{\pi i}{4} \operatorname{sgn} A} \int_{\mathbb{R}^n} \hat{f}(\xi) e^{-\frac{i}{2t}\langle B\xi, \xi\rangle} d\xi. \qquad (15.7)$$

The proof is via diagonalization as before. We may make an orthogonal change of coordinates relative to which A becomes diagonal. Then if f is a product function

$$f(x_1, \ldots, x_n) = f(x_1) \cdot f(x_2) \cdots f(x_n)$$

the formula reduces to the one dimensional formula we have already proved. Since the linear combination of these functions are dense, the formula is true in general.

15.5 A local version of stationary phase.

In order to conform with standard notation let us set $t = \hbar^{-1}$ in (15.7). The right hand side of (15.7) becomes

$$\hbar^{\frac{n}{2}} |\det A|^{-\frac{1}{2}} e^{\frac{\pi i}{4}\operatorname{sgn} A} a(\hbar)$$

where

$$a(\hbar) = \int_{\mathbb{R}^n} \hat{f}(\xi) e^{-i\frac{\hbar}{2}\langle B\xi,\xi\rangle} d\xi.$$

Let us now use the Taylor formula for the exponential:

$$\left| e^{ix} - \sum_{k=0}^{m} \frac{(ix)^k}{k!} \right| \le \frac{|x|^{m+1}}{(m+1)!}.$$

Thus the function a can be estimated by the sum

$$\sum_{k=0}^{m} \frac{1}{k!} \left(-\frac{i\hbar}{2}\right)^k \int_{\mathbb{R}^n} \langle B\xi,\xi\rangle^k \hat{f}(\xi) d\xi$$

with an error that is bounded by

$$\frac{1}{(m+1)!} \left(\frac{\hbar}{2}\right)^{m+1} \int_{\mathbb{R}^n} \left| \langle B\xi,\xi\rangle^{m+1} \hat{f}(\xi) \right| d\xi.$$

In the "Taylor expansion"

$$a(\hbar) = \sum a_k \hbar^k$$

we can interpret the coefficient

$$a_k = \left(-\frac{i}{2}\right)^k \int_{\mathbb{R}^n} \langle B\xi,\xi\rangle^k \hat{f}(\xi) d\xi$$

as follows: Let $b(D)$ be the constant coefficient differential operator

$$b(D) := \sum b_{k\ell} D_k D_\ell$$

where

$$D_k = \frac{1}{i} \frac{\partial}{\partial x_k}.$$

Then $\langle B\xi,\xi\rangle^k \hat{f}(\xi)$ is the Fourier transform of the function $b(D)^k f$. So by the Fourier inversion formula,

$$(b(D)^k f)(0) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} \langle B\xi,\xi\rangle^k \hat{f}(\xi) d\xi.$$

We can thus state our local version of the stationary phase formula as follows:

Theorem 15.5.1. *If* $f \in \mathcal{S}(\mathbb{R}^n)$ *and*

$$I(\hbar) := \int_{\mathbb{R}^n} f(x) e^{i\frac{\langle Ax,x\rangle}{2\hbar}} dx$$

then

$$I(\hbar) = \left(\frac{\hbar}{2\pi}\right)^{\frac{n}{2}} \gamma_A a(\hbar)$$

where

$$\gamma_A = |\det A|^{-\frac{1}{2}} e^{\frac{\pi i}{4}\operatorname{sgn} A}$$

and $a \in C^\infty(\mathbb{R})$. *Furthermore* a *has the asymptotic expansion*

$$a(\hbar) \sim \left(\exp(-i\frac{\hbar}{2} b(D) f\right)(0).$$

The next step in our program is to use Morse's lemma.

15.6 The formula of stationary phase.

15.6.1 Critical points.

Let M be a smooth compact n-dimensional manifold, and let ψ be a smooth real valued function defined on M. Recall that a point $p \in M$ is called a *critical point* of ψ if $d\psi(p) = 0$. This means that $(X\psi)(p) = 0$ for any vector field X on M, and if X itself vanishes at p then $X\phi$ vanishes at p "to second order" in the sense that $YX\psi$ vanishes at p for any vector field Y. Thus $(YX\psi)(p)$ depends only on the value $X(p)$. Furthermore

$$(XY\psi)(p) - (YX\psi)(p) = ([X,Y]\psi)(p) = 0$$

so we get a well defined symmetric bilinear form on the tangent space TM_p called the **Hessian** of ψ at p and denoted by $d_p^2\psi$. For any pair of tangent vectors $v, w \in TM_p$ it is given by

$$d_p^2\psi(p)(v,w) := (XY\psi)(p)$$

where X and Y are any vector fields with

$$X(p) = v, \quad Y(p) = w.$$

Recall that a critical point p is called **non-degenerate** if this symmetric bilinear form is non-degenerate. We can then talk of the signature of the quadratic form $d_p^2\psi$ – i.e. the number of +'s minus the number of -'s when we write $d_p^2\psi$ in canonical form as a sum of $\pm(x^i)^2$ where the x^i form an appropriate basis of TM_p^*. We will write this signature as $\operatorname{sgn} d_p^2\psi$ or more simply as $\operatorname{sgn}_p \psi$. The symmetric bilinear form $d_p^2\psi$ determines a symmetric bilinear form on all the exterior powers of TM_p, in particular on the highest exterior power, $\wedge^n TM_p$.

This then in turn defines a density at p, assigning to every basis $v_1, \ldots, v_n$ of TM_p the number

$$|d_p^2(\psi)(v_1 \wedge \cdots \wedge v_n, v_1 \wedge \cdots \wedge v_n)|^{\frac{1}{2}}.$$

Replacing $v_1, \ldots, v_n$ by $Av_1, \ldots, Av_n$ has the effect of multiplying the above number by $|\det A|$ which is the defining property of a density. In particular, if we are given some other positive density at p the quotient of these two densities is a number, which we will denote by

$$|\det d_p^2\psi|^{\frac{1}{2}},$$

the second density being understood. The reason for this somewhat perverse notation is as follows: Suppose, as we always can, that we have introduced coordinates $y^1, \ldots, y^n$ at p such that our second density assigns the number one to the the basis

$$v_1 = \left(\frac{\partial}{\partial y^1}\right)_p, \ \ldots \ \ v_n = \left(\frac{\partial}{\partial y^n}\right)_p.$$

Then

$$d_p^2(\psi)(v_1 \wedge \cdots \wedge v_n, v_1 \wedge \cdots \wedge v_n) = \det\left(\frac{\partial^2\psi}{\partial y^i \partial y^j}\right)(p)$$

so

$$|\det d_p^2\psi|^{\frac{1}{2}} = \left|\det\left(\frac{\partial^2\psi}{\partial y^i \partial y^j}\right)(p)\right|^{\frac{1}{2}}.$$

15.6.2 The formula.

With these notations let us first state a preliminary version of the formula of stationary phase. Suppose we are given a positive density, Ω, on M and that all the critical points of ψ are non-degenerate (so that there are only finitely many of them). Then for any smooth function a on M we have

$$\int_M e^{i\tau\psi} a\Omega = \left(\frac{2\pi}{\tau}\right)^{\frac{n}{2}} \sum_{p|d\psi(p)=0} e^{\frac{1}{4}\pi i \operatorname{sgn}_p \psi} \frac{e^{i\tau\psi(p)}a(p)}{|\det d_p^2\psi|^{\frac{1}{2}}} + O(\tau^{-\frac{n}{2}-1}) \tag{15.8}$$

as $\tau \to \infty$.

In fact, we can be more precise. Around every critical point we can introduce coordinates such that the Hessian of ψ is given by a quadratic form. We can also write $\Omega = b(y)dy$ for some smooth function b. We can also pull out the factor $e^{i\tau\psi(p)}$ and set $\tau^{-1} = \hbar$. We may then get the complete asymptotic expansion as given by Theorem 15.5.1.

We will prove the stationary phase formula by a series of reductions. Given any finite cover of M by coordinate neighborhoods, we may apply a partition of unity to reduce our integral to a finite sum of similar integrals, each with the function a supported in one of these neighborhoods.

By partition of unity, our proof of the stationary phase formula thus reduces to estimating integrals over Euclidean space of the form

$$\int e^{i\tau\psi(y)}a(y)dy$$

where a is a smooth function of compact support and where either

1. $d\psi \neq 0$ on supp a so that

$$|d\psi|^2 := \left(\frac{\partial\psi}{\partial y^1}\right)^2 + \cdots + \left(\frac{\partial\psi}{\partial y^n}\right)^2 > \epsilon > 0$$

   on supp a, or

2. ψ is a non-degenerate quadratic form, which, by Sylvester's theorem in linear algebra, we may take to be of the form

$$\psi(y) = \frac{1}{2}\left((y^1)^2 + \cdots (y^k)^2 - (y^{k+1})^2 - \cdots - (y^n)^2\right)$$

   (with, of course, the possibility that $k = 0$ in which case all the signs are negative and $k = n$ in which case all the signs are positive). The number $2k - n$ is the signature of the quadratic form ψ and is what we have denoted by $\mathrm{sgn}(d_0^2\psi)$ in the stationary phase formula.

We treat each of these two cases separately:

The case of no critical points.

In this case we will show that

$$\int e^{i\tau\psi}ady = O(\tau^{-k}) \tag{15.9}$$

for any k.

Consider the vector field

$$X := \frac{\partial\psi}{\partial y^1}\frac{\partial}{\partial y^1} + \cdots + \frac{\partial\psi}{\partial y^n}\frac{\partial}{\partial y^n}.$$

This vector field does not vanish, and in fact

$$X\left(e^{i\tau\psi}\right) = i\tau|d\psi|^2 e^{i\tau\psi}.$$

So we can write

$$\int e^{i\tau\psi}ady = \frac{1}{i\tau}\int X(e^{i\tau\psi})\frac{a}{|d\psi|^2}dy = \frac{1}{\tau}\int e^{i\tau\psi}bdy$$

where

$$b = iX\left(\frac{a}{|d\psi|^2}\right)$$

by integration by parts. Repeating this integration by parts argument proves(15.9). This takes care of the case where there are no critical points.

The case near a critical point.

We assume that p is an isolated critical point, and we have chosen coordinates y about p such that p has coordinates $y = 0$ and that $\psi = \psi(p) + \frac{1}{2}Q(y)$ in these coordinates where $Q(y)$ is a diagonal quadratic form. We now have a single summand on the right of (15.8) and by pulling out the factor $e^{i\tau\psi(p)}$ we may assume that $\psi(p) = 0$. Now apply Theorem 15.5.1. □

15.6.3 The clean version of the stationary phase formula.

Suppose now that the phase function, ψ, on the left hand side of (15.8) is a Bott-Morse function: i.e. satisfies

1. The critical set,
$$C_\psi = \{p \in M\,, d\psi(p) = 0\}$$
is a submanifold of M, and

2. For every $p \in C_\psi$ the quadratic form $d^2\psi_2$ on the normal space N_pC_ψ is non-degenerate.

Then for every connected component, W of C_ψ the restriction of ψ to W has to be constant, and we will denote this constant by γ_W.

Also as explained in §14.6.1 The Hessian, $d^2\psi_2$, gives rise to a density on N_pW. Hence since
$$T_pM = T_pW \oplus N_pW$$
the quotient of the density $\Omega(p)$ by this density is now a density $|\det d^2\psi_p|^{-\frac{1}{2}}\Omega_W(p)$ on T_pW. The clean version of stationary phase asserts that for Bott-Morse functions the integral
$$\int_M e^{i\tau\psi} s\, d\Omega$$
on the left hand side of (15.8) is equal to the sum over the connected components, W of C_ψ of the expressions
$$\left(\frac{2\pi}{\tau}\right)^{\frac{n_W}{2}} \left(e^{\frac{1}{4}\pi i\,\mathrm{sgn}(W)} e^{i\tau\gamma_W} \int_W |\det\, d^2\psi|^{-\frac{1}{2}} a\Omega_W + O(\tau^{-1})\right) \qquad (15.10)$$
where n_W is the codimension of W and $\mathrm{sgn}(W)$ the signature of $d^2\,\psi_p$ at points, $p \in W$.

Remark: As in (15.8) one can replace the $O(\tau^{-1})$ by an asymptotic expansion
$$\tau^{-1}\sum_{i=0}^{\infty} a_{i,W}\tau^{-i}$$
where the $a_{i,W}\tau^{-i}$'s are integrals over W of derivatives of a.

Proof. By localizing we can assume, as above, that $M = \mathbb{R}^n$, that W is defined by the equation $x_{k+1} = x_n = 0$ and that $\Omega = dx_1 \cdots dx_n$. Then by integration by parts

$$\int e^{i\tau\psi} a\Omega = \int dx_{k+1} \dots dx_n \left(\int e^{i\tau\psi} a dx_1 \dots dx_k \right)$$

and (15.10) follows by applying the version of stationary phase proved in §14.5 to the inner integral.

□

We now turn to various applications of the formula of stationary phase.

15.7 Group velocity.

In this section we describe one of the most important applications of stationary phase to physics. Let $\hbar$ be a small number (eventually we will take $\hbar = h/2\pi$ where h is Planck's constant, but for the moment we want to think of $\hbar$ as a parameter which approaches zero, so that $\tau := (1/\hbar) \to \infty$). We want to consider a family of "traveling waves"

$$e^{-(i/\hbar)(E(p)t - p\cdot x)}.$$

For simplicity in exposition we will take p and x to be scalars, but the discussion works as well for x a vector in three (or any) dimensional space and p a vector in the dual space. For each such wave, and for each fixed time t, the wave number of the space variation is h/p. Since we allow E to depend on p, each of these waves will be traveling with a possibly different velocity. Suppose we superimpose a family of such waves, i.e. consider an integral of the form

$$\int a(p) e^{-(i/\hbar)(E(p)t - p\cdot x)} dp. \tag{15.11}$$

Furthermore, let us assume that the function $a(p)$ has its support in some neighborhood of a fixed value, p_0. Stationary phase says that the only non-negligible contributions to the above integral will come from values of p for which the derivative of the exponent with respect to p vanishes, i.e. for which

$$E'(p)t - x = 0.$$

Since $a(p)$ vanishes unless p is close to p_0, this equation is really a constraint on x and t. It says that the integral is essentially zero except for those values of x and t such that

$$x = E'(p_0)t \tag{15.12}$$

holds approximately. In other words, the integral looks like a little blip called a *wavepacket* when thought of as a function of x, and this blip moves with velocity $E'(p_0)$ called the *group velocity*.

Let us examine what kind of function E can be of p if we demand invariance under (the two dimensional version of) all Lorentz transformations, which are all linear transformations preserving the quadratic form $c^2t^2 - x^2$. Since (E, p) lies in the dual space to (t, x), the dual Lorentz transformation sends $(E, p) \mapsto (E', p')$ where

$$E^2 - c^2p^2 = (E')^2 - c^2(p')^2$$

and given any (E, p) and (E', p') satisfying this condition, we can find a Lorentz transformation which sends one into the other. Thus the only invariant relation between E and p is of the form

$$E^2 - (pc)^2 = \text{constant}.$$

Let us call this constant m^2c^4 so that $E^2 - (pc)^2 = m^2c^4$ or

$$E(p) = \left((pc)^2 + m^2c^4\right)^{1/2}.$$

Then

$$E'(p) = \frac{pc^2}{E(p)} = \frac{p}{M}$$

where M is defined by

$$E(p) = Mc^2 \quad \text{or} \quad M = \left(m^2 + \left(\frac{p}{c}\right)^2\right)^{1/2}.$$

Notice that if p/c is small in comparison with m then $M \doteq m$. If we think of M as a *mass*, then the relationship between the group velocity $E'(p)$ and p is precisely the relationship between velocity and momentum in classical mechanics. In this way have associated a wave number $k = p/h$ to the momentum p and if we think of E as energy we have associated the frequency $\nu = E/h$ to energy. We have established the three famous formulas

$$\begin{array}{llllll} E & = & c^2\left(m^2 + \left(\frac{p}{c}\right)^2\right)^{1/2} & \doteq & mc^2 & \text{Einstein's mass energy formula} \\ \lambda & = & \frac{1}{k} & = & \frac{h}{p} & \text{de Broglie's formula} \\ E & = & h\nu & & & \text{Einstein's energy frequency formula.} \end{array}$$

In these formulas we have been thinking of h or $\hbar$ as a small parameter tending to zero. The great discovery of quantum mechanics is that h should not tend to zero but is a fundamental constant of nature known as *Planck's constant*. In the energy frequency formula it occurs as a conversion factor from inverse time to energy, and hence has units energy $\times$ time. It is given by

$$h = 6.626 \times 10^{-34}\text{J s}.$$

15.8 The Fourier inversion formula.

We used the Fourier transform and the Fourier inversion formula to derive the lemma of stationary phase. But if we knew stationary phase then we could derive the fourier inversion formula as follows:

Consider the function $p = p(x, \xi)$ on $\mathbb{R}^n \oplus \mathbb{R}^n$ given by

$$p(x, \xi) = x \cdot (\xi - \eta)$$

where $\eta \in \mathbf{R}^n$. This function has only one critical point, at

$$x = 0, \xi = \eta$$

where its signature is zero. We conclude that for any such function $a = a(x, \xi) \in \mathcal{S}(\mathbb{R}^n \oplus \mathbb{R}^n$ we have

$$\int\int e^{i\tau x\cdot(\xi-\eta)} a(x, \xi) dx d\xi = \left(\frac{2\pi}{\tau}\right)^n a(0, \eta) + O(\tau^{-(n+1)}).$$

Let us choose $a(x, \xi) = f(x)g(\xi)$ where f and g are smooth functions vanishing rapidly with their derivatives at infinity. We get

$$\left(\frac{1}{\tau^n}\right) f(0)g(\eta) = \frac{1}{(2\pi)^n} \int\int e^{i\tau x\cdot(\xi-\eta)} f(x)g(\xi) dx d\xi + O(\tau^{-(n+1)}).$$

Let us set $u = \tau x$ in the integral, so that $dx = \tau^{-n} du$. Multiplying by τ^n we get

$$f(0)g(\eta) = \frac{1}{(2\pi)^n} \int\int f\left(\frac{u}{\tau}\right) g(\xi) e^{iu\cdot(\xi-\eta)} du d\xi + O(\tau^{-1}).$$

So if we define

$$\hat{g}(u) := \frac{1}{(2\pi)^{n/2}} \int g(\xi) e^{i\xi\cdot u} d\xi$$

we have proved that

$$f(0)g(\eta) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}} f\left(\frac{u}{\tau}\right) \hat{g}(u) e^{iu\cdot\eta} du + O(\tau^{-1}).$$

If we choose f such that $f(0) = 1$ and let $\tau \to \infty$ we obtain the Fourier inversion formula:

$$g(\eta) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}} \hat{g}(u) e^{iu\cdot\eta} du.$$

15.9 Fresnel's version of Huygen's principle.

15.9.1 The wave equation in one space dimension.

As a warm up to the study of spherical waves in three dimensions we study the homogeneous wave equation

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0$$

where $u = u(x,t)$ with x and t are real variables.

If we make the change of variables $p = x+t, q = x-t$ this equation becomes

$$\frac{\partial^2 u}{\partial p \partial q} = 0$$

and so by integration

$$u = u_1(p) + u_2(q)$$

where u_1 and u_2 are arbitrary differentiable functions. Reverting to the original coordinates this becomes

$$u(x,t) = u_1(x+t) + u_2(x-t). \tag{15.13}$$

Any such function is clearly a solution. The function $u_2(x-t)$ can be thought of dynamically: At each instant of time t, the graph of $x \mapsto u_2(x-t)$ is given by the graph of $x \mapsto u_2(x)$ displaced t units to the right. We say that $u_2(x-t)$ represents a **traveling wave** moving without distortion to the right with unit speed.

Thus the most general solution of the homogeneous wave equation in one space dimension is given by the superposition of two undistorted traveling wave, one moving to the right and the other moving to the left.

15.9.2 Spherical waves in three dimensions.

In three space dimensions the wave equation (in spherical coordinates) is

$$\frac{\partial^2 u}{\partial t^2} = \frac{1}{r^2}\frac{\partial}{\partial r} r^2 \frac{\partial u}{\partial r} + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\sin\theta \frac{\partial u}{\partial \theta} + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 u}{\partial \phi^2}.$$

If $u = u(r,t)$ the last two terms on the right disappear while

$$\frac{1}{r^2}\frac{\partial}{\partial r} r^2 \frac{\partial u}{\partial r} = \frac{1}{r}\left[2\frac{\partial u}{\partial r} + r\frac{\partial^2 u}{\partial r^2}\right] = \frac{1}{r}\frac{\partial^2 (ru)}{\partial r^2}.$$

Thus $v := ru$ satisfies the wave equation in one space variable, and so the general spherically symmetric solution of the wave equation in three space dimensions is given by

$$u(r,t) = \frac{f(r+t)}{r} + \frac{g(r-t)}{r}.$$

The first term represents and incoming spherical wave and the second term an outgoing spherical wave. In particular, if we take $f = 0$ and $g(s) = e^{iks}$ then

$$w_k(r,t) := \frac{e^{ik(r-t)}}{r}$$

is an outgoing spherical sinusoidal wave of frequency k.

15.9.3 Helmholtz's formula

Recall Green's second formula (a consequence of Stokes' formula) which says that if u and v are smooth functions on a bounded region $V \subset \mathbb{R}^3$ with piecewise smooth boundary ∂V then

$$\int_{\partial V} (u \star dv - v \star du) = \int_V (v\Delta u - u\Delta v) dx \wedge dy \wedge dz.$$

In particular, if u and v are both solutions of the reduced wave equation $\Delta\phi - k^2\phi = 0$ the right hand side vanishes, and we get

$$\int_{\partial V} (u \star dv - v \star du) = 0.$$

Now

$$d\left(\frac{e^{ikr}}{r}\right) = \frac{e^{ikr}}{r}\left[ik - \frac{1}{r}\right] dr. \tag{15.14}$$

Let D be a bounded domain with piecewise smooth boundary, let r_P denote the distance from a point P interior to D, and take V to consist of those points of D exterior to a small sphere about P. Then if v is a solution to the reduced wave equation and we take $u = e^{ikr}/r$ we obtain Helmhotz's formula

$$v(P) = \frac{1}{4\pi}\int_{\partial D}\left[\frac{e^{ikr_P}}{r_P} \star dv - v \star d\frac{e^{ikr_P}}{r_P}\right]$$

by shrinking the small sphere to zero.

Green's formula also implies that if P is exterior to D the integral on the right vanishes.

Now let D consist of all points exterior to a surface S but inside a ball of radius R centered at P. If Σ_R denotes the sphere of radius R centered at P, then the contribution to Helmhotz's formula coming from integrating over σ_R will be the integral over the unit sphere

$$\int e^{ikr}\left[r\left(\frac{\partial v}{\partial r} - iku\right) + v\right]\Big|_{r=R} d\omega$$

where $d\omega$ is the area element of the unit sphere. This contribution will go to zero if the **Sommerfeld radiation conditions**

$$\int |v| d\omega = o(1), \quad \text{and} \qquad \int \left|\frac{\partial v}{\partial r} - iku\right| d\omega = o(R^{-1})$$

are satisfied (where the integrals are evaluated at $r = R$).

Assuming these conditions, we see that if P is exterior to S then

$$v(P) = \frac{1}{4\pi}\int_S\left[\frac{e^{ikr_P}}{r_P} \star dv - v \star d\frac{e^{ikr_P}}{r_P}\right]. \tag{15.15}$$

while the integral vanishes if P is inside S.

Huyghens had the idea that propagated disturbances in wave theory could be represented as the superposition of secondary disturbances along an intermediate surface such as S. But he did not have an adequate explanation as to why there was no “backward wave”, i.e. why the propagation was only in the outward direction. Fresnel believed that if all the original sources of radiation were inside S, the integrand in Helmholtz’s formula would vanish due to interference. The above argument due to Helmholtz was the first rigorous mathematical treatment of the problem, and shows that the internal cancellation is due to the total effect of the boundary.

However, we will see, by using stationary phase, that Fresnel was right up to order $1/k$.

15.9.4 Asymptotic evaluation of Helmholtz’s formula

We will assume that near S the v that enters into (15.15) has the form

$$v = ae^{ik\phi}$$

where a and ϕ are smooth and $\|\operatorname{grad}\phi\| \equiv 1$. For example, if v represent radiation from some point Q interior to S then this would hold with $\phi = r_Q$.

We assume that P is sufficiently far from S so that $1/r_P$ is negligible in comparison with k, and we also assume that a and da are negligible in comparison with k. As P will be held fixed, we will write r for r_P. Then inserting (15.14) into (15.15) shows that the leading term in(15.15) (in powers of k) is

$$\frac{ik}{4\pi}\int_S \frac{a}{r}e^{ik(\phi+r)}(\star d\phi - \star dr).$$

We want to apply stationary phase to this integral. The critical points are those points y on S at which the restriction of $d\phi + dr$ to S vanishes. This says that the projection of $\operatorname{grad}\phi(y)$ onto the tangent space to S at y is the negative of the projection of $\operatorname{grad} r(y)$ onto this tangent space. Since $\|\operatorname{grad}\phi\| = \|\operatorname{grad} r\| = 1$, this implies that the projections of $\operatorname{grad}\phi(y)$ and $\operatorname{grad} r(y)$ onto the normal have the same absolute value. There are thus two possibilities:

1. $\operatorname{grad}\phi(y) = -\operatorname{grad} r(y)$. In this case $\star d\phi(y) = -\star dr(y)$ when restricted to the tangent space to S at y.

2. $\operatorname{grad}\phi(y) = 2(\operatorname{grad}\phi(y), n)n - \operatorname{grad} r(y)$. In this case $\star d\phi(y) = \star dr(y)$ when restricted to the tangent space to S at y.

Let us assume for the moment that the critical points are non-degenerate. (We will discuss this condition below.)

Suppose we are in case 2). Then the leading term in the integral in (15.15) vanishes, and hence the contribution from (15.15) is of order $1/k$. If S wre convex and grad ϕ pointed outward, then for any P insised S we would be in case 2). This justifies Fresnel’s view that there is local cancellation of the backward wave (at least up to terms of order $1/k$).

15.9.5 Fresnel's hypotheses.

Suppose we are in case 1). Then the leading term in (15.15) is

$$\frac{ik}{4\pi}\int_S \frac{a}{r} e^{ik(\phi+r)} \star dr.$$

This shows that up to terms of order $1/k$ the "induced secondary radiation" coming from S behaves as if it

- has amplitude equal to $1/\lambda$ times the amplitude of the primary wave where $\lambda = 2\pi/k$ is the wave length, and
- has phas one quarter of a period ahead of the primary wave. (This is one way of interpreting the factor i.)

Fresnel made these two assumptions directly in his formulation of Huyghen's principle leading many to regard them as *ad hoc*. We see that it is a consequence of Helmholtz's formula and stationary phase.

15.10 The lattice point problem.

Let D be a domain in the plane with piecewise smooth boundary. The high school method of computing the area of D is to superimpose a square grid on the plane and count the number of squares "associated" with D. Since some squares may intersect D but not be contained in D, we must make a choice: let us choose to count all squares which intersect $\overline{D}$. Furthermore, in order to avoid unnecessary notation, let us assume that D is taken to include its boundary, i.e. D is closed: $D = \overline{D}$. If we let $\mathbf{Z}^2$ denote the lattice determined by the corners of our grid, then our procedure is to count the number of points in

$$D \cap \mathbf{Z}^2.$$

Of course this is only an approximation to the area of D. To get better and better approximations we would shrink the size of the grid. Our problem is to find an estimate for the error in this procedure.

For notational reasons, it is convenient to keep the lattice fixed, and dilate the domain D. That is, we want to count the number of lattice points in λD where λ is a (large) positive real number. So we set

$$N_D^\sharp(\lambda) := \#(\lambda D \cap \mathbf{Z}^2). \tag{15.16}$$

Equally well, if χ^D denotes the indicator function (sometimes called the characteristic function) of D:

$$\chi^D(x) = 1 \ \text{ if } x \in D, \qquad \chi^D(x) = 0 \text{ if } x \notin D,$$

then

$$N_D^\sharp(\lambda) = \sum_{\nu\in\mathbf{Z}^2} \chi_\lambda^D(\nu), \tag{15.17}$$

where

$$\chi_\lambda(x) := \chi\left(\frac{x}{\lambda}\right).$$

(Frequently, in what follows, we will drop the D when D is fixed. Also, we will pass from 2 to n with the obvious minor changes in notation.)

Now it is clear that

$$N_D^\sharp(\lambda) = \lambda^2 \cdot \text{Area}(D) + \text{ error.}$$

Our problem is to estimate the error. Without any further assumptions, it is relatively easy to see that we can certainly say that the error can be estimated by a constant times λ where the constant involves only the length of ∂D. In general, we can not do better, especially if the boundary of D contains straight line segments of rational slope: For the worst possible scenario, consider the case where D is a square centered at the origin. Then every time that λ is such that the vertices of λD lie in $\mathbf{Z}^2$, then the number of boundary points lying in $\mathbf{Z}^2$ will be proportional to λ times the length of the perimeter of D. But a slightly larger or small value of λ will yield no boundary points in $\mathbf{Z}^2$. We might expect that if the boundary is curved everywhere, we can improve on the estimate of the error.

The main result of this section, due to Van der Corput [VDC], asserts that if D is convex, with smooth boundary whose curvature is everywhere positive (we will give more precise definitions later) then we can estimate the error terms as being

$$O(\lambda^{\frac{2}{3}}).$$

In fact, Van der Corput shows that this result is sharp if we allow all such strongly convex smooth domains, although we will not establish this result here.

15.10.1 The circle problem.

Suppose that we take D to be the unit disk. In this case

$$N_D^\sharp(\lambda) = N(\lambda)$$

where

$$N(\lambda) = \#\{\nu = (m,n) \in \mathbf{Z}^2 | m^2 + n^2 \leq \lambda^2\}. \tag{15.18}$$

In this case, there will only be lattice points on the boundary of λD if λ^2 is an integer which can be represented as a sum of two squares, and the number of points on the boundary will be the number of ways of representing λ^2 as a sum of two squares.

The number of ways of representing an integer N as the sum of two integer squares is closely related to the number of prime factors of N of the form $4k+1$ and the number of prime square factors of the form $4k+3$. In fact, as we shall remind you later on, if $r(N)$ denotes the number of ways of writing N as a sum of two squares then $r(N)$ can be evaluated as follows: Suppose we factorize N into prime powers, collect all the powers of 2, collect all the primes congruent to 1 (mod 4), and collect all the primes which are congruent to 3 (mod 4). In other words, we write

$$N = 2^f N_1 N_2 \tag{15.19}$$

where

$$N_1 = \prod p^r \quad p \equiv 1 (\mathrm{mod}\ 4)$$

and

$$N_2 = \prod q^s \quad q \equiv 3 (\mathrm{mod}\ 4).$$

Then $r(N) = 0$ if any s is odd. If all the s are even, then

$$r(N) = 4d(N_1). \tag{15.20}$$

So there are relatively few points on the boundary of λD when D is the unit disk, and we might expect special results in this case. Of course our problem is to estimate the number of lattice points close to a given circle, not necessarily exactly on it.

Let us set

$$t := \lambda^2, \tag{15.21}$$

as the square of λ is the parameter used frequently in the number theoretical literature. Let us define $R(t)$ as the error in terms of t, so

$$\sum_{n \le t} r(n) = \pi t + R(t). \tag{15.22}$$

Then the result of Van der Corput cited above asserts that

$$R(t) = O(t^{\frac{1}{3}}). \tag{15.23}$$

In fact, work of Van der Corput himself in the twenties and early thirties, involving the theory of "exponent pairs" improves upon this estimate. For example, one consequence of the method of "exponent pairs" is that

$$R(t) = O(t^{\frac{27}{82}}). \tag{15.24}$$

In fact, the long standing conjecture (going back to Gauss, I believe) has been that

$$R(t) = O(t^{\frac{1}{4}+\epsilon}) \quad \text{for any} \quad \epsilon > 0. \tag{15.25}$$

Notice the sequence of more and more refined results: trivial arguments, valid for any region with piecewise smooth boundary give an estimate $R(t) = O(t^\rho)$ where $\rho = \frac{1}{2}$. The Van der Corput method valid for all smooth strongly convex domains gives $\rho = \frac{1}{3}$. The method of exponent pairs gives $\rho = (k+\ell)/(2k+2)$ whenever (k,ℓ) is an exponent pair, but although this method improved on $\frac{1}{3}$, it did not yield the desired conjecture - that we may take $\rho = \frac{1}{4} + \epsilon$ for any $\epsilon > 0$.

15.10.2 The divisor problem.

Let $d(n)$ denote the number of divisors of the positive integer n. Using elementary arguments, Dirichlet (1849) showed that

$$\sum_{n\leq t} d(n) = t(\log t + 2\gamma - 1) + O(t^{\frac{1}{2}}) \tag{15.26}$$

where γ is Euler's constant

$$\gamma := \lim_{N\to\infty}\left(\sum_{n\leq N}\frac{1}{n} - \log N\right).$$

Dirichlet's argument is as follows: First of all observe that we can regard the divisor problem as a lattice point counting problem. Indeed, consider the region, T_t, in the (x, y) plane bounded by the hyperbola $xy = t$ and the straight line segments from $(1, 1)$ to $(1, t)$ and from $(1, 1)$ to $(t, 1)$. So T_t is a "triangle" with the hypotenuse replaced by a hyperbola. Then $d(n)$ is the number of lattice points on the "integer hyperbola" $xy = n,\ n \leq t$, and so $\sum_{n\leq t} d(n)$ is the total number of lattice points in T_t. The area of T_t is $t\log t - t + 1$, which has the same leading term as above. To count the number of lattice points in T_t, observe that T_t is symmetric about the line $y = x$, and there are $[\sqrt{t}]$ lattice points in T_t on this line. For each integer $d \leq [\sqrt{t}]$ the number of lattice points on the horizontal line) $y = d$ in T_t to the right of the diagonal is

$$\left[\frac{t}{d}\right] - d$$

so

$$\sum_{n\leq t} d(n) = 2\sum_{d\leq\sqrt{t}}\left(\left[\frac{t}{d}\right] - d\right) + \left[\sqrt{t}\right].$$

Since $[s] = s + O(1)$ we can write this as

$$2t\sum_{d\leq\sqrt{t}}\frac{1}{d} - 2\cdot\frac{\sqrt{t}(\sqrt{t}+1)}{2} + O(\sqrt{t}).$$

The formula leading to Euler's constant has error term $1/s$:

$$\sum_{n\leq s}\frac{1}{d} = \log s + \gamma + O(\frac{1}{s}) \tag{15.27}$$

as follows from Euler MacLaurin (see later on). So setting $s = \sqrt{t}$ in the above we get (15.26).

Once again we may ask if this estimate can be improved: Define

$$\Delta(t) := \sum_{n\leq t} d(n) - t(\log t + 2\gamma - 1) \tag{15.28}$$

and ask for better σ such that

$$\Delta(t) = O(t^\sigma) \tag{15.29}$$

It turns out, that the method of exponent pairs yields the same answer as in the circle problem case: If (k,ℓ) is an "exponent pair" then

$$\sigma = (k+\ell)/(2k+2)$$

is a suitable exponent in (15.29). Once again, the conjectured theorem has been that we may take $\sigma = \frac{1}{4} + \epsilon$ for any positive ϵ.

These "lattice point problems" are closely related to studying the growth of the Riemann zeta function on the critical line, i.e. to obtain power estimates for $\zeta(\frac{1}{2}+it)$. Furthermore, the Riemann hypothesis itself is known to be closely related to somewhat deeper "approximation" problems. See, for example, the book *Area, Lattice Points, and Exponential Sums* by M.N Huxley, page 15.

15.10.3 Using stationary phase.

Van der Corput revolutionized the study of the lattice point problem in the 1920's by bringing to bear two classical tools of analysis - the Poisson summation formula and the method of stationary phase.

Our application will be of the following nature: Recall that a subset of $\mathbf{R}^n$ is convex if it is the intersection of all the half spaces containing it. Suppose that D is a (compact) convex domain with smooth boundary, containing the origin and that u is a unit vector. Then the function $y \mapsto u\cdot y$ achieves a maximum m^+ and a minimum m^- on D and the condition that these be taken on at exactly one point each is what is usually meant by saying that D is strictly convex. We want to impose the stronger condition that restriction of the function $y \mapsto u \cdot y$ to the boundary is non-degenerate having only two critical points, the maximum and the minimum, for all unit vectors. This has the following consequence: Let K be a compact subset of $\mathbf{R}^n - \{0\}$ and consider the Fourier transform of the indicator function $\chi = \chi^D$ evaluated at τx for $x \in K$:

$$\hat{\chi}(\tau x) = \int_D e^{i\tau x\cdot y} dy.$$

(For today it will be convenient to use this definition of the Fourier transform so that

$$\hat{\chi}(0) = \text{vol } (D)$$

without the factors of 2π.)

Holding x fixed, we have (as differential forms in y)

$$d\left(e^{i\tau x\cdot y} x^1 dy^2 \wedge \cdots \wedge dy^n\right) = i\tau (x^1)^2 e^{i\tau x\cdot y} dy^1 \wedge \cdots dy^n$$

so

$$e^{i\tau x\cdot y} dy = e^{i\tau x\cdot y} dy^1 \wedge \cdots \wedge dy^n = \frac{1}{i\tau |x|^2} d\left(e^{i\tau x\cdot y}\omega\right)$$

where

$$\omega := x^1 dy^2 \wedge \cdots \wedge dy^n - x^2 dy^1 \wedge dy^3 \cdots \wedge dy^n + \cdots \pm x^n dy^1 \wedge \cdots \wedge dy^{n-1}.$$

By Stokes,

$$\hat{\chi}(\tau x) = \frac{1}{i\tau|x|^2} \int_{\partial D} e^{i\tau x\cdot y}\omega.$$

The integral on the right is $O(\tau^{-\frac{n-1}{2}})$ by stationary phase, and hence

$$\hat{\chi}(\tau x) = O(\tau^{-\frac{n+1}{2}}) \tag{15.30}$$

uniformly for $x \in K$ where K is any compact subset of $\mathbf{R}^n - \{0\}$. As this is the property we will use, we might as well take this as the definition of a **strongly convex region**.

15.10.4 Recalling Poisson summation.

The second theorem from classical analysis that goes into the proof of Van der Corput's theorem is the Poisson summation formula. This says that if f is a smooth function vanishing rapidly with its derivatives at infinity on $\mathbf{R}^n$ then (in the current notation)

$$\sum_{\mu\in\mathbf{Z}^n} \hat{f}(2\pi\mu) = \sum_{\nu\in\mathbf{Z}^n} f(\nu). \tag{15.31}$$

We recall the elementary proof of this fact :
Set

$$h(x) := \sum_{\nu\in\mathbf{Z}^n} f(x+\nu)$$

so that h is a smooth periodic function with period the unit lattice, $\mathbf{Z}^n$. By definition

$$h(0) = \sum_{\nu\in\mathbf{Z}^n} f(\nu).$$

Since h is periodic, we may expand it into a Fourier series

$$h(x) = \sum_{\mu\in\mathbf{Z}^n} c_\mu e^{-2\pi i\mu\cdot x}$$

where

$$c_\mu = \int_0^1 \cdots \int_0^1 h(x) e^{2\pi i\mu\cdot x} dx = \int_0^1 \cdots \int_0^1 \sum_{\nu\in\mathbf{Z}^n} f(x+\nu) e^{2\pi i\mu\cdot x} dx.$$

We may interchange the order of summation and integration and make the change of variables $x + \nu \mapsto x$ to obtain

$$c_\mu = \hat{f}(2\pi\mu).$$

Setting $x = 0$ in the Fourier series

$$h(x) = \sum_{\mu \in \mathbf{Z}^n} \hat{f}(2\pi\mu) e^{-2\pi i \mu \cdot x}$$

gives

$$h(0) = \sum_{\mu \in \mathbf{Z}^n} \hat{f}(2\pi\mu).$$

Equating the two expressions for $h(0)$ is (15.31).

15.11 Van der Corput's theorem.

In n-dimensions this says:

Theorem 15.11.1. *Let D be a strongly convex domain. Then*

$$N_D^\sharp(\lambda) = \lambda^n \operatorname{vol}(D) + O(\lambda^{n-2+\frac{2}{n+1}}) \tag{15.32}$$

Proof. Let $\chi = \chi^D$ be the indicator function of D so that χ_λ defined by

$$\chi_\lambda(y) := \chi(\frac{y}{\lambda})$$

is the indicator (characteristic) function of λD. Thus

$$N^\sharp(\lambda) = \sum_{\nu \in \mathbf{Z}^n} \chi_\lambda(\nu)$$

where we have written $N^\sharp$ for $N_D^\sharp$. The Fourier transform of χ_λ is given in terms of the Fourier transform of χ by

$$\hat{\chi_\lambda}(x) = \lambda^n \hat{\chi}(\lambda x).$$

Furthermore,

$$\hat{\chi}(0) = \operatorname{vol}(D).$$

If we could apply the Poisson summation formula directly to χ_λ then the contribution from 0 would be $\lambda^n \operatorname{vol}(D)$, and we might hope to control the other terms using (15.30). (For example, if we could brutally apply (15.30) to control *all* the remaining terms in the case of the circle, we would be able to estimate the error in the circle problem as $\lambda^{2-3/2} = \lambda^{1/2}$ which is the circle conjecture.) But this will not work directly since χ_λ is not smooth. We must first regularize χ_λ and the clever idea will be to choose this regularization to depend the right way on λ.

So let ρ be a non-negative smooth function on $\mathbf{R}^n$ supported in the unit ball with integral one. Let

$$\rho_\epsilon(y) = \frac{1}{\epsilon^n} \rho\left(\frac{y}{\epsilon}\right)$$

so ρ_ϵ is supported in the ball of radius ϵ and has total integral one. Thus

$$\hat{\rho}_\epsilon(x) = \hat{\rho}(\epsilon x)$$

and

$$\hat{\rho}(0) = 1.$$

Define

$$N^\sharp_\epsilon(\lambda) = \sum_{\nu \in \mathbf{Z}^n} (\chi_\lambda \star \rho_\epsilon)(\nu)$$

where $\star$ denotes convolution. If ν lies a distance greater than ϵ from the boundary of λD, then $(\chi_\lambda \star \rho_\epsilon)(\nu) = \chi_\lambda(\nu)$. Thus

$$N^\sharp_\epsilon(\lambda - C\epsilon) \le N^\sharp(\lambda) \le N^\sharp_\epsilon(\lambda + C\epsilon)$$

where C is some constant depending only on D. Suppose we could prove that $N^\sharp_\epsilon$ satisfies an estimate of the type (15.32). Then we could conclude that

$$(\lambda - C\epsilon)^n \operatorname{vol}(D) + O(\lambda^{n-2+\frac{2}{n+1}}) \le N^\sharp(\lambda) \le (\lambda + C\epsilon)^n + O(\lambda^{n-2+\frac{2}{n+1}}).$$

Suppose we set

$$\epsilon = \lambda^{-1+\frac{2}{n+1}}. \tag{15.33}$$

Then

$$(\lambda \pm C\epsilon)^n = \lambda^n + O(\lambda^{n-2+\frac{2}{n+1}})$$

and we obtain the Van der Corput estimate for $N^\sharp(\lambda)$. So it is enough to prove the analogue of (15.32) with $N^\sharp_\epsilon$ watching out for the dependence on ϵ.

Since $\chi_\lambda \star \rho_\epsilon$ is smooth and of compact support, and since

$$(\chi_\lambda \star \rho_\epsilon)\hat{} = \hat{\chi}_\lambda \cdot \hat{\rho}_\epsilon$$

we may apply the Poisson summation formula to conclude that

$$N^\sharp_\epsilon(\lambda) = \lambda^n \operatorname{vol}(D) + \sum_{\nu \in \mathbf{Z}^n - \{\mathbf{0}\}} \lambda^n \hat{\chi}(2\pi\lambda\nu)\hat{\rho}(2\pi\epsilon\nu)$$

and we must estimate the sum on the right hand side. Now since ρ is of compact support its Fourier transform vanishes faster than any inverse power of $(1+|x|^2)$. So, using (15.30) we can estimate this sum by

$$\lambda^{n-\frac{n+1}{2}} \sum_{\nu \in \mathbf{Z}^n - \{0\}} |\nu|^{-\frac{n+1}{2}} (1 + |\epsilon\nu|^2)^{-K}$$

were K is large, or, what is the same by

$$\lambda^{\frac{n-1}{2}} \int \frac{1}{|x|^{\frac{n+1}{2}}} (1 + |\epsilon x|^2)^{-K} dx$$

where K is large. Making the change of variables $x = \epsilon z$ this becomes

$$\lambda^{\frac{n-1}{2}}\epsilon^{-\frac{n-1}{2}}\int \frac{1}{|z|^{\frac{n+1}{2}}}(1+|z|)^{-K}dz.$$

The integral does not depend on anything, and if we substitute (15.33) for ϵ, the power of λ that we obtain is

$$\frac{n-1}{2}-\frac{n-1}{2}\left(-1+\frac{2}{n+1}\right)=\frac{n-1}{2}+\frac{n-1}{2}-\frac{n+1}{n+1}+\frac{2}{n+1}=n-2+\frac{2}{n+1}$$

proving (15.32). $\square$

Chapter 16

The Weyl Transform.

A fundamental issue lying at the interface of classical and quantum mechanics is to choose a means of associating an operator **H** on a Hilbert space, the "quantum Hamiltonian", to a function H, the "classical Hamiltonian" on phase space. The celebrated Groenwald - van-Hove theorem shows that Dirac's original idea - to associate operators to all functions in such a way that Poisson brackets go over into operator brackets - can not work. Indeed, if the phase space is a symplectic vector space, and if one insists that linear functions are "quantized" in such a way that the Heisenberg commutation relations hold, then these determine how to "quantize" all polynomials of degree two or less (the metaplectic representation) but we can not add any polynomial of higher degree to our collection of functions we wish to "quantize" without running into a violation of Dirac's prescription. The method of "geometric quantization" is to take the Dirac prescription as primary, but apply it to a Lie subalgebra of the algebra of all functions (under Poisson bracket), a subalgebra which will not include all linear functions. For the physicist faced with the problem of finding a quantum model corresponding to a classical approximation given by a Hamiltionian H, this involves finding an appropriate (and sufficiently large) group of symmetries (canonical transformations) whose Lie algebra contains H.

Another approach, suggested by Hermann Weyl is to take the Heisenberg commutation relations as primary, and give up on the Dirac program.

The Weyl transform thus associates to "any" function (or generalized function) on phase space an operator on Hilbert space. To describe its structure, consider the following: If ϱ is a unitary representation of a (locally compact, Hausdorff, toplogical) group G on a Hilbert space $\mathfrak{H}$, and ϕ is a continuous function of compact support on G then we can define

$$\varrho(\phi) := \int_G \varrho(g)\phi(g)dg$$

where dg is Haar measure. This associates an operator $R(\phi)$ to each continuous function of compact support on G in such a way that convolution goes over into operator multiplication: $R(\phi \star \psi) = R(\phi)R(\psi)$.

For the Weyl transform, the group G is the Heisenberg group $V \times \mathbb{R}$ or $V \times (\mathbb{R}/(2\pi\mathbb{Z})$. For a more detailed description of these groups see Section 16.4 below. The Haar measure has the form $\mu \times dt$ where μ is the Liouville measure.

For each non-zero value of $\hbar$ there is a unique (up to equivalence) irreducible representation $\varrho_\hbar$ characterized by the image of the center. This is the Stone - von-Neumann theorem, originally conjectured by Weyl, see below. Let $p_1, \ldots, p_n, q_1, \ldots, q_n$ a symplectic basis of V, so we can write the most general element of V as

$$\xi p + \eta q = \xi^1 p_1 + \cdots \xi^n p_n + \eta^1 q_1 + \cdots \eta^n q_n.$$

Then the Weyl transform is given by

$$W(\phi) = \int_V \varrho(\xi p + \eta q)\tilde{\phi}(\xi, \eta) d\xi d\eta \tag{16.1}$$

where $\tilde{\phi}$ is the Fourier transform of ϕ and we have have suppressed the dependence on $\hbar$. In other words, instead of $\varrho(\phi)$ we have something that looks like $\varrho(\tilde{\phi})$ except that the integral is over V and not over all of G.

Unfortunately, this is not how the Weyl transform is written either in the physics or in the mathematics literature.

16.1 The Weyl transform in the physics literature.

The representation ϱ induces a representation $\dot{\varrho}$ of the Lie algebra $\mathfrak{g}$ of G which can be identified with $V \times \mathbb{R}$. Let $\exp : \mathfrak{g} \to G$ denote the exponential map. So

$$\xi p + \eta q = \exp(\xi P + \eta Q)$$

where $(P, Q) = (p, q)$ but thought of as elements of the Lie algebra $\mathfrak{g}$. Then

$$\varrho(\xi p + \eta q) = \exp(\xi \dot{\varrho}(P) + \eta \dot{\varrho}(Q))$$

where the exponential on the right is the exponential of skew adjoint operators in Hilbert space.

The physicists like self-adjoint operators rather than skew adjoint operators, so set

$$\hat{p} := \frac{1}{i}\dot{\varrho}(P), \qquad \hat{q} := \frac{1}{i}\dot{\varrho}(Q).$$

Then (16.1) can be written as

$$W(\phi) = \int \exp[i(\xi \hat{p} + \eta \hat{q})]\tilde{\phi}(\xi, \eta) d\xi d\eta. \tag{16.2}$$

16.1.1 The Weyl transform and the Weyl ordering.

Let us apply (16.2) to the generalized function q^2p in two dimensions whose Fourier transform is (up to factors of 2π and ± 1 in front of the i depending on convention)

$$\left(i\frac{\partial}{\partial \xi}\right)\left(i\frac{\partial}{\partial \eta}\right)^2 \delta.$$

Then (16.2) with $\hbar = 1$ gives

$$\left(-i\frac{\partial}{\partial \xi}\right)\left(-i\frac{\partial}{\partial \eta}\right)^2 \exp[i(\xi\hat{p} + \eta\hat{q})\Big|_{\xi=0,\eta=0}.$$

Only the cubic term in the expansion of the exponential contributes, and we get

$$W(q^2p) = \frac{1}{3}\left[\hat{q}^2\hat{p} + \hat{q}\hat{p}\hat{q} + \hat{p}\hat{q}^2\right].$$

This (and its generalization to an arbitrary monomial) is a version of the famous Weyl ordering.

In fact, the Weyl ordering in the physics literature is also presented somewhat differently, e.g.

$$W(q^2p) = \frac{1}{4}\left[\hat{q}^2\hat{p} + 2\hat{q}\hat{p}\hat{q} + \hat{p}\hat{q}^2\right].$$

But straightforward manipulations of the commutation relations shows that this definition of $W(q^2p)$ is the same as that given above, and that this is true for arbitrary polynomials in p and q.

16.2 Definition of the semi-classical Weyl transform.

In the mathematical literature, especially in the literature of semi-classical analysis, the Weyl transform is usually defined as follows: Assume (temporarily) that $\sigma \in \mathcal{S}(\mathbb{R}^{2n})$. Define the Weyl transform $\mathrm{Weyl}_{\sigma,\hbar}$ acting on $\mathcal{S}(\mathbb{R}^n)$ by

$$\left(\mathrm{Weyl}_{\sigma,\hbar}\,\phi\right)(x) = \frac{1}{(2\pi\hbar)^n}\int e^{\frac{i}{\hbar}(x-y)\cdot\xi}\sigma\left(\frac{x+y}{2},\xi\right)\phi(y)dyd\xi. \tag{16.3}$$

When $\hbar = 1$ we will sometimes write Weyl_σ instead of $\mathrm{Weyl}_{\sigma,1}$. We will also use various other notations (as found in the literature) for $\mathrm{Weyl}_{\sigma,\hbar}$. We will see below in Section 16.11.2 that this is in fact the same as (16.1), see, in particular, equation (16.25).

16.3 Group algebras and representations.

16.3.1 The group algebra.

If G is a locally compact Hausdorff topological group with a given choice of Haar measure, we define the convolution of two continuous functions of compact support on G by

$$(\phi_1 \star \phi_2)(g) := \int_G \phi_1(u)\phi_2(u^{-1}g)du.$$

If ψ is another continuous function on G we have

$$\int_G (\phi_1 \star \phi_2)(g)\psi(g)dg = \int_{G\times G} \phi_1(u)\phi_2(h)\psi(uh)dudh.$$

This right hand side makes sense if G is a Lie group, ϕ_1 and ϕ_2 are distributions of compact support and ψ is smooth. Also the left hand side makes sense if ϕ_1 and ϕ_2 belong to $L_1(G)$ and ψ is bounded, etc.

16.3.2 Representing the group algebra.

If we have a continuous unitary representation τ of G on a Hilbert space $\mathfrak{H}$, we can define

$$\tau(\phi) := \int_G \phi(g)\tau(g)dg$$

which means that for u and $v \in H$

$$(\tau(\phi)u, v) = \int_G \phi(g)(\tau(g)u, v)dg. \tag{16.4}$$

This integral makes sense if ϕ is continuous and of compact support, or if G is a Lie group, if u is a C^∞ vector in the sense that $\tau(g)u$ is a C^∞ function of g and ϕ is a distribution. In either case we have

$$\tau(\phi_1 \star \phi_2) = \tau(\phi_1)\tau(\phi_2).$$

If the left invariant measure is also invariant under the map $g \mapsto g^{-1}$ and so right invariant, and if we define

$$\phi^*(g) := \overline{\phi(g^{-1})} \tag{16.5}$$

then

$$\tau(\phi^*) = \tau(\phi)^*. \tag{16.6}$$

A group whose Haar measure is both left and right invariant is called **unimodular**.

16.3.3 Application that we have in mind.

We are going to want to apply this construction to the case where G is the Heisenberg group and where $\tau = \rho_\hbar$ is the Schrödinger representation (see Section 16.10) associated with the parameter $\hbar$ (thought of as "Planck's constant"). So we need to make some definitions:

16.4 The Heisenberg algebra and group.

16.4.1 The Heisenberg algebra.

Let V be a symplectic vector space. So V comes equipped with a skew symmetric non-degenerate bilinear form ω. We make

$$\mathbf{h} := V \oplus \mathbb{R}$$

into a Lie algebra by defining

$$[X, Y] := \omega(X, Y)E$$

where $E = 1 \in \mathbb{R}$ and

$$[E, E] = 0 = [E, X] \quad \forall X \in V.$$

The Lie algebra $\mathbf{h}$ is called the **Heisenberg algebra**. It is a nilpotent Lie algebra. In fact, the Lie bracket of any three elements is zero.

16.4.2 The Heisenberg group.

We will let N denote the simply connected Lie group with this Lie algebra. We may identify the $2n+1$ dimensional vector space $V + \mathbb{R}$ with N via the exponential map, and with this identification the multiplication law on N reads

$$\exp(v + tE)\exp(v' + t'E) = \exp\left(v + v' + (t + t' + \frac{1}{2}\omega(v, v'))E\right). \qquad (16.7)$$

Let dv be the Euclidean (Lebesgue) measure on V. Then the measure $dvdt$ is invariant under left and right multiplication. So the group N is unimodular.

It will be useful to record a commutator computation in N: Let $x, y \in V$ Then

$$\exp(-x)(\exp y) = \exp(y - x - \frac{1}{2}\omega(x, y)E)$$

while

$$\exp(y)\exp(-x) = \exp(y - x - \frac{1}{2}\omega(y, x)E)$$

so, since ω is antisymmetric, we get

$$(\exp(-x))(\exp y) = (\exp y)(\exp(-x))\exp(-\omega(x, y)E). \qquad (16.8)$$

16.4.3 Special representations.

Schur's lemma says that if τ is an irreducible (unitary) representation of a group G on a Hilbert space $\mathfrak{H}$ and $T : \mathfrak{H} \to \mathfrak{H}$ is a bounded operator such that

$$T\tau(g) = \tau(g)T \quad \forall\ g \in G$$

then T must be a scalar multiple of the identity.

For the Heisenberg group, this implies that any irreducible unitary representation must send the elements $\exp(tE)$ into scalar multiples of the identity where the scalar has absolute value one. So there are two alternatives, either this scalar is identically one, or not. It turns out that the first case corresponds to certain finite dimensional represetntions. It is the second case that is interesting:

Let $\hbar$ be a non-zero real number. So we are interested in unitary representations of N which have the property that

$$\exp(tE) \mapsto e^{\hbar it}\mathrm{Id}.$$

The **Stone-von-Neumann theorem** asserts that for each non-zero $\hbar$ **there exists a unique such irreducible representation** $\rho_\hbar$ up to unitary equivalence. This theorem was conjectured by Hermann Weyl in the 1920's and proved (independently) by Stone and von-Neumann in the early 1930's.

16.5 The Stone-von-Neumann theorem.

In fact, to be more precise, the theorem asserts that any unitary representation of N such that

$$\exp(tE) \mapsto e^{\hbar it}\mathrm{Id}$$

must be isomorphic to a **multiple** of $\rho_\hbar$ in the following sense:

Let $\mathfrak{H}_1$ and $\mathfrak{H}_2$ be Hilbert spaces. We can form their tensor product as vector spaces, and this tensor product inherits a scalar product determined by

$$(u \otimes v, x \otimes y) = (u, x)(v, y).$$

The completion of this (algebraic) tensor product with respect to this scalar product will be denoted by $\mathfrak{H}_1 \hat{\otimes} \mathfrak{H}_2$ and will be called the (Hilbert space) tensor product of $\mathfrak{H}_1$ and $\mathfrak{H}_2$. If we have a representation τ of a group G on $\mathfrak{H}_1$ we get a representation

$$g \mapsto \tau(g) \otimes \mathrm{Id}_{\mathfrak{H}_2}$$

on $\mathfrak{H}_1 \hat{\otimes} \mathfrak{H}_2$ which we call a multiple of the representation τ.

Theorem 16.5.1. [The Stone-von-Neumann theorem.] *Let $\hbar$ be a non-zero real number. Up to unitary equivalence there exists a unique irreducible unitary representation $\rho_\hbar$ satisfying*

$$\rho_\hbar(e^{itE}) = e^{i\hbar t} Id. \tag{16.9}$$

Any representation such that $\exp(tE) \mapsto e^{i\hbar t}\mathrm{Id}$ *is isomorphic to a multiple of* $\rho_\hbar$.

Here is an outline of the proof: The first step is to explicitly construct a model for the representation $\rho_\hbar$ by the method of induced representations. The second step is to prove that it is irreducible by showing that the image of the group algebra will contain all Hilbert-Schmidt operators. From this the rest of the theorem will follow. We follow the presentation in [**?**].

We will do the first step now and postpone the second step until later in this chapter.

16.6 Constructing $\rho_\hbar$.

Fix $\hbar \neq 0$. If ℓ is a Lagrangian subspace of V, then $\ell \oplus \mathbb{R}$ is an Abelian subalgebra of **h**, and in fact is maximal Abelian. Similarly

$$L := \exp(\ell \oplus \mathbb{R})$$

is a maximal Abelian subgroup of N.

Define the function

$$f = f_\hbar : \quad N \to \mathbb{T}^1$$

(where $\mathbb{T}^1$ is the unit circle) by

$$f(\exp(v + tE)) := e^{i\hbar t}. \tag{16.10}$$

We have

$$f\left((\exp(v + tE))(\exp(v' + t'E))\right) = e^{i\hbar(t+t'+\frac{1}{2}\omega(v,v'))}. \tag{16.11}$$

Therefore

$$f(h_1h_2) = f(h_1)f(h_2)$$

for

$$h_1, \quad h_2 \in L.$$

We say that the restriction of f to L is a **character** of L.

Consider the quotient space

$$N/L$$

which has a natural action of N (via left multiplication). In other words N/L is a homogeneous space for the Heisenberg group N. Let ℓ' be a Lagrangian subspace transverse to ℓ. Every element of N has a unique expression as

$$(\exp y)(\exp(x + sE)) \quad \text{where } y \in \ell' \quad x \in \ell.$$

This allows us to make the identification

$$N/L \sim \ell'$$

and the Euclidean measure dv' on ℓ' then becomes identified with the (unique up to scalar multiple) measure on N/L invariant under N.

Consider the space of continuous functions ϕ on N which satisfy

$$\phi(nh) = f(h)^{-1}\phi(n) \quad \forall\ n \in N \quad h \in L \tag{16.12}$$

and which in addition have the property that the function on N/L

$$n \mapsto |\phi(n)|$$

(which is well defined on N/L on account of (16.12)) is square integrable on N/L. We let $\mathfrak{H}(\ell, \hbar)$ denote the Hilbert space which is the completion of this space of continuous functions relative to this L_2 norm. So $\phi \in \mathfrak{H}(\ell, \hbar)$ is a "function" on N satisfying (16.12) with norm

$$\|\phi\|^2 = \int_{N/L} |\phi|^2 d\dot{n}$$

where $d\dot{n}$ is left invariant measure on N/L.

Define the representation $\rho_{\ell,\hbar}$ of N on $\mathfrak{H}(\ell)$ by left translation:

$$(\rho_{\ell.\hbar}(m)\phi)(n) := \phi\left(m^{-1}n\right). \tag{16.13}$$

This is an example of the standard method of constructing an induced representation from a character of a subgroup.

For the rest of this section we will keep ℓ and $\hbar$ fixed, and so may write $\mathfrak{H}$ for $\mathfrak{H}(\ell, \hbar)$ and ρ for $\rho_{\ell,\hbar}$. Since $\exp tE$ is in the center of N, we have

$$\rho(\exp tE)\phi(n) = \phi\left((\exp -tE)n\right) = \phi\left(n(\exp -tE)\right) = e^{i\hbar t}\phi(n).$$

In other words

$$\rho\left(\exp tE\right) = e^{i\hbar t}\mathrm{Id}_H. \tag{16.14}$$

Suppose we choose a complementary Lagrangian subspace ℓ' and then identify N/L with ℓ' as above. Condition (16.12) becomes

$$\phi\left((\exp y)(\exp(x))(\exp tE)\right) = \phi(\exp y)e^{-i\hbar t}.$$

So $\phi \in \mathfrak{H}$ is completely determined by its restriction to $\exp \ell'$. In other words the map

$$\phi \mapsto \psi, \quad \psi(y) := \phi(\exp y)$$

defines a unitary isomorphism

$$R : H \to L_2(\ell')$$

and if we set

$$\sigma := R\rho R^{-1}$$

then

$$\begin{array}{lclr} [\sigma\,(\exp x)\,\psi](y) & = & e^{i\hbar\omega(x,y)}\psi(y) & x\in\ell,\ y\in\ell' \\ [\sigma(\exp u)\psi](y) & = & \psi(y-u) & y,u\in\ell' \\ \sigma(\exp(tE)) & = & e^{i\hbar t}\mathrm{Id}_{L_2(\ell')}. & \end{array} \tag{16.15}$$

The first of these equations follows from (16.8) and the definition (16.13) and the last two follow immediately from (16.13).

We can regard the three equations of (16.15) as an “integrated version” of the Heisenberg commutation relations.

16.7 The “twisted convolution”.

Let Φ denote the collection of continuous functions on N which satisfy

$$\phi(n\exp tE) = e^{-i\hbar t}\phi(n).$$

Let

$$B = B_\hbar := N/\Gamma_\hbar$$

where

$$\Gamma_\hbar = \{\exp kE,\ k\in(2\pi/\hbar)\mathbb{Z}\|.$$

The effect of replacing N by B is to replace the center of N which is $\mathbb{R}$ with the circle $\mathbb{T}=\mathbb{T}^1_\hbar=\mathbb{R}/(2\pi/\hbar)\mathbb{Z}$.

Every $\phi\in\Phi$ can be considered as a function on B, and every $n\in B$ has a unique expression as $n=(\exp v)(\exp tE)$ with $v\in V$ and $t\in\mathbb{T}$. We take as our left invariant measure on B the measure $dvdt$ where dv is Lebesgue measure on V and dt is the invariant measure on the circle $\mathbb{T}$ with total measure one. The set of elements of Φ are then determined by their restriction to $\exp(V)$. Then for $\phi_1,\phi_2\in\Phi$ of compact support (as functions on B) we have (with $\star$ denoting convolution on B)

$$(\phi_1\star\phi_2)(\exp v)$$

$$\begin{aligned} &= \int_V\int_T \phi_1((\exp u)(\exp tE))\phi_2((-\exp u)(\exp(-tE))(\exp v))dudt \\ &= \int_V \phi_1(\exp u)\phi_2((\exp -u)(\exp v))du \\ &= \int_V \phi_1(\exp u)\phi_2(\exp(v-u)\exp(-\frac{1}{2}\omega(u,v)E))du \\ &= \int_V \phi_1(\exp u)\phi_2(\exp(v-u))e^{\frac{1}{2}i\hbar\omega(u,v)}du. \end{aligned}$$

So if we use the notation

$$\psi(u)=\phi(\exp u)$$

and $\psi_1\star\psi_2$ for the ψ corresponding to $\phi_1\star\phi_2$ we have

$$(\psi_1\star\psi_2)(v) = \int_V \psi_1(u)\psi_2(v-u)e^{\frac{1}{2}i\hbar\omega(u,v)}du. \tag{16.16}$$

We thus get a “twisted” convolution on V.

16.8 The group theoretical Weyl transform.

If $\phi \in \Phi$ and if we define ϕ^* as in (16.5), then $\phi^* \in \Phi$ and the corresponding transformation on the ψ's is

$$\phi^*(\exp v) = \overline{\psi(-v)}.$$

We now define

$$W_\tau(\psi) = \tau(\phi) = \int_B \phi(b)\tau(b)db = \int_V \psi(v)\tau(\exp v)dv.$$

The last equation holds because of the opposite transformation properties of τ and $\phi \in \Phi$.

If $\phi \in \Phi$ then $\delta_m \star \phi$ is given by

$$(\delta_m \star \phi)(n) = \phi(m^{-1}n)$$

which belongs to Φ if ϕ does and if $m = \exp(w)$ then

$$(\delta_m \star \phi)(\exp u) = e^{\pi i \omega(w,u)}\psi(u - w).$$

Similarly,

$$(\phi \star \delta_m)(\exp u) = e^{-\pi i \omega(w,u)}\psi(u - w).$$

Let us write $w \star \psi$ for the function on V corresponding to $\delta_m \star \phi$ under our correspondence between elements of Φ and functions on V.

Then the facts that we have proved such as

$$\tau(\phi_1 \star \phi_2) = \tau(\phi_1)\tau(\phi_2)$$

translate into

$$W_\tau(\psi_1 \star \psi_2) = W_\tau(\psi_1)W_\tau(\psi_2) \tag{16.17}$$
$$W_\tau(\psi^*) = W_\tau(\psi)^* \tag{16.18}$$
$$W_\tau(w \star \psi) = \tau(\exp w)W_\tau(\psi) \tag{16.19}$$
$$W_\tau(\psi \star w) = W_\tau(\psi)\tau(\exp w). \tag{16.20}$$

We now temporarily to leave this group theoretical side of the Weyl transform and turn our original subject which is the semi-classical Weyl transform. For the completion of the proof of the Stone - von-Neumann theorem, the reader can skip ahead to Section 16.16.

16.9 Two two by two matrices.

In studying semi-classical Weyl transform we will be frequently making certain changes of variables, so let us put these up front:

We have

$$\begin{pmatrix} 1 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and both matrices on the left have determinant -1. So if we define the operators T and T^{-1} on $L^2(\mathbb{R}^{2n}) = L^2(\mathbb{R}^n)\hat{\otimes}L^2(\mathbb{R}^n)$ by

$$(TF)(x,y) := F\left(x + \frac{y}{2}, x - \frac{y}{2}\right), \quad (T^{-1}F)(x,y) = F\left(\frac{x+y}{2}, x-y\right)$$

then T and T^{-1} are inverses of one another and are both unitary.

16.10 Schrödinger representations.

Define

$$(R^\hbar)(q,p,t)(f)(x) = e^{i\hbar(q\cdot x + \frac{1}{2}q\cdot p + \frac{1}{4}t)}f(x+p).$$

It is easy to check that this is a representation of the Heisenberg group where the symplectic form on $\mathbb{R}^n \oplus \mathbb{R}^n$ is

$$\omega\left((q,p),(q',p')\right) = 2(q'\cdot p - q\cdot p')$$

and that it is unitary and irreducible. So it is a model for the Stone - von-Nuemann representation with parameter $\hbar/4$.

We will let

$$\varrho_\hbar(q,p) := R^\hbar(q,p,0)$$

and $V_\hbar(f,g)(q,p) = 1/(2\pi)^{n/2}\times$ the matrix element of $\varrho_\hbar$ for $f,g \in L^2(\mathbb{R}^n)$ so

$$V_\hbar(f,g)(q,p) = \frac{1}{(2\pi)^{n/2}}\int e^{i\hbar(q\cdot x + \frac{1}{2}q\cdot p)}f(x+p)\overline{g(x)}dx.$$

Under the change of variables $y = x + \frac{p}{2}$ this becomes

$$V_\hbar(f,g)(q,p) := \frac{1}{(2\pi)^{n/2}}\int e^{i\hbar q\cdot y}f\left(y+\frac{p}{2}\right)\overline{g\left(y-\frac{p}{2}\right)}dy. \tag{16.21}$$

We let $W_\hbar = W_\hbar(x,\xi) = W_\hbar(f,g)(x,\xi)$ denote the Fourier transform of $V_\hbar(f,g)$ (in $2n$ variables) so

$$W_\hbar(x,\xi) = \frac{1}{(2\pi)^{3n/2}}\int\int\int e^{-ix\cdot q - i\xi\cdot p + i\hbar q\cdot y}f\left(y+\frac{p}{2}\right)\overline{g\left(y-\frac{p}{2}\right)}dydqdp.$$

Doing the q integration first (with the usual distributional justification) this gives

$$\int \delta(x-\hbar y)e^{-i\xi\cdot p}f\left(y+\frac{p}{2}\right)\overline{g\left(y-\frac{p}{2}\right)}dydp = \hbar^{-n}\int e^{-ip\cdot\xi}f\left(\frac{x}{\hbar}+\frac{p}{2}\right)\overline{g\left(\frac{x}{\hbar}-\frac{p}{2}\right)}dp.$$

So let $\mathcal{D}_\hbar$ denote the unitary operator on $L^2(\mathbb{R}^n)$

$$(\mathcal{D}_\hbar f)(x) := \hbar^{-n/2} f\left(\frac{x}{\hbar}\right)$$

and set $p = \hbar p'$. Then the above equation gives

$$W_\hbar(f,g)(x,\xi) = W(\mathcal{D}_\hbar f, \mathcal{D}_\hbar g)\left(x, \frac{\xi}{\hbar}\right) \tag{16.22}$$

where we have written W for W_1. So we can work with $\hbar = 1$. We will work with a slightly different "rescaling" later. In any event, we will work for the moment with

$$W(f,g)(x,\xi) = \frac{1}{(2\pi)^{n/2}} \int e^{-i\xi\cdot p} f\left(x + \frac{p}{2}\right) \overline{g\left(x - \frac{p}{2}\right)} dp. \tag{16.23}$$

A direct computation using Plancherel shows that if $f_1, g_1, f_2, g_2 \in \mathcal{S}(\mathbb{R}^n)$ then $W(f_1, g_1)$ and $W(f_2, g_2)$ are in $\mathcal{S}(\mathbb{R}^{2n})$ and

$$(W(f_1,g_1), W(f_2,g_2))_{L^2(\mathbb{R}^{2n}} = (f_1,f_2)_{L^2(\mathbb{R}^n)}\overline{(g_1,g_2)_{L^2(\mathbb{R}^n}}$$

so W extends to a map

$$L^2(\mathbb{R}^n) \times L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^{2n}).$$

16.11 The Weyl transform.

16.11.1 Repeat of the definition of the semi-classical Weyl transform.

Assume (temporarily) that $\sigma \in \mathcal{S}(\mathbb{R}^{2n})$. We defined the Weyl transform $\mathrm{Weyl}_{\sigma,\hbar}$ acting on $\mathcal{S}(\mathbb{R}^n)$ by (16.3):

$$\left(\mathrm{Weyl}_{\sigma,\hbar}\,\phi\right)(x) = \frac{1}{(2\pi\hbar)^n} \int e^{\frac{i}{\hbar}(x-y)\cdot\xi} \sigma\left(\frac{x+y}{2}, \xi\right) \phi(y) dy d\xi.$$

When $\hbar = 1$ we will sometimes write Weyl_σ instead of $\mathrm{Weyl}_{\sigma,1}$. We will also use various other notations (as found in the literature) for $\mathrm{Weyl}_{\sigma,\hbar}$.

16.11.2 Weyl_σ and the Schrödinger representation of the Heisenberg group.

By definition,

$$\left(\mathrm{Weyl}_\sigma(\phi)\right)(x) = \frac{1}{(2\pi)^n} \int\int e^{i(x-y)\cdot\xi} \sigma\left(\frac{x+y}{2}, \xi\right) \phi(y) dy d\xi.$$

We claim that the matrix coefficients of Weyl_σ are given by

$$(\mathrm{Weyl}_\sigma f, g) = \int\int \sigma(x,\xi) W(f,g)(x,\xi) dx d\xi. \tag{16.24}$$

Indeed, the double integral on the right is the triple integral

$$\frac{1}{(2\pi)^{n/2}} \int\int\int \sigma(x,\xi) e^{-i\xi\cdot p} f\left(x+\frac{p}{2}\right) \overline{g\left(x-\frac{p}{2}\right)} dp dx d\xi$$

(where we have interchanged the order of integration). If we set $u = x + \frac{p}{2},\ v = x - \frac{p}{2}$ (see our two by two matrices above) this becomes

$$\frac{1}{(2\pi)^{n/2}} \int\int\int \sigma\left(\frac{u+v}{2}, \xi\right) e^{i(v-u)\cdot\xi} f(u)\overline{g(v)} du d\xi dv$$

proving (16.24).

Since $W(f,g)$ is the Fourier transform of the matrix coefficient of $\varrho(q,p) = \varrho_1(q,p)$ we can use the theorem

$$\int \hat{F} G = \int F \hat{G}$$

(in $2n$ dimensions) to conclude that

$$(\mathrm{Weyl}_\sigma f, g) = \frac{1}{(2\pi)^n} \left(\left(\int \hat{\sigma}(q,p) \varrho(q,p) dq dp \right)(f), g \right). \tag{16.25}$$

In other words, we see that the Weyl transform is the extension to $\mathcal{S}(\mathbb{R}^{2n})$ of the Schrödinger representation applied to the Fourier transform of σ:

$$\mathrm{Weyl}_\sigma = \frac{1}{(2\pi)^n} \int \hat{\sigma}(q,p) \rho_{1/4}(q,p) dq dp. \tag{16.26}$$

We will see by suitable "rescaling" that the Weyl transform $\mathrm{Weyl}_{\sigma,\hbar}$ is associated to the Stone - von-Neumann representation with parameter $\hbar/4$.

Also, we can use the right hand side of (16.24) to define the Weyl transformation of an element of $\mathcal{S}'(\mathbb{R}^{2n})$ as a map from $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}'(\mathbb{R}^n)$: For $f \in \mathcal{S}(\mathbb{R}^n)$ we define $\mathrm{Weyl}_\sigma(f) \in \mathcal{S}'(\mathbb{R}^n)$ by

$$(\mathrm{Weyl}_\sigma(f))(g) = \frac{1}{(2\pi)^{n/2}} \sigma\left(W(f,\overline{g})\right), \quad g \in \mathcal{S}(\mathbb{R}^n).$$

In particular this applies when σ is a symbol. We will want to define various subspaces of $\mathcal{S}'(\mathbb{R}^{2n})$ and describe the properties of the corresponding operators.

16.12 Weyl transforms with symbols in $L^2(\mathbb{R}^{2n})$.

Again we are working with a fixed $\hbar$ and so may assume that $\hbar = 1$. We wish to show that the set of all Weyl transforms with symbols $\sigma \in L^2(\mathbb{R}^{2n})$ coincides with the set of all Hilbert Schmidt operators on $L^2(\mathbb{R}^n)$. For the definition and elementary properties of Silbert-Schmidt operators see Section 16.16 below.

We will let $\mathcal{F}$ denote the Fourier transform on $\mathbb{R}^n$ and $\mathcal{F}_1,\ \mathcal{F}_2$ denote the partial Fourier transforms on $L^2(\mathbb{R}^{2n})$ with respect to the first and second variables so that

$$\mathcal{F}_1(f\otimes g) = (\mathcal{F}(f))\otimes g, \quad \mathcal{F}_2(f\otimes g) = f\otimes(\mathcal{F}(g)).$$

Since the linear combinations of the $f\otimes g$ are dense in $L^2(\mathbb{R}^{2n})$, these equations determine $\mathcal{F}_1$ and $\mathcal{F}_2$. If we go back to the definition of the operator T in Section 16.9 and the definition (16.23) of W we see that

$$W(f,g) = \mathcal{F}_2 T(f\otimes\overline{g}).$$

So if $\sigma\in L^2(\mathbb{R}^{2n})$ then (16.24) says that

$$(\mathrm{Weyl}_\sigma f, g) = \frac{1}{(2\pi)^{n/2}}(W(f,g),\overline{\sigma})_{L^2(\mathbb{R}^{2n}} = \frac{1}{(2\pi)^{n/2}}\left(\mathcal{F}_2 T(f\otimes\overline{g}),\overline{\sigma}\right)_{L^2(\mathbb{R}^{2n}}$$

$$= \frac{1}{(2\pi)^{n/2}}(f\otimes\overline{g}, T^{-1}\mathcal{F}_2^{-1}(\overline{\sigma})_{L^2(\mathbb{R}^{2n}})$$

This shows that W_σ is given by the integral kernel $K_\sigma\in L^2(\mathbb{R}^{2n})$ where

$$K_\sigma(x,y) = \frac{1}{(2\pi)^{n/2}}T^{-1}\mathcal{F}_2\sigma(y,x)$$

and hence is Hilbert-Schmidt. Since all this is reversible, we see that every Hilbert-Schmidt operator comes in this fashion from a Weyl transform.

16.13 Weyl transforms associated to linear symbols and their exponentials.

16.13.1 The Weyl transform associated to ξ^α is $(\hbar D)^\alpha$.

When $\alpha = 0$ this says that

$$u(x) = \frac{1}{(2\pi\hbar)^n}\int\int e^{i\frac{(x-y)\cdot\xi}{\hbar}}u(y)dyd\xi.$$

Under the change of variables $\xi = \hbar\eta$ the right hand side becomes $u(x)$ by the inversion formula for the Fourier transform.

Differentiating under the integral sign then proves the formula stated in the title of this subsection.

16.13.2 The Weyl transform associated to $a = a(x)$ is multiplication by a.

This again follows from the Fourier inversion formula.

16.13.3 The Weyl transform associated to a linear function.

If $\ell = (j,k) \in (\mathbb{R}^n)^* \oplus \mathbb{R}^n = (\mathbb{R}^n \oplus (\mathbb{R}^n)^*)^*$ then combining the two previous results we see that the Weyl transform associated to ℓ is the first order linear differential operator

$$u(x) \mapsto j(x)u(x) + (k(\hbar D)u)(x).$$

We will write this as

$$L = \ell(x, \hbar D)$$

where we are using A to denote the Weyl operator $\mathrm{Weyl}_{a,\hbar}$ associated to a.

Another notation in use (and suggested by the above formulas) is

$$a(x, \hbar D)$$

for $\mathrm{Weyl}_{a,\hbar}$ for a general a.

16.13.4 The composition $L \circ B$.

We want to prove the following formula

$$L \circ B = C$$

where

$$c = \ell b + \frac{\hbar}{2i}\{\ell, b\} \tag{16.27}$$

where $\{a, b\}$ denotes Poisson bracket on $\mathbb{R}^n \oplus (\mathbb{R}^n)^*$:

$$\{a, b\} = \langle \partial_\xi a, \partial_x b\rangle - \langle \partial_x a, \partial_\xi b\rangle$$

For the case that $a = \ell = (j,k)$ is a linear function the Poisson bracket becomes

$$\{\ell, b\} = k(\partial_x b) - j(\partial_\xi b).$$

We will prove (16.27) under the assumption that $b \in \mathcal{S}(\mathbb{R}^{2n})$. It will then follow that it is true for any tempered function on $\mathbb{R}^{2n}$. It suffices to prove (16.27) separately for the cases $k = 0$ and $j = 0$ since the general result follows by linearity.

- $k = 0$. In this case L is the operator of multiplication by the linear function $j = j(x)$ so

$$((L \circ B)u)(x) = \frac{1}{(2\pi\hbar)^n} \int j(x) e^{\frac{i}{\hbar}(x-y)\cdot\xi} b\left(\frac{x+y}{2}.\xi\right) u(y) dy d\xi.$$

Write

$$j(x) = j\left(\frac{x+y}{2}\right) + j\left(\frac{x-y}{2}\right).$$

The first term has the effect of replacing b by ℓb. As to the second term, we have

$$\frac{x-y}{2}e^{\frac{i}{\hbar}(x-y)\cdot\xi} = \frac{\hbar}{2i}\partial_\xi e^{\frac{i}{\hbar}(x-y)\cdot\xi}$$

so integration by parts gives (16.27).

- $j = 0$ so $L = k(\hbar D)$. Differentiation under the integral sign gives (16.27). □

16.14 The one parameter group generated by L.

Let $\ell = (j,k)$ as above and consider the operators $U_\ell(t)$ on $\mathcal{S}(\mathbb{R}^n)$ defined by

$$(U_\ell(t)\psi)(x) := e^{\frac{i}{\hbar}t\langle j,\frac{1}{2}tk-x\rangle}\psi(x-tk).$$

A direct check shows that

$$U_\ell(s+t) = U_\ell(s)\circ U_\ell(t)$$

and

$$i\hbar\frac{d}{dt}U_\ell(t) = L\circ U_\ell(t).$$

So as operators we can write

$$U_\ell(t) = \exp\left(-\frac{i}{\hbar}tL\right).$$

Also, it is clear that the $U_\ell(t)$ are unitary with respect to the L^2 norm on $\mathcal{S}(\mathbb{R}^n)$ and hence extend uniquely to a one parameter group of unitary transformations on $L^2(\mathbb{R}^n)$. By Stone's theorem this shows that L (with domain $\mathcal{S}(\mathbb{R}^n)$) is essentially self adjoint and so extends to a unique self adjoint operator on $L^2(\mathbb{R}^n)$ which we can continue to write as L.

On the other hand, consider the operator associated to the symbol $e^{-\frac{it}{\hbar}\ell}$, call it temporarily $V_\ell(t)$. Then

$$i\hbar\left(\frac{d}{dt}V_t\right)\psi(x) = \frac{1}{(2\pi\hbar)^n}\int\int e^{\frac{i}{\hbar}(x-y)\cdot\xi)}\ell\left(\frac{x+y}{2},\xi\right)e^{-\frac{it}{\hbar}\ell\left(\frac{x+y}{2},\xi\right)}\psi(y)dyd\xi.$$

Since

$$\left\{\ell, e^{\frac{i}{\hbar}\ell}\right\} = 0$$

we see from (16.27) that this is $L\circ V_\ell(t)\psi$ so $V_\ell(t) = U_\ell(t)$.

In other words, the operator associated to $e^{-\frac{i}{\hbar}\ell}$ is $e^{-\frac{i}{\hbar}L}$. Since $L = L(x,\hbar D) = j(x) + k(\hbar D)$ we see from taking $t = 1$ in the definition of $U_\ell(t)$ that

$$e^{-\frac{i}{\hbar}L} = \mu\left(e^{-\frac{i}{2\hbar}\langle j,x\rangle}\right)\circ T_k\circ\mu\left(e^{-\frac{i}{2\hbar}\langle j.x\rangle}\right). \tag{16.28}$$

Here μ denotes the operator of multiplication: $\mu(f)u = fu$ and T denotes the translation operator:

$$T_k u(x) = u(x-k).$$

From this we see that

$$e^{-i\frac{i}{\hbar}L} \circ^{-i\frac{i}{\hbar}M} = e^{\frac{i}{2\hbar}\{\ell,m\}} e^{-i\frac{i}{\hbar}(L+M)}$$

which brings us back to a Schrödinger representation of the Heisenberg group.

Let me here follow (approximately and temporarily) the conventions of Dimassi-Sjöstrand and Evans-Zworsky and define the $\hbar$ Fourier transform of $a \in \mathcal{S}(\mathbb{R}^{2n})$ by

$$\hat{a}_h(\ell) = \int\int e^{-\frac{i}{\hbar}\ell(x,\xi)} a(x,\xi) dx d\xi.$$

Writing $z = (x, \xi)$ this shortens to

$$\hat{a}_\hbar(\ell) = \int e^{-\frac{i}{\hbar}\ell \cdot z} a(z) dz.$$

So the Fourier inversion formula gives

$$a(z) = \frac{1}{(2\pi\hbar)^{2n}} \int e^{\frac{i}{\hbar}\langle \ell, z\rangle} \hat{a}_\hbar \ell d\ell.$$

So we get the Weyl quantization A of a as the superposition

$$A = \frac{1}{(2\pi\hbar)^{2n}} \int \hat{a}_\hbar(\ell) e^{\frac{i}{\hbar}L} d\ell. \tag{16.29}$$

D-S and E-Z write this as

$$a^w(x, \hbar D) = \frac{1}{(2\pi\hbar)^{2n}} \int \hat{a}_\hbar(\ell) e^{\frac{i}{\hbar}\ell(x,\hbar D)} d\ell.$$

Since the $e^{\frac{i}{\hbar}L}$ are unitary, this convergence is also in the operator norm on $L^2(\mathbb{R}^n$ and we conclude that

$$\|A\|_2 \leq \frac{1}{(2\pi\hbar)^{2n}} \|\hat{a}\|_{L^1(\mathbb{R}^{2n})}. \tag{16.30}$$

We shall make some major improvements on this estimate.

16.15 Composition.

The decomposition (16.29) allows us to (once again) get the formula for the composition of two Weyl operators by "twisted convolution":

$$A \circ B = C$$

where

$$\hat{c}_\hbar(r) = \frac{1}{(2\pi\hbar)^{2n}} \int_{\ell+m=r} \hat{a}_\hbar(\ell) \hat{b}_\hbar(m) e^{\frac{i}{\hbar}\{\ell,m\}} d\ell. \tag{16.31}$$

This can also be expressed as follows: Let $z = (x, \xi)$ and similarly z_1, w_1, w_2 denote points of $\mathbb{R}^{2n}$. The claim is that

$$c(z) = \frac{1}{(4\pi\hbar)^{4n}} \int_{\mathbb{R}^{2n}} \int_{\mathbb{R}^{2n}} e^{\frac{i}{\hbar}(\ell(z)+m(z)+\frac{1}{2}\{\ell,m\})} \hat{a}(\ell)\hat{b}(m) d\ell dm. \qquad (16.32)$$

To check that this is so, we need to check that the Fourier transform of the c given by (16.32) is the $\hat{c}$ given in (16.31). Taking the Fourier transform of the c given by (16.32) and interchanging the order of integration gives the following function of r:

$$\frac{1}{(2\pi\hbar)^{2n}} \int\int \left(\frac{1}{(2\pi\hbar)^{2n}} \int e^{\frac{i}{\hbar}(\ell(z)+m(w)-r(z)} dz \right) e^{\frac{i}{\hbar}(\{\ell,m\}} d\ell dm.$$

The inner integral is just $\delta(\ell + m - r)$ giving (16.31) as desired. If we insert the definition of the Fourier transform into (16.32) we get

$$\frac{1}{(2\pi\hbar)^{4n}} \int_{\mathbb{R}^{2n}} \int_{\mathbb{R}^{2n}} \int_{\mathbb{R}^{2n}} \int_{\mathbb{R}^{2n}} e^{\frac{i}{\hbar}(\ell(z-w_1)+m(z-w_2)+\frac{1}{2}\{\ell,m\})} a(w_1)b(w_2) d\ell dm dw_1 dw_2.$$

We will make some changes of variable in this four-fold integral. First set $w_3 = z - w_1, w_4 = z - w_2$ so we get

$$\frac{1}{(2\pi\hbar)^{4n}} \int_{\mathbb{R}^{2n}} \int_{\mathbb{R}^{2n}} \int_{\mathbb{R}^{2n}} \int_{\mathbb{R}^{2n}} e^{\frac{i}{\hbar}(\ell(w_3)+m(w_4)+\frac{1}{2}\omega(\ell,m))} a(z-w_3)b(z-w_4) d\ell dm dw_3 dw_4.$$

Next write the symplectic form in terms of the standard dot product on $\mathbb{R}^{2n}$

$$\omega(\ell, m) = \ell \cdot Jm.$$

So

$$\ell(w_3) + \frac{1}{2}\omega(\ell, m) = \ell \cdot (w_3 + Jm).$$

So doing the integral with respect to ℓ gives

$$(2\pi\hbar)^{2n} \delta_{w_3 + \frac{1}{2}m}.$$

The integral with respect to m becomes

$$(2\pi\hbar)^{2n} \int_{\mathbb{R}^{2n}} e^{\frac{i}{2} m \cdot w_4} \delta_{w_3 + \frac{1}{2}Jm} dm.$$

Make the change of variables $m' = w_3 + \frac{1}{2}Jm$ in the above integral. We get

$$(2\pi\hbar)^{2n} \int_{\mathbb{R}^{2n}} e^{\frac{i}{\hbar}(2J(w_3 - m') \cdot w_4} \delta(m') dm'$$

where now the delta function is at the origin. So this integral becomes

$$(2\pi\hbar)^{2n} 2^{2n} e^{\frac{i}{\hbar} 2Jw_3 \cdot w_4} = (2\pi\hbar)^{2n} 2^{2n} e^{-\frac{2i}{\hbar}\omega(w_3, w_4)}.$$

Putting this back into the four-fold integral above and replacing $3, 4$ by $1, 2$ gives

$$a\sharp b(z) = \frac{1}{(\pi\hbar)^{2n}} \int_{\mathbb{R}^{2n}} \int_{\mathbb{R}^{2n}} e^{-\frac{2i}{\hbar}\omega(w_1,w_2)} a(z-w_1)b(z-w_2)dw_1dw_2, \quad (16.33)$$

where we let $a\sharp b$ denote the c such that $C = A \circ B$.

BACK TO THE STONE-VON-NEUMANN THEOREM.

16.16 Hilbert-Schmidt Operators.

Let $\mathfrak{H}$ be a separable Hilbert space. An operator A on $\mathfrak{H}$ is called **Hilbert-Schmidt** if in terms of some orthonormal basis $\{e_i\}$ we have

$$\sum \|Ae_i\|^2 < \infty.$$

Since

$$Ae_i = \sum (Ae_i, e_j)e_j$$

this is the same as the condition

$$\sum_{ij} |(Ae_i, e_j)|^2 < \infty$$

or

$$\sum |a_{ij}|^2 < \infty$$

where

$$a_{ij} := (Ae_i, e_j)$$

is the matrix of A relative to the orthonormal basis. This condition and sum does not depend on the orthonormal basis and is denoted by

$$\|A\|_{HS}^2.$$

This norm comes from the scalar product

$$(A, B)_{HS} = \operatorname{tr} B^*A = \sum (B^*Ae_i, e_i) = \sum (Ae_i, Be_i).$$

Indeed,

$$\begin{aligned}
(A^*Ae_i, e_i) &= (Ae_i, Ae_i) \\
&= \left(\sum_j (Ae_i, e_j)e_j, Ae_i \right) \\
&= \sum_j (Ae_i, e_j)(e_j, Ae_i) \\
&= \sum_j a_{ij}\overline{a_{ij}} \\
&= \sum_j |a_{ij}|^2,
\end{aligned}$$

and summing over i gives $\|A\|_{HS}^2$.

The rank one elements

$$E_{ij}, \qquad E_{ij}(x) := (x, e_j)e_i$$

form an orthonormal basis of the space of Hilbert-Schmidt operators. We can identify the space of Hilbert-Schmidt operators with the completed tensor product $\mathfrak{H}\hat{\otimes}\overline{\mathfrak{H}}$ where $\overline{\mathfrak{H}}$ is the space $\mathfrak{H}$ with scalar multiplication and product given by the complex conjugate, e.g multiplication by $c \in \mathbf{C}$ is given by multiplication by $\overline{c}$ in $\mathfrak{H}$.

If $\mathfrak{H} = L_2(M, dm)$ where (M, dm) can be any measure space with measure dm, we can describe the space of Hilbert-Schmidt operators as being given by integral operators with L_2 kernels: Indeed, let $\{e_i\}$ be an orthonormal basis of $\mathfrak{H} = L_2(M, dm)$ so that the $e_{ij} \in L_2(M \times M)$

$$e_{ij}(x, y) := e_i(x)\overline{e_j(y)}$$

form an orthonormal basis of $L_2(M \times M)$. Consider the rank one operators E_{ij} introduced above. Then

$$(E_{ij}\psi)(x) = (\psi, e_j)e_i(x) = \int_V \psi(y)\overline{e_j(y)}e_i(x)dy$$

$$= \int_Y K_{ij}(x, y)\psi(y)dy$$

where

$$K_{ij}(x, y) = e_i(x)\overline{e_j(y)}.$$

This has norm one in $L_2(M \times M)$ and hence the most general Hilbert-Schmidt operator A is given by the $L_2(M \times M)$ kernel

$$K = \sum a_{ij}K_{ij}$$

with a_{ij} the matrix of A as above.

16.17 Proof of the irreducibility of $\rho_{\ell,\hbar}$.

We go back to our earlier notation.

Let us consider the case where $\tau = \rho = \rho_{\ell,\hbar}$. We claim that the map W_ρ defined on the elements of Φ of compact support extends to an isomorphism from $L_2(V)$ to the space of all Hilbert-Schmidt operators on $\mathfrak{H}(\ell)$. Indeed, write

$$W_\rho(\psi) = \int_V \psi(v)\rho(\exp v)dV$$

and decompose

$$V = \ell \oplus \ell'$$

$$v = y + x, \quad s \in \ell, \; y \in \ell'$$

so

$$\exp(y+x) = \exp(y)\exp(x)\exp(-\frac{1}{2}\omega(y,x))$$

so

$$\rho(\exp(y+x)) = \rho(y)\rho(x)e^{-i\frac{1}{2}\hbar\omega(y,x)}$$

and hence

$$W_\rho(\psi) = \int\int \psi(y+x)\rho(\exp y)\rho(\exp x)e^{-\frac{1}{2}\hbar i\omega(y,x)}dxdy.$$

So far the above would be true for any τ, not necessarily ρ. Now let us use the explicit realization of ρ as σ on $L_2(\mathbb{R}^n)$ in the form given in (16.15).

We obtain

$$[W_\sigma(\psi)(f)](\xi) = \int\int e^{-\frac{1}{2}i\hbar\omega(y,x)}\psi(y+x)e^{\hbar i\omega(x,\xi-y)}f(\xi-y)dxdy.$$

Making the change of variables $y \mapsto \xi - y$ this becomes

$$\int\int e^{-\frac{1}{2}i\hbar\omega(\xi-y,x)}e^{i\hbar\omega(x,y)}\psi(\xi-y+x)f(y)dxdy.$$

so if we define

$$K_\psi(\xi,y) := \int e^{\frac{1}{2}i\hbar\omega(x,y+\xi)}\psi(\xi-y+x)dx$$

we have

$$[W_\sigma(\psi)f)](\xi) = \int K_\psi(\xi,y)f(y)dy.$$

Here we have identified ℓ' with $\mathbb{R}^n$ and $V = \ell' + \ell$ where ℓ is the dual space of ℓ' under ω. So if we consider the partial Fourier transform

$$\mathcal{F}_x : L_2(\ell' \oplus \ell) \to L_2(\ell' \oplus \ell')$$

$$(\mathcal{F}_x\psi)(y,\xi) = \int e^{-2\pi i\omega(x,\xi)}\psi(y+x)dx$$

(which is an isomorphism) we have

$$K_\psi(\xi,y) = (\mathcal{F}_x\psi)(\xi-y, -\frac{1}{2}(y+\xi)).$$

We thus see that the set of all K_ψ is the set of all Hilbert-Schmidt operators on $L_2(\mathbb{R}^n)$.

Now if a bounded operator C commutes with all Hilbert-Schmidt operators on a Hilbert space, then $CE_{ij} = E_{ij}C$ implies that $c_{ij} = c\delta_{ij}$, i.e. $C = c\mathrm{Id}$. So we have proved that every bounded operator that commutes with all the $\rho_\ell(n)$ must be a constant. Thus $\rho(\ell)$ is irreducible.

16.18 Completion of the proof.

We fix ℓ, ℓ' as above, and have the representation ρ realized as σ on $L_2(\ell')$ which is identified with $L_2(\mathbb{R}^n)$ all as above. We want to prove that any representation τ satisfying (16.14) is isomorphic to a multiple of σ.

We consider the "twisted convolution" (16.16) on the space of Schwartz functions $\mathcal{S}(\mathcal{V})$. If $\psi \in \mathcal{S}(V)$ then its Weyl kernel $K_\psi(\xi, y)$ is a rapidly decreasing function of (ξ, y) and we get all operators with rapidly decreasing kernels as such images of the Weyl transform W_σ sending ψ into the kernel giving $\sigma(\phi)$.

Consider some function $u \in \mathcal{S}(\ell')$ with

$$\|u\|_{L_2(\ell')} = 1.$$

Let P_1 be the projection onto the line through u, so P_1 is given by the kernel

$$p_1(x, y) = \overline{u(y)}u(x).$$

We know that it is given as

$$p_1 = W_\sigma(\psi) \quad \text{for some} \quad \psi \in \mathcal{S}(V).$$

We have $P_1^2 = P_1, P_1^* = P_1$ and

$$P_1\sigma(n)P_1 = \alpha(n)P_1 \quad \text{with} \quad \alpha(n) = (\sigma(n)u, u).$$

Recall that $\phi \mapsto \sigma(\phi)$ takes convolution into multiplication, and that K_ψ is the kernel giving the operator $W_\sigma(\psi) = \sigma(\phi)$ where $\phi \in \Phi$ corresponds to $\psi \in \mathcal{S}(V)$. Then in terms of our twisted convolution $\star$ given by (16.16) the above three equations involving P_1 get translated into

$$\psi \star \psi = \psi, \quad \psi^* = \psi, \quad \psi \star n \star \psi = \alpha(n)\psi. \tag{16.34}$$

Now let τ be any unitary representation of N on a Hilbert space $\mathfrak{H}$ satisfying (16.14). We can form $W_\tau(\psi)$.

Lemma 16.18.1. *The set of linear combinations of the elements*

$$\tau(n)W_\tau(\psi)x, \quad x \in H, \quad n \in N$$

is dense in H.

Proof. Suppose that $y \in H$ is orthogonal to all such elements and set $n = \exp w$. Then for any $x \in H$

$$0 = (y, \tau(n)W_\tau(\psi)\tau(n)^{-1}x) = \int_V (y, \tau(\exp w)\tau(\exp(v)\tau(\exp(-w)\psi(v)dv$$

$$= \int_V (y, \tau(\exp(v + \omega(w, v)E)x)\psi(v)dv = \int_V (y, \tau(\exp v)x)e^{-2\pi i\omega(w,v)}\psi(v)dv$$

$$= \mathcal{F}[(y, \tau(\exp v)x)\psi].$$

The function is square brackets whose Fourier transform is being taken is continuous and rapidly vanishing. Indeed, x and y are fixed elements of H and τ is unitary, so the expression $(y, \tau(\exp v)x)$ is bounded by $\|y\|\|x\|$ and is continuous, and ψ is a rapidly decreasing functions of v. Since the Fourier transform of the function

$$v \mapsto (y, \tau(\exp(v))x)\psi(v)$$

vanishes, the function itself must vanish. Since ψ does not vanish everywhere, there is some value v_0 with $\psi(v_0) \neq 0$, and hence

$$(y, \tau(\exp v_0)x) = 0 \quad \forall x \in H.$$

Writing $x = \tau(\exp v_0)^{-1}z$ we see that y is orthogonal to all of H and hence $y = 0$. QED

Now from the first two equations in (16.34) we see that $W_\tau(\psi)$ is an orthogonal projection onto a subspace, call it $\mathfrak{H}_1$ of $\mathfrak{H}$. We are going to show that $\mathfrak{H}$ is isomorphic to $\mathfrak{H}(\ell) \otimes \mathfrak{H}_1$ as a Hilbert space and as a representation of N.

We wish to define

$$I : \mathfrak{H}(\ell) \otimes \mathfrak{H}_1 \to H, \quad \rho(n)u \otimes b \mapsto \tau(n)b$$

where $b \in \mathfrak{H}_1$.

We first check that if

$$b_1 = W_\tau(\psi)x_1 \quad \text{and} \quad b_2 = W_\tau(\psi)x_2$$

then for any $n_1, n_2 \in N$ we have

$$(\tau(n_1)W_\tau(\psi)x_1, \tau(n_2)W_\tau(\psi)x_2)_{\mathfrak{H}} = (\rho(n_1)u, \rho(n_2)u)_{\mathfrak{H}(\ell)} \cdot (b_1, b_2)_{\mathfrak{H}_1}. \quad (16.35)$$

Proof. Since $\tau(n)$ is unitary and $W_\tau(\psi)$ is self-adjoint, we can write the left hand side of (16.35) as

$$(\tau(n_1)W_\tau(\psi)x_1, \tau(n_2)W_\tau(\psi)x_2)_H = (W_\tau(\psi)\tau(n_2^{-1}n_1)W_\tau(\psi)x_1, x_2)_H$$

and by the last equation in (16.34) this equals

$$= \alpha(n_2^{-1}n_1)(W_\tau(\psi)x_1, x_2)_{\mathfrak{H}}.$$

From the definition of α we have

$$\alpha(n_2^{-1}n_1) = (\rho(n_2^{-1}n_1)u, u)_{\mathfrak{H}(\ell)} = (\rho(n_1)u, \rho(n_2)u)_\ell$$

since $\rho(n_2)$ is unitary. This is the first factor on the right hand side of (16.35).

Since $W_\tau(\psi)$ is a projection we have

$$(W_\tau(\psi)x_1, x_2)_{\mathfrak{H}} = (W_\tau(\psi)x_1, W_\tau(\psi)x_2)_{\mathfrak{H}} = (b_1, b_2)_{\mathfrak{H}_1},$$

which is the second factor on the right hand side of (16.35). We have thus proved (16.35). $\square$

Now define

$$I: \quad \sum_{i=1}^{N} \rho(n_i)u \otimes b_i \mapsto \sum \tau(n_i)b_i.$$

This map is well defined, for if

$$\sum_{i=1}^{N} \rho(n_i)u \otimes b_i = 0$$

then

$$\| \sum_{i=1}^{N} \rho(n_i)u \otimes b_i \|_{H(\ell)\otimes H_1} = 0$$

and (16.35) then implies that

$$\| \sum_{i=1}^{N} \rho(n_i)u \otimes b_i \|_{H(\ell)\otimes H_1} = \| \sum_{i=1}^{N} \tau(n_i)b_i \|_H = 0.$$

Equation (16.35) also implies that the map I is an isometry where defined. Since ρ is irreducible, the elements $\sum_{i=1}^{N} \rho(n_i)u$ are dense in $\mathfrak{H}(\ell)$, and so I extends to an isometry from $\mathfrak{H}(\ell) \otimes \mathfrak{H}_1$ to $\mathfrak{H}$. By Lemma 16.18.1 this map is surjective. Hence I extends to a unitary isomorphism (which clearly is also a morphism of N modules) between $\mathfrak{H}(\ell) \otimes \mathfrak{H}_1$ and $\mathfrak{H}$. This completes the proof of the Stone - von Neumann Theorem.

Bibliography

[AM] Abraham, R. and Marsden, J. *Foundation of mechanics* Perseus Publishing (1978).

[Ben] Benenti,S. and Tulczjew, W.M. "Relazioni lineari binarie simplettiche", *Memorie dell Accademia delle scienze di Torino* **5** (1977-85) pp. 71-140.

[Can] Cannas Da Silva, A., *Lectures on Symplectic Geometry* Springer (2001).

[Col] Colin de Verdière, Y. *Méthodes Semi-Classiques et Théorie Spectral*, in preparation. pdf file on web.

[Davies] Davies, E.B. *Spectral theory and differential operators (Cambridge Studies in Advanced Mathematics)*, Cambridge University Press (1995).

[Did] Robert, D. *Autour de l'approximation semi-classique*, Birkhauser (1987).

[DiSj] Dimassi,M . and Sjostrand, J. *Spectral asymptotics in the semi-classical limit*, Cambridge University Press (1999).

[Fr] Friedrichs, K.O. "Spektraltheorie halbbeschrankter Operatoren und Anwendun auf die Spektralzerlegung von Differential operatoren". *Math. Ann.* Teil 1, **109**: 465-87; Teil 2,**109**: 685-713; Berichtigung, **110**: 777-79. (1934)

[GKRS] Gross, B.Kostant, B, Raymond, P., Sternberg, S. "The Weyl character formula, the half-spin representations, and equal rank subgroups," *Proc. Natl. Acad. Sci. USA* **95** (1998), no. 15, 8441-8442 (electronic).

[GSGA] Victor Guillemin and Shlomo Sternberg. *Geometric Asymptotics* Math. Surveys **14** (1977) American Math. Soc.

[GSIG] Guillemin, V. and Sternberg, S. "Some problems in integral geometry and some related problems in microlocal analysis", *Amer. J. Math.* **101** (1979), pp. 915-955.

[GSSyT] Victor Guillemin and Shlomo Sternberg, *Symplectic techniques in physics* Cambridge University Press (1984).

[GSMMR] Guillemin, V. and Sternberg, S. "The moment map revisited" em J. Diff. Geom. **69** (2005) 137-162.

[GW] Guillemin, V. and Wnag,Z., "Semiclassical spectral invariants for the Schroedinger operator." Journalfof Differential Geometry, to appear.

[Ko] Kostant, B "A cubic Dirac operator and the emergence of Euler number multiplets of representations for equal rank subgroups", *Duke Math. J.* **100** (1999), no. 3, 447501.

[Mart] Martinez, A. *An Introduction to Semiclassical and Microlocal Analysis* Springer (2002).

[PU] Paul, T. and Uribe,A. " Sur la formule semiclassique des traces", C. R. Acad. Sci. Paris, série I **313** (1991) 217-222.

[VDC] van der Corput, J. G. , "Zur Methode der stationaren Phase" I, em Compositio Math. **1** (1935), 15-38; II, ibid.**3** (1936), 328-372.

[Weil] Weil, A. *Introduction l'étude des variétés khlériennes* Publications de l'lnstitutde Mathematique de l'Universite de Nancago VI (Hermann, Paris),(1958).

[Wein81] Weinstein, A. "Symplectic geometry", *Bulletin Amer. Math. Soc.* (new series) **5** (1981), 113.

[Wein82] Weinstein, A. "The symplectic "category"", *Lect. Notes Math.* **905** Springer (1982), 4550.

[Wein11] Weinstein, A. "A note on the Wehrheim-Woodward category" arXiv:1012.0105v2.

[Zwor] Zworski, M. *Semiclassical Analysis* Amer. Math. Soc. (2012)

Index

$I^k(X, \Lambda)$, 180
$I^k(X, \Lambda)$ in terms of a generating function, 180
$N_D^\sharp(\lambda)$, 421
$\Gamma_2 \star \Gamma_1$, 59
$\mathrm{sgn}_p\, \psi$, 411
$d(n)$, 424
$r(N)$, 422

canonical form on the cotangent bundle, 3
canonical relation, 69
canonical two form on the cotangent bundle, 4
functional calculus for Weyl operators, 376
functional calculus on manifolds, 273
infinitesimal generator of a semigroup, 355
Lagrangian subspace, 28
Plancks's constant, 416
second resolvent identity, 348
symbol of an element of $I^k(X, \Lambda)$, 185
Weyl integration theorem , 324

almost analytic extension, 255
almost holomorphic extensions, 383

billiard map, 139
Bohr–Sommerfeld condition, 308

canonical one form on the cotangent bundle, 3
canonical relation associated to a fibration, 103
canonical relation of a map, 78
canonical relation, linear, 59
category, 51
category of sets and relations, 54
caustic, 8
circle conjecture, 423
clean composition of Fourier integral operators, 224
clean composition of canonical relations, 72
clean generating function, 105
clean intersection, 70, 71
closed linear transformations, 344
coisotropic subspace, 27
composition and the sum of generating functions, 190
conormal bundle, 77
consistent Hermitian structures, 28
contravariant functor, 52
covariant functor, 52

Darboux theorem, 36, 38
Darboux-Weinstein theorems, 35
de Broglie's formula, 416
densities of order 1, 156
densities on manifolds, 154
densities, elementary properties of, 147
densities, linear algebra of, 145
densities, pullback and pushforward, 149
density of order α, 146, 147, 154
density of states, 20
diagonal, 54
differential operators on manifolds, on functions, 17
differential operators on manifolds, on half-densities, 19
differential operators on manifolds, on sections of vector bundles, 18
divisor problem, 424
Donnelly's theorem, 294

Dynkin-Helffer-Sjöstrand formula, 256, 351, 378, 383

eikonal, 2
eikonal equation, 2
eikonal equation, local solution of, 8
Einstein's energy frequency formula, 416
enhancing a fibration, 167
enhancing an immersion, 167
enhancing the symplectic "category, 161
envelope, 81
Euler vector field, 110
Euler's constant, 424
Euler's theorem, 110
exact Lagrangian submanifolds, 99
exact square, 60
exact symplectic category, 100
Exact symplectic manifolds, 98
exterior differential caclulus, 388

fiber product, 60
FinRel, 54
first resolvent identity, 346
formal theory of symbols, 250
functional calculus and the spectral theorem, 366
functor, 52

Gaussian integrals, 405
generating function, 46, 105
generating function of a composition, 112
generating function, existence of, 120
generating function, local description, 106
generating function, reduced, 119
geodesically convex, 109
graph of a linear transformation, 343
group velocity, 415
Gutzwiller formula, 287

Hörmander Morse lemma, 126
Hörmander moves, 125
Hörmander-Morse lemma, 125
half-densities, 9
Hamiltonian vector field, 5, 35
Hamiltonian vector fields., 5
Heisenberg algebra, 435
Heisenberg group, 435
Hille Yosida theorem, 362
horizontal Lagrangian submanifold, 7
hyperbolic differential operator, 3
hyperbolicity, 3

indicator function, 421
integral symplectic category, 101, 297
involutive functor, 53
involutory functor, 53
isotropic embedding theorem, 38
isotropic submanifold, 5
isotropic subspace, 27

Kantorovitz's non-commutative Taylor formula, 379
Kantorovitz's non-commutative Taylor's formula, 378
kinetic energy, 109
Kirillov character formula, 326

Lagrangian complements, 29
Lagrangian Grassmannian, 41
Lagrangian submanifolds, 7
Lagrangian subspaces, existence of, 28
Lefschetz symplectic linear transformations, 151
Legendre transform, 124
length spectrum, 286
linear canonical relation, image of, 61
linear canonical relation, kernel of, 61
linear symplectic category, 59
local symbol calculus, 209

mapping torus, 284
Maslov bundle, 132
Maslov cocycle, 133
Maslov line bundle, 45
mass energy formula, 416
microlocality, 205
moment Lagrangian, 84
moment map, 83
moment map, classical, 83
moment map, in general, 84

morphism, 53
Morse lemma, 402
Moser trick, 400
moves on generating functions, 132

natural transformation, 53
normal form for a symplectic vector space, 28

period spectrum of a symplectomorphism 282
Poisson summation, 426
Poisson summation formula, 426
polyhomogeneous pseudo-differential operators, 233
principal series representations, 157
principal symbol, 2, 18
pseudolocality, 241
pull-back of a density under a diffeomorphism, 155
pullback of densities, 150
pushforward of densities, 150
pushforward of Lagrangian submanifolds of the cotangent bundle, 78

quadratic generating functions, 142

recovering the potential well, 266
reduced generating function, 119
reductions are coisotropics., 65
relation, 54
resolvent, 344
resolvent set, 344

Schrödinger operators with magnetic fields, 268
self-adjoint operators , 349
Semi-classical differential operators, 19
semi-classical Fourier integral operators, 183, 184, 275
semi-classical Fourier integral operators, comosition of, 202
semi-classical Fourier integral operators, composition of, 184, 191
semi-classical Fourier integral operators, symbol of, 188
semi-classical pseudo-differential operators, 231
spectral invariants, 262
spectral theorem, 366
spectral theorem, multiplicative version, 368
spectrum, 344
stationary phase, 277, 412, 425
stationary phase, abstract version, 227
Stone's theorem, 354
Stone-von-Neumann theorem, 436
strongly convex, 426
sub-principal symbol, 13
sub-principal symbol of a differential operator on half-densities, 19
superalgebras, 387
support of a density, 155
symbol calculus, 243
symbol calculus, left, 209
symbol calculus, right, 209
symbol calculus, Weyl, 209
symbol, functoriality of, 189
Symmetric operators, 350
symplectic form, 4
symplectic manifold, 4, 35
symplectic reduction, 310
symplectic subspace, 27
symplectic vector field, 35
symplectic vector space, 27
symplectomorphism, 4, 35

total symbol, 12
transport equation, first order , 9
transport equations, 8
transport operator, 17
transport operator, local expression, 14
transport operator, semi-classical, 206
transpose, 12, 58
transpose of a differential operator on half-densities, 19
transverse composition of canonical relations, 73
transverse generating function, 105

van der Corput's theorem, 423, 427
volumes of spheres and balls, 24

wave packet, 415
Weil's formula, 392, 394
Weil's formula, general version, 397, 398
Weyl character formula, 319
Weyl identity, 4
Weyl ordering, 433
Weyl transform, 432, 433
Weyl transform, group theoretical, 440
Weyl transform, semi-classical, 433, 440, 442
Weyl transforms with symbols in L^2, 443
Weyl's law, 19
Weyl's law for the harmonic oscillator, 23